



DISSERTATION

Regular and Singular AR and ARMA Models: The Single and The Mixed Frequency Case

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Doktors der technischen Wissenschaften
unter der Leitung von

Em.o.Univ.-Prof. Dipl.-Ing. Dr.techn. Manfred Deistler

E105

Institut für Wirtschaftsmathematik

eingereicht an der Technischen Universität Wien
bei der Fakultät für Mathematik und Geoinformation

von

Dipl.-Ing. Elisabeth Felsenstein

Matrikelnummer: 0404807

Rabengasse 17/19

1030 Wien

Deutsche Kurzfassung

Diese Dissertation befasst sich mit AR und ARMA Modellen, die zu den wichtigsten Modellklassen der Ökonometrie gehören. Ein Schwerpunkt liegt auf der Analyse singulärer AR und ARMA Modelle. Insbesondere werden Schätzverfahren für die Modellparameter im singulären Fall betrachtet.

Ein zentrales Thema der Dissertation ist der so genannte mixed frequency Fall, d.h. dass die einzelnen Komponenten einer multivariaten Zeitreihe zu unterschiedlichen Abtastraten vorliegen. Die Analyse des mixed frequency Falls beschränkt sich auf Beobachtungen von AR Modellen. Es werden zwei verschiedene Ansätze zur Schätzung der Modellparameter beschrieben, die beide unterschiedliche Schätzer liefern und auch unterschiedliche Einsichten in die Problemstellung gewähren. Beide ermöglichen generisch eine konsistente Schätzung der AR Parameter. Der erste Ansatz beruht auf erweiterten Yule Walker Gleichungen und besticht durch einen simplen Schätzer. Der zweite Ansatz besteht im Blocken der beobachteten Daten. Dadurch werden alle vorhandenen Informationen in den Schätzer einbezogen.

Der letzte Teil der Dissertation befasst sich mit generalisierten linearen dynamischen Faktormodellen, die zur Modellierung hochdimensionaler Zeitreihen eingesetzt werden. Singuläre AR und ARMA Modelle sind von besonderer Bedeutung für diese Modelle, weil sie dort für die Modellierung so genannter statischer Faktoren eingesetzt werden können. Von besonderem Interesse ist hier der mixed frequency Fall, d.h. dass die einzelnen Komponenten der Beobachtungen des Faktormodells zu unterschiedlichen Abtastraten vorliegen. Es kann ein Verfahren angegeben werden, dass es erlaubt, einen mixed frequency statischen Faktor zu bestimmen. Dieser Faktor kann nun mittels eines AR Systems beschrieben werden, dessen Parameter mit den zuvor diskutierten Methoden bestimmt werden können.

Abstract

In this thesis we are concerned with AR and ARMA models which are some of the most important model classes in econometric modeling. We focus on singular AR and ARMA models and consider estimation procedures for the system and noise parameters for the singular case.

A central part of the thesis is concerned with time series whose univariate component series are available at different sampling frequencies. We call this the mixed frequency case. Our analysis is restricted to mixed frequency observations of AR processes. We consider two approaches for showing generic identifiability of the AR parameters. These two approaches give different insights into the problem and lead to different estimation procedures. The first approach is based on extended Yule Walker equations and yields a simple estimator. The second approach is based on blocking all observed data and thus the corresponding estimator exploits all available information.

In the last part of the thesis we consider generalized linear dynamic factor models which are used to model high dimensional time series. Singular AR and ARMA models are of particular importance for these factor models as they can be used to model so-called static factors. We focus on the case that the univariate components of the observations of the factor model are mixed frequency. We propose a procedure for consistently estimating a mixed frequency static factor. This static factor can then be modeled by an AR system. The system and noise parameters of this AR system can be determined with methods developed in the thesis.

Acknowledgment

First and foremost, I want to express my sincere gratitude to Prof. Manfred Deistler for giving me the great opportunity to be a project assistant and to work on my doctoral thesis. I want to thank him for challenging me, for always having an open ear for all of my questions, and for giving me the time to grow. I learned so much from our discussions which I could not have learned from books.

I also want to thank Prof. Brian D.O. Anderson for so many enlightening discussions during our research visits at the ANU Canberra and his visits in Vienna. The collaboration between our research groups broadened my horizon. I was not only given the opportunity to view our research from a different angle but I was also allowed to experience a different academic environment.

Finally, I want to thank my colleagues Dipl.-Ing. Lukas Kölbl and Mag. Alexander Braumann for all their support throughout the years, for many entertaining office days, and for not getting impatient with me and my recurring questions on programming.

Contents

Deutsche Kurzfassung	i
Abstract	iii
Acknowledgment	v
Introduction	1
Part 1. Single Frequency	3
Chapter 1. Regular and Singular AR and ARMA Models	5
1.1. AR Systems and Solutions	5
1.2. ARMA Systems	14
1.3. Rational Spectra and Transfer Functions	16
1.4. Yule Walker Equations	20
1.5. State Space Representations and Subspace Procedures for ARMA Processes	25
Part 2. Mixed Frequency	45
Introduction - Mixed Frequency AR Models	47
Chapter 2. Identifiability Results Using Extended Yule Walker Equations	51
2.1. Derivation of the Extended Yule Walker Equations for Mixed Frequency Data	51
2.2. Generic Identifiability of System Parameters	53
2.3. Generic Identifiability of the Noise Parameters	54
2.4. General Remarks	55
2.5. Flow Variables and More General Aggregation Schemes	56
2.6. Generic Identifiability for Prescribed Column Degrees	57
Chapter 3. Identifiability Results Using Blocking	61
3.1. Generic Identifiability for AR(1) Systems	61
3.2. Generic Identifiability for AR(p) Systems	66
3.3. Generic Identifiability for Prescribed Column Degrees	79

3.4. Generic Identifiability for AR(p) Systems with General Sampling Frequency N	87
Part 3. GDFMs Single and Mixed Frequency	93
Chapter 4. Generalized Linear Dynamic Factor Models (GDFMs)	95
4.1. GDFMs: Definitions and Assumptions	96
4.2. Minimal Static Factors of the Latent Variables	98
4.3. Modeling the Minimal Static Factors	101
Chapter 5. Mixed Frequency GDFMs	105
5.1. Mixed Frequency Static Factors	105
5.2. Simulation Example	108
Bibliography	111
Curriculum Vitae	115

Introduction

Modeling high dimensional time series and mixed frequency data are important topics in current econometric modeling. In this thesis, we deal with these two topics and we highlight the interface of both.

This thesis consists of three parts. Firstly, we consider AR and ARMA systems with particular emphasis on singular AR and ARMA systems. These models are important as models for static factors of generalized linear dynamic factor models which are widely used for modeling high dimensional time series. A focus here is identifiability of the parameters from covariance data and the corresponding estimation procedures. For regular AR system, we shortly consider the well-known Yule Walker equations. For singular AR systems, we consider an approach also based on Yule Walker equations described in Deistler et al. [2011]. For ARMA processes, we want to determine the parameters of a state space model which also determine the ARMA parameters. For the regular case, we discuss the subspace procedure presented in Deistler et al. [1995]. Subsequently, we adapt the procedure of Deistler et al. [1995] for the singular case.

In the second part, which is in a certain sense the central part of the thesis, we are concerned with mixed frequency observations of an AR process: We assume that part of the observations are available at every time point and part of the observations are available at every N th time point, $N \in \mathbb{N}$. The main focus is on identifiability of the parameters of the AR system generating all outputs at the highest frequency. We consider two approaches which both enable us to show generic identifiability of the AR parameters, but will give different insights into the identifiability problem.

The first approach, based on extended Yule Walker equations (see Chen and Zadrozny [1998]) is presented in Chapter 2. The second approach is based on the technique of blocking is presented in Chapter 3. Chapter 2 and most of Chapter 3 is joint work with the coauthors Manfred Deistler, Brian D. O. Anderson, Bernd Funovits, Lukas Kölbl and Mohsen Zamani and has been taken from Anderson et al. [2014]. In addition, in this thesis, we extend some results of the paper mentioned above, namely results concerning the blocking approach in Sections

3.2, 3.3, and 3.4 to more general cases as considered in Anderson et al. [2014]. Furthermore we give an alternative proof of Theorem 8 in Anderson et al. [2014].

In the third part, we are concerned with modeling high dimensional time series. We concisely outline GDFMs following the seminal papers Forni et al. [2000], Forni and Lippi [2001], Forni et al. [2004]. Finally, we assume that we have mixed frequency observations of our GDFM. Applying the ideas of Hallin and Liska [2007] to our mixed frequency setting we propose a procedure for consistently estimating the minimal static factor on a mixed frequency level. This static factor can then be modeled using the techniques developed in the second part of the thesis.

Part 1

Single Frequency

CHAPTER 1

Regular and Singular AR and ARMA Models

In the first part of this thesis we are considering AR and ARMA models. These models are some of the most important model classes in econometrics and have been extensively treated in respective literature, see e.g. Hannan [1970], Brockwell and Davis [1987], Reinsel [1993], Hannan and Deistler [2012], Deistler [2007].

We especially dedicate our attention to singular AR and ARMA systems which are not as thoroughly explored as regular AR and ARMA systems. We are particularly interested in singular AR and ARMA models as they can be used as models for the so-called static factor in generalized linear dynamic factor models. Further fields of applications for singular AR and ARMA models are definitional equations and dynamic stochastic general equilibrium models.

A main focus of this thesis is the problem of *identifiability*, i.e. question if the (real-valued) parameters of the AR or ARMA model can be uniquely determined from population second moments which can be obtained from observed variables. However, if we are not able to show identifiability we can often show identifiability on a generic subset of the parameter space where a subset of the parameter space Θ is called *generic* if it contains a subset that is open and dense in Θ .

1.1. AR Systems and Solutions

In the following we are going to use z both for the backward shift on \mathbb{Z} , i.e. $z(y_t)_{t \in \mathbb{Z}} = (y_{t-1})_{t \in \mathbb{Z}}$, as well as for a complex variable, $z \in \mathbb{Z}$. It will become clear from the context how z is used.

DEFINITION 1.1.1. A vector *autoregressive process* is a (wide sense) stationary solution $(y_t)_{t \in \mathbb{Z}}$ of an *AR(p) system*

$$(1.1.1) \quad y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p} + b \varepsilon_t$$

with $a_j \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^{n \times q}$ and where $a(z) = I - a_1 z - \cdots - a_p z^p$ fulfills the stability assumption $\det a(z) \neq 0$, $|z| \leq 1$ and where ε_t is white noise with variance covariance matrix $\mathbb{E} \varepsilon_t \varepsilon_t' = I_q$ and b has full column rank.

If $\Sigma = bb' > 0$, we say that the AR system is *regular*, whereas for $\text{rk}\Sigma = \text{rk}b = q < n$ we call the AR system *singular*.

In this definition we restrict ourselves to $a_0 = I$, but it is no restriction of generality to introduce this normalization, since the set of solutions is not changed by premultiplying the AR equation by a constant non-singular matrix.

Note that a $n \times q$ matrix b with $q < n$ is commonly called tall.

Our interest in singular AR system was sparked by the occurrence of such models as models for the so-called latent variables in generalized linear dynamic factor models which will be introduced in Chapter 4.

Equation (1.1.1) can be written as

$$(1.1.2) \quad a(z)y_t = b\varepsilon_t.$$

We will call (a_1, \dots, a_p) the system parameters and b the noise parameters of the AR system. The transfer function from $(\varepsilon_t)_{t \in \mathbb{Z}}$ to $(y_t)_{t \in \mathbb{Z}}$ can be written as $k(z) = a^{-1}(z)b$. Note that there can be many different AR systems corresponding to the same transfer function. The next definition will be helpful in many situations to determine a particular AR system such that $k(z) = a^{-1}(z)b$:

DEFINITION 1.1.2. The polynomial matrix $(a(z), b(z))$ is called *left coprime* if all polynomial matrices $u(z)$ satisfying

$$(a(z), b(z)) = u(z) (\hat{a}(z), \hat{b}(z))$$

where $\hat{a}(z), \hat{b}(z)$ are polynomial too, are unimodular, i.e. $\det u(z) = \text{const.} \neq 0$.

As can be shown, left coprimeness is equivalent to $(a(z), b(z))$ being of full row rank $\forall z \in \mathbb{C}$, see e.g. Hannan and Deistler [2012] Lemma 2.2.1. For an AR system, left coprimeness means that $(a(z), b)$ has to be of full row rank which is always fulfilled for regular AR systems. For a singular AR system, it can be shown that we can always find a coprime AR realization of $k(z)$, see e.g. Anderson et al. [2012a] Theorem 1.

System (1.1.2) can be written in companion form as a state space system:

$$(1.1.3) \quad \underbrace{\begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}}_{x_{t+1}} = \underbrace{\begin{pmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p \\ I_n & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \end{pmatrix}}_{\mathcal{A}} \underbrace{\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}}_{x_t} + \underbrace{\begin{pmatrix} b \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{\mathcal{B}} \varepsilon_t$$

$$y_t = \underbrace{\begin{pmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p \end{pmatrix}}_{\mathcal{C}} x_t + \underbrace{b}_{\mathcal{D}} \varepsilon_t$$

The vector x_t is called the state of (1.1.3). The transfer function from $(\varepsilon_t)_{t \in \mathbb{Z}}$ to $(y_t)_{t \in \mathbb{Z}}$ can be written in the parameters of the state space system as $k(z) = (\mathcal{C}(I - \mathcal{A}z)^{-1}\mathcal{B}z + \mathcal{D})\varepsilon_t$.

The corresponding (discrete time) Lyapunov equation is

$$(1.1.4) \quad \Gamma_p = \mathcal{A}\Gamma_p\mathcal{A}' + \mathcal{B}\mathcal{B}'$$

where $\Gamma_p = \mathbb{E}x_t x_t'$.

If a_p is non-singular, the degree of $\det a(z)$ is np and all eigenvalues λ_j of \mathcal{A} are the inverse of the zeros of $\det a(z)$. If a_p is singular of rank $m < n$, the degree of $\det a(z)$ is less than np . We say that $\det a(z)$ has an infinite zero. Clearly, \mathcal{A} has eigenvalue 0 of multiplicity $n - m$. Thus under the stability assumption holds $|\lambda_j| < 1$.

As is easily seen, $\Gamma_p = \sum_{j=0}^{\infty} \mathcal{A}^j \mathcal{B} \mathcal{B}' (\mathcal{A}^j)'$ is a solution of (1.1.4) and under the stability assumption the solution is unique for given $(\mathcal{A}, \mathcal{B})$. (If there were any other solution, the difference of both solutions would fulfill $\Gamma^D = \mathcal{A}\Gamma^D\mathcal{A}'$ which cannot hold since for all eigenvalues λ_j of \mathcal{A} $|\lambda_j| < 1$, see e.g. Inouye [1983].)

The pair $(\mathcal{A}, \mathcal{B})$ is called *controllable* if

$$\text{rk}(\mathcal{B}, \mathcal{A}\mathcal{B}, \dots, \mathcal{A}^{np-1}\mathcal{B}) = np$$

and a pair $(\mathcal{C}, \mathcal{A})$ is called *observable* if

$$\text{rk} \begin{pmatrix} \mathcal{C} \\ \mathcal{C}\mathcal{A} \\ \vdots \\ \mathcal{C}\mathcal{A}^{np-1} \end{pmatrix} = np.$$

A system is called *minimal* if the state dimension for x_t in (1.1.3) is minimal among all systems with given transfer function $k(z)$. The system (1.1.3) is minimal if and only if it is controllable and observable, see e.g. Hannan and Deistler [2012], Chapter 2.

For showing observability or controllability, the PBH (Popov Belevitch Hautus) test proves very helpful, see e.g. Kailath [1980] p 135:

THEOREM 1.1.3 (PBH Test). • A pair $(\mathcal{A}, \mathcal{B})$ is controllable if and only if there is no left eigenvector of \mathcal{A} that is orthogonal to \mathcal{B} .
• A pair $(\mathcal{C}, \mathcal{A})$ is observable if and only if there is no right eigenvector of \mathcal{A} that is orthogonal to \mathcal{C} .

LEMMA 1.1.4. $(\mathcal{C}, \mathcal{A})$ in (1.1.3) is observable if and only if a_p is non-singular.

PROOF. If a_p is singular, there exists an eigenvector $c = \begin{pmatrix} 0 & \cdots & 0 & c'_p \end{pmatrix}'$ of \mathcal{A} corresponding to the eigenvalue $\lambda = 0$ which is obviously in the right kernel of \mathcal{C} and thus $(\mathcal{C}, \mathcal{A})$ is not observable.

Note that if a_p is non-singular observability of $(\mathcal{C}, \mathcal{A})$ is equivalent to observability of $\left(\begin{pmatrix} I_n & 0 & \cdots & 0 \end{pmatrix}, \mathcal{A} \right)$. If a_p is non-singular, any eigenvector $c = \begin{pmatrix} c'_1 & \cdots & c'_p \end{pmatrix}' \neq 0$ of \mathcal{A} corresponding to the eigenvalue $\lambda \neq 0$ has to fulfill

$$\begin{aligned} \sum_{i=1}^p a_i c_i &= \lambda c_1 \\ c_1 &= \lambda c_2 \\ &\vdots \\ c_{p-1} &= \lambda c_p \end{aligned}$$

see e.g. Anderson et al. [2012b] Lemma 2. Thus $c_1 \neq 0$ and obviously c is not in the right kernel of $\begin{pmatrix} I_n & 0 & \cdots & 0 \end{pmatrix}$ and thus $\left(\begin{pmatrix} I_n & 0 & \cdots & 0 \end{pmatrix}, \mathcal{A} \right)$ and equivalently $(\mathcal{C}, \mathcal{A})$ is observable. \square

LEMMA 1.1.5. Controllability of $(\mathcal{A}, \mathcal{B})$ is equivalent to $\Gamma_p > 0$.

PROOF.

$$\Gamma_p = \sum_{j=0}^{\infty} \mathcal{A}^j \mathcal{B} \mathcal{B}' (\mathcal{A}^j)' = \begin{pmatrix} \mathcal{B} & \mathcal{A}\mathcal{B} & \mathcal{A}^2\mathcal{B} & \cdots \end{pmatrix} \begin{pmatrix} \mathcal{B}' \\ \mathcal{B}'\mathcal{A}' \\ \mathcal{B}'(\mathcal{A}^2)'\ \\ \vdots \end{pmatrix}$$

Using the Cayley-Hamilton Theorem it is easy to see that $(\mathcal{A}, \mathcal{B})$ is controllable if and only if Γ_p is non-singular. \square

The next lemma relates controllability of the pair $(\mathcal{A}, \mathcal{B})$ and the AR system $(a(z), b)$.

LEMMA 1.1.6. *Controllability of $(\mathcal{A}, \mathcal{B})$ is equivalent to left coprimeness of $(a(z), b)$ and (a_p, b) being of full row rank.*

PROOF. The PBH Test tells us that controllability of $(\mathcal{A}, \mathcal{B})$ is equivalent to $(\lambda I - \mathcal{A}, \mathcal{B})$ being of full row rank for all eigenvalues λ of \mathcal{A} . As is easily seen, see e.g. Anderson et al. [2012b] Lemma 1, α is a left eigenvectors of \mathcal{A} if and only if

$$\begin{aligned} \alpha'_1 a_1 + \alpha'_2 &= \lambda \alpha'_1 \\ \alpha'_1 a_2 + \alpha'_3 &= \lambda \alpha'_2 \\ &\vdots \\ \alpha'_1 a_{p-1} + \alpha'_p &= \lambda \alpha'_{p-1} \\ \alpha'_1 a_p &= \lambda \alpha'_p. \end{aligned}$$

Thus $(\lambda I - \mathcal{A}, \mathcal{B})$ being of full row rank is equivalent to $(\lambda^p I - \lambda^{p-1} a_1 - \dots - a_p, b)$ being of full row rank. For eigenvalues $\lambda \neq 0$, this is equivalent to $(a(\lambda^{-1}), b)$ being of full row rank which is - together with $a(0) = I$ - equivalent to left coprimeness of $(a(z), b)$. For the eigenvalue $\lambda = 0$, $(\lambda^p I - \lambda^{p-1} a_1 - \dots - a_p, b)$ being of full row rank is equivalent to (a_p, b) being of full row rank. \square

Thus we can formulate the meaning of minimality of (1.1.3) in terms of the parameters of the underlying AR system:

LEMMA 1.1.7. *The state space system with parameter matrices $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is minimal if and only if $(a(z), b)$ is left coprime and a_p is non-singular.*

Observability is generic in Θ as for an open and dense subset of Θ a_p is non-singular. Obviously, controllability is always fulfilled for regular AR systems. However, for singular AR systems controllability is still generic in Θ , see Wonham [1985] p 44.

A main focus of this thesis is the recovery of the AR parameters using second moments of observed data. We explicitly define the parameter space although it has already been defined implicitly. Note that we assume n , p and q to be given.

DEFINITION 1.1.8. The *parameter space* of the AR system (1.1.2) is the set

$$(1.1.5) \quad \Theta = \{(a_1, \dots, a_p) \mid \det a(z) \neq 0, |z| < 1\} \times \{\Sigma \mid \Sigma = \Sigma^T, \Sigma \geq 0, \text{rk}\Sigma = q\}.$$

We assume that there are no cross-restrictions between system and noise parameters. Note that we choose this particular definition of the parameter space since without any further

restrictions b can only be determined up to postmultiplication by an orthogonal constant matrix. To specify a unique choice, we can consider LQ decompositions and assume that b is a quasi lower triangular matrix, see Filler [2010], Proposition 3.1.2. Nevertheless, we will use both (a_1, \dots, a_p, b) and $(a_1, \dots, a_p, \Sigma)$ for the system and noise parameters of the AR system.

Also note that the corresponding set of system parameters

$$S = \{(A_1, \dots, A_p) \mid \det(a(z)) \neq 0, |z| \leq 1\}$$

is open in $\mathbb{R}^{n \times np}$.

Given the parameters of an AR(p) system and a white noise process ε_t , $\Sigma_\varepsilon = I_q$, we can ask for the set of solutions $(y_t)_{t \in \mathbb{Z}}$ fulfilling the AR equation.

It is well known, see e.g. Hannan [1970], p 14, that the set of solutions of a system of linear vector difference equations $a(z)y_t = b\varepsilon_t$ is

$$(y_t^p)_{t \in \mathbb{Z}} + \left\{ (y_t^h)_{t \in \mathbb{Z}} \mid a(z)y_t^h = 0 \right\}$$

where $(y_t^p)_{t \in \mathbb{Z}}$ is a particular solution solving $a(z)y_t^p = b\varepsilon_t$ and $\left\{ (y_t^h)_{t \in \mathbb{Z}} \mid a(z)y_t^h = 0 \right\}$ is the set of all homogenous solutions. The latter set is a linear vector space and the homogeneous solutions of linear vector difference equations are of the form

$$(1.1.6) \quad y_t^h = \sum_{j=0}^m \left(\sum_{k=0}^{n_j-1} c_k t^k \right) z_j^{-t} v_j$$

where m is the number of distinct zeros z_j of $\det a(z)$ and n_j is the multiplicity of the zero z_j , v_j are elements of the kernel of $a(z_j)$ and c_k are arbitrary coefficients.

We are interested in stationary solutions only.

For finding a particular solution, we use the so-called z -transform, see e.g. Deistler [1975]. Using z now as a complex variable, the stability assumption on $a(z)$ guarantees that $a^{-1}(z)$ can be expanded as a power series on a closed disk containing the unit circle. Thus we have the causal transfer function

$$k(z) = a^{-1}(z)b = \sum_{j=0}^{\infty} k_j z^j.$$

It is easy to see that there is an isomorphism with respect to multiplication between power series in the backward shift z and in the complex variable z respectively. Note that $\|k_j\| < c\rho_0^{-j}$, $|\rho_0| > 1$ holds where $|\rho_0| < |z_i|$, $\forall z_i$ such that $\det a(z_i) = 0$ and where $\|\cdot\|$ denotes the Euclidean norm. Thus $\sum_{j=0}^{\infty} \|k_j\| < \infty$.

Therefore $k(z)$ is stable, i.e. for bounded inputs the output is bounded. Moreover, for stationary inputs, the output is stationary. A stationary particular solution thus is the steady state solution starting in the infinite past

$$(1.1.7) \quad y_t^p = \sum_{j=0}^{\infty} k_j \varepsilon_{t-j}$$

This solution is called steady state and in particular only depends on inputs. For every homogeneous solution, the stability assumption on $a(z)$ implies $\lim_{t \rightarrow \infty} y_t^h = 0$, which can easily be seen from (1.1.6). Hence the only stationary homogeneous solution is 0 and the steady state solution (1.1.7) is the unique stationary solution of (1.1.2).

In some situations it might make sense to look at AR systems where the stability assumption does not hold.

If the stability assumption does not hold, for singular AR systems, the transfer function $k(z) = a^{-1}(z)b$ might still have a convergent causal power series expansion and $y_t^p = k(z)\varepsilon_t$ might still be a stationary particular solution of $a(z)y_t = b\varepsilon_t$. Note that this can only occur for non left coprime systems $(a(z), b)$.

We quote the following lemma and its proof from Deistler et al. [2011] (Proposition 2).

LEMMA 1.1.9. *The AR equation $a(z)y_t = b\varepsilon_t$ with $\det a(z) \neq 0$, $|z| < 1$ and $rk b = q < n$ has a stationary solution if and only if there is a $\tilde{a}(z)$ such that $\det \tilde{a}(z) \neq 0$, $|z| \leq 1$ where $a^{-1}b = \tilde{a}^{-1}b$ and $(\tilde{a}(z), b)$ is left coprime.*

PROOF. If $\det \tilde{a}(z) \neq 0$, $|z| \leq 1$, then there exists the stationary solution $y_t^p = a^{-1}(z)b\varepsilon_t$. On the other hand, if $a(z)y_t = b\varepsilon_t$ has a stationary solution, y_t say, then the state equation of (1.1.3) must be fulfilled and thus also the Lyapunov equation (1.1.4) holds for this solution. If \mathcal{A} is not stable, i.e. it has at least one eigenvalue λ such that $|\lambda| \geq 1$, then for a corresponding eigenvector there must hold

$$(1 - |\lambda|^2) x' \Gamma_p x = x' \mathcal{B} \mathcal{B}' x \geq 0.$$

Thus for $|\lambda| \geq 1$ there has to hold $x' \Gamma_p x = 0$ and therefore $x' \mathcal{B} \mathcal{B}' x = 0$ which implies that $x' \mathcal{B} = 0$ for $|\lambda| \geq 1$. Subsequently, $(I - \mathcal{A} \lambda^{-1}, \mathcal{B})$ is not of full rank or equivalently $(a(\lambda^{-1}), b)$ is not of full rank and thus not left coprime. Let $u(z)$ be a (polynomial) greatest common left divisor of $(a(z), b)$ such that $(a(z), b) = u(z)(\bar{a}(z), \bar{b}(z))$, with $(\bar{a}(z), \bar{b}(z))$ left coprime, then $\bar{b}(z)$ is of full column rank for all $z \in \mathbb{C}$ since $a^{-1}(z)b = \bar{a}^{-1}(z)\bar{b}(z)$. Thus we find a left coprime pair $(\tilde{a}(z), b)$ (see Deistler et al. [2010]) with $a^{-1}(z)b = \bar{a}^{-1}(z)\bar{b}(z) = \tilde{a}^{-1}(z)b$. \square

The lemma establishes that $y_t^p = k(z)\varepsilon_t = \tilde{a}^{-1}(z)b = \sum_{j=0}^{\infty} k_j\varepsilon_{t-j}$ is a particular solution of $a(z)y_t = b\varepsilon_t$. But unlike in the case of strict stability of $a(z)$ there can be non-trivial stationary homogeneous solutions.

If we do not impose strict stability, but instead allow for zeros on the unit circle, i.e. $\det a(z) \neq 0, |z| < 1$, there are non-trivial stationary homogeneous solutions which belong to a special class of stationary processes, compare Felsenstein [2010], Section 2.4:

LEMMA 1.1.10. *Stationary (real valued) homogeneous solutions of (1.1.2) can only exist if $\det a(z)$ has zeros on the unit circle. The solutions then are harmonic processes of the form*

$$(1.1.8) \quad y_t^h = \sum_{j=0}^m e^{-i\lambda_j t} v_j$$

where the frequencies λ_j are determined by $\det a(e^{i\lambda_j t}) = 0$, $\lambda_j \in [-\pi, \pi]$, and the weights v_j have to fulfill $a(e^{i\lambda_j t})v_j = 0$ (and for y_t^h to be real valued $-\lambda_{1+j} = \lambda_{h-j}$ and $v_{1+j} = \bar{v}_{h-j}$).

PROOF. Recall that homogeneous solutions of linear vector difference equations are of the form

$$y_t^h = \sum_{j=0}^m \left(\sum_{k=0}^{n_j-1} c_k t^k \right) z_j^{-t} v_j$$

where m is the number of distinct zeros z_j of $\det a(z)$ and n_j is the multiplicity of the zero z_j , v_j are elements of the kernel of $a(z_j)$ and c_k are arbitrary coefficients, see e.g. Hannan [1970], p 14.

As is easily seen, for $|z_j| \neq 1$ or $c_k \neq 0, k > 0$ $\mathbb{E} (y_t^h)' y_t^h$ would not be bounded and $(y_t^h)_{t \in \mathbb{Z}}$ would therefore not be stationary. So possible stationary homogeneous solutions take the form

$$y_t^h = \sum_{j=0}^m e^{-i\lambda_j t} v_j$$

with $v_j : \Omega \rightarrow \mathbb{C}^n$, $\mathbb{E} v_j = \begin{cases} 0 & \lambda_j \neq 0 \\ \mathbb{E} y_t^h & \lambda_j = 0 \end{cases}$, $\mathbb{E} v_j^* v_j < \infty$ and $\mathbb{E} v_j v_l^* = 0, j \neq l$ and are real valued

random variables if $-\lambda_{1+j} = \lambda_{h-j}$ and $v_{1+j} = \bar{v}_{h-j}$. Thus y_t^h has to be a harmonic process which has to fulfill

$$a(z)y_t^h = \sum_{j=0}^m (e^{-i\lambda_j t} v_j - a_1 e^{-i\lambda_j(t-1)} v_j - \dots - a_p e^{-i\lambda_j(t-p)} v_j) = \sum_{j=0}^m e^{-i\lambda_j t} a(e^{i\lambda_j t}) v_j = 0.$$

This implies that $a(e^{i\lambda_j t})v_j = 0$ because of the linear independence of $e^{i\lambda}$, $\lambda \in [-\pi, \pi]$ and thus we have $\det a(e^{i\lambda_j t}) = 0$, $\lambda_j \in [-\pi, \pi]$ and $a(e^{i\lambda_j t})v_j = 0, \forall j$. \square

Note that m is determined by the number of (distinct) zeros on the unit circles of $\det a(z)$.

If we have $y_{t-1}^h, \dots, y_{t-m}^h$ and the zeros on the unit circles of $\det a(z)$, we can determine the weights v_j in 1.1.8, compare Deistler et al. [2011]:

$$\begin{pmatrix} y_{t-1}^h \\ \vdots \\ y_{t-m}^h \end{pmatrix} = \underbrace{\begin{pmatrix} e^{-i\lambda_1(t-1)}I & \dots & e^{-i\lambda_m(t-1)}I \\ \vdots & & \vdots \\ e^{-i\lambda_1(t-m)}I & \dots & e^{-i\lambda_m(t-m)}I \end{pmatrix}}_{V'} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

where V' is the transpose of a generalized Vandermonde matrix and as such non-singular. If we have the covariances $\gamma^h(0), \dots, \gamma^h(m-1)$ of the homogeneous solution, we can determine the variances F_1, \dots, F_m of the random weights v_1, \dots, v_m

$$\begin{pmatrix} \gamma^h(0) \\ \vdots \\ \gamma^h(m-1) \end{pmatrix} = \underbrace{\begin{pmatrix} e^{-i\lambda_1(t-1)}I & \dots & e^{-i\lambda_m(t-1)}I \\ \vdots & & \vdots \\ e^{-i\lambda_1(t-m)}I & \dots & e^{-i\lambda_m(t-m)}I \end{pmatrix}}_{V'} \begin{pmatrix} F_1 \\ \vdots \\ F_m \end{pmatrix}.$$

Above we described a particular solution $(y_t^p)_{t \in \mathbb{Z}}$ and for the singular case with zeros of $\det a(z)$ on the unit circle the set of homogeneous solutions.

For $(y_t)_{t \in \mathbb{Z}} = (y_t^p + y_t^h)_{t \in \mathbb{Z}}$ to be stationary, we have to postulate that $\mathbb{E}\varepsilon_t v'_j = 0, \forall j = 1, \dots, m, \forall t \in \mathbb{Z}$. It is easy to see that this condition is both necessary and sufficient, see e.g. Felsenstein [2010]. Under this initial conditions y_t^p and y_t^h themselves are orthogonal.

THEOREM 1.1.11. *The decomposition $y_t = y_t^p + y_t^h = a^{-1}(z)b\varepsilon_t + \sum_{j=0}^m e^{-i\lambda_j t} v_j$ corresponds to the Wold decomposition of y_t where $a^{-1}(z)b\varepsilon_t$ is the linearly regular part and $\sum_{j=0}^m e^{-i\lambda_j t} v_j$ is the linearly singular part.*

PROOF. The process $(y_t^p)_{t \in \mathbb{Z}}$ is a linearly regular process. Let \mathcal{L}_2 denote the Hilbert space of all scalar square integrable random variables over an underlying probability space $(\Omega, \mathcal{A}, \mathcal{P})$. Let $\mathbb{H}_x(t) \subseteq \mathcal{L}_2$ be the Hilbert space spanned by the components of $x_s, s \leq t$. Then (1.1.7) implies that $\mathbb{H}_{y^p}(t) \subseteq \mathbb{H}_\varepsilon(t)$ and the AR equation (1.1.2) implies that $\mathbb{H}_{y^p}(t) \supseteq \mathbb{H}_\varepsilon(t)$. Thus $\mathbb{H}_{y^p}(t) = \mathbb{H}_\varepsilon(t)$. Thus we have that (1.1.7) is the Wold representation of $(y_t^p)_{t \in \mathbb{Z}}$.

On the other hand, since y_t^h is a harmonic process it is linearly singular and it is orthogonal to y_t^p .

Finally, since $\mathbb{H}_{y^p}(t) = \mathbb{H}_\varepsilon(t) \subseteq \mathbb{H}_y(t)$, where the inclusion trivially holds because of the AR equation (1.1.2), and $\mathbb{H}_{y^h}(t) \subseteq \mathbb{H}_y(t)$ because of $\underbrace{y_t - y_t^p}_{\in \mathbb{H}_y(t)} = y_t^h$ the decomposition $y_t = y_t^p + y_t^h$ coincides with the Wold decomposition. \square

Note that we will only consider linearly regular causal solutions of the AR equation obtained via the transfer function, i.e. the steady state solutions.

For stationary y_t ,

$$y_t = \underbrace{a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p}}_{y_{t|t-1, \dots, t-p}} + b \varepsilon_t$$

$y_{t|t-1, \dots, t-p}$ is the orthogonal projection of y_t on its finite past y_{t-1}, \dots, y_{t-p} , a_1, \dots, a_p are the coefficients of the projection and $b \varepsilon_t$ is the approximation (one-step-ahead prediction) error.

THEOREM 1.1.12. *A regular stationary process $(y_t)_{t \in \mathbb{Z}}$ is an AR process if and only if the best linear least squares projection onto the finite past $\overline{\text{span}} \left\{ y_{t-1}^{(1)}, \dots, y_{t-1}^{(n)}, \dots, y_{t-p}^{(1)}, \dots, y_{t-p}^{(n)} \right\}$, $p > 0$, is the same as the projection on the infinite past $\mathbb{H}_y(t-1)$.*

PROOF. \Leftarrow If the best linear least squares projection of y_t onto the infinite past y_{t-1}, \dots uses only a finite number of regressors with nonzero coefficients, the process is AR. Stability can be proven as follows: (y_t) has a transfer function of the form $a^{-1}(z)b$ where $a(z)$ is a polynomial matrix and bb' with b of full column rank is the variance covariance matrix of the regression residuals. Similarly to the proof of Lemma 1.1.9, we argue that if $\det a(z)$ had a zero $|z| \leq 1$, then there would be a polynomial matrix $\tilde{a}(z)$, $\det \tilde{a}(z) \neq 0$, $|z| \leq 1$, such that $\tilde{a}^{-1}(z)b = a^{-1}(z)b$ and therefore (y_t) is an AR process with stable $\tilde{a}(z)$.

\Rightarrow For an AR process, the projection onto the finite past is the best linear prediction from the infinite past. This can be easily seen since ε_t is orthogonal to $\mathbb{H}_y(t-1)$ and thus not only to y_{t-1}, \dots, y_{t-p} , but also to the infinite past y_{t-1}, \dots and therefore $y_{t|t-1, \dots, t-p} = y_{t|t-1}$, where $y_{t|t-1}$ denotes the projection on the infinite past. \square

REMARK 1.1.13. In the following we will restrict ourselves to the case that $\gamma(0) > 0$. If $\gamma(0)$ were singular, a case which can only occur for singular AR systems, y_t would contain linearly dependent components and we could select all linearly independent components of y_t , Sy_t . The random vector y_t could be recovered from Sy_t by a linear transformation T , $y_t = TSy_t$. The selected process Sy_t would still be an AR process which can easily be seen by considering that the Hilbert space spanned by the components of y_s , $t-p \leq s \leq t-1$ is the same as the Hilbert space spanned by Sy_s , $t-p \leq s \leq t-1$.

1.2. ARMA Systems

In this section we introduce the class of ARMA systems which includes the class of AR systems as a proper subset.

In order to define ARMA systems we first give a mathematically proper definition of poles and zeros of a rational matrix via the so-called Smith McMillan form, e.g. Hannan and Deistler [2012]:

LEMMA 1.2.1 (Smith McMillan form). *Every $n \times r$ rational matrix of rank q can be written as $k = u\Omega v$ where u and v are unimodular matrices and Ω is diagonal of the form*

$$\Omega(z) = \left(\begin{array}{ccc|c} \omega_1(z) & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & \omega_q(z) & 0 \\ \hline 0 & \dots & 0 & 0 \end{array} \right)$$

with $\omega_i = \phi_i \psi_i^{-1}$, where ϕ_i and ψ_i are relatively prime monic polynomials, ϕ_i divides ϕ_{i+1} and ψ_{i+1} divides ψ_i , $i = 1, \dots, q-1$.

Ω is then called Smith McMillan form of k . Note that the Smith McMillan form is unique for k .

DEFINITION 1.2.2. Let k be a rational matrix. $\lambda \in \mathbb{C}$ is called a

- zero of k if there is a numerator polynomial ϕ_i in its Smith McMillan form such that $\phi_i(\lambda) = 0$.
- pole of k if there is a denominator polynomial ψ_i in its Smith McMillan form such that $\psi_i(\lambda) = 0$.

As can be easily seen, a value λ can be both a zero and a pole of the rational matrix k . A rational matrix is called zeroless if it has no zeros. In this case the numerator polynomials ϕ_i in its Smith McMillan form are all equal to one.

DEFINITION 1.2.3. A vector *autoregressive moving average process* is a n -dimensional (wide sense) stationary solution $(y_t)_{t \in \mathbb{Z}}$ of an *ARMA*(p, v) *system*

$$(1.2.1) \quad y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + b_0 \varepsilon_t + \dots + b_v \varepsilon_{t-rv}$$

with $a_j \in \mathbb{R}^{n \times n}$, $b_j \in \mathbb{R}^{n \times q}$ where b_0 has full column rank q . The stability assumption, i.e. $\det a(z) \neq 0$, $|z| \leq 1$, for $a(z) = I - a_1 z - \dots - a_p z^p$, holds as well as the strict miniphase assumption, i.e. $b(z) = b_0 - \dots - b_v z^v$ has no zeros for $|z| \leq 1$. The process $(\varepsilon_t)_{t \in \mathbb{Z}}$ is white noise with variance covariance matrix $\mathbb{E} \varepsilon_t \varepsilon_t' = I_q$.

If $\Sigma = b_0 b_0' > 0$, we say that the ARMA system is regular, whereas for $\text{rk} \Sigma = \text{rk} b_0 = q < n$ we call the ARMA system singular.

Equivalently, (1.2.1) can be written as

$$(1.2.2) \quad a(z)y_t = b(z)\varepsilon_t$$

Obviously, the class of ARMA(p, v) systems contains as a subset the class of AR(p) systems.

In this thesis, we are mainly interested in singular ARMA systems which will appear as models for the so-called static factor in Chapter 4. As we will show, generically we can model this factor as an AR process, but for the sake of completeness we will analyze a procedure for modeling singular ARMA processes in Section 1.5 where we estimate a state space model for the ARMA process.

Note that for ARMA systems we only consider linearly regular stationary solutions $(y_t)_{t \in \mathbb{Z}}$ which exist because of the stability assumption:

$$y_t = a^{-1}(z)b(z)\varepsilon_t = \sum_{j=0}^{\infty} k_j \varepsilon_{t-j}.$$

The strict miniphase assumption enables us to find a causal left inverse of $b(z)$, $b^-(z)b(z) = I_q$, such that

$$\varepsilon_t = b^-(z)a(z)y_t = \sum_{j=0}^{\infty} h_j y_{t-j}.$$

Therefore $(\varepsilon_t)_{t \in \mathbb{Z}}$ are the innovations and we have $\mathbb{H}_y(t) = \mathbb{H}_\varepsilon(t)$. Thus this is the Wold representation.

REMARK 1.2.4. For an ARMA process $(y_t)_{t \in \mathbb{Z}}$ which is not an AR process, the best linear least squares prediction is the orthogonal projection of y_t onto its infinite past, $\mathbb{H}_y(t-1)$, the projection onto finite past gives an approximation where the variance of the residuals is larger.

1.3. Rational Spectra and Transfer Functions

1.3.1. Rational Spectra - ARMA. Every linearly regular process $(y_t)_{t \in \mathbb{Z}}$ with Wold decomposition $y_t = \sum_{j=0}^{\infty} k_j \varepsilon_{t-j} = k(z)\varepsilon_t$, $\sum_{j=0}^{\infty} \|k_j\|^2 < \infty$, has a spectral density

$$f_y(\lambda) = \frac{1}{2\pi} k(e^{-i\lambda})k^*(e^{-i\lambda})$$

where $*$ denotes transposition and complex conjugation. Thus the spectral density of a linearly regular process can be easily obtained via its transfer function. The converse problem, which is finding a transfer function from a spectral density of a linearly regular process, is the so-called spectral factorization see e.g. Rozanov [1967].

It is well known that $f_y(\lambda)$ is rational if and only if $(y_t)_{t \in \mathbb{Z}}$ is an ARMA process, see e.g. Hannan and Deistler [2012].

The spectral density of an ARMA process is

$$f_y(\lambda) = \frac{1}{2\pi} a^{-1}(e^{-i\lambda}) b(z) b^*(z) a^{-*}(e^{-i\lambda}).$$

Obviously, the continuation of the spectral density f_y on \mathbb{C} ,

$$f_y(z) = \frac{1}{2\pi} k(z) k^*(z) = \frac{1}{2\pi} a^{-1}(z) b(z) b^*(z) a^{-*}(z),$$

is singular for singular ARMA systems and non-singular for regular ARMA systems. The normal rank of $f_y(z)$ is q .

Conversely, by spectral factorization of a rational spectral density we obtain a rational transfer function from the spectral density, see e.g. Rozanov [1967], Chapter I.10 or Hannan [1970]:

THEOREM 1.3.1 (Spectral Factorization). *Every $n \times n$ rational spectral density $f(\lambda)$, $\lambda \in [-\pi, \pi]$, of rank q a.e. which has no poles in $[-\pi, \pi]$ can be factorized as*

$$f(\lambda) = \frac{1}{2\pi} k(e^{-i\lambda}) k^*(e^{-i\lambda})$$

where $k(z)$, $z \in \mathbb{C}$ is a $n \times q$ real rational matrix of full normal column rank which has no poles $|z| \leq 1$ and no zeros for $|z| < 1$.

If we in addition assume that the spectral density has no zero on the unit circle, the transfer function $k(z)$ has no zeros on the unit circle either. Then it can be shown that $k(z)$ is unique up to postmultiplication by constant orthogonal matrices. This is the case we are interested in from now on.

$k(z)$ is often called a spectral factor of $f(\lambda)$. The transfer function $k(z)$ is called causal and stable if it has no poles for $|z| \leq 1$ and it is called strictly miniphase if it has no zeros for $|z| \leq 1$.

The causal, stable and strictly miniphase rational transfer function $k(z)$ can be realized by a stable and strictly miniphase ARMA system, see Hannan and Deistler [2012].

1.3.2. Zeroless Rational Spectra - AR. The particular solution $(y_t^p)_{t \in \mathbb{Z}}$ of the AR system (1.1.2) we are considering is a regular process obtained via $k(z) = a^{-1}(z)b$ and thus has a rational spectral density

$$f_{y^p}(\lambda) = \frac{1}{2\pi} a^{-1}(e^{-i\lambda}) b b' a^{-*}(e^{-i\lambda}).$$

The continuation of the spectral density f_{y^p} on \mathbb{C} ,

$$f_{y^p}(z) = \frac{1}{2\pi} k(z) k^*(z) = \frac{1}{2\pi} a^{-1}(z) b b' a^{-*}(z),$$

is singular for singular AR systems and non-singular for regular AR systems. f_{y^p} has no zeros, i.e. for no value $z \in \mathbb{C}$ the rank of $f_{y^p}(z)$ is less than q , the normal rank, which means that f_{y^p} is of constant rank q . Thus the causal, stable and strictly miniphase spectral factor $k(z)$ of f_{y^p} is zeroless.

As the next theorem and its proof from Deistler et al. [2010] show, every zeroless stable and strictly miniphase rational transfer function can be realized by an AR system.

THEOREM 1.3.2. *The following three statements for (y_t^p) are equivalent:*

- (1) *The spectral factors $k(z)$ of the spectral density f_{y^p} of (y_t^p) as defined in Theorem 1.3.1 are zeroless.*
- (2) *There exists a polynomial left inverse $k^-(z)$ of $k(z)$.*
- (3) *(y_t^p) is a stationary solution of a stable AR system (1.1.2) where $\det a(z) \neq 0$, $|z| \leq 1$ and where ε_t is white noise with variance covariance matrix $\mathbb{E}\varepsilon_t\varepsilon_t' = \Sigma_\varepsilon = I_n$ and b has full column rank q .*

PROOF. 1. \rightarrow 2. Let Ω be the Smith McMillan form of $k = u\Omega v$. It is easy to see that $k^- = v^{-1}(\Omega'\Omega)^{-1}\Omega'u^{-1}$ is a left inverse of k and that k^- has no poles and no zeros for $|z| \leq 1$. Since all numerator polynomials in Ω are equal to one and u and v are unimodular k^- is polynomial.

2. \rightarrow 1. is straightforward.

3. \rightarrow 2. Premultiplying (1.1.2) by $(b'b)^{-1}b'$ yields a k^- of the desired form.

1. \rightarrow 3. We commence from an ARMA representation for y_t^p : $\tilde{a}(z)y_t^p = \tilde{b}(z)\varepsilon_t$ where $\tilde{a}(z)$ is stable and we can choose $\tilde{a}(z)$ and $\tilde{b}(z)$ left coprime. As $k(z) = \tilde{a}(z)^{-1}\tilde{b}(z)$ is zeroless, because of left coprimeness the same holds for $\tilde{b}(z)$. As is well known every zeroless tall polynomial matrix $\tilde{b}(z)$ can be completed by suitable choice of a polynomial matrix $g(z)$ to a unimodular matrix $u(z) = \begin{pmatrix} \tilde{b}(z)g(z) \\ \tilde{b}(z) \end{pmatrix}$ by extending the Smith McMillan form of $\tilde{b}(z) = \tilde{u}\tilde{\Omega}\tilde{v}$ to

$$\begin{pmatrix} \tilde{b}(z)g(z) \\ \tilde{b}(z) \end{pmatrix} = \tilde{u} \begin{pmatrix} \tilde{\Omega} & \begin{pmatrix} 0 \\ I \end{pmatrix} \end{pmatrix} \begin{pmatrix} \tilde{v} & 0 \\ 0 & I \end{pmatrix}.$$

Then $\tilde{a}(z)y_t^p = u(z) \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}$ and $u(z)^{-1}\tilde{a}(z)y_t^p = \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}$. By premultiplication of the last equation

$\tilde{a}^{-1}(0)u(0)u(z)^{-1}\tilde{a}(z)y_t^p = \tilde{a}^{-1}(0)u(0) \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}$ gives the autoregressive representation in 3. \square

The homogeneous solution $(y_t^h)_{t \in \mathbb{Z}}$ of the AR system (1.1.2) is a singular process, in fact a harmonic process with a spectral distribution function which is a monotonic, non-negative, right continuous step function and thus does not have a spectral density. The

spectral distribution function is

$$F_{y^h}(\lambda) = \sum_{j: \lambda_j \leq \lambda} F_j$$

with $F_j = \mathbb{E}v_j v_j^*$ which has its jumps at the frequencies λ_j corresponding to the zeros of $\det a(z)$ on the unit circle. Note that there is a one-to-one relationship between the spectral distribution and the (non-central) covariance function.

1.3.3. Observational Equivalence. In the last sections, we argued that a spectral density is rational if and only if it is the density of an ARMA process and that a rational spectral density is zeroless on \mathbb{C} if and only if it is the density of an AR process. But it may be that this ARMA or AR process is not uniquely determined by the spectral density.

First recall that we already stated that the spectral factor $k(z)$ is only unique up to postmultiplication by a constant orthogonal matrix O . To specify a unique choice, we can consider LQ decompositions and assume that b_0 is a quasi lower triangular matrix, see Filler [2010], Proposition 3.1.2.

The next definition introduces the class of ARMA (or AR) systems realizing the same spectral density. Since we consider only innovations with unit variance, i.e. $\mathbb{E}\varepsilon_t \varepsilon_t' = I_q$, the definition given here is slightly different than the definition in standard literature:

DEFINITION 1.3.3. Two ARMA systems $(a(z), b(z))$ and $(\tilde{a}(z), \tilde{b}(z))$, where $a(z)$ and $\tilde{a}(z)$ are $n \times n$ and where $b(z)$ and $\tilde{b}(z)$ are $n \times q$, are called *observationally equivalent* if for the spectral densities there holds $a^{-1} b b' a^{-*} = \tilde{a}^{-1} \tilde{b} \tilde{b}' \tilde{a}^{-*}$.

REMARK 1.3.4. Note that we have identifiability in the parameter space Θ if Θ contains no observationally equivalent systems.

Obviously, the systems (a, b) and $(\tilde{a}, \tilde{b}) = u(z) (a, b)$ where $u(z)$ is (non-singular) polynomial are observationally equivalent. A way to restrict the classes of observationally equivalent systems is to only consider left coprime system. Let us consider the class of observationally equivalent left coprime ARMA(p, v) systems:

LEMMA 1.3.5. *For a left coprime ARMA(p, v) system, $(a(z), b(z))$ with $a(0) = I$ the class of observationally equivalent left coprime ARMA(p, v) systems $(\tilde{a}(z), \tilde{b}(z))$ with $\tilde{a}(0) = I$ is given by*

$$(\tilde{a}(z), \tilde{b}(z)) = u(z) (a(z), b(z)O)$$

where O is an orthogonal matrix and $u(z)$ is unimodular, $u(0) = I$, and such that for the degrees hold $\delta(u(z)a(z)) \leq p$ and $\delta(u(z)b(z)) \leq v$.

If $\text{rk}(a_p, b_v) = n$ and b_0 is restricted to be quasi lower triangular, it can be shown that $u(z) = I$.

PROOF. The proof of the first part is trivial. Therefore we only show the second part which is straightforward. If b_0 is restricted to be quasi lower triangular, $O = I$. If $u(z) \neq I$, there would be a coefficient matrix $u_i \neq 0$ of z^i in $u(z)$, but since the degrees of $\tilde{a}(z)$ and $\tilde{b}(z)$ are restricted $u_i(a_p, b_v) = 0$ had to hold which is a contradiction to $\text{rk}(a_p, b_v) = n$. \square

1.4. Yule Walker Equations

We are interested in determining the AR parameters (a_1, \dots, a_p, b) from the second moments of observed data.

1.4.1. Regular AR Systems. For regular AR systems, the renowned Yule Walker equations are obtained by postmultiplying (1.1.2) by lagged y_{t-1}, \dots, y'_{t-p} and by y_t and taking expectation:

$$(1.4.1) \quad (\gamma(1), \dots, \gamma(p)) = (a_1, \dots, a_p) \underbrace{\begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(p-1) \\ \gamma(-1) & \gamma(0) & & \\ \vdots & & \ddots & \\ \gamma(-p+1) & & & \gamma(0) \end{pmatrix}}_{\Gamma_p}$$

$$(1.4.2) \quad \Sigma = bb' = \gamma(0) - (a_1, \dots, a_p) (\gamma(1), \dots, \gamma(p))'$$

where $\gamma(h) = \mathbb{E}y_{t+h}y'_t$.

For regular AR systems, i.e. $n = q$, the state of the companion form (1.1.3) x_t contains only linearly independent components which implies $\Gamma_p = \mathbb{E}x_t x'_t > 0$ and thus we have identifiability, i.e. the parameters can be uniquely determined from the second moments.

REMARK 1.4.1. If the population second moments are replaced by consistent estimators in the Yule Walker equations (1.4.1), we obtain consistent estimators $(\hat{a}_1, \dots, \hat{a}_p, \hat{b})$ for the parameters.

Moreover, we know that the AR system with parameters $(\hat{a}_1, \dots, \hat{a}_p, \hat{b})$ estimated by Yule Walker equations is always stable, for a proof for $n = 1$ see Brockwell and Davis [1987].

REMARK 1.4.2. Note that given $(\gamma(0), \dots, \gamma(p))$ the Yule Walker equations allow us to uniquely continue the covariance sequence since $\gamma(p+h) = (a_1, \dots, a_p) \begin{pmatrix} \gamma(p+h-1) \\ \vdots \\ \gamma(h) \end{pmatrix}$.

1.4.2. Singular AR Systems. For singular AR systems, Γ_p might be non-singular or singular. If $\Gamma_p > 0$ holds, the Yule Walker equations still give a unique solution (a_1, \dots, a_p, b) . For singular Γ_p , the set of solutions of the Yule Walker equations (1.4.1) for each row of (a_1, \dots, a_p) consists of one particular solution plus the left kernel of Γ_p . The noise parameters bb' are still unique as the variance covariance matrix of the perpendiculars of the orthogonal projection of y_t on its past and are determined by (1.4.2) for any solution (a_1, \dots, a_p) .

Despite the non-uniqueness of (a_1, \dots, a_p) for singular Γ_p , the next lemma from Deistler et al. [2011] shows that the transfer function $a^{-1}(z)b$ is always unique:

LEMMA 1.4.3. *All solutions (a_1, \dots, a_p, b) of the Yule Walker equations determine the same causal transfer function $k(z) = a^{-1}(z)b$.*

PROOF. Let the systems $(a_1(z), b)$ and $(a_2(z), b)$ denote two different solutions of the Yule Walker equations. The process $(y_t)_{t \in \mathbb{Z}}$ is a solution to both systems. By Theorem 1.1.11 $a_1^{-1}(z)b\varepsilon_t = a_2^{-1}(z)b\varepsilon_t$ and thus $a_1^{-1}(z)b = a_2^{-1}(z)b$ holds. \square

Thus the particular linearly regular process $(y_t^p)_{t \in \mathbb{Z}}$ is independent of the choice of the solution of the Yule Walker equations.

REMARK 1.4.4. Also for the singular case, the Yule Walker equations enable us to uniquely continue the covariance sequence: Form the point of view of orthogonal projections, the AR equation

$$y_t = \underbrace{a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p}}_{y_{t|t-1}} + b\varepsilon_t$$

can be used to interpret y_t as the sum of its projection on its own past, $y_{t|t-1}$, plus its perpendicular, $b\varepsilon_t$. The coefficients of this projection might not be unique, but the projection $y_{t|t-1}$ is. Thus

$$\gamma(p+h) = \mathbb{E}y_{t|t-1}y'_{t-p-h} = (a_1, \dots, a_p) \begin{pmatrix} \gamma(p+h-1) \\ \vdots \\ \gamma(h) \end{pmatrix}$$

is the same for all (a_1, \dots, a_p) .

. For the case of singular Γ_p , a procedure is described in Deistler et al. [2011] to determine a canonical form for the solution (a_1, \dots, a_p) of the Yule Walker equations:

Since $\Gamma_p = \mathbb{E}x_t x_t'$ holds, it is immediate to see that the dependence structure of the rows of Γ_p is the same as the dependence structure of the one dimensional components of the state x_t . Thus selecting the first basis from the elements of x_t for the Hilbert space spanned by the

one dimensional components of x_t corresponds to selecting a first basis for the row space of Γ_p .

Note that if $y_{t-h-1}^{(i)}$ is part of the first basis, so is $y_{t-h}^{(i)}$ since if $y_{t-h}^{(i)}$ can be expressed by preceding basis elements then $y_{t-h-1}^{(i)}$ can be expressed by the same linear combination of the corresponding shifted components and thus also by its preceding basis elements. If S denotes a matrix selecting the first basis of x_t , it is easy to see that $\mathbb{E}Sx_t x_t' S' = S\Gamma_p S' > 0$.

Set in (a_1, \dots, a_p) the columns of a_i not corresponding to components of the first basis are set equal to zero. Then

$$\begin{aligned} (\gamma(1), \dots, \gamma(p)) S' &= (a_1, \dots, a_p) S' (S\Gamma_p S') \\ &= (\bar{a}_1, \dots, \bar{a}_p) (S\Gamma_p S') \end{aligned}$$

defines a unique solution for $(\bar{a}_1, \dots, \bar{a}_p) = (a_1, \dots, a_p) S'$, i.e. the columns of (a_1, \dots, a_p) corresponding to the first basis, are obtained.

In a second step, the components of Sx_t might be further reduced by deleting more zero columns in $(a_1, \dots, a_p) S'$: If the i th column of a_{g+h} , $h \geq 0$ is zero, the component $y_{t-g}^{(i)}$ is not needed for any predictor and will be deleted from Sx_t . (Note that if the component $y_{t-g}^{(i)}$ is not needed for any predictor also the component $y_{t-g-1}^{(i)}$ is not needed.)

We denote the state obtained by deleting all components as described above by $S_e x_t$. The state $S_e x_t$ contains the components $y_{t-1}^{(1)}, \dots, y_{t-p_1}^{(1)}, \dots, y_{t-1}^{(n)}, \dots, y_{t-p_n}^{(n)}$. The corresponding nonzero system parameters are $(a_1, \dots, a_p) S'_e$. The corresponding polynomial matrix has column degrees (p_1, \dots, p_n) .

REMARK 1.4.5. Note that the canonical solution $(a_1, \dots, a_p) S'_e$ defined by this procedure gives us the solution of the Yule Walker equations with the smallest possible column degrees and thus by this solution we have defined a predictor using the most recent past.

The procedure described above motivates the following state space system

$$(1.4.3) \quad \begin{aligned} \bar{x}_{t+1} &= \bar{\mathcal{A}}\bar{x}_t + \bar{\mathcal{B}}\varepsilon_t \\ y_t &= (\bar{a}_1, \dots, \bar{a}_p) \bar{x}_t + b\varepsilon_t \end{aligned}$$

where $\bar{x}_t = Sx_t$, $\bar{\mathcal{A}} = SAS^T$, $\bar{\mathcal{B}} = SB$. Equation $\bar{\mathcal{A}}$ has been called the quasi companion form in Deistler et al. [2011]. Let $\bar{\Gamma}_p = \mathbb{E}\bar{x}_t \bar{x}_t'$.

Next we are interested in the stability of the canonical solution which was shown in Deistler et al. [2011], Theorem 4.1 and Corollary 4.2.

LEMMA 1.4.6. *The canonical solution $(\bar{a}_1, \dots, \bar{a}_p)$ is stable if and only if the set of solutions of the Yule Walker equations contain a stable solution.*

PROOF. If the canonical solution $(\bar{a}_1, \dots, \bar{a}_p)$ is stable, then the set of solutions of the Yule Walker equations trivially contain a stable solution.

Conversely, if the set of solutions of the Yule Walker equations contain a stable solution we can show that the canonical solution $(\bar{a}_1, \dots, \bar{a}_p)$ is stable.

If the Yule Walker equations contain a stable solution, the corresponding AR process must be linearly regular with block Toeplitz variance covariance matrix Γ_p . Since Remark 1.4.3 tells us that any solution of the Yule Walker equations determines the same transfer function and therefore the same particular process we know that Γ_p is the block Toeplitz variance covariance matrix of any particular process. Thus \bar{x}_t is also linearly regular and $\bar{\Gamma}_p = (\bar{\mathcal{B}}, \bar{\mathcal{A}}\bar{\mathcal{B}}, \dots) (\bar{\mathcal{B}}, \bar{\mathcal{A}}\bar{\mathcal{B}}, \dots)' > 0$. Hence $(\bar{\mathcal{A}}, \bar{\mathcal{B}})$ is controllable.

Since (1.1.4) also $\bar{\Gamma}_p$ fulfills a Lyapunov equation:

$$\bar{\Gamma}_p = \bar{\mathcal{A}}\bar{\Gamma}_p\bar{\mathcal{A}}' + \bar{\mathcal{B}}\bar{\mathcal{B}}'$$

If $\bar{\mathcal{A}}$ were not stable, i.e. there was an eigenvalue $\lambda \geq 1$ of $\bar{\mathcal{A}}$ with eigenvector x such that

$$(1 - |\lambda|^2) x' \bar{\Gamma}_p x = x' \bar{\mathcal{B}}\bar{\mathcal{B}}' x.$$

Since $x' \bar{\mathcal{B}}\bar{\mathcal{B}}' x \geq 0$ it follows that $x' \bar{\mathcal{B}} = 0$ which is a contradiction to the PBH Test for controllability, see Theorem 1.1.3.

That the system $(a(z), b)$ is stable easily follows since the eigenvalues of $\bar{\mathcal{A}}$ are the reciprocals of the zeros of $\det a(z)$. \square

From this point on we only consider the case where the Yule Walker equations have a stable solution (a_1, \dots, a_p) and therefore the unique solution $(y_t)_{t \in \mathbb{Z}}$ of the AR system $(a(z), b)$ with noise $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a linearly regular process.

REMARK 1.4.7. The system (1.4.3) is not necessarily minimal.

By definition we have $\bar{\Gamma}_p = \mathbb{E} \bar{x}_t \bar{x}_t' = (\bar{\mathcal{B}}, \bar{\mathcal{A}}\bar{\mathcal{B}}, \dots) (\bar{\mathcal{B}}, \bar{\mathcal{A}}\bar{\mathcal{B}}, \dots)' > 0$ and therefore controllability of (1.4.3) is fulfilled.

Observability is not always fulfilled. The observability matrix

$$\begin{pmatrix} (\bar{a}_1, \dots, \bar{a}_p) \\ (\bar{a}_1, \dots, \bar{a}_p) \bar{\mathcal{A}} \\ \vdots \\ (\bar{a}_1, \dots, \bar{a}_p) \bar{\mathcal{A}}^{\sum_i p_i - 1} \end{pmatrix}$$

for the case of nonzero column degrees, $p_i > 0, \forall i$, is the same as and in the case that there is an i such that $p_i = 0$ contains as a submatrix

$$\begin{pmatrix} (I_{n-s}, 0, \dots, 0) \\ (I_{n-s}, 0, \dots, 0) \bar{\mathcal{A}} \\ \vdots \\ (I_{n-s}, 0, \dots, 0) \bar{\mathcal{A}}^{\sum_i p_i - 1} \end{pmatrix} \bar{\mathcal{A}}$$

where s is the number of zero column degrees.

Let

$$a_i(z) = e_i - [a_1]_{\cdot, i} z \dots - [a_p]_{\cdot, i} z^p$$

denote the i th column of $a(z)$ and let

$$C_E = \left([a_{p_1}]_{\cdot, 1}, \dots, [a_{p_n}]_{\cdot, n} \right)$$

be the so-called column end matrix of $a(z)$.

For $p_i > 0, \forall i$, it is easy to see that $(\bar{\mathcal{A}}, (I_n, 0, \dots, 0))$ is observable if and only if the column end matrix of $a(z)$ is non-singular. If the column end matrix of $a(z)$ is non-singular, then all eigenvalues λ of $\bar{\mathcal{A}}$ are nonzero and all right eigenvectors α have to fulfill

$$\begin{aligned} \bar{\mathcal{A}}\alpha &= \lambda\alpha \\ S_2 S_1' \alpha_1 &= \lambda\alpha_2 \\ S_3 S_2' \alpha_2 &= \lambda\alpha_3 \\ &\vdots \\ S_p S_{p-1}' \alpha_{p-1} &= \lambda\alpha_p \end{aligned}$$

where $\alpha = (\alpha'_1, \dots, \alpha'_p)'$ and $S = \text{diag}(S_i)$ are partitioned conformable to the columns of $(\bar{a}_1, \dots, \bar{a}_p)$ (note that $p_i > 0$ implies that $S_1 = I_n$ and α_1 is $n \times 1$), which implies $\alpha_1 \neq 0$ and therefore $(I_n, 0, \dots, 0)\alpha \neq 0$. If the column end matrix is singular, then $\bar{\mathcal{A}}$ is singular and therefore observability does not hold.

Thus in case of nonzero column degrees (1.4.3) is minimal if and only if the column end matrix of $a(z)$ is non-singular and $\bar{\Gamma}_p > 0$. In case that there is an i such that $p_i = 0, \bar{\Gamma}_p > 0$ and a non-singular column end matrix of $S_1 a(z)$ imply minimality of (1.4.3).

. Let $\Theta_{(p_1, \dots, p_n)}$ denote the subspace of Θ where additionally the highest degree of the respective i th column of $a(z)$ is bounded by p_i and let $(\bar{a}_1, \dots, \bar{a}_p)$ denote the submatrix consisting of all columns of (a_1, \dots, a_p) which are not prescribed to be zero.

Considering the canonical form described in this section, we see that if appropriate maximum column degrees p_1, \dots, p_n of $a(z)$ rather than the degree of $a(z)$ are prescribed we can obtain identifiability in the restricted parameter space $\Theta_{(p_1, \dots, p_n)}$. Note that, as we see from the second step of the procedure for obtaining the canonical form, the actual column degrees of $a(z)$ might be smaller than the prescribed column degrees for guaranteeing uniqueness.

. Another solution of the Yule Walker equations is the so-called minimum norm solution which is presented in Deistler et al. [2010] and in Chen et al. [2011].

It is given by

$$(\tilde{a}_1, \dots, \tilde{a}_p) = (\gamma(1), \dots, \gamma(p)) \Gamma_p^\#$$

where $\Gamma_p^\#$ denotes the Moore-Penrose pseudo inverse of Γ_p and as such gives the row-wise minimum norm solution among all solutions

REMARK 1.4.8. In Deistler et al. [2011] it is shown that the minimum norm solution in the set of all solutions of the Yule Walker equations has the least number of zeros on the unit circle and the largest number of zeros outside of the unit circle and no zeros inside the unit circle. Thus if there is a stable solution in the set of all solutions of the Yule Walker equations, the minimum norm solution is also stable. If the minimum norm solution is not stable, then there is no stable solution of the Yule Walker equations.

1.5. State Space Representations and Subspace Procedures for ARMA Processes

Theorem 1.3.1 states that a rational transfer function $k(z)$ with no poles and zeros for $|z| \leq 1$ can be realized by a stable and strictly miniphase left coprime ARMA system. Analogously, such a rational transfer function can also be realized as a minimal, stable and miniphase state space system.

Here we will have a closer look at the definition of a state and the meaning of minimality.

Starting from an observed ARMA process $(y_t)_{t \in \mathbb{Z}}$ we want to find a minimal state space system with output $(y_t)_{t \in \mathbb{Z}}$ and the innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ as inputs.

In Subsection 1.5.1 we discuss a so-called subspace procedure which is built along a realization procedure where the estimation of the state is based on canonical correlation analysis which was described in detail in Larimore [1983], Deistler et al. [1995]. This subspace procedure is defined for the case that output dimension n and input dimension q are the same, i.e. $(y_t)_{t \in \mathbb{Z}}$ is the solution of a regular ARMA system. For the case that $n > q$, i.e. $(y_t)_{t \in \mathbb{Z}}$ is the solution of a singular ARMA system, we have to modify the algorithm in Subsection 1.5.2.

An n -dimensional ARMA process $(y_t)_{t \in \mathbb{Z}}$ with q -dimensional innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$, which are white noise with $\mathbb{E}\varepsilon_t \varepsilon_t' = I_q$, $q \leq n$, and a rational causal, stable and strictly miniphase transfer function $k(z)$ can also be modeled by a state space system

$$(1.5.1) \quad \begin{aligned} x_{t+1} &= Ax_t + B\varepsilon_t \\ y_t &= Cx_t + D\varepsilon_t \end{aligned}$$

where x_t is the, m -dimensional say, state and $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{m \times q}$, $C \in \mathbb{R}^{n \times m}$ and $D \in \mathbb{R}^{n \times q}$ and where the system is stable, i.e. for the eigenvalues λ_i of A holds $|\lambda_j| < 1$, and strictly miniphase, i.e. the system matrix

$$(1.5.2) \quad M(z) = \begin{pmatrix} I_m - Az & B \\ -Cz & D \end{pmatrix}$$

has no zeros for $|z| \leq 1$.

As we already stated in Section 1.1, a state space system (A, B, C, D) is minimal if and only if the pair (A, B) is controllable, i.e. the rank of $\mathcal{C} = (B, AB, A^2B, \dots)$ is equal to m , and the pair (C, A) is observable, i.e. the rank of $\mathcal{O}' = (C', A'C', (A^2)'\ C', \dots)$ is equal to m .

The transfer function $k(z)$ then can be written as

$$(1.5.3) \quad y_t = \underbrace{(C(I - zA)^{-1}Bz + D)}_{k(z)} \varepsilon_t = \sum_{j=1}^{\infty} \underbrace{CA^{j-1}B}_{k_j} \varepsilon_{t-j} + \underbrace{D}_{k_0} \varepsilon_t.$$

Let $(a(z), b(z))$ be a left coprime ARMA realization of $k(z)$. Thus

$$k(z) = a^{-1}(z)b(z).$$

It is easy to show that the system matrix $M(z)$ of a minimal state space system for $(y_t)_{t \in \mathbb{Z}}$ and the matrix $b(z)$ of a left coprime ARMA system have the same (finite) zeros and thus the notions of a miniphase left coprime ARMA system and a miniphase minimal state space system are both reflected in the corresponding transfer function being miniphase. Note that the miniphase assumption implies $b_0 = D$ to be of full column rank. The following lemma and its proof are a slight modification of Lemma 9.2.7 in Filler [2010], see also Kailath [1980] p 448.

LEMMA 1.5.1. *Let $k(z)$ be an (nxq) rational transfer function such that $y_t = k(z)\varepsilon_t$ holds. Let (A, B, C, D) be a minimal state space system for y_t with $k(z) = C(I_m - Az)^{-1}zB + D$ and let $(a(z), b(z))$ be a left coprime ARMA realization of $k(z) = a^{-1}(z)b(z)$ where $b(z)$ is of dimension (nxq) and $a(z)$ is of dimension $(n \times n)$.*

Then the matrices $\begin{pmatrix} (I_m - Az) & B \\ -Cz & D \end{pmatrix}$ and $\begin{pmatrix} I_m & 0 \\ 0 & b(z) \end{pmatrix}$ have the same Smith McMillan form, i.e. they have the same (finite) zeros.

PROOF. By definition we have

$$k(z) = a^{-1}(z)b(z) = C(I_m - Az)^{-1}zB + D$$

with $(a(z), b(z))$ left coprime where $a(z)$ is of dimension $(n \times n)$ and $b(z)$ is of dimension $(n \times q)$. We define $w(z) = C(I_m - Az)^{-1}z + \tilde{D} = \tilde{a}^{-1}(z)\tilde{b}(z)$ with $(\tilde{a}(z), \tilde{b}(z))$, $\tilde{a}(z)$ of dimension $(n \times n)$ and $\tilde{b}(z)$ of dimension $(n \times m)$ left coprime such that $b(z) = \tilde{b}(z)B$ (which implies $D = k(0) = b(0) = \tilde{b}(0)B = \tilde{D}B$) holds. Note that a full rank condition on B is not needed. We know that since $(\tilde{a}(z), \tilde{b}(z))$ is left coprime there exist polynomial matrices $\tilde{X}(z), \tilde{Y}(z)$ of dimensions $(m \times n)$ and $(n \times n)$ such that

$$\tilde{b}(z)\tilde{X}(z) + \tilde{a}(z)\tilde{Y}(z) = I_n$$

holds, see e.g. Hannan and Deistler [2012] Lemma 2.2.1. We know further that observability of (C, A) - or equivalently via PBH Test full rank $\forall z$ of $\begin{pmatrix} (I_m - Az) \\ C \end{pmatrix}$ - implies full rank $\forall z$ of $\begin{pmatrix} (I_m - Az) \\ Cz + \tilde{D}(I_m - Az) \end{pmatrix}$ (since according to the PBH Test we only have to test for left eigenvectors of A , and thus $\tilde{D}(I_m - Az)$ is irrelevant) and thus $\begin{pmatrix} (I_m - Az) \\ Cz + \tilde{D}(I_m - Az) \end{pmatrix}$ has full column rank $\forall z$ which is equivalent to right coprimeness. Therefore there exist polynomial matrices $X(z), Y(z)$ of dimensions $(m \times m)$ and $(m \times n)$ such that

$$X(z)(I_m - Az) + Y(z)(Cz + \tilde{D}(I_m - Az)) = I_m$$

holds. As $\tilde{b}(z)(I_m - Az) - \tilde{a}(z)(Cz + \tilde{D}(I_m - Az)) = 0$ holds too we have

$$(1.5.4) \quad \begin{pmatrix} X(z) & -Y(z) \\ \tilde{b}(z) & \tilde{a}(z) \end{pmatrix} \begin{pmatrix} (I_m - Az) & \tilde{X}(z) \\ -Cz - \tilde{D}(I_m - Az) & \tilde{Y}(z) \end{pmatrix} = \begin{pmatrix} I_m & -Q(z) \\ 0 & I_n \end{pmatrix}$$

where $Q(z) = -X(z)\tilde{X}(z) + Y(z)\tilde{Y}(z)$. As the two square block matrices on the left hand side of 1.5.4 are polynomial they are unimodular as the matrix on the right hand side of 1.5.4

is unimodular. Thus we have

$$\begin{aligned} & \begin{pmatrix} X(z) & -Y(z) \\ \tilde{b}(z) & \tilde{a}(z) \end{pmatrix} \begin{pmatrix} I & 0 \\ -\tilde{D} & I \end{pmatrix} \begin{pmatrix} (I_m - Az) & B \\ -Cz & D \end{pmatrix} \\ &= \begin{pmatrix} X(z) & -Y(z) \\ \tilde{b}(z) & \tilde{a}(z) \end{pmatrix} \begin{pmatrix} (I_m - Az) & B \\ -Cz - \tilde{D}(I_m - Az) & 0 \end{pmatrix} = \begin{pmatrix} I_m & X(z)B \\ 0 & \tilde{b}(z)B \end{pmatrix} \end{aligned}$$

As matrices which are related by a multiplication of a unimodular matrix have the same Smith McMillan form, the matrices $\begin{pmatrix} I_m & X(z)B \\ 0 & \tilde{b}(z)B \end{pmatrix} = \begin{pmatrix} I_m & X(z)B \\ 0 & b(z) \end{pmatrix}$ and $\begin{pmatrix} (I_m - Az) & B \\ -Cz & D \end{pmatrix}$ have the same Smith McMillan form. As

$$\begin{pmatrix} I_m & X(z)B \\ 0 & b(z) \end{pmatrix} \begin{pmatrix} I_m & -X(z)B \\ 0 & I_q \end{pmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & b(z) \end{bmatrix}$$

holds, and the second matrix on the left hand side of the equation above is unimodular, the result follows. \square

Let us have a closer look at the concept of a minimal state:

By the results obtained by Akaike [1974] and Kalman [1963, 1965, 1974] (see also Hannan and Deistler [2012] Chapter 2) we have the following state construction:

Let $\mathbb{H}_y^+(t)$ denote the Hilbert space spanned by all one-dimensional components of the future random variables $y_{t+h}^{(i)}$, $h > 0$, $\mathbb{H}_y^+(t) = \overline{\text{span}} \{y_{t+1}^1, \dots, y_{t+1}^n, y_{t+2}^1, \dots, y_{t+2}^n, \dots\}$ and let $\mathbb{H}_y^-(t)$ denote the Hilbert space spanned by all components of the present and past variables $y_{t-j}^{(i)}$, $j \geq 0$, $\mathbb{H}_y^-(t) = \overline{\text{span}} \{y_t^1, \dots, y_t^n, y_{t-1}^1, \dots, y_{t-1}^n, \dots\}$. The state space is the space spanned by the projections of the elements of $\mathbb{H}_y^+(t)$ onto $\mathbb{H}_y^-(t)$. This state space is finite dimensional if and only if the spectral density of $(y_t)_{t \in \mathbb{Z}}$ is rational, and in this case every basis of the state space forms a minimal state of a stable and miniphase state space system with $(y_t)_{t \in \mathbb{Z}}$ as outputs and the innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ as inputs. Thus if x_t is a minimal state all other minimal states are obtained as Qx_t where Q is a non-singular matrix. The dimension of the state space is the so-called *McMillan degree* of the transfer function corresponding to such a stable and miniphase system.

For the AR case, the minimal state x_t consists of finite linear combinations of $y_s^{(i)}$, $s < t$, $i = 1, \dots, n$, see Theorem 1.1.12. If $(y_t)_{t \in \mathbb{Z}}$ is not an AR process, this is not the case since contrary to the AR case the best linear predictor of $(y_t)_{t \in \mathbb{Z}}$ is the projection on the infinite past, see Remark 1.2.4.

The last considerations help us in obtaining a procedure for determining the state from $(y_t)_{t \in \mathbb{Z}}$. We are using a slightly different notation here than in Deistler et al. [1995] to accommodate also the case of singular ARMA systems.

First let us transform the system (1.5.1) into a system with input $(y_t)_{t \in \mathbb{Z}}$ and output $(\varepsilon_t)_{t \in \mathbb{Z}}$. Let D^- denote the (generalized left) inverse of D , i.e. $D^-D = I$, then we have

$$(1.5.5) \quad \begin{aligned} x_{t+1} &= (A - BD^-C)x_t + BD^-y_t \\ \varepsilon_t &= -D^-Cx_t + D^-y_t. \end{aligned}$$

and call $\mathcal{K} = (BD^-, (A - BD^-C)BD^-, (A - BD^-C)^2BD^-, \dots)$ the controllability matrix of this system. As easily can be seen \mathcal{K} is of full column rank if and only if \mathcal{O} is of full column rank. (In fact it is easy to see that (1.5.5) is minimal if and only if (1.5.1) is minimal.)

Let us define the vector of all future variables $y^+(t) = (y'_{t+1}, y'_{t+2}, y'_{t+3}, \dots)'$ and the vector of all past variables $y^-(t) = (y'_t, y'_{t-1}, y'_{t-2}, \dots)'$ and accordingly the vector of future innovations $\varepsilon^+(t) = (\varepsilon'_{t+1}, \varepsilon'_{t+2}, \varepsilon'_{t+3}, \dots)'$ and the vector of past innovations $\varepsilon^-(t) = (\varepsilon'_t, \varepsilon'_{t-1}, \varepsilon'_{t-2}, \dots)'$. Obviously, the components of $y^+(t)$ span $\mathbb{H}_y^+(t)$ and the components of $y^-(t)$ span $\mathbb{H}_y^-(t)$. Note that because we assume the system (1.5.1) to be miniphase we have that $\mathbb{H}_y^-(t) = \mathbb{H}_\varepsilon(t)$ where $\mathbb{H}_\varepsilon(t)$ is spanned by the components of $\varepsilon^-(t)$. The vector $\varepsilon^+(t)$ is orthogonal to both $\varepsilon^-(t)$ and $y^-(t)$.

Clearly,

$$\begin{pmatrix} y_{t+1} \\ y_{t+2} \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \\ k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}^k} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \end{pmatrix} + \underbrace{\begin{pmatrix} k_0 \\ k_1 & k_0 \\ k_2 & k_1 & k_0 \\ \vdots & \vdots & \ddots \end{pmatrix}}_{\mathcal{I}} \begin{pmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+2} \\ \vdots \end{pmatrix}$$

or in short since $\mathcal{H}^k = \mathcal{O}\mathcal{C}$

$$y^+(t) = \mathcal{O}\mathcal{C}\varepsilon^-(t) + \mathcal{I}\varepsilon^+(t).$$

From (1.5.1) and (1.5.5) we also have

$$x_t = \mathcal{C}\varepsilon^-(t) = \mathcal{K}y^-(t)$$

and thus

$$(1.5.6) \quad y^+(t) = \mathcal{O}\mathcal{K}y^-(t) + \mathcal{I}\varepsilon^+(t).$$

Note that because of observability the row space of $\mathcal{O}\mathcal{K}$ is the same as the row space of \mathcal{K} .

1.5.1. A Subspace Procedure for Regular ARMA Systems. Now we shortly describe the subspace procedure presented in Deistler et al. [1995] which has first been proposed

in Larimore [1983] for estimating a minimal state and subsequently estimating a minimal state space system (A, B, C, D) for the case that $(y_t)_{t \in \mathbb{Z}}$ is the solution of a regular ARMA system, i.e. $n = q$. In Deistler et al. [1995] consistency of the estimated procedure was proved up to basis transformation.

The “key” equation for the procedure is (1.5.6). For a given finite data set (y_1, \dots, y_T) , analogously to (1.5.6), we consider a finite regression

$$(1.5.7) \quad y_h^+(t) = \beta_h y_h^-(t) + v(t)$$

where the dependent variables are $y_h^+(t) = (y'_{t+1}, y'_{t+2}, \dots, y'_{t+h})'$, the regressors are $y_h^-(t) = (y'_t, y'_{t-1}, \dots, y'_{t-h+1})'$ and $v(t)$ is the orthogonal residual. The index h is determined by approximating $(y_t)_{t \in \mathbb{Z}}$ by a long autoregression. The autoregressive order is estimated by the BIC criterion where this criterion is minimized over $1 \leq h \leq (\ln T)^\alpha$, $\alpha < \infty$. For an AR process, h will be finite whereas for an ARMA process h will increase with T , but much more slowly. It is shown in Deistler et al. [1995] that β_h of the finite regression model (1.5.7) (under suitable assumptions on $(\varepsilon_t)_{t \in \mathbb{Z}}$) converges to $\mathcal{O}_h \mathcal{K}_h$ for $h \rightarrow \infty$ where \mathcal{O}_h and \mathcal{K}_h consist of the first nh rows of \mathcal{O} and columns of \mathcal{K} respectively. We assume that $h \geq m$.

If we denote the variance covariance matrices of $y_h^-(t)$ and $y_h^+(t)$ and the cross covariance matrix of $y_h^-(t)$ and $y_h^+(t)$ as

$$\begin{aligned} \Gamma_h^- &= \mathbb{E} y_h^-(t) y_h^-(t)', \\ \Gamma_h^+ &= \mathbb{E} y_h^+(t) y_h^+(t)', \end{aligned}$$

and

$$\mathcal{H}_h^\gamma = \mathbb{E} y_h^+(t) y_h^-(t)',$$

then

$$\beta_h = \mathcal{H}_h^\gamma (\Gamma_h^-)^{-1}.$$

Note that for h large enough $\text{rk} \beta_h = m$ holds. This can be seen as follows. As $\gamma(j) = \mathbb{E} y_j y_0' = \sum_{i=0}^{\infty} k_{j+i} k_i'$ we have

$$\underbrace{\begin{pmatrix} \gamma(1) & \gamma(2) & \gamma(3) & \cdots \\ \gamma(2) & \gamma(3) & \gamma(4) & \cdots \\ \gamma(3) & \gamma(4) & \gamma(5) & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}^\gamma} = \underbrace{\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \\ k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}^k} \begin{pmatrix} k'_0 & & & \\ k'_1 & k'_0 & & \\ k'_2 & k'_1 & k'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix}$$

where the second matrix on the right hand side has full row rank. Thus the rank of the Hankel matrix \mathcal{H}^γ of covariances of $(y_t)_{t \in \mathbb{Z}}$ is the same as the rank of the Hankel matrix \mathcal{H}^k of a causal, stable and miniphase transfer function of $(y_t)_{t \in \mathbb{Z}}$, which has finite rank m , since. Therefore for h large enough \mathcal{H}_h^γ has rank m .

On the sample level we have the following: Call

$$Y^+ = \begin{pmatrix} y_2 & y_3 & \cdots & y_{T+1} \\ y_3 & y_4 & \cdots & y_{T+2} \\ \vdots & \vdots & & \vdots \\ y_{h+1} & y_{h+2} & \cdots & y_{T+h} \end{pmatrix}$$

and

$$Y^- = \begin{pmatrix} y_1 & y_2 & \cdots & y_T \\ y_0 & y_1 & \cdots & y_{T-1} \\ \vdots & \vdots & & \vdots \\ y_{-h+2} & y_{-h+3} & \cdots & y_{T-h+1} \end{pmatrix}$$

where $y_s = 0$ for $s \leq 0$ and $s > T$ the data matrices. Thus the sample counterparts of Γ_h^-, Γ_h^+ and \mathcal{H}_h^γ can be written as

$$\hat{\Gamma}_h^- = \frac{1}{T} Y^- (Y^-)',$$

$$\hat{\Gamma}_h^+ = \frac{1}{T} Y^+ (Y^+)',$$

and

$$\hat{\mathcal{H}}_h^\gamma = \frac{1}{T} Y^+ (Y^-)'$$

respectively. Then the least squares estimator of β_h is

$$\hat{\beta}_h = \hat{\mathcal{H}}_h^\gamma (\hat{\Gamma}_h^-)^{-1}.$$

Whereas β_h has rank m it is clear that $\hat{\beta}_h$ will “typically” be of full rank nh . Thus the authors in Deistler et al. [1995] propose a reduced rank regression: First let us transform $\hat{\beta}_h$

$$\left(\hat{\Gamma}_h^+\right)^{-1/2} \hat{\beta}_h \left(\hat{\Gamma}_h^-\right)^{1/2} = \left(\hat{\Gamma}_h^+\right)^{-1/2} \hat{\mathcal{H}}_h^\gamma \left(\hat{\Gamma}_h^-\right)^{-1/2}$$

where $\hat{\Gamma}_h^\pm = \left(\hat{\Gamma}_h^\pm\right)^{1/2} \left(\hat{\Gamma}_h^\pm\right)^{1/2}$ and let

$$U\Lambda V' = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{pmatrix} V_1' \\ V_2' \end{pmatrix}$$

be a singular value decomposition of this transformation where Λ_1 contains on its diagonal the m largest singular values. Thus we have an estimator for \mathcal{K}_h , the first nh columns of \mathcal{K} ,

$$\hat{\mathcal{K}}_h = \hat{\Lambda}_1^{1/2} V_1' \left(\hat{\Gamma}_h^-\right)^{-1/2},$$

and an estimator the state x_t ,

$$\hat{x}_t = \hat{\mathcal{K}}_h y_h^-(t).$$

Note that the transformation of $\hat{\beta}_h$ is used because of its statistical properties. The m singular values in Λ_{11} are also the first canonical correlation coefficients between the row spaces of Y^- and Y^+ , see e.g. Brillinger [1981].

Now that we have an estimator \hat{x}_t , $t = 1, \dots, T$ we can compute an estimate of C from

$$y_t = \hat{C} \hat{x}_t + \hat{\nu}_t$$

by least squares regression

$$\hat{C} = \left(\frac{1}{T} \left(\sum_{t=1}^T y_t \hat{x}_t' \right) \right) \left(\frac{1}{T} \sum_{t=1}^T \hat{x}_t \hat{x}_t' \right)^{-1}$$

where the residual $\hat{\nu}_t$ is an estimate of $D\varepsilon_t$. Note that in Deistler et al. [1995] D was restricted to $D = I_n$ and instead the variance covariance matrix of ε_t was free. Since we restrict $\mathbb{E}\varepsilon_t \varepsilon_t' = I_q$, D is free, but can only be determined up to postmultiplication by an orthogonal constant matrix. A unique choice of D can be specified using LQ decomposition, see Filler [2010], Proposition 3.1.2. We can give an estimator of D from the eigenvalue decomposition of

$$\frac{1}{T} \sum_{t=1}^T \hat{\nu}_t \hat{\nu}_t' = O\Delta O'$$

and determine \hat{D} as the lower triangular matrix of the LQ decomposition of $O\Delta^{1/2}$. An estimate of ε_t can be obtained as

$$\hat{\varepsilon}_t = \hat{D}^{-1} \hat{\nu}_t.$$

Estimators of A and B can be obtained from

$$\hat{x}_{t+1} = \hat{A}\hat{x}_t + \hat{B}\hat{\varepsilon}_t + u_t$$

by least squares regression as

$$\hat{A} = \left(\frac{1}{T} \left(\sum \hat{x}_{t+1} \hat{x}_t' \right) \right) \left(\frac{1}{T} \sum \hat{x}_t \hat{x}_t' \right)^{-1}$$

and

$$\hat{B} = \left(\frac{1}{T} \left(\sum \hat{x}_{t+1} \hat{\varepsilon}_t' \right) \right) \left(\frac{1}{T} \sum \hat{\varepsilon}_t \hat{\varepsilon}_t' \right)^{-1}.$$

Note that the estimated system is always stable but not necessarily miniphase. In Deistler et al. [1995] consistency of the procedure is shown.

1.5.2. A Subspace Procedure for Singular ARMA Systems. For the case that $(y_t)_{t \in \mathbb{Z}}$ is the solution of a singular ARMA system, i.e. $q < n$, (1.5.6) still holds, but $y^-(t)$ and $y^+(t)$ both contain linearly dependent components. Thus we first consider a procedure for determining the first bases of linearly independent components of $y^-(t)$ and $y^+(t)$.

Let us consider the spectral density of our ARMA(p, v) process $(y_t)_{t \in \mathbb{Z}}$

$$f_y(z) = \underbrace{a^{-1}(z)b(z)}_{k(z)} \underbrace{b(z^{-1})'a^{-1}(z^{-1})'}_{k^*(z)}$$

which clearly has rank q . We are interested in a polynomial $(n - q) \times n$ matrix $w(z)$ in the backward shift z whose rows form a basis of the left kernel of $f_y(z)$. As is easily seen,

$$w(z)f_y(z) = 0$$

immediately implies

$$w(z)y_t = 0.$$

In particular, we are interested in a minimal polynomial basis of the left kernel of $f_y(z)$, i.e. a polynomial basis with the smallest sum of row degrees v_i (or smallest order), see Forney jr. [1975]. As is shown in Forney jr. [1975], all minimal polynomial bases have the same row degrees up to reordering. Furthermore, in Theorem 2 of Forney jr. [1975] it is stated that all minimal polynomial bases have the same unique echelon form, where a $(n - q) \times n$ matrix is said to be in (row polynomial) echelon form if

- it is row reduced with decreasing row degrees, i.e. $v_1 \geq v_2 \geq \dots \geq v_{n-q}$

(A polynomial $(n - q) \times n$ matrix is row reduced if and only if there is a $(n - q) \times (n - q)$ submatrix whose determinant has degree $\sum v_i$, see e.g. Hannan and Deistler [2012] p 42.)

- for row i with $1 \leq i \leq (n - q)$ there is a pivot index p_i such that the (i, p_i) element is monic with degree v_i and for all elements (i, j) , $j > p_i$ the degree is less than v_i
- for $v_i = v_j$, $i < j$ it follows $p_i < p_j$ and
- the (j, p_i) , $j \neq i$, elements have degree less than v_i .

Note that here we follow the definition of an echelon form given in Anderson et al. [2012a], where the difference to the more standard definition in Forney jr. [1975] is the descending order of row degrees instead of increasing. This difference, however, is immaterial for the cited results, but by the descending order we obtain the following:

The unique echelon form gives us the first basis of the Hilbert space of the past variables, $\mathbb{H}_y^-(t)$, i.e. the first components of $y^-(t) = (y'_t, y'_{t-1}, \dots)'$ forming a basis of $\mathbb{H}_y^-(t)$:

Consider the i th row of the echelon form of $w(z)$, $w_i(z)$. Then since $w_i(z)y_t = 0$ the component $y_{t-v_i}^{(p_i)}$ can be expressed as a linear combination of preceding components $y_{t-s}^{(l)}$, $l = 1, \dots, n$, $s = 0, \dots, v_i - 1$ and $y_{t-v_i}^{(l)}$, $l < p_i$. Because of the last bullet point this linear combination does not include components $y_{t-s}^{(p_j)}$, $s \geq v_j$, $j \neq i$ which themselves can be expressed as linear combinations of preceding components. Those q components without a corresponding pivot index in the echelon form of $w(z)$ can never be expressed as linear combinations of preceding components.

Thus the row degrees of the echelon form give us the first basis of $\mathbb{H}_y^-(t)$.

Next we discuss the problem of giving an upper bound for the degree of the (unique) echelon form of a minimal polynomial basis by the (assumed to be known) integer model class parameters p , v and q .

Obviously, the spectral density $f_y(z)$ has in its kernel

$$k^\perp(z) = b^\perp(z)a(z)$$

with $b^\perp(z)b(z) = 0$, where the rows of $b^\perp(z)$ form a minimal polynomial basis of the left kernel of $b(z)$. Theorem 3 and its Corollary of Forney jr. [1975] establishes that the sum of row degrees of a minimal polynomial basis of the space spanned by the rows of $b^\perp(z)$ is equal to the sum of column degrees of a minimal polynomial basis of the space spanned by the columns of $b(z)$. Therefore we can give a bound for the degree of $b^\perp(z)$, the maximum row degree, namely vq , the bound of the sum of the v column degrees of $b(z)$.

Hence we can give a bound on the degree of the unique echelon form of a minimal polynomial basis, which is the bound on the degree of $k^\perp(z) = b^\perp(z)a(z)$, which is $p + vq$.

As is easily seen, if $w(z)$ is the basis of a polynomial left kernel of $f_y(z)$ in the backward shift z then $w^*(z) = w'(z^{-1})$ is the basis of a polynomial right kernel of $f_y(z)$ in the forward

shift z^{-1} . Then

$$f_y(z)w'(z^{-1}) = 0$$

immediately implies

$$\begin{aligned} y'_t w'(z^{-1}) &= 0 \\ w(z^{-1})y_t &= 0. \end{aligned}$$

Therefore, completely analogously to the case of the left kernel in the backward shift z , the unique echelon form of a minimal polynomial basis of $w(z^{-1})$ in the forward shift z^{-1} gives us the first basis of Hilbert space of the future variables, $\mathbb{H}_y^+(t)$, i.e. the first components of $y^+(t) = (y'_{t+1}, y'_{t+2}, \dots)'$ forming a basis of $\mathbb{H}_y^+(t)$.

Thus summarizing we have established the following lemma:

LEMMA 1.5.2. *Let the n -dimensional process $(y_t)_{t \in \mathbb{Z}}$ be the stationary solution of a singular ARMA(p, v) system with q -dimensional innovations. Then we can find one selector matrix S selecting the first bases of $y^-(t)$ and $y^+(t)$, $Sy^-(t)$ and $Sy^+(t)$ respectively.*

If a component $y_{t-p-vq}^{(i)}$ is in $Sy^-(t)$, then also all components $y_{t-g}^{(i)}$, $g > p+pq$ are elements of $Sy^-(t)$. Likewise, if a component $y_{t+p+vq+1}^{(i)}$ is in $Sy^+(t)$ then also all components $y_{t+h}^{(i)}$, $h > p + pq + 1$ are elements of $Sy^-(t)$.

Since the first basis of components in $y^-(t)$ is the same as the first basis of the row space of $\Gamma^- = \mathbb{E}y^-(t)y^-(t)'$ we can determine the infinite matrix S from the finite matrix Γ_{p+vq+1}^- : We determine the first basis of Γ_{p+vq+1}^- . The rows selected from the last block of rows of Γ_{p+vq+1}^- will also be selected from all further blocks for a first basis of Γ_g^- , $g > p + pq + 1$ and hence we are able to determine S .

As in the regular case we want to estimate a minimal state by projecting the future space $\mathbb{H}_y^+(t)$ onto the past space $\mathbb{H}_y^-(t)$ in order to select a basis of the resulting subspace. A basis for the resulting subspace will be a minimal state x_t . In order to avoid complications with invertibility of covariance matrices and to avoid the computation of redundant equations we are projecting a basis of $\mathbb{H}_y^+(t)$ onto a basis of $\mathbb{H}_y^-(t)$. Thus similar to (1.5.6) we obtain

$$\begin{aligned} Sy^+(t) &= SOKP'Sy^-(t) + S\mathcal{I}\varepsilon^+(t). \\ x_t &= KP'Sy^-(t) \end{aligned}$$

where P' is lower triangular such that $y^-(t) = P'Sy^-(t)$.

Since the McMillan degree of the transfer function is m , we can conclude that $S\mathcal{O}$ has full column degree m and therefore the row space of $SOKP'$ is the same as the row space of KP' .

Thus an estimate of x_t may be constructed from an estimate of $\mathcal{K}P'$. As in the regular case we consider a truncated finite regression

$$S_h y_h^+(t) = \beta_h S_h y_h^-(t) + v(t)$$

where h is chosen analogously to the regular case. The matrix S_h selects the first basis of $y_h^-(t)$, and $y_h^+(t)$ (and let us also define P'_h as a submatrix of P' such that $y_h^-(t) = P'_h S_h y_h^-(t)$).

The procedure for estimating the state x_t and the parameter matrices is analogous to the regular case:

The least squares estimator of β_h is given by

$$\hat{\beta}_h = S_h \hat{\mathcal{H}}_h^\gamma S'_h \left(S_h \hat{\Gamma}_h^- S'_h \right)^{-1}.$$

Now “typically” $\hat{\beta}_h$ has full rank whereas β_h has rank m . Therefore we perform a reduced rank step:

Let

$$U \Lambda V' = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{pmatrix} V'_1 \\ V'_2 \end{pmatrix}$$

be the singular value decomposition of

$$\left(S_h \hat{\Gamma}_h^+ S'_h \right)^{-1/2} \hat{\beta}_h \left(S_h \hat{\Gamma}_h^- S_h \right)^{1/2} = \left(S_h \hat{\Gamma}_h^+ S'_h \right)^{-1/2} S_h \hat{\mathcal{H}}_h^\gamma S'_h \left(S_h \hat{\Gamma}_h^- S'_h \right)^{-1/2}$$

Then $\mathcal{K}_h P'_h$, i.e. the first h block columns of $\mathcal{K}P'$, is estimated as

$$\hat{\mathcal{K}}_h P'_h = \Lambda_1^{1/2} V'_1 \left(S_h \hat{\Gamma}_h^- S'_h \right)^{-1/2}$$

and the state as

$$\hat{x}_t = \hat{\mathcal{K}}_h P'_h S_h y_h^-(t).$$

For the estimate for the state, the estimates of the system matrices (A, B, C, D) are then computed by least squares regression. From

$$(1.5.8) \quad y_t = \hat{C} \hat{x}_t + \hat{v}_t$$

we obtain

$$\hat{C} = \left(\frac{1}{T} \left(\sum_{t=1}^T y_t \hat{x}'_t \right) \right) \left(\frac{1}{T} \sum_{t=1}^T \hat{x}_t \hat{x}'_t \right)^{-1}.$$

The residuals $\hat{\nu}_t$, an estimate of $D\varepsilon_t$, will consist of n linearly independent components, not only q . Thus we look at the eigenvalue decomposition of

$$\frac{1}{T} \sum_{t=1}^T \hat{\nu}_t \hat{\nu}_t' = O\Delta O = O_1\Delta_1 O_1' + O_2\Delta_2 O_2'$$

to obtain an estimator \hat{D} for D as the unique quasi lower triangular matrix in the LQ decomposition of $O_1\Lambda_1^{1/2}$, see Filler [2010], Proposition 3.1.2., with Moore-Penrose pseudo inverse \hat{D}^- . We then estimate ε_t as

$$\hat{\varepsilon}_t = \hat{D}^- \hat{\nu}_t.$$

Estimators of A and B can be obtained from

$$\hat{x}_{t+1} = \hat{A}\hat{x}_t + \hat{B}\hat{\varepsilon}_t + u_t$$

by least squares regression as

$$\hat{A} = \left(\frac{1}{T} \left(\sum_{t=1}^T \hat{x}_{t+1} \hat{x}_t' \right) \right) \left(\frac{1}{T} \sum_{t=1}^T \hat{x}_t \hat{x}_t' \right)^{-1}$$

$$\hat{B} = \left(\frac{1}{T} \left(\sum_{t=1}^T \hat{x}_{t+1} \hat{\varepsilon}_t' \right) \right) \left(\frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t' \right)^{-1}$$

1.5.3. Simulation Results. In this thesis, we are not investigating the analytical properties of the subspace procedure for the singular ARMA case described in the last subsection. Instead, in this subsection, we are considering Monte Carlo simulations. We consider a singular stable and miniphase ARMA system with 3-dimensional outputs and 2-dimensional normally distributed innovations. For sample sizes $T = 1000, 2000, 4000, 8000,$ and 16000 we repeatedly, i.e. 100 times, simulate observations for this ARMA system and estimate the system and noise parameters. We compare the echelon form of the estimated and the true parameters, $\hat{\theta}$ and θ respectively, see Hannan and Deistler [2012] Theorem 2.5.2, and compute the mean squared error, $\frac{1}{100} \sum_i \hat{\theta}_i^{(j)} - \theta^{(j)}$ of the unrestricted parameters.

Note that we use a different estimator for the block Toeplitz matrices Γ_h^- , and Γ_h^+ and the Hankel matrix \mathcal{H}_h^γ by truncating the data matrices

$$Y^+ = \begin{pmatrix} y_{h+1} & y_{h+2} & \cdots & y_{T-h+1} \\ y_{h+2} & y_{h+3} & \cdots & y_{T-h+2} \\ \vdots & \vdots & & \vdots \\ y_{2h} & y_{2h+1} & \cdots & y_T \end{pmatrix}$$

and

$$Y^- = \begin{pmatrix} y_h & y_{h+1} & \cdots & y_{T-h} \\ y_{h-1} & y_h & \cdots & y_{T-h-1} \\ \vdots & \vdots & & \vdots \\ y_1 & y_2 & \cdots & y_{T-2h+1} \end{pmatrix}$$

where all elements are observed. We sacrifice the efficiency of the estimator of the second moments in order to preserve the dependence structure of the estimators $\hat{\Gamma}_h^-$, $\hat{\Gamma}_h^+$ and $\hat{\mathcal{H}}_h^\gamma$. Using these estimators, it is easy to determine the selector matrix S_h .

As in Deistler et al. [1995], we determine the number of blocks h by AIC where we use a form of the likelihood of singular ARMA systems described in Srivastava and von Rosen [2002].

Let us first consider a singular stable and miniphase ARMA(2, 1) system.

$$(1.5.9) \quad y_t - \begin{pmatrix} 2.46313 & 0.97628 & 0.61516 \\ -1.66592 & -0.26521 & -1.17880 \\ -0.80264 & 0.17668 & 0.82098 \end{pmatrix} y_{t-1} - \begin{pmatrix} -1.35121 & -0.80471 & -0.49328 \\ 1.32774 & 0.85409 & 0.93859 \\ 0.55288 & -0.19481 & -0.08439 \end{pmatrix} y_{t-2} \\ = \begin{pmatrix} 1.61266 & 0 \\ 1.01399 & 1.50797 \\ 2.27170 & -2.91294 \end{pmatrix} \varepsilon_t + \begin{pmatrix} 3.42768 & 3.18331 \\ 0.28016 & 0.66802 \\ 0.61889 & -4.00152 \end{pmatrix} \varepsilon_{t-1}$$

The McMillan degree m of the corresponding transfer function, which is the number of linearly independent rows of the Hankel matrix \mathcal{H}^k , is 6. The so-called Kronecker indices m_i , $i = 1, \dots, 3$ indicate how often the i th row of the 3-dimensional blocks in \mathcal{H}^k is part of the first basis of rows, see Hannan and Deistler [2012] Chapter 2. Clearly, $\sum_i m_i = m$ holds. In our example we have $m_i = 2$.

The parameter matrices of the state space system in echelon form are

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -1.35121 & 2.46313 & -0.80471 & 0.97628 & -0.49328 & 0.61516 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1.32774 & -1.66592 & 0.85409 & -0.26521 & 0.93859 & -1.17880 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0.55288 & -0.80264 & -0.19481 & 0.17668 & -0.08439 & 0.82098 \end{pmatrix}$$

$$B = \begin{pmatrix} 9.78724 & 2.86360 \\ 15.60734 & 7.12211 \\ -5.35319 & 3.70187 \\ -11.35901 & 0.02357 \\ 1.36869 & -6.12656 \\ -7.17541 & -6.72211 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, D = \begin{pmatrix} 1.61266 & 0 \\ 1.01399 & 1.50797 \\ 2.27170 & -2.91294 \end{pmatrix},$$

compare Hannan and Deistler [2012], Theorem 2.6.2. The absolute values of the zeros of the corresponding transfer function lie in (1.18, 1.54) and the absolute values of the poles lie in (1.15, 1.3). The mean squared errors of the unrestricted parameters of the echelon form are summarized in Table 1. We see that with increasing sample size the mean squared errors seem to converge to zero. Note that we consider here 35 unrestricted parameters in Table 1 even though the corresponding echelon ARMA(2, 1) system has only 29 free parameters, i.e. 18 AR parameters, 6 MA parameters and 5 noise parameters, compare Hannan and Deistler [2012] Chapter 2.

As a second example we consider an ARMA(2, 1) model with the same MA part but the zeros of the AR part are closer to the unit circle.

$$(1.5.10) \quad \begin{aligned} & y_t - \begin{pmatrix} -2.20527 & 7.17390 & 3.96550 \\ -0.15429 & 1.38884 & 0.23676 \\ -2.47919 & 5.48425 & 3.98930 \end{pmatrix} y_{t-1} - \begin{pmatrix} 1.81410 & -4.81866 & -2.51307 \\ 0.23189 & -0.85009 & -0.29477 \\ 1.47996 & -3.24972 & -2.02613 \end{pmatrix} y_{t-2} \\ & = \begin{pmatrix} 1.61266 & 0 \\ 1.01399 & 1.50797 \\ 2.27170 & -2.91294 \end{pmatrix} \varepsilon_t + \begin{pmatrix} 3.42768 & 3.18331 \\ 0.28016 & 0.66802 \\ 0.61889 & -4.00152 \end{pmatrix} \varepsilon_{t-1}. \end{aligned}$$

The absolute values of the zeros of the corresponding transfer function lie again in (1.18, 1.54) and the absolute values of the poles lie in (1.56, 1.74). The echelon state space form has parameter matrices

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1.81410 & -2.20527 & -4.81866 & 7.17390 & -2.51307 & 3.96550 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0.23189 & -0.15429 & -0.85009 & 1.38884 & -0.29477 & 0.23676 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1.47996 & -2.47919 & -3.24972 & 5.48425 & -2.02613 & 3.98930 \end{pmatrix}$$

$$B = \begin{pmatrix} 16.15400 & 2.45005 \\ 15.48193 & -19.63442 \\ 1.97747 & 2.07266 \\ 1.75861 & 0.33662 \\ 11.24424 & -7.35204 \\ 10.14145 & -23.03509 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, D = \begin{pmatrix} 1.61266 & 0 \\ 1.01399 & 1.50797 \\ 2.27170 & -2.91294 \end{pmatrix}.$$

As to be expected, the mean squared errors of the AR parameters for this model are larger than for the last model considered but also seem to converge with growing sample size, see Table 2.

Finally, we consider a singular stable and miniphase ARMA(3, 3) system with the Kronecker indices of the corresponding Hankel matrix of the transfer function $m_1 = 3$, $m_2 = 2$, and $m_3 = 1$. The parameter matrices of the state space echelon form of this system are

$$(1.5.11) \quad A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -0.30800 & -0.19358 & -0.96632 & -0.10290 & -0.23966 & -0.65652 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0.60403 & 0.42846 & 1.56074 & 0.39152 & 0.56827 & 0.90352 \\ -0.23963 & -0.21281 & 0 & -0.16083 & -0.16131 & -0.18751 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.03670 & -0.31239 \\ 1.90743 & 0.75658 \\ -0.03840 & 0.15236 \\ -0.46565 & 0.35749 \\ 0.01974 & -0.07971 \\ -1.01956 & -0.19552 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 2.07500 & 0 \\ 0.44091 & 1.55924 \\ 0.45089 & -1.52729 \end{pmatrix}.$$

Note that in this case the free echelon ARMA parameters are exactly those considered in Table 3. Also for this model the mean squared errors of the free parameters seem to converge to zero.

Thus the results of our Monte Carlo simulation seem encouraging for the procedure we proposed.

TABLE 1. Mean squared errors for model (1.5.9)

T	1000	2000	4000	8000	16000
a_{21}	1.39604	0.59406	0.24119	0.13154	0.07452
a_{22}	0.01310	0.00673	0.00260	0.00137	0.00078
a_{23}	0.44414	0.19124	0.07740	0.04216	0.02374
a_{24}	0.32940	0.13452	0.05437	0.02959	0.01689
a_{25}	0.50332	0.21463	0.08644	0.04688	0.02676
a_{26}	0.01546	0.00552	0.00247	0.00145	0.00073
a_{41}	0.20657	0.07587	0.03049	0.02090	0.00923
a_{42}	0.00236	0.00095	0.00039	0.00024	0.00012
a_{43}	0.06620	0.02458	0.00985	0.00668	0.00297
a_{44}	0.04545	0.01642	0.00655	0.00458	0.00196
a_{45}	0.07420	0.02733	0.01089	0.00741	0.00331
a_{46}	0.00178	0.00059	0.00024	0.00020	0.00007
a_{61}	2.02843	0.68686	0.29598	0.18415	0.06824
a_{62}	0.02378	0.00814	0.00420	0.00207	0.00088
a_{63}	0.65948	0.22085	0.09565	0.05916	0.02195
a_{64}	0.43869	0.15056	0.06147	0.04025	0.01488
a_{65}	0.72436	0.24470	0.10435	0.06555	0.02422
a_{66}	0.01806	0.00647	0.00248	0.00175	0.00065
b_{11}	0.07289	0.02836	0.01372	0.00668	0.00377
b_{12}	0.00828	0.00412	0.00157	0.00083	0.00029
b_{21}	0.33429	0.14625	0.06035	0.03329	0.01640
b_{22}	0.15604	0.08761	0.02840	0.01555	0.00674
b_{31}	0.03112	0.01818	0.00932	0.00421	0.00205
b_{32}	0.01158	0.00602	0.00280	0.00145	0.00061
b_{41}	0.13078	0.06646	0.02830	0.01160	0.00741
b_{42}	0.05387	0.02718	0.01137	0.00426	0.00263
b_{51}	0.04867	0.02857	0.01521	0.00609	0.00358
b_{52}	0.03555	0.01716	0.00967	0.00461	0.00245
b_{61}	0.12458	0.04516	0.02685	0.01461	0.00833
b_{62}	0.06763	0.03698	0.02161	0.00807	0.00586
d_{11}	0.00157	0.00069	0.00033	0.00016	0.00009
d_{21}	0.00302	0.00119	0.00072	0.00033	0.00016
d_{22}	0.00129	0.00054	0.00029	0.00016	0.00007
d_{31}	0.01042	0.00658	0.00334	0.00163	0.00065
d_{32}	0.00483	0.00201	0.00107	0.00061	0.00025

TABLE 2. Mean squared errors for model (1.5.10)

T	1000	2000	4000	8000	16000
a_{21}	12.26395	7.70273	1.65868	0.97747	0.09165
a_{22}	6.13558	17.63984	1.17842	0.49677	0.03365
a_{23}	43.38748	85.32671	4.78501	2.37126	0.30552
a_{24}	176.98326	451.69339	33.08721	13.23878	0.80607
a_{25}	26.80685	17.52294	3.90321	2.21355	0.19628
a_{26}	0.57935	6.93584	0.14766	0.03691	0.00287
a_{41}	1.01008	0.70584	0.10223	0.05567	0.00662
a_{42}	0.40794	1.49361	0.06776	0.03049	0.00324
a_{43}	3.68878	15.04505	0.32806	0.15269	0.02428
a_{44}	9.19908	45.75105	1.83640	0.76375	0.06979
a_{45}	2.12246	1.16545	0.23590	0.12609	0.01440
a_{46}	0.03865	0.75635	0.00872	0.00277	0.00040
a_{61}	73.52685	15.44782	0.76441	0.33869	0.10806
a_{62}	17.02306	25.37387	0.63922	0.24087	0.04811
a_{63}	337.59338	329.22158	3.06262	0.89766	0.32390
a_{64}	282.58066	852.20290	21.52508	7.66332	1.05689
a_{65}	148.83477	23.49256	1.84560	0.80465	0.23614
a_{66}	0.96825	13.82036	0.08043	0.02619	0.00483
b_{11}	0.16621	0.06987	0.03337	0.01646	0.00927
b_{12}	0.00674	0.00352	0.00128	0.00068	0.00024
b_{21}	0.80584	0.50762	0.21805	0.09076	0.04838
b_{22}	0.49949	0.22157	0.10498	0.05691	0.03517
b_{31}	0.01068	0.00432	0.00221	0.00123	0.00060
b_{32}	0.00767	0.00392	0.00150	0.00075	0.00036
b_{41}	0.01410	0.00678	0.00242	0.00175	0.00093
b_{42}	0.01010	0.00498	0.00262	0.00128	0.00083
b_{51}	0.12594	0.08207	0.03956	0.01651	0.00934
b_{52}	0.04652	0.02203	0.01217	0.00584	0.00304
b_{61}	0.72838	0.43031	0.23047	0.08688	0.05302
b_{62}	0.48965	0.22836	0.13221	0.05941	0.03500
d_{11}	0.00151	0.00068	0.00033	0.00016	0.00009
d_{12}	0.00000	0.00000	0.00000	0.00000	0.00000
d_{21}	0.00297	0.00119	0.00072	0.00032	0.00016
d_{22}	0.00125	0.00051	0.00029	0.00016	0.00007
d_{31}	0.01042	0.00662	0.00335	0.00164	0.00065
d_{32}	0.00466	0.00191	0.00107	0.00061	0.00025

TABLE 3. Mean squared errors for model (1.5.11)

T	1000	2000	4000	8000	16000
a_{31}	0.00453	0.00157	0.00083	0.00027	0.00016
a_{32}	0.02415	0.00880	0.00399	0.00127	0.00103
a_{33}	0.30890	0.15302	0.03663	0.01780	0.00841
a_{34}	0.01768	0.00788	0.00389	0.00127	0.00083
a_{35}	0.14745	0.06996	0.01886	0.00812	0.00397
a_{36}	0.05350	0.01888	0.00854	0.00387	0.00270
a_{51}	0.01074	0.00329	0.00159	0.00068	0.00038
a_{52}	0.05711	0.02151	0.00859	0.00357	0.00277
a_{53}	0.71374	0.36526	0.08523	0.04253	0.01894
a_{54}	0.04435	0.02010	0.00974	0.00354	0.00239
a_{55}	0.28333	0.14217	0.03334	0.01529	0.00715
a_{56}	0.12289	0.04417	0.01852	0.00980	0.00675
a_{61}	0.00077	0.00043	0.00017	0.00009	0.00005
a_{62}	0.00184	0.00072	0.00040	0.00021	0.00008
a_{64}	0.00147	0.00070	0.00035	0.00013	0.00008
a_{65}	0.00997	0.00359	0.00162	0.00065	0.00039
a_{66}	0.00612	0.00291	0.00116	0.00065	0.00034
b_{11}	0.00444	0.00235	0.00094	0.00044	0.00021
b_{12}	0.00388	0.00227	0.00071	0.00042	0.00022
b_{21}	0.00702	0.00302	0.00139	0.00071	0.00040
b_{22}	0.00168	0.00063	0.00032	0.00016	0.00010
b_{31}	0.00746	0.00370	0.00175	0.00066	0.00034
b_{32}	0.00413	0.00216	0.00093	0.00040	0.00025
b_{41}	0.00287	0.00147	0.00070	0.00033	0.00021
b_{42}	0.00213	0.00103	0.00045	0.00022	0.00019
b_{51}	0.00207	0.00141	0.00054	0.00034	0.00018
b_{52}	0.00089	0.00051	0.00028	0.00012	0.00006
b_{61}	0.00346	0.00160	0.00069	0.00049	0.00020
b_{62}	0.00193	0.00105	0.00047	0.00018	0.00016
d_{11}	0.00236	0.00106	0.00052	0.00027	0.00014
d_{21}	0.00253	0.00115	0.00067	0.00031	0.00014
d_{22}	0.00128	0.00054	0.00031	0.00017	0.00007
d_{31}	0.00227	0.00130	0.00070	0.00034	0.00014
d_{32}	0.00123	0.00051	0.00030	0.00016	0.00007

Part 2

Mixed Frequency

Introduction - Mixed Frequency AR Models

In the second part of this thesis we consider the case that the individual component series of a multivariate time series are available at different sampling frequencies. We call this setting mixed frequency.

In many cases when dealing with high dimensional time series, the individual one dimensional component series are mixed frequency. For example, in economic applications, some time series may be available monthly, such as unemployment data, whereas others may be available quarterly, such as GNP data. Another example of mixed frequency data are time series containing both real and financial data, where financial data, such as interest rates, are typically sampled much more frequently than real data such as industrial production.

As the setting of mixed frequency data is quite common, a lot of approaches have been considered in the literature, see e.g. Zdrozny [1990], Chen and Zdrozny [1998], Ghysels et al. [2006], Ghysels [2012]. Here we follow the approach of Chen and Zdrozny [1998].

We restrict our analysis of mixed frequency data to the case that there is an AR model at the highest frequency generating all the outputs, but we observe these outputs only partially, i.e. we observe some outputs at the highest frequency but others only at an integer submultiple of the highest frequency. Our main focus will be identifiability of the system and noise parameters of the underlying high frequency AR model.

We will consider two different approaches to this problem which will show that generically identifiability can be ensured using second moments which can be observed in principle. In Chapter 2 we consider extended Yule Walker equations which were proposed in Chen and Zdrozny [1998]. In Chapter 3 we consider the technique of blocking which is commonly used in signal processing.

We consider these two different approaches as they both give different insights into the problem of identifiability and they provide two different estimation procedures for $\theta \in \Theta$. A vantage point for the extended Yule Walker equations is their simplicity as they are linear in the second moments. However, they do not use all available second moments. By blocking we are using all second moments which can be observed in principle.

These approaches are both “constructive” in the sense that we present “realization algorithms”, i.e. algorithms which give the parameters from the second moments which can be observed in principle. In addition it is shown that the problem is well-posed in the sense that the parameters depend continuously on these second moments. Thus, when these population second moments are replaced by their sample counterparts, this leads to consistent estimation procedures.

Clearly, the setting of mixed frequency is more likely to occur the higher the dimension of the time series considered in an application. A way of dealing with high dimensional mixed frequency data is modeling the data by a mixed frequency factor model as described in Chapter 5.

Except for Sections 3.3 and 3.4, the content of Chapters 2, and 3 is joint work with the coauthors Manfred Deistler, Brian D. O. Anderson, Bernd Funovits, Lukas Kölbl and Mohsen Zamani and has been presented in Anderson et al. [2014]. Note that the realization algorithm considered here in Section 3.2 works under slightly more general conditions than that of Anderson et al. [2014]. Furthermore we give an alternative proof of Theorem 3.2.3.

The description of the mixed frequency setting is the same as in Anderson et al. [2014]:

For simplicity, we consider the case where the components of the observed process are available at two sampling frequencies, a so-called high frequency and a slow frequency which is an integer multiple N of the high frequency.

Except for Section 2.5, we consider the case where there are stock variables only. We partition the process

$$y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix}$$

into the n_f -dimensional so-called fast component y_t^f which is observed at the highest frequency $t \in \mathbb{Z}$ and the n_s -dimensional so-called slow component y_t^s which is observed only for $t \in N\mathbb{Z}$ ($N \in \mathbb{N}$, $N > 1$), i.e. for every N th time point. Throughout we assume $n_f \geq 1$.

We assume that the *high frequency system*, i.e. the system generating all data at the highest frequency, is a vector autoregression (1.5.2) of order p with parameters (a_1, \dots, a_p, b) where $a(z)$ fulfills the stability assumption $\det a(z) \neq 0$, $|z| \leq 1$ and b has full column rank.

The matrices

$$a_i = \begin{pmatrix} a_{ff}(i) & a_{fs}(i) \\ a_{sf}(i) & a_{ss}(i) \end{pmatrix}, \Sigma = bb' = \begin{pmatrix} \sigma_{ff} & \sigma_{fs} \\ \sigma_{sf} & \sigma_{ss} \end{pmatrix}$$

are partitioned according to y_t^f and y_t^s .

The *population second moments which can be observed in principle*, i.e. the population second moments which can be obtained from the observed components of the process $(y_t)_{t \in \mathbb{Z}}$, are

$$\begin{aligned}
 \gamma^{ff}(h) &= \mathbb{E} \left(y_{t+h}^f (y_t^f)' \right), \quad h \in \mathbb{Z}; \\
 \gamma^{sf}(h) &= \mathbb{E} \left(y_{t+h}^s (y_t^f)' \right), \quad h \in \mathbb{Z}; \\
 \gamma^{ss}(h) &= \mathbb{E} \left(y_{t+h}^s (y_t^s)' \right), \quad h \in N\mathbb{Z}.
 \end{aligned}
 \tag{1.5.12}$$

We focus on the central question of *identifiability*, i.e. whether, for given parameter space Θ , defined in (1.1.5), the parameters a_i and $\Sigma = bb'$ of the high frequency system are uniquely determined by those second moments which can be observed in principle.

Let us introduce the 'blocked observed' process $(\tilde{y}_t)_{t \in N\mathbb{Z}}$,

$$\tilde{y}_t = \begin{pmatrix} y_t \\ y_{t-1}^f \\ \vdots \\ y_{t-N+1}^f \end{pmatrix},$$

whose covariances are exactly the second moments described above.

Note that if identifiability holds and if in addition there is an algorithm for obtaining the parameters of the high frequency system from the population second moments which can be observed in principle, we can reconstruct the missing moments $\gamma^{ss}(h) = \mathbb{E} (y_{t+h}^s (y_t^s)')$, $h \in N\mathbb{Z} - j$, $j \in \{1, \dots, N-1\}$ and thus all $\gamma(h) = \mathbb{E} (y_{t+h} (y_t)')$, $h \in \mathbb{Z}$. Then linear least squares methods for forecasting, nowcasting, and interpolation of non-observed output variables can be applied. If such an algorithm defines a continuous function and thus the problem is well-posed, the algorithm may be applied to sample second moments, in order to yield consistent estimators of the system and noise parameters of the underlying high frequency system and thus of missing second moments. In other words, identifiability and well-posedness are important in obtaining consistent estimators of the system and noise parameters, a_i and Σ respectively. In this case, nowcasts, forecasts and interpolations of high frequency observations based on mixed frequency data are available.

Identifiability Results Using Extended Yule Walker Equations

2.1. Derivation of the Extended Yule Walker Equations for Mixed Frequency Data

By postmultiplying equation (1.1.2) by y'_{t-j} , $j > 0$ and forming expectations, we obtain the extended Yule Walker equations

(2.1.1)

$$\left(\begin{array}{cc|ccc} \gamma^{ff}(1) & \gamma^{fs}(1) & \cdots & \gamma^{ff}(p) & \gamma^{fs}(p) & \cdots \\ \gamma^{sf}(1) & \gamma^{ss}(1) & \cdots & \gamma^{sf}(p) & \gamma^{ss}(p) & \cdots \end{array} \right) = \left(\begin{array}{cc|ccc} a_{ff}(1) & a_{fs}(1) & \cdots & a_{ff}(p) & a_{fs}(p) \\ a_{sf}(1) & a_{ss}(1) & \cdots & a_{sf}(p) & a_{ss}(p) \end{array} \right) \cdot \left(\begin{array}{cc|ccc} \gamma^{ff}(0) & \gamma^{fs}(0) & \cdots & \gamma^{ff}(p-1) & \gamma^{fs}(p-1) & \cdots \\ \gamma^{sf}(0) & \gamma^{ss}(0) & \cdots & \gamma^{sf}(p-1) & \gamma^{ss}(p-1) & \cdots \\ \hline \gamma^{ff}(-p+1) & \gamma^{fs}(-p+1) & \cdots & \gamma^{ff}(0) & \gamma^{fs}(0) & \cdots \\ \gamma^{sf}(-p+1) & \gamma^{ss}(-p+1) & \cdots & \gamma^{sf}(0) & \gamma^{ss}(0) & \cdots \end{array} \right)$$

The problem with equation (2.1.1) is that matrices on both the left and right hand side contain unobserved second moments. In order to overcome this problem, we postmultiply equation (1.1.2) by $(y_{t-j}^f)'$, $j > 0$ and form expectations. Thereby we obtain extended Yule Walker equations (XYW, see Chen and Zadrozny [1998]) as a subsystem of equations of (2.1.1) as

$$(2.1.2) \quad \mathbb{E} \left(y_t \left((y_{t-1}^f) ', (y_{t-2}^f) ', \dots \right) \right) = (A_1, \dots, A_p) \mathbb{E} \left(\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} \left((y_{t-1}^f) ', (y_{t-2}^f) ', \dots \right) \right).$$

Let

$$(2.1.3) \quad K := \mathbb{E} \left(x_t \left(y_{t-1}^f \right)' \right) = \mathbb{E} \left(\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} \left(y_{t-1}^f \right)' \right) = \begin{pmatrix} \gamma^{ff}(0) \\ \gamma^{sf}(0) \\ \vdots \\ \gamma^{ff}(-p+1) \\ \gamma^{sf}(-p+1) \end{pmatrix} = \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

From equation (1.1.3), i.e. $x_{t+1} = \mathcal{A}x_t + \mathcal{B}\varepsilon_t$, we have that $x_t = \sum_{i=0}^{\infty} \mathcal{A}^i \mathcal{B}\varepsilon_{t-i-1}$ and $x_{t+s} = \mathcal{A}^s x_t + \sum_{i=0}^{s-1} \mathcal{A}^i \mathcal{B}\varepsilon_{t+s-i-1}$. The block columns of the second matrix on the right hand side of (2.1.2) are of the form

$$\begin{aligned} \mathbb{E} \left(x_t \left(y_{t-j-1}^f \right)' \right) &= \mathbb{E} \left(x_{t+j} \left(y_{t-1}^f \right)' \right) = \mathbb{E} \left(\left(\mathcal{A}^j x_t + \sum_{i=0}^{j-1} \mathcal{A}^i \mathcal{B}\varepsilon_{t+j-i-1} \right) \left(y_{t-1}^f \right)' \right) \\ &= \mathcal{A}^j \mathbb{E} \left(x_t \left(y_{t-1}^f \right)' \right) = \mathcal{A}^j K, \quad j \geq 0. \end{aligned}$$

Thus the rightmost matrix in the extended Yule Walker equations (2.1.2) can be written as $(K, \mathcal{A}K, \mathcal{A}^2K, \dots)$. From the Cayley-Hamilton Theorem and since $\mathcal{A} \in \mathbb{R}^{np \times np}$, we see that the second matrix on the right hand side of (2.1.2) has full row rank if and only if the matrix consisting of the first np blocks has full row rank. In this way we have obtained our XYW equations which are of the form

$$(2.1.4) \quad \mathbb{E} \left[y_t \left((y_{t-1}^f)' , \dots , (y_{t-np}^f)' \right) \right] = (A_1, \dots, A_p) \mathbb{E} \underbrace{\left[\begin{array}{c} (y_{t-1}) \\ \vdots \\ (y_{t-p}) \end{array} \left((y_{t-1}^f)' , \dots , (y_{t-np}^f)' \right) \right]}_{=Z}.$$

The crucial point is that the matrix Z can be written as

$$(2.1.5) \quad Z = (K, \mathcal{A}K, \mathcal{A}^2K, \dots, \mathcal{A}^{np-1}K),$$

and therefore has the structure of a controllability matrix.

Clearly, the system parameters (a_1, \dots, a_p) of (1.1.2) are identifiable if Z has full row rank np , or equivalently, and in the language of linear system theory, the pair (\mathcal{A}, K) is controllable. Note, however, that contrary to usual controllability matrices compare Lee and Markus [1967] and Wonham [1985], here K depends on \mathcal{A} , which makes the task of verifying controllability more demanding.

REMARK 2.1.1. As will be shown in Section 3.1 full row rank of the matrix Z is a sufficient condition for identifiability of the system parameters, but it is not necessary.

REMARK 2.1.2. Note that since the XYW equations (2.1.4) only use covariances $\gamma^{ff}(h)$ and $\gamma^{sf}(h)$, $h \geq 0$, they can also handle mixed frequency data with more than two sampling frequencies.

REMARK 2.1.3. The advantage of commencing from the XYW equations is that they immediately give linear and consistent estimators.

2.2. Generic Identifiability of System Parameters

The parameter space in this section is the set Θ of all $((a_1, \dots, a_p), \Sigma)$ where $(a_1, \dots, a_p) \in S$ and Σ has rank q as defined in (1.1.5). We analyze identifiability of system parameters first.

The next theorem, which is a central result, shows that the matrix Z in equation (2.1.4) is generically of full row rank and thus we have generic identifiability for (a_1, \dots, a_p) . Note that this holds both for regular and singular AR systems, for all sampling frequency ratios N , and all $n^f \geq 1$.

THEOREM 2.2.1. *The matrix Z in the extended Yule Walker equations (2.1.4) has full row rank np on a generic subset of the parameter space Θ , and thus the system parameters are generically identifiable.*

PROOF. The proof uses the following well known result, see e.g. Lee and Markus [1967], Wonham [1985], Bochnak et al. [1998]:

Let $f : \Theta \rightarrow \mathbb{R}$ be a polynomial function. If there exists a $\theta^* \in \Theta$ such that $f(\theta^*) \neq 0$, then the set of zeros of f is a proper algebraic set and in particular its complement in Θ is generic.

In a first step, we have to show that Z is a rational function of $\theta \in \Theta$. It follows immediately that Z is rational if we can show that K is a rational function of $\theta \in \Theta$. Vectorizing the Lyapunov equation (1.1.4) we obtain

$$\text{vec}\Gamma_p = (\mathcal{A} \otimes \mathcal{A}) \text{vec}\Gamma_p + \text{vec}\mathcal{B}\mathcal{B}'$$

and thus

$$(2.2.1) \quad \text{vec}\Gamma_p = (I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))^{-1} \text{vec}\mathcal{B}\mathcal{B}'.$$

Note that the absolute value of all eigenvalues λ_j of \mathcal{A} is smaller than one by the stability assumption $\det a(z) \neq 0, |z| \leq 1$. Therefore the same holds for the eigenvalues of $(\mathcal{A} \otimes \mathcal{A})$ since the eigenvalues of $(\mathcal{A} \otimes \mathcal{A})$ are $\lambda_i \lambda_j, i, j = 1, \dots, np$ and thus $(I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))$ is non-singular. This implies that $\text{vec}\Gamma_p$ is a rational function in $((a_1, \dots, a_p), \Sigma)$ having no poles in Θ . Thus K and $\mathcal{A}^j K$ and subsequently Z are rational in $((a_1, \dots, a_p), \Sigma)$ on Θ . Without loss of generality we may restrict ourselves to the case where K is a vector and thus Z is square. Multiplying Z by $\det(I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))$ we obtain a polynomial in the entries of $((a_1, \dots, a_p), \Sigma)$ since $\det(I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))$ has no zeros. Thus the set of zeros of the determinant of the polynomial matrix $\det(I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A})) Z$ is the same as the set of zeros of the determinant of Z and thus is an algebraic set in Θ , compare Bochnak et al. [1998] page 23.

Now consider a point θ^* in Θ given by

$$(2.2.2) \quad \mathcal{A} = \begin{pmatrix} 0 & \cdots & 0 & \rho C \\ I_n & & & \\ & \ddots & & \\ & & I_n & 0 \end{pmatrix}, \quad \mathcal{B} = \mathcal{E}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $\rho \in (0, 1)$ and

$$C = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

is a so-called circulant matrix and define $e_1 \in \mathbb{R}^n$, where the first component is one and all others are zero. We will show that for this point in the parameter space, $\det Z \neq 0$ holds.

Note that for an AR(p) process w_t of a system with parameters θ^* the covariances $\gamma(jp)$, $j \in \mathbb{Z}$ are diagonal and all other covariances are zero which is easily seen by looking at the Wold decomposition $w_t = \sum_{j=0}^{\infty} \rho^j C^j e_1 \varepsilon_{t-jp}$. Obviously, $\gamma(0) = \sum_{j=0}^{\infty} \rho^{2j} C^j e_1 e_1' (C^j)'$ is non-singular. Thus $\Gamma_p > 0$ holds and this implies that $(\mathcal{B}, \mathcal{A}\mathcal{B}, \mathcal{A}^2\mathcal{B}, \dots)$ is of full row rank, see Lemma 1.1.5.

Now it is immediate that Z is of full row rank since, as Γ_p is diagonal,

$$Z = (\Gamma_p \mathcal{E}_1, \mathcal{A}\Gamma_p \mathcal{E}_1, \mathcal{A}^2\Gamma_p \mathcal{E}_1, \dots)$$

is a multiple of $(\mathcal{B}, \mathcal{A}\mathcal{B}, \mathcal{A}^2\mathcal{B}, \dots)$. Thus $\det Z \neq 0$ holds.

Thus the set of zeros of $\det Z$ is a proper algebraic set, i.e., an algebraic set of dimension smaller than the dimension of Θ . Therefore its complement in the parameter space, which corresponds to all controllable pairs, is the complement of a proper algebraic set and hence is open and dense in the parameter space. \square

2.3. Generic Identifiability of the Noise Parameters

THEOREM 2.3.1. *The noise parameters Σ are generically identifiable in Θ .*

PROOF. We commence from identifiable system parameters (a_1, \dots, a_p) .

Let us define $\mathcal{G} = (I_n, 0, \dots, 0)$. Through columnwise vectorization of

$$\gamma(0) = \mathbb{E}(y_t y_t') = \mathcal{G} \Gamma_p \mathcal{G}'$$

we obtain

$$\text{vec} \gamma(0) = (\mathcal{G} \otimes \mathcal{G}) \text{vec} \Gamma_p.$$

This together with (2.2.1) gives

$$(2.3.1) \quad \text{vec}\gamma(0) = (\mathcal{G} \otimes \mathcal{G}) (I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))^{-1} (\mathcal{G}' \otimes \mathcal{G}') \text{vec}\Sigma.$$

where we used that $\mathcal{B}\mathcal{B}' = \mathcal{G}'\Sigma\mathcal{G}$.

Note that $(I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))$ is non-singular. For $a_1 = \dots = a_p = 0$, the matrix $(I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))^{-1}$ is triangular with ones on its diagonal. Thus, in view of the particular form of \mathcal{G} , the matrix $(\mathcal{G} \otimes \mathcal{G}) (I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))^{-1} (\mathcal{G}' \otimes \mathcal{G}')$ is a principal submatrix of $(I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))^{-1}$ with the same property and is therefore non-singular. $(\mathcal{G} \otimes \mathcal{G}) (I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))^{-1} (\mathcal{G}' \otimes \mathcal{G}')$ is a function rational in (a_1, \dots, a_p) having no poles. Thus the set of zeros of this function is a proper algebraic set on Θ not depending on Σ . On the complement of this proper algebraic set we have

$$(2.3.2) \quad \text{vec}\Sigma = \left((\mathcal{G} \otimes \mathcal{G}) (I_{(np)^2} - (\mathcal{A} \otimes \mathcal{A}))^{-1} (\mathcal{G}' \otimes \mathcal{G}') \right)^{-1} \text{vec}\gamma(0).$$

□

2.4. General Remarks

REMARK 2.4.1. From Theorems 2.2.1 and 2.3.1 we see that the system and noise parameters are generically identifiable, i.e. identifiable on the intersection of the sets described in the proof of Theorem 2.2.1 and Theorem 2.3.1. Note that the results shown in the proofs above are stronger than the genericity results, because the set where Z has not full row rank np is a proper algebraic set and the same statement holds for the case of noise parameters.

REMARK 2.4.2. Note that the property that Z has full row rank np depends on (a_1, \dots, a_p) as well as on Σ whereas the uniqueness of Σ obtained via (2.3.2) depends only on (a_1, \dots, a_p) . The first assertion is easy to see by considering a two-dimensional AR example: Consider the special cases $a_{ff} \neq a_{ss}$, $a_{fs} = a_{sf} = 0$. If Σ is diagonal, Z has rank 1 otherwise Z has rank 2.

REMARK 2.4.3. We have not been able to give an explicit description of those elements in Θ which are not identifiable or those parameters where Z is not of full row rank np .

If the system (1.1.3) is not controllable, i.e. if Γ_p is singular, then clearly we have non-identifiability even for high frequency data, as the Yule Walker equations then have no unique solution. Note however, that, as will be shown in Section 2.6, in such a situation identifiability might be obtained by suitably prescribing the column degrees in $a(z)$.

2.5. Flow Variables and More General Aggregation Schemes

In the previous sections only stock variables have been considered. Here we deal with the case where the process $(y_t^s)_{t \in \mathbb{Z}}$ consists of flow variables or variables aggregated by more general schemes. For flow variables, the aggregation to the corresponding observed process, $(w_t)_{t \in N\mathbb{Z}}$ say, is of the form

$$(2.5.1) \quad w_t = y_t^s + y_{t-1}^s + \cdots + y_{t-N+1}^s = (1 + z + \cdots + z^{N-1})y_t^s, \quad t \in N\mathbb{Z}.$$

Remember that z denotes the backward shift on \mathbb{Z} .

Note that the second moments required in the extended Yule Walker equations are the autocovariances $\mathbb{E} \left(y_{t+h}^f (y_t^f)' \right)$, $h \in \mathbb{Z}$ and the cross covariances $\mathbb{E} \left(y_{t+h}^s (y_t^f)' \right)$, $h \in \mathbb{Z}$. We now show how these cross covariances can be retrieved from the cross covariances $\mathbb{E} \left(w_{t+h} (y_t^f)' \right)$, $h \in \mathbb{Z}$ of the observations.

To show this, assume for the moment that w_t is available $\forall t \in \mathbb{Z}$ and that the inverse of the linear transformation (2.5.1) exists for $t \in \mathbb{Z}$, i.e.

$$(2.5.2) \quad y_t^s = \text{l.i.m.}_{M \rightarrow \infty} \sum_{j=0}^M h_j^{(M)} w_{t-j}, \quad h_j^{(M)} \in \mathbb{R}^{n_s \times n_s}, \quad t \in \mathbb{Z}$$

where l.i.m. denotes the limit in mean squares. Then

$$(2.5.3) \quad \gamma_{sf}(h) = \mathbb{E} \left(y_{t+h}^s (y_t^f)' \right) = \lim_{M \rightarrow \infty} \sum_{j=0}^M h_j^{(M)} \underbrace{\mathbb{E} \left(w_{t+h-j} (y_t^f)' \right)}_{\gamma_{wyf}(h-j)}.$$

Note that for our purposes the inverse of the linear transformation (2.5.2) only has to exist for the special input $(w_t)_{t \in \mathbb{Z}}$. In order to show the existence of the inverse transformation (2.5.2), it is more convenient to use the frequency domain rather than the time domain, see Rozanov [1967], Hannan [1970]. Let

$$f_{y^s y^s}(\lambda) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \gamma_{ss}(h) e^{-i\lambda h}$$

and

$$f_{ww}(\lambda) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \mathbb{E} w_{t+h} (w_t)^T e^{-i\lambda h}$$

denote the spectral density of $(y_t^s)_{t \in \mathbb{Z}}$ and $(w_t)_{t \in \mathbb{Z}}$, respectively.

As is well known, then the spectral density $f_{ww}(\lambda)$ of $(w_t)_{t \in \mathbb{Z}}$ satisfies

$$f_{ww}(\lambda) = \left(1 + e^{-i\lambda} + \dots + e^{-i(N-1)\lambda}\right) I_{n_s} f_{y^s y^s}(\lambda) I_{n_s} \left(1 + e^{i\lambda} + \dots + e^{i(N-1)\lambda}\right)$$

and thus

$$\begin{aligned} \int \left(1 + e^{-i\lambda} + \dots + e^{-i(N-1)\lambda}\right)^{-1} I_{n_s} f_{ww}(\lambda) I_{n_s} \left(1 + e^{i\lambda} + \dots + e^{i(N-1)\lambda}\right)^{-1} d\lambda \\ = \int f_{y^s y^s}(\lambda) d\lambda < \infty \end{aligned}$$

and therefore each row of $\left(1 + e^{-i\lambda} + \dots + e^{-i(N-1)\lambda}\right)^{-1} I_{n_s}$ is an element of the frequency domain $\mathcal{L}_2(f_{ww} d\lambda)$ of f_{ww} and by the isomorphism between the frequency and the time domain the inverse transformation (2.5.2) is therefore well defined. From (2.5.3) we then obtain

$$(2.5.4) \quad f_{y^s y^f}(\lambda) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \gamma_{sf}(h) e^{-i\lambda h} = \left(1 + e^{-i\lambda} + \dots + e^{-i(N-1)\lambda}\right)^{-1} I_{n_s} f_{w y^f}(\lambda)$$

and thus $\gamma_{sf}(h)$, $h \in \mathbb{Z}$. In this way, we get all covariances in the extended Yule Walker equations. Note that whereas for the case of stock variables these covariances can be directly observed in principle, in the case considered here, they have to be reconstructed as described above.

A completely analogous derivation holds if we replace (2.5.1) by the more general aggregation scheme

$$(2.5.5) \quad w_t = k_0 y_t^s + k_1 y_{t-1}^s + \dots + k_{N-1} y_{t-N+1}^s, \quad k_0 \text{ nonsingular.}$$

Thus, taking into account that generically Z has row rank equal to np , we obtain

THEOREM 2.5.1. *Given the aggregation scheme (2.5.5) for the slow variables $(w_t)_{t \in \mathbb{N}\mathbb{Z}}$, the system and noise parameters of the high frequency system (1.1.2) are generically identifiable from $\gamma_{ff}(h)$ and $\gamma_{wy^f}(h)$, $h \in \mathbb{Z}$.*

Note that if we set $k_0 = I$ and $k_j = 0$, $j = 1, \dots, N-1$, we have the case of stock variables. Thus Theorems 2.2.1 and 2.3.1 are special cases of Theorem 2.5.1.

As is immediately seen, Theorem 2.5.1 also covers the case where the slow variables are formed by a mixture of stock and flow variables.

2.6. Generic Identifiability for Prescribed Column Degrees

In this section we are interested in identifiability of AR systems from mixed frequency data for the case that the column degrees of $a(z)$ rather than the degree of $a(z)$ are prescribed.

As in the high frequency case described in Section 1.4.2, prescribing column degrees may help in obtaining identifiability of the system and noise parameters of singular AR systems in the mixed frequency case.

Singular AR systems are of particular interest for us as they serve as models for latent variables or for static factors in generalized linear dynamic factor models, see Section 1.3.2. In Chapter 5 we will consider mixed frequency factor models where the mixed frequency latent variables or the mixed frequency static factor will be modeled by singular AR systems. Therefore we deal with the problem of identifiability of singular AR systems from mixed frequency data in this section.

As in Section 1.4.2, let p_1, \dots, p_n denote these prescribed column degrees and again let $\Theta_{(p_1, \dots, p_n)}$ denote the subspace of Θ where additionally the highest degree of the respective i th column of $a(z)$ is bounded by p_i and let $(\bar{a}_1, \dots, \bar{a}_p)$ denote all columns of (a_1, \dots, a_p) which are not prescribed to be zero.

Remember that for the regular AR case Γ_p is always non-singular, whereas for the singular case it might be non-singular or singular. For both, the high frequency and the mixed frequency case, prescribing only the degree of $a(z)$ is not enough to ensure identifiability when Γ_p is singular as the (high frequency) Yule Walker equations (1.4.1) do not have a unique solution. As we discussed in Section 1.4.2, in the high frequency case, if Γ_p is singular, identifiability in a restricted parameter space can be obtained by selecting a basis for the row space of Γ_p consisting of the first basis rows. As in Section 1.4.2 let $S\Gamma_p$ denote the matrix formed by these basis rows where S is a $\sum p_i \times np$ selector matrix.

We shall first consider the case $p_i > 0, i = 1, \dots, n$. Let us recall the so-called quasi companion form (1.4.3). Note that in the case of nonzero column degrees in forming $\bar{\mathcal{A}}$ and $\bar{\mathcal{B}}$ only “structural” zeros and ones are deleted in \mathcal{A} and \mathcal{B} . Thus in this case of prescribed column degrees, identifiability of (a_1, \dots, a_p) and Σ is equivalent to the identifiability of $(\bar{a}_1, \dots, \bar{a}_p) = (a_1, \dots, a_p)S'$ and Σ .

We recall that the observation equation of the quasi companion form (1.4.3) is

$$(2.6.1) \quad y_t = (\bar{a}_1, \dots, \bar{a}_p)\bar{x}_t + b\varepsilon_t$$

which we can use to derive the modified XWY equations completely analogously to Section 2.1:

$$(2.6.2) \quad \begin{aligned} \mathbb{E} \left(y_t \left((y_{t-1}^f)', \dots, (y_{t-np+s}^f)' \right) \right) &= (\bar{a}_1, \dots, \bar{a}_p) \mathbb{E} \left(\bar{x}_t \left((y_{t-1}^f)', \dots, (y_{t-np+s}^f)' \right) \right) \\ &= (\bar{a}_1, \dots, \bar{a}_p) \underbrace{(\bar{K}, \bar{\mathcal{A}}\bar{K}, \bar{\mathcal{A}}^2\bar{K}, \dots, \bar{\mathcal{A}}^{np-s-1}\bar{K})}_{=\bar{Z}} \end{aligned}$$

where s is the number of prescribed zero columns in (a_1, \dots, a_p) and $\bar{\mathcal{A}} \in \mathbb{R}^{(np-s) \times (np-s)}$ and where $\bar{\Gamma}_p = \mathbb{E} \bar{x}_t \bar{x}_t'$ and $\bar{K} = \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix}$.

Now obviously the parameter matrices $(\bar{a}_1, \dots, \bar{a}_p)$ are identifiable if the matrix \bar{Z} has full row rank.

In an analogous way as in the two preceding sections we obtain:

THEOREM 2.6.1. *For prescribed nonzero column degrees p_1, \dots, p_n , the system parameters $(\bar{a}_1, \dots, \bar{a}_p)$ and the noise parameters Σ_ν are identifiable on a generic subset of the parameter space $\Theta_{(p_1 \dots p_n)}$. Moreover this statement remains true for more general aggregation schemes (2.5.5).*

PROOF. The proof is along the same lines as the proof of Theorem 2.2.1. A point $\theta^* \in \Theta_{(p_1 \dots p_n)}$ where \bar{Z} has full row rank is constructed as follows: Let

$$a_i(z) = e_i - [a_1]_{\cdot, i} z \dots - [a_p]_{\cdot, i} z^p$$

denote the i th column of $a(z)$ and let

$$C_E = \left([a_{p_1}]_{\cdot, 1}, \dots, [a_{p_n}]_{\cdot, n} \right)$$

be the column end matrix of $a(z)$, compare Remark 1.4.7. Then we take

$$C_E = \rho C, \quad \rho \in (0, 1)$$

where C is the circulant defined in the proof of Theorem 2.2.1 and

$$[A_k]_{\cdot, i} = 0, \quad 0 < k < p_i; \quad i = 1, \dots, n$$

and $b = e_1$. Then again, $\bar{\Gamma}_p = S \Gamma_p S'$ can be shown to be diagonal and non-singular and thus $\det \bar{Z} \neq 0$ holds. Once the system parameters are unique, Σ_ν is obtained in the same way as in the proof of Theorem 2.3.1. \square

We now consider the case where there is at least one i such that $p_i = 0$. In this case we define two subprocesses of $(y_t)_{t \in \mathbb{Z}}$: Let $(y_t^r)_{t \in \mathbb{Z}}$ contain all components of y_t , $y_t^{(i)}$, with $p_i > 0$, $y_t^r = S_1 y_t$, and let $(y_t^z)_{t \in \mathbb{Z}}$ contain all components $y_t^{(i)}$ with $p_i = 0$, $y_t^z = S_z y_t$. It is easy to see that $(y_t^r)_{t \in \mathbb{Z}}$ is again an AR process. We obtain the following theorem:

THEOREM 2.6.2. *For prescribed column degrees p_1, \dots, p_n , the system and noise parameters $(\bar{a}_1, \dots, \bar{a}_p)$ and Σ are identifiable on a generic subset of the parameter space $\Theta_{(p_1 \dots p_n)}$ if $(y_t^r)_{t \in \mathbb{Z}}$ contains at least one fast component. Moreover this statement remains true for more general aggregation schemes (2.5.5).*

PROOF. We first consider the AR process $(y_t^r)_{t \in \mathbb{Z}}$ with parameters not prescribed zero $(\bar{a}_1^r, \dots, \bar{a}_p^r) = S_1(\bar{a}_1, \dots, \bar{a}_p)$. Let us define \bar{x}_t^r as the state of a quasi-companion form of y_t^r . Obviously, using Theorem 2.6.1 we have generic identifiability of $(\bar{a}_1^r, \dots, \bar{a}_p^r)$ from mixed frequency data since

$$\begin{aligned} \bar{Z}^r &= \mathbb{E} \left(\bar{x}_t^r \left((y_{t-1}^{rf})', \dots, (y_{t-np+s}^{rf})' \right) \right) \\ &= \left(\bar{K}^r, \bar{\mathcal{A}}^r \bar{K}^r, (\bar{\mathcal{A}}^r)^2 \bar{K}^r, \dots, (\bar{\mathcal{A}}^r)^{np-s-1} \bar{K}^r \right), \end{aligned}$$

where $\bar{K}^r = \bar{\Gamma}_p^r \begin{pmatrix} I^{(n_r)_f} \\ 0 \end{pmatrix}$ and $\bar{\Gamma}_p^r = \mathbb{E} \bar{x}_t^r (\bar{x}_t^r)'$, is generically of full column rank.

Thus we are left to show generic identifiability of the remaining rows $(\bar{a}_1^z, \dots, \bar{a}_p^z) = S_z(\bar{a}_1, \dots, \bar{a}_p)$, which is easily done since $y_t^z = (\bar{a}_1^z, \dots, \bar{a}_p^z) \bar{x}_t^r + b^z \varepsilon_t$, where b^z are the rows of b corresponding to y_t^z , and thus

$$\mathbb{E} \left(y_t^z \left((y_{t-1}^{rf})', \dots, (y_{t-np+s}^{rf})' \right) \right) = (\bar{a}_1^z, \dots, \bar{a}_p^z) \underbrace{\mathbb{E} \left(\bar{x}_t^r \left((y_{t-1}^{rf})', \dots, (y_{t-np+s}^{rf})' \right) \right)}_{=\bar{Z}^r}.$$

Therefore $(\bar{a}_1, \dots, \bar{a}_p)$ is identifiable if \bar{Z}^r has full row rank which is generic.

Once the system parameters are unique, Σ is obtained in the same way as in the proof of Theorem 2.3.1. \square

CHAPTER 3

Identifiability Results Using Blocking

In this chapter for the case of stock variables only we consider the technique of blocking for the mixed frequency setting. Blocking has been used in signal processing for a number of purposes, see Bittanti et al. [1988], Vaidyanathan [1993], Chen and Francis [1995]. For blocking in case of mixed frequency data see Filler [2010], Ghysels [2012], Zamani [2014].

As in Chapter 2 we obtain generic identifiability results. We present this alternative approach here for three reasons: First, it provides additional insights into the structure of the problem; second, all second moments which are directly observed are used (note that in the XYW equations the available autocovariances of the slow process have not been used), and third, it leads to an alternative estimation procedure.

First we consider the special case of two-dimensional AR(1) systems for which we are able to give an explicit description of the set of all identifiable systems in Section 3.1. In Section 3.2 we consider the more general case of AR(p) systems and in Section 3.3 we consider the case of prescribed column degrees (p_1, \dots, p_n) for $a(z)$ for which we also can show generic identifiability of the parameters $\theta \in \Theta$. For mostly notational convenience, we assume throughout Sections 3.1, 3.2 and 3.3 that $N = 2$ holds. In Section 3.4 we consider the method of blocking for $N > 2$.

3.1. Generic Identifiability for AR(1) Systems

In this section we restrict the analysis of identifiability for mixed frequency data to the case $p = 1$. In addition, we restrict ourselves to the case $N = 2$, $n_f = n_s = 1$. Furthermore, we assume throughout this section that the AR(1) system is regular. As will be shown below, this analysis yields special results; in particular the subset, Θ_I say, of Θ , where identifiability is obtained, can be described explicitly. The complement of this set is a so-called semi-algebraic set, and so we also conclude that for generic parameter values identifiability occurs. This special case is dealt with before we treat the general case in order to give an example and to illustrate the comparative intricacy of identifying the semi-algebraic set on which identifiability cannot be achieved. The AR(1) case has been described in detail in Anderson et al. [2012b].

We first consider the case where in addition $\Sigma = bb'$ is diagonal. Let $\nu_t = b\varepsilon_t$. Using a self-evident notation, we can write

$$(3.1.1) \quad \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} = \underbrace{\begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \end{pmatrix}}_{\mathcal{A}} \begin{pmatrix} y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} + \begin{pmatrix} \nu_t^f \\ \nu_t^s \end{pmatrix}.$$

Now, the one-step-ahead predictor for y_{t-1}^f , $t-1$ odd, based on observed outputs is obtained from the following equation, also easily derived:

$$y_{t-1}^f = a_{ff}y_{t-2}^f + a_{fs}y_{t-2}^s + \nu_{t-1}^f$$

and the two-step-ahead predictor of y_t , t even, is obtained from

$$y_t = \mathcal{A}^2 y_{t-2} + \mathcal{A} \nu_{t-1} + \nu_t.$$

Combining both equations gives a three-dimensional system on $2\mathbb{Z}$:

$$(3.1.2) \quad \underbrace{\begin{pmatrix} y_t^f \\ y_t^s \\ y_{t-1}^f \end{pmatrix}}_{\tilde{y}_t} = \underbrace{\begin{pmatrix} \mathcal{A}^2 & 0 \\ a_{ff} & a_{fs} & 0 \end{pmatrix}}_{\tilde{\mathcal{A}}} \begin{pmatrix} y_{t-2}^f \\ y_{t-2}^s \\ y_{t-3}^f \end{pmatrix} + \underbrace{\begin{pmatrix} \mathcal{A} \nu_{t-1} + \nu_t \\ \nu_{t-1}^f \end{pmatrix}}_{\tilde{\nu}_t}$$

Note that (3.1.2) is an AR(1) system on $2\mathbb{Z}$ whose outputs \tilde{y}_t are the observed variables and thus may serve as a model for the MF data.

The parameter matrices $\tilde{\mathcal{A}}$ and $\Sigma_{\tilde{\nu}} = \mathbb{E} \tilde{\nu}_t \tilde{\nu}_t^T$ are uniquely determined from $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$; however not all entries in $\tilde{\mathcal{A}}$, $\Sigma_{\tilde{\nu}}$ are free, as

$$(3.1.3) \quad \tilde{\mathcal{A}} = \begin{pmatrix} a_{ff}^2 + a_{fs}a_{sf} & a_{ff}a_{fs} + a_{fs}a_{ss} & 0 \\ a_{sf}a_{ff} + a_{ss}a_{sf} & a_{sf}a_{fs} + a_{ss}^2 & 0 \\ a_{ff} & a_{fs} & 0 \end{pmatrix}$$

$$(3.1.4) \quad \Sigma_{\tilde{\nu}} = \begin{pmatrix} \sigma_{ff} & 0 & 0 \\ 0 & \sigma_{ss} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sigma_{ff} & 0 \\ 0 & \sigma_{ss} \end{pmatrix} \begin{pmatrix} a_{ff} & a_{sf} & 1 \\ a_{fs} & a_{ss} & 0 \end{pmatrix}$$

hold.

Here the high frequency system has 6 free parameters, whereas a general AR(1) system for $n = 3$ has 15 free parameters. As the components of \tilde{y}_{t-2} are linearly independent by the regularity assumption, $\tilde{\mathcal{A}}$ and $\Sigma_{\tilde{\nu}}$ are uniquely determined from the second moments which can be observed in principle. In order to analyze identifiability we solve (3.1.3), (3.1.4) for given

$\tilde{\mathcal{A}}, \Sigma_{\tilde{\nu}}$ for the high frequency parameters \mathcal{A} and Σ_{ν} . We see that if a_{fs} and a_{sf} are both zero, then only a_{ss}^2 is unique, otherwise \mathcal{A} and Σ_{ν} are unique and thus we have non-identifiability if and only if $a_{fs} = a_{sf} = 0$ and $a_{ss} \neq 0$ hold.

It is interesting to note that here we have identifiability whenever the two component processes $(y_t^f)_{t \in \mathbb{Z}}$ and $(y_t^s)_{t \in \mathbb{Z}}$ are not orthogonal.

If we drop the assumption $\sigma_{sf} = 0$, we obtain

THEOREM 3.1.1. *Assume that $p = 1$, $n_f = n_s = 1$, $\Sigma_{\nu} > 0$ and $N = 2$. The system and noise parameters $\begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \end{pmatrix}$, σ_{ff} , σ_{sf} and σ_{ss} are not identifiable if and only if they satisfy the equations*

$$(3.1.5) \quad \begin{aligned} a_{fs} &= 0 \\ a_{sf} + \frac{\sigma_{sf}}{\sigma_{ff}}(a_{ss} - a_{ff}) &= 0 \\ a_{ss} &\neq 0. \end{aligned}$$

The complement of the set of solutions of (3.1.5) is a superset of an open and dense set (with respect to the whole parameter space).

PROOF. If σ_{sf} is not necessarily equal to zero, then (3.1.3) remains unchanged and (3.1.4) is changed to

$$(3.1.6) \quad \Sigma_{\tilde{\nu}} = \begin{pmatrix} \sigma_{ff} & \sigma_{sf} & 0 \\ \sigma_{sf} & \sigma_{ss} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sigma_{ff} & \sigma_{sf} \\ \sigma_{sf} & \sigma_{ss} \end{pmatrix} \begin{pmatrix} a_{ff} & a_{sf} & 1 \\ a_{fs} & a_{ss} & 0 \end{pmatrix}.$$

Thus a_{ff} , a_{fs} and σ_{ff} are unique for given $\tilde{\mathcal{A}}, \Sigma_{\tilde{\nu}}$.

We are left with the problem to uniquely solve equation systems (3.1.3) and (3.1.6) in the variables a_{sf} , a_{ss} , σ_{sf} , and σ_{ss} . Thereto we distinguish two cases, namely the case $a_{fs} = 0$ and the case $a_{fs} \neq 0$, considering that we already know a_{fs} .

We start with the case $a_{fs} \neq 0$. It is easy to see that the missing parameters a_{sf} , a_{ss} , σ_{sf} can be recovered using (3.1.3) and (3.1.6). Subsequently σ_{ss} can be recovered using equation (3.1.6).

In the event that $a_{fs} = 0$, then \mathcal{A}^2 is lower triangular, with (2, 2) entry a_{ss}^2 . First, observe that the (2, 3) and (2, 1) entries of $\Sigma_{\tilde{\nu}}$ are $\sigma_{ff}a_{sf} + \sigma_{sf}a_{ss}$ and $a_{ff}(\sigma_{ff}a_{sf} + \sigma_{sf}a_{ss}) + \sigma_{sf}$, respectively. It is immediate that σ_{sf} is available.

Next, if $a_{ss} = 0$, something which is immediately known from the (2, 2) entry of \mathcal{A}^2 , then the (2, 3) entry of $\Sigma_{\tilde{\nu}}$ is simply $\sigma_{ff}a_{sf}$ and since σ_{ff} is obviously nonzero, the value of a_{sf} can

be obtained. Also, the (2,2) entry of $\Sigma_{\tilde{\nu}}$ is $a_{sf}(\sigma_{ff}a_{sf} + \sigma_{sf}a_{ss}) + \sigma_{ss}$ and one immediately has σ_{ss} .

It therefore remains to consider the situation where $a_{ss} \neq 0$. The following quantities α and β , corresponding to the (2,1) entry of \mathcal{A}^2 and the (2,3) entry of $\Sigma_{\tilde{\nu}}$, are known:

$$a_{sf}a_{ff} + a_{ss}a_{sf} = \alpha$$

$$\sigma_{ff}a_{sf} + \sigma_{sf}a_{ss} = \beta$$

By eliminating a_{sf} , we obtain

$$-a_{ss}^2\sigma_{sf} + (\beta - a_{ff}\sigma_{sf})a_{ss} + a_{ff}\beta = \alpha\sigma_{ff}$$

Using this equation and the value for a_{ss}^2 available from \mathcal{A}^2 , it follows that a_{ss} is uniquely determined if and only if

$$\beta - a_{ff}\sigma_{sf} \neq 0$$

Introducing the expression above for β , this yields:

$$a_{sf}\sigma_{ff} + a_{ss}\sigma_{sf} - a_{ff}\sigma_{sf} \neq 0$$

To sum up, identification is possible except for parameters satisfying

$$a_{fs} = 0$$

$$a_{ss} \neq 0$$

$$a_{sf}\sigma_{ff} + a_{ss}\sigma_{sf} - a_{ff}\sigma_{sf} = 0$$

The set of non-identifiable points as described by equations (3.1.5) is a so-called semi-algebraic set, see Bochnak et al. [1998] p 24, Definition 2.1.4, i.e. a set of (multivariate) polynomial zeros where in addition inequalities are imposed. Here, in particular, the set of all identifiable parameters, which is a complement of the semi-algebraic set above, contains a generic subset of the parameter space, viz. the complement of the set defined by the zeros of the polynomial equalities alone. \square

REMARK 3.1.2. An interesting interpretation of Theorem 3.1.1 is the following. The parameters of the underlying high frequency model cannot be obtained if and only if there is a static linear transformation such that the transformed model has a diagonal innovation variance and the transformed system matrix is diagonal with nonzero (2,2) entry. Note that such a transformation must be of the form $T = \begin{pmatrix} 1 & 0 \\ -\sigma_{sf}\sigma_{ff}^{-1} & 1 \end{pmatrix}$ and that for given T the conditions non-identifiability arising are exactly the same as in (3.1.5). Thus identifiability for systems

with non-diagonal innovation variance can be traced back to identifiability for systems with diagonal innovation variance.

REMARK 3.1.3. Consider again the two-dimensional AR(1) process of Section 3.1. In this case the matrix Z in equation (2.1.4) is rank deficient for $\tilde{a}_{ss} = 0$, $\tilde{a}_{sf} = 0$ even if $\tilde{a}_{fs} \neq 0$ which shows that the condition that $rk(Z) = np$ hold, is not necessary for identifiability. For the example discussed in the previous section, where for the case $\sigma_{sf} = a_{fs} = a_{sf} = 0$ the classes of observationally equivalent parameters consist of two points (corresponding to the two choices for the square root of a_{ss}^2), the solution set of the XYW is a nontrivial affine subset. This shows that the XYW do not use the full information contained in the second moments which are in principle observed.

Note that equations (3.1.3), (3.1.4) may also be used for identifiability analysis for $n = q > 2$ though dealing with the various special cases is more intricate since scalars are replaced by matrices. The analysis above cannot be extended to the case $p > 1$; its advantage on the other hand is that the subset of identifiable parameters is explicitly given, and the genericity property stands out clearly.

To clarify the restrictions of this method, we consider a two dimensional AR(2) example

$$y_t = A_2 y_{t-2} + \nu_t \quad t \in \mathbb{Z}$$

where A_2 is non-diagonal with Wold representation $y_t = \sum_{j=0}^{\infty} A_2^j \nu_{t-2j}$. Obviously, there holds $\mathbb{E}y_{t+2h-1}y_t' = 0, \forall h \in \mathbb{Z}$. Thus the blocked process on $t \in 2\mathbb{Z}$

$$\begin{pmatrix} y_t^f \\ y_t^s \\ y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} = \begin{pmatrix} A_2 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} y_{t-2}^f \\ y_{t-2}^s \\ y_{t-3}^f \\ y_{t-3}^s \end{pmatrix} + \begin{pmatrix} \nu_t \\ \nu_{t-1} \end{pmatrix}$$

consists of two orthogonal subprocesses, namely $(y_t)_{t \in 2\mathbb{Z}}$ and $(y_{t-1})_{t \in 2\mathbb{Z}}$, which are AR(1) processes on $2\mathbb{Z}$. Therefore in the mixed frequency case, when we do not observe $(y_t^s)_{t \in 2\mathbb{Z}}$, to determine whether $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ is an AR process, we only have to look at the subspectrum for $(y_t)_{t \in 2\mathbb{Z}}$, whose spectral density may look like this:

$$f(z^2) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{(az^2+1)(cz^2+1)} & 0 \\ 0 & \frac{1}{(az^2+1)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

where z^2 is a symbol of its own. The subspectrum for $(y_t^s)_{t \in 2\mathbb{Z}}$ is $\frac{1}{(az^2+1)(cz^2+1)} + \frac{1}{(az^2+1)}$. Clearly, this subspectrum has a zero at $z^2 = -\frac{2}{c}$ and therefore is the spectrum of an ARMA, not an AR process.

The fact that AR systems are not closed under marginalization is well known, see e.g. Amemiya and Wu [1972], Tiao [1972].

3.2. Generic Identifiability for AR(p) Systems

In the last section we already discussed the special case of two-dimensional AR(1) systems. In this section we discuss AR(p) systems $p \geq 1$. We restrict ourselves to the case that $\Gamma_p > 0$.

We commence from the high frequency process $(y_t)_{t \in \mathbb{Z}}$. In blocked form this process can be written as $(Y_t)_{t \in 2\mathbb{Z}}$ where $Y_t = \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}$.

Note that the state spaces for even t of $(y_t)_{t \in \mathbb{Z}}$ and $(Y_t)_{t \in 2\mathbb{Z}}$ are the same, see Section 1.5, and both processes are AR processes with minimal state dimension smaller than or equal to np and equal to np if and only if a_p is non-singular. Throughout this section, we assume that a_p is non-singular. In particular, we have from (1.1.3) the state space representation

$$(3.2.1) \quad x_{t+1} = \underbrace{\mathcal{A}^2}_{A_b} x_{t-1} + \underbrace{(\mathcal{B}, \mathcal{A}\mathcal{B})}_{B_b} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}$$

$$(3.2.2) \quad Y_t = \begin{pmatrix} \left(\begin{array}{cccc} I_n & 0 & 0 & \cdots & 0 \end{array} \right) \mathcal{A}^2 \\ \left(\begin{array}{cccc} I_n & 0 & 0 & \cdots & 0 \end{array} \right) \mathcal{A} \end{pmatrix} x_{t-1} + \begin{pmatrix} b & a_1 b \\ 0 & b \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}$$

for $t \in 2\mathbb{Z}$, where, if for notational simplicity we consider the case of p even, the minimal state is given as

$$x_{t+1} = \begin{pmatrix} Y_t \\ \vdots \\ Y_{t-p+2} \end{pmatrix}.$$

Now, for the mixed frequency case, for stock variables and $N = 2$, we use the blocked observed process as $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$, $\tilde{y}_t = \begin{pmatrix} y_t \\ y_{t-1}^f \end{pmatrix}$ exactly as in Section 3.1. Note that the second moments of $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ are precisely those second moments (1.5.12) which can be observed in principle.

From (3.2.1), (3.2.2) we obtain the following state space representation for $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$

$$(3.2.3) \quad x_{t+1} = A_b x_{t-1} + B_b \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}$$

$$(3.2.4) \quad \tilde{y}_t = \underbrace{\begin{pmatrix} \left(\begin{matrix} I_n & 0 & 0 & \cdots & 0 \end{matrix} \right) \mathcal{A}^2 \\ \left(\begin{matrix} I_{n_f} & 0 & 0 & \cdots & 0 \end{matrix} \right) \mathcal{A} \end{pmatrix}}_{C_b} x_{t-1} + \underbrace{\begin{pmatrix} b & a_1 b \\ 0 & (I_{n_f}, 0_{n_f \times n_s}) b \end{pmatrix}}_{D_b} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}.$$

Whereas in (3.2.1), (3.2.2), as easily can be seen, $\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}_{t \in 2\mathbb{Z}}$ are innovations for $(Y_t)_{t \in 2\mathbb{Z}}$,

$\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}_{t \in 2\mathbb{Z}}$ are not innovations for \tilde{y}_t , as is already clear for the case $n = q$ from considering the dimensions of the respective vectors.

The spectral density of $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$, using (3.2.3), (3.2.4), can be represented as

$$(3.2.5) \quad f_{\tilde{y}}(z^2) = \left(C_b \left(I_{np} (z^2)^{-1} - A_b \right)^{-1} B_b + D_b \right) \left(B'_b \left(I_{np} z^2 - A'_b \right)^{-1} C'_b + D'_b \right)$$

where the spectral factor on the r.h.s. of (3.2.5) is “fat” and not miniphase. On the other hand the spectral density of $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ can be written as

$$(3.2.6) \quad f_{\tilde{y}}(z^2) = k(z^2) k'(z^{-2})$$

where $k(z^2)$ is a stable and miniphase spectral factor. Therefore for a suitable quadruple $(\bar{A}_b, \bar{B}_b, \bar{C}_b, \bar{D}_b)$ denoting a minimal, stable and miniphase state space system we have

$$(3.2.7) \quad k(z^2) = \left(\bar{C}_b \left(I_{np} (z^2)^{-1} - \bar{A}_b \right)^{-1} \bar{B}_b + \bar{D}_b \right).$$

The question of the relation of the state dimensions of minimal, stable and miniphase state space systems for $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ and $(Y_t)_{t \in 2\mathbb{Z}}$ arises. The next theorem states that despite the fact that the unobserved outputs have been omitted in \tilde{y}_t , generically the McMillan degrees in (3.2.1), (3.2.2) and in $k(z^2) = \left(\bar{C}_b \left(I_{np} (z^2)^{-1} - \bar{A}_b \right)^{-1} \bar{B}_b + \bar{D}_b \right)$ are the same.

For the proof of the theorem, we will need the following lemma:

LEMMA 3.2.1. *If for eigenvalues $\lambda_i \neq 0$ of A $\lambda_i \neq \lambda_j$ implies $\lambda_i^2 \neq \lambda_j^2$, then the kernels of $(A - \lambda_i I)^m$ and $(A^2 - \lambda_i^2 I)^m$ are the same $\forall m \in \mathbb{N}$.*

PROOF. To show

$$\ker(A - \lambda I)^m \subseteq \ker(A^2 - \lambda^2 I)^m$$

is trivial since $(A^2 - \lambda^2 I)^m = (A + \lambda I)^m (A - \lambda I)^m$. To show

$$\ker(A - \lambda I)^m \supseteq \ker(A^2 - \lambda^2 I)^m$$

consider a vector $v \notin \ker(A - \lambda I)^m$. Thus

$$(A^2 - \lambda^2 I)^m v = (A + \lambda I)^m \underbrace{(A - \lambda I)^m v}_{\neq 0} \neq 0$$

since $(A + \lambda I)$ is non-singular because of our assumption $-\lambda$ cannot be an eigenvalue of A . \square

THEOREM 3.2.2. *For $((a_1, \dots, a_p), \Sigma_\nu) \in \Theta$, if a_p non-singular, $\Gamma_p > 0$, and if for eigenvalues of \mathcal{A} such that $\lambda_i \neq \lambda_j$ it follows that $\lambda_i^2 \neq \lambda_j^2$ holds, the McMillan degree of a causal and miniphase spectral factor $k(z^2)$ of $f_{\tilde{y}}(z^2)$ is equal to np .*

PROOF. First note that the McMillan degree of $k(z^2)$ is equal to the rank of the Hankel matrix of covariances \mathcal{H}^γ :

Let $k(z^2) = \sum_{j=0}^{\infty} k_{2j} z^{2j}$ be the power series expansion of $k(z^2)$. As $\tilde{\gamma}(2j) = \mathbb{E} \tilde{y}_{2j} \tilde{y}'_0 = \sum_{i=0}^{\infty} k_{2j+2i} k'_{2i}$ we have

$$\underbrace{\begin{pmatrix} \tilde{\gamma}(2) & \tilde{\gamma}(4) & \tilde{\gamma}(6) & \cdots \\ \tilde{\gamma}(4) & \tilde{\gamma}(6) & \tilde{\gamma}(8) & \cdots \\ \tilde{\gamma}(6) & \tilde{\gamma}(8) & \tilde{\gamma}(10) & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}^\gamma} = \underbrace{\begin{pmatrix} k_2 & k_4 & k_6 & \cdots \\ k_4 & k_6 & k_8 & \cdots \\ k_6 & k_8 & k_{10} & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}^k} \begin{pmatrix} k'_0 & & & \\ k'_2 & k'_0 & & \\ k'_4 & k'_2 & k'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix}$$

where the second matrix on the right hand side is of full row rank since $k(z^2)$ is miniphase. Thus we have that $\text{rk} \mathcal{H}^\gamma = \text{rk} \mathcal{H}^k$ holds. Since a spectral miniphase factor of the spectral density $f_Y(z^2)$ of $(Y_t)_{t \in 2\mathbb{Z}}$ has McMillan degree np , the McMillan degree of $k(z^2)$ must be smaller than or equal to np .

Thus it remains to prove that a finite submatrix of \mathcal{H}^γ has rank np :

$$\begin{aligned}
\mathcal{H}_{np}^\gamma &= \mathbb{E} \begin{pmatrix} \tilde{y}_{t+2} \\ \vdots \\ \tilde{y}_{t+2np} \end{pmatrix} \begin{pmatrix} \tilde{y}'_t & \cdots & \tilde{y}'_{t-2(np-1)} \end{pmatrix} \in \mathbb{R}^{(n+n_f)np \times (n+n_f)np} \\
&= \mathbb{E} \begin{pmatrix} y_{t+2} \\ y_{t+1}^f \\ y_{t+4} \\ y_{t+3}^f \\ \vdots \\ y_{t+2np} \\ y_{t+2np-1}^f \end{pmatrix} \begin{pmatrix} y'_t & (y_{t-1}^f)' & \cdots & y'_{t-2np+2} & (y_{t-2np+1}^f)' \end{pmatrix} \\
&= \begin{pmatrix} \tilde{\gamma}(2) & \tilde{\gamma}(4) & \cdots & \tilde{\gamma}(2np) \\ \tilde{\gamma}(4) & \tilde{\gamma}(6) & \cdots & \\ \vdots & \vdots & \ddots & \\ \tilde{\gamma}(2np) & & & \tilde{\gamma}(4np-2) \end{pmatrix} \\
&= \left(\begin{array}{cc|cc|c} \gamma(2) & [\gamma(3)]_{\cdot,1:n_f} & \gamma(4) & [\gamma(5)]_{\cdot,1:n_f} & \cdots \\ [\gamma(1)]_{1:n_f,\cdot} & \gamma^{ff}(2) & [\gamma(3)]_{1:n_f,\cdot} & \gamma^{ff}(4) & \cdots \\ \hline \gamma(4) & [\gamma(5)]_{\cdot,1:n_f} & \gamma(6) & [\gamma(7)]_{\cdot,1:n_f} & \cdots \\ [\gamma(3)]_{1:n_f,\cdot} & \gamma^{ff}(4) & [\gamma(5)]_{1:n_f,\cdot} & \gamma^{ff}(6) & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \\ \hline \gamma(2np) & [\gamma(2np+1)]_{\cdot,1:n_f} & \gamma(2np+2) & [\gamma(2np+3)]_{\cdot,1:n_f} & \cdots \\ [\gamma(2np-1)]_{1:n_f,\cdot} & \gamma^{ff}(2np) & [\gamma(2np+1)]_{1:n_f,\cdot} & \gamma^{ff}(2np+2) & \cdots \\ \hline \cdots & \gamma(2np) & [\gamma(2np+1)]_{\cdot,1:n_f} & & \\ \cdots & [\gamma(2np-1)]_{1:n_f,\cdot} & \gamma^{ff}(2np) & & \\ \hline \cdots & \gamma(2np+2) & [\gamma(2np+3)]_{\cdot,1:n_f} & & \\ \cdots & [\gamma(2np+1)]_{1:n_f,\cdot} & \gamma^{ff}(2np+2) & & \\ \hline & \vdots & \vdots & & \\ \hline \cdots & \gamma(4np-2) & [\gamma(4np-1)]_{\cdot,1:n_f} & & \\ \cdots & [\gamma(4np-3)]_{1:n_f,\cdot} & \gamma^{ff}(4np-2) & & \end{array} \right)
\end{aligned}$$

$$= \left(\begin{array}{cc|cc|cc}
(I_n & 0) \mathcal{A}^2 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^3 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^4 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^5 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots \\
(I_{n_f} & 0) \mathcal{A} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^2 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^3 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^4 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots \\
\hline
(I_n & 0) \mathcal{A}^4 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^5 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^6 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^7 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \\
(I_{n_f} & 0) \mathcal{A}^3 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^4 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^5 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^6 \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \\
\hline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
\hline
(I_n & 0) \mathcal{A}^{2np} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{2np+1} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{2np+2} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{2np+3} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots \\
(I_{n_f} & 0) \mathcal{A}^{2np-1} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^{2np} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^{2np+1} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^{2np+2} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots \\
\hline
\dots & (I_n & 0) \mathcal{A}^{2np} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{2np+1} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & & & \\
\dots & (I_{n_f} & 0) \mathcal{A}^{2np-1} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^{2np} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & & & \\
\hline
(I_n & 0) \mathcal{A}^{2np+2} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{2np+3} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & & & & \\
(I_{n_f} & 0) \mathcal{A}^{2np+1} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^{2np+2} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & & & & \\
\hline
\vdots & \vdots & & & & & \\
\hline
\dots & (I_n & 0) \mathcal{A}^{4np-2} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{4np-1} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & & & \\
\dots & (I_{n_f} & 0) \mathcal{A}^{4np-3} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \mathcal{A}^{4np-2} \Gamma_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & & &
\end{array} \right)$$

We look at the following submatrix of \mathcal{H}_{np}^γ

$$= \begin{pmatrix}
(I_n & 0) \mathcal{A}^2 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^4 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^6 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & (I_n & 0) \mathcal{A}^{2np} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \\
(I_n & 0) \mathcal{A}^4 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^6 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^8 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & (I_n & 0) \mathcal{A}^{2np+2} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \\
(I_n & 0) \mathcal{A}^6 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^8 \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{10} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \\
(I_n & 0) \mathcal{A}^{2np} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \mathcal{A}^{2np+2} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & & & & (I_n & 0) \mathcal{A}^{4np-2} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix}
\end{pmatrix} \\
= \begin{pmatrix}
\begin{pmatrix} I_n & 0 \\ I_n & 0 \\ I_n & 0 \\ \vdots \\ I_n & 0 \end{pmatrix} \mathcal{A}^2 & \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \mathcal{A}^2 \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \mathcal{A}^4 \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & \mathcal{A}^{2np-2} \begin{pmatrix} I_n \\ 0 \end{pmatrix}
\end{pmatrix} = \mathcal{O}\mathcal{C} = \mathcal{H}$$

Since we assumed that the eigenvalues are nonzero, $\lambda_i \neq 0$, and for eigenvalues $\lambda_i \neq \lambda_j$ of \mathcal{A} $\lambda_i^2 \neq \lambda_j^2$ holds, it is easy to see from Lemma 3.2.1 that q_i is an eigenvector of \mathcal{A}^2 if and only if q_i is an eigenvector of \mathcal{A} .

Observe that \mathcal{O} is of full column rank. To show this we are using the PBH Test, Theorem 1.1.3, and the fact that for any right eigenvector q_i of \mathcal{A} the first n components are not all equal to zero, as shown in Anderson et al. [2012b], Lemma 2:

$$\begin{pmatrix} (\mathcal{A}^2 - \lambda_i^2 I) \\ (I_n \quad 0) \mathcal{A}^2 \end{pmatrix} q_i = \begin{pmatrix} 0 \\ [\lambda_i^2 q_i]_{1:n} \neq 0 \end{pmatrix}$$

Also observe that \mathcal{C} is of full row rank if and $\Gamma_p > 0$. Again we are using the PBH Test, Theorem 1.1.3:

We have to test for all left eigenvectors q'_i of \mathcal{A}^2 or equivalently \mathcal{A} that

$$q'_i \left((\mathcal{A}^2 - \lambda_i^2 I), \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \right) = \left(0, q'_i \begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix} \right) \neq 0$$

Thus if q_i is orthogonal to $\begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix}$ also

$$q'_i \mathcal{A}^j \begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix} = 0, \forall j \in \mathbb{N},$$

holds which implies $q'_i \Gamma_p = 0$ which is in contradiction to $\Gamma_p > 0$.

Therefore for all eigenvectors q_i of \mathcal{A}^2

$$\left(0, q'_i \begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix} \right) \neq 0.$$

Now, according to Theorem 2.3.2 in Hannan and Deistler [2012] \mathcal{H} has rank np . \square

The next theorem can be used to determine the relation between (\bar{A}_b, \bar{C}_b) and (A_b, C_b) . It is stated in a general form, for spectral factors which are rational in z , but, as is easily seen, it also holds for the “special case” that the spectral factors are rational in z^2 .

We define the McMillan degree of a spectral density f_y to be twice the McMillan degree of a stable and miniphase spectral factor which is equivalent to twice the rank of the Hankel matrix of covariances of (y_t) .

THEOREM 3.2.3. *Let $f_y(z)$ be a rational spectral density with McMillan degree $2m$ and let $k(z)$ and $\tilde{k}(z)$ be spectral factors of $f_y(z)$ with the same McMillan degree m but not necessarily the same dimensions. Then for minimal realizations of $k(z)$ and $\tilde{k}(z)$ the matrices (A, C) and (\tilde{A}, \tilde{C}) are the same up to basis change, i.e.*

$$\begin{aligned}\tilde{A} &= T^{-1}AT \\ \tilde{C} &= CT\end{aligned}$$

i.e. for a suitable chosen non-singular $m \times m$ matrix T .

PROOF. First we show how to determine a minimal state of $(y_t)_{t \in \mathbb{Z}}$ and a corresponding minimal state space realization of $k(z)$ from the Hankel matrix of the transfer function $k(z)$, see Ho and Kalman [1966], Akaike [1974] or Hannan and Deistler [2012] Chapter 2. Let (ε_t) be the inputs of the transfer function $k(z)$ such that $y_t = k(z)\varepsilon_t$. Clearly, we have

$$\begin{pmatrix} y_t \\ y_{t+1} \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \\ k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}^k} \begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix} + \begin{pmatrix} k_0 & & & \\ k_1 & k_0 & & \\ k_2 & k_1 & k_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t+1} \\ \vdots \end{pmatrix}.$$

A minimal state x_t of $(y_t)_{t \in \mathbb{Z}}$ is a basis of the space spanned by the projection of the future of $(y_t)_{t \in \mathbb{Z}}$ onto the space spanned by the past of $(y_t)_{t \in \mathbb{Z}}$:

$$\begin{pmatrix} y_{t|t-1} \\ y_{t+1|t-1} \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \\ k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}^k} \begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix}$$

Let S be a selector matrix selecting a basis of this projection (where S does not necessarily select the first basis).

$$x_t = S \begin{pmatrix} y_{t|t-1} \\ y_{t+1|t-1} \\ \vdots \end{pmatrix} = S \underbrace{\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \\ k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}_\alpha^k} \begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix}$$

A state equation of a minimal state space system can be obtained as follows:

$$\begin{aligned}
x_{t+1} &= S \begin{pmatrix} y_{t+1|t} \\ y_{t+2|t} \\ \vdots \end{pmatrix} \\
&= S \underbrace{\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \\ k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}_\alpha^k} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \end{pmatrix} \\
&= S \underbrace{\begin{pmatrix} k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ k_4 & k_5 & k_6 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\mathcal{H}_{\alpha+n}^k} \begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix} + S \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ \vdots \end{pmatrix} \varepsilon_t \\
&= A S \underbrace{\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \\ k_2 & k_3 & k_4 & \cdots \\ k_3 & k_4 & k_5 & \cdots \\ \vdots & \vdots & & \ddots \end{pmatrix}}_{\substack{\mathcal{H}_\alpha^k \\ x_t}} \begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix} + S \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ \vdots \end{pmatrix} \varepsilon_t.
\end{aligned}$$

Thus the state transition matrix A can be uniquely obtained as a solution of the linear equation system

$$(3.2.8) \quad \mathcal{H}_{\alpha+n}^k = A \mathcal{H}_\alpha^k.$$

An observation equation can be obtained as follows:

$$\begin{aligned} y_t &= \begin{pmatrix} k_1 & k_2 & k_3 & \cdots \end{pmatrix} \begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix} + k_0 \varepsilon_t \\ &= \underbrace{C \mathcal{H}_\alpha^k}_{x_t} \begin{pmatrix} \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \end{pmatrix} + k_0 \varepsilon_t. \end{aligned}$$

Therefore C can be determined as the unique solution of the linear equation system

$$(3.2.9) \quad \begin{pmatrix} k_1 & k_2 & k_3 & \cdots \end{pmatrix} = C \mathcal{H}_\alpha^k.$$

In the next step we discuss the connection of the row spaces of the Hankel matrices of the transfer functions $k(z)$ and $\tilde{k}(z)$. We use the relation of the Hankel matrices of the transfer functions $k(z)$ and $\tilde{k}(z)$ and \mathcal{H}^γ , the Hankel matrix of the covariances of $(y_t)_{t \in \mathbb{Z}}$:

$$\mathcal{H}^\gamma = \mathcal{H}^k \begin{pmatrix} k'_0 & & & \\ k'_1 & k'_0 & & \\ k'_2 & k'_1 & k'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix}$$

and

$$\mathcal{H}^\gamma = \mathcal{H}^{\tilde{k}} \begin{pmatrix} \tilde{k}'_0 & & & \\ \tilde{k}'_1 & \tilde{k}'_0 & & \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix}.$$

Note that the left kernels of \mathcal{H}^k and $\mathcal{H}^{\tilde{k}}$ are subsets of the left kernel of \mathcal{H}^γ . Thus we have the following: Let S be a selector matrix selecting a basis of the row space of \mathcal{H}^γ . Then S also selects bases of the row spaces of \mathcal{H}^k and $\mathcal{H}^{\tilde{k}}$.

Let $\mathcal{H}_\alpha^k = S \mathcal{H}^k$ and $\mathcal{H}_{\alpha+n}^k$ be defined as above. Define $\mathcal{H}_\alpha^{\tilde{k}} = S \mathcal{H}^{\tilde{k}}$ and $\mathcal{H}_{\alpha+n}^{\tilde{k}}$ accordingly. Then

$$\mathcal{H}_\alpha^k \begin{pmatrix} k'_0 & & & \\ k'_1 & k'_0 & & \\ k'_2 & k'_1 & k'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix} = \mathcal{H}_\alpha^{\tilde{k}} \begin{pmatrix} \tilde{k}'_0 & & & \\ \tilde{k}'_1 & \tilde{k}'_0 & & \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix}$$

and

$$\mathcal{H}_{\alpha+n}^k \begin{pmatrix} k'_0 \\ k'_1 & k'_0 \\ k'_2 & k'_1 & k'_0 \\ \vdots & \vdots & \ddots \end{pmatrix} = \mathcal{H}_{\alpha+n}^{\tilde{k}} \begin{pmatrix} \tilde{k}'_0 \\ \tilde{k}'_1 & \tilde{k}'_0 \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 \\ \vdots & \vdots & \ddots \end{pmatrix}$$

hold. Therefore

$$\mathcal{H}_{\alpha+n}^k \begin{pmatrix} k'_0 \\ k'_1 & k'_0 \\ k'_2 & k'_1 & k'_0 \\ \vdots & \vdots & \ddots \end{pmatrix} = A \mathcal{H}_{\alpha}^k \begin{pmatrix} k'_0 \\ k'_1 & k'_0 \\ k'_2 & k'_1 & k'_0 \\ \vdots & \vdots & \ddots \end{pmatrix}$$

with transition matrix A from (3.2.8) implies

$$\mathcal{H}_{\alpha+n}^{\tilde{k}} \begin{pmatrix} \tilde{k}'_0 \\ \tilde{k}'_1 & \tilde{k}'_0 \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 \\ \vdots & \vdots & \ddots \end{pmatrix} = A \mathcal{H}_{\alpha}^{\tilde{k}} \begin{pmatrix} \tilde{k}'_0 \\ \tilde{k}'_1 & \tilde{k}'_0 \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

W.l.o.g. let $\tilde{k}(z)$ be miniphase then \tilde{k}'_0 has full column rank and the last equation above implies

$$\mathcal{H}_{\alpha+n}^{\tilde{k}} = A \mathcal{H}_{\alpha}^{\tilde{k}}.$$

Thus A is also the state transition matrix of a minimal state space realization of $\tilde{k}(z)$.

Analogously, we treat the observation equation. Since

$$\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \end{pmatrix} \begin{pmatrix} k'_0 \\ k'_1 & k'_0 \\ k'_2 & k'_1 & k'_0 \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \tilde{k}_1 & \tilde{k}_2 & \tilde{k}_3 & \cdots \end{pmatrix} \begin{pmatrix} \tilde{k}'_0 \\ \tilde{k}'_1 & \tilde{k}'_0 \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 \\ \vdots & \vdots & \ddots \end{pmatrix}$$

we have that

$$\begin{pmatrix} k_1 & k_2 & k_3 & \cdots \end{pmatrix} \begin{pmatrix} k'_0 \\ k'_1 & k'_0 \\ k'_2 & k'_1 & k'_0 \\ \vdots & \vdots & \ddots \end{pmatrix} = C \mathcal{H}_{\alpha}^k \begin{pmatrix} k'_0 \\ k'_1 & k'_0 \\ k'_2 & k'_1 & k'_0 \\ \vdots & \vdots & \ddots \end{pmatrix}$$

with C as in (3.2.9) implies

$$\begin{pmatrix} \tilde{k}_1 & \tilde{k}_2 & \tilde{k}_3 & \cdots \end{pmatrix} \begin{pmatrix} \tilde{k}'_0 & & & \\ \tilde{k}'_1 & \tilde{k}'_0 & & \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix} = C \mathcal{H}_\alpha^{\tilde{k}} \begin{pmatrix} \tilde{k}'_0 & & & \\ \tilde{k}'_1 & \tilde{k}'_0 & & \\ \tilde{k}'_2 & \tilde{k}'_1 & \tilde{k}'_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix}.$$

Thus since we assumed w.l.o.g. that $\tilde{k}(z)$ is miniphase we have that

$$\begin{pmatrix} \tilde{k}_1 & \tilde{k}_2 & \tilde{k}_3 & \cdots \end{pmatrix} = C \mathcal{H}_\alpha^{\tilde{k}}.$$

Therefore any \tilde{A} and \tilde{C} of a minimal state space realization of $\tilde{k}(z)$ are related to A and C of a minimal state space realization of $k(z)$ via a basis change.

□

Thus we have that

$$(3.2.10) \quad \bar{A}_b = T^{-1} A_b T$$

$$(3.2.11) \quad \bar{C}_b = C_b T$$

REMARK 3.2.4. Note that Theorem 3.2.2 is essential for (3.2.10), (3.2.11) because it ensures that \bar{A}_b and A_b are of the same dimension. Also note that the result of Theorem 3.2.3 holds despite the fact that the states in (3.2.3), (3.2.4) and in the minimal, stable and miniphase system corresponding to (3.2.7) are not the same not even up to basis change.

These considerations lead us to the following procedure for obtaining the high frequency parameters (a_1, \dots, a_p) :

We need the following assumptions:

- a_p is non-singular.
- $\Gamma_p > 0$.
- For eigenvalues $\lambda_i \neq \lambda_j$ of \mathcal{A} , $\lambda_i^2 \neq \lambda_j^2$ holds.
- The pair $\left(\begin{pmatrix} I_{n_f} & 0 \end{pmatrix}, \mathcal{A} \right)$ is observable (which is generic, see Anderson et al. [2012b]).

Note that the subset defined by the assumptions listed above is generic in Θ .

We commence from the population spectral density for the observed blocked process $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$. Then we determine a stable and miniphase spectral factor. For factorization of rational spectral densities see Rozanov [1967], Hannan [1970], Hannan and Deistler [2012]. Realizing this factor gives us matrices \bar{A}_b and \bar{C}_b which are similar to A_b and C_b according to Theorems 3.2.2 and 3.2.3.

We are left with the task to find the root $\bar{A} = T^{-1}\mathcal{A}T$ of the matrix $\bar{A}_b = \bar{A}^2$ and to find the transformation T corresponding to basis change to get $\mathcal{A} = T\bar{A}T^{-1}$.

For this purpose, we write \mathcal{A} in Jordan normal form as $\mathcal{A} = Q_1\Lambda_1Q_1^{-1}$ where $\Lambda_1 = \text{diag}(J_1, \dots, J_{np})$, J_i are the Jordan blocks of \mathcal{A} and $Q_1 = (q_1, \dots, q_{np})$ where q_i are the eigenvectors or generalized eigenvectors respectively. In the exact same manner we write \mathcal{A}^2 in Jordan normal form as $\mathcal{A}^2 = Q_2\Lambda_2Q_2^{-1}$. Note that since A_p is non-singular and for eigenvalues $\lambda_i \neq \lambda_j$ of \mathcal{A} $\lambda_i^2 \neq \lambda_j^2$ holds the block sizes of the Jordan blocks of \mathcal{A} and \mathcal{A}^2 are the same which follows from Lemma 3.2.1. Therefore we know which columns of Q_2 are eigenvectors both of \mathcal{A} and \mathcal{A}^2 .

Since we computed in the first step $\bar{A}_b = T^{-1}Q_2\Lambda_2Q_2^{-1}T$ we can determine $T^{-1}Q_2$. Therefore we can compute

$$(3.2.12) \quad \begin{aligned} \bar{C}_b T^{-1}Q_2 &= \left(\begin{array}{cccc} I_n & 0 & 0 & \cdots & 0 \\ I_{n_f} & 0 & 0 & \cdots & 0 \end{array} \right) \begin{array}{l} \mathcal{A}^2 \\ \mathcal{A} \end{array} T T^{-1}Q_2 \\ &= \left(\begin{array}{cccc} I_n & 0 & \cdots & 0 \\ I_{n_f} & 0 & \cdots & 0 \end{array} \right) \begin{array}{l} \mathcal{A}^2 Q_2 \\ \mathcal{A} Q_2 \end{array} \end{aligned}$$

Since we know which columns of Q_2 are eigenvectors of both \mathcal{A} and \mathcal{A}^2 by looking at the submatrix

$$\begin{pmatrix} \left(\begin{array}{cccc} I_{n_f} & 0 & \cdots & 0 \end{array} \right) \mathcal{A}^2 Q_2 \\ \left(\begin{array}{cccc} I_{n_f} & 0 & \cdots & 0 \end{array} \right) \mathcal{A} Q_2 \end{pmatrix}$$

of (3.2.12) we can determine the eigenvalues λ_i of \mathcal{A} : For a column q_i^2 of Q_2 which is an eigenvector of both \mathcal{A} and \mathcal{A}^2 , we obtain

$$\left(\begin{array}{cccc} I_{n_f} & 0 & \cdots & 0 \end{array} \right) \mathcal{A}^2 q_i^2 = \left(\begin{array}{cccc} I_{n_f} & 0 & \cdots & 0 \end{array} \right) \lambda_i^2 q_i^2$$

and

$$\left(\begin{array}{cccc} I_{n_f} & 0 & \cdots & 0 \end{array} \right) \mathcal{A} q_i^2 = \left(\begin{array}{cccc} I_{n_f} & 0 & \cdots & 0 \end{array} \right) \lambda_i q_i^2.$$

Observability of $\left(\left(\begin{array}{c} I_{n_f} \\ 0 \end{array} \right), \mathcal{A} \right)$ guarantees that the first n_f components of q_i^2 are not all zero and thus we obtain λ_i .

So far we described how to determine Λ_1 . To determine $\bar{A} = T^{-1}Q_1\Lambda_1Q_1^{-1}T$ we are left with the task to determine $T^{-1}Q_1$. Note first that the Jordan normal form of Λ_1^2 is

$\Lambda_1^2 = Q_3 \Lambda_2 Q_3^{-1}$. It is easy to see that

$$\begin{aligned} \mathcal{A}^2 &= Q_1 \Lambda_1^2 Q_1^{-1} \\ &= Q_1 Q_3 \Lambda_2 Q_3^{-1} Q_1^{-1} \\ &= Q_2 \Lambda_2 Q_2^{-1} \end{aligned}$$

and therefore a basis Q_1 of appropriate eigenvectors and generalized eigenvectors of \mathcal{A} is transformed through basis change with transformation matrix Q_3 into a basis Q_2 of appropriate eigenvectors and generalized eigenvectors of \mathcal{A}^2 . Thus since we can compute $T^{-1}Q_2$ and Q_3 we can compute

$$T^{-1}Q_1 = T^{-1}Q_2Q_3^{-1}.$$

Let us partition T as

$$T = \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_p \end{pmatrix}, T_i \in \mathbb{R}^{n \times np}.$$

$\bar{\mathcal{A}}$ can be calculated as follows:

$$(3.2.13) \quad \bar{\mathcal{A}} = T^{-1}P_1\Lambda_1P_1^{-1}T = T^{-1}\mathcal{A}T$$

From (3.2.13) we obtain

$$\begin{aligned} a_1T_1 + a_2T_2 + \dots + a_pT_p &= T_1\bar{\mathcal{A}} \\ T_1 &= T_2\bar{\mathcal{A}} \\ T_2 &= T_3\bar{\mathcal{A}} \\ &\vdots \\ T_{p-1} &= T_p\bar{\mathcal{A}}. \end{aligned}$$

Using the fact that we know T_1 from

$$(3.2.14) \quad \bar{C}_b\bar{\mathcal{A}}_b^{-1} = \begin{pmatrix} I_n & 0 & 0 & \dots & 0 \\ 0 & (I_{n_f}, 0_{n_f \times n_s}) & 0 & \dots & 0 \end{pmatrix} \mathcal{A}^2 T T^{-1} \mathcal{A}^{-2} T = \begin{pmatrix} T_1 \\ * \end{pmatrix}$$

we can calculate the remaining T_i from $T_i = T_1\bar{\mathcal{A}}^{-i+1}$, $i = 2, \dots, p$. Finally we obtain the desired companion form $\mathcal{A} = T\bar{\mathcal{A}}T^{-1}$ where the free system parameters are in the first n rows and are uniquely determined.

Thus we have shown:

THEOREM 3.2.5. *Under the assumptions of Theorem 3.2.2 and the additional assumptions that the pair $\left(\begin{pmatrix} I_{n_f} & 0 & \cdots & 0 \end{pmatrix}, \mathcal{A}\right)$ is observable the system parameters (a_1, \dots, a_p) are uniquely determined from those population second moments which can be observed in principle.*

Note that the assumptions of Theorem 3.2.5 do not determine a largest set where identifiability holds. Consider again the set described in Theorem 3.1.1 for the two dimensional AR(1) case. For $a_{ss} = 0$, $a_{sf} = 0$, the assumption that a_p is non-singular is violated.

Given the system parameters, the noise parameters can be determined as in Subsection 2.3.

The approach via blocking is again constructive as is the approach via XYW, as we have proposed an algorithm for obtaining a unique parameter which is identifiable in Θ . It is straightforward to show that θ depends on the second moments of $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ in a continuous way: In an analogous way as in Hannan and Deistler [2012], Chapter 2, the continuous dependence of \bar{A}_b and \bar{C}_b on these second moments can be seen and by the procedure described above, θ is continuously dependent on \bar{A}_b and \bar{C}_b .

Note that (\bar{A}_b, \bar{C}_b) is contained in an Euclidean space of dimension $np(np + n + n_f)$ and the corresponding innovation covariance of $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ has $(n + n_f)(n + n_f + 1)/2$ free parameters whereas θ has $n^2p + n(n + 1)/2$ free parameters see Section 3.1.

We have not been able to describe a relation between the generic set corresponding to XYW and the generic set corresponding to blocking.

One advantage for blocking, as has been said already, seems to be that all second moments which can be observed in principle are used in contrast to the case of XYW. Accordingly, one could hope for better results when working with real data.

3.3. Generic Identifiability for Prescribed Column Degrees

In this section we consider blocking for the case that Γ_p and a_p may be singular. As we already mentioned the case that Γ_p is singular is of special interest for us as singular AR models can be used as models for static factors in GDFMs.

As in Section 2.6 we are interested in identifiability of mixed frequency AR systems for the case that the column degrees of $a(z)$ rather than the degree of $a(z)$ are prescribed.

Again, let p_1, \dots, p_n denote these prescribed column degrees and again let $\Theta_{(p_1, \dots, p_n)}$ denote the subspace of Θ where additionally the highest degree of the respective i th column of $a(z)$ is bounded by p_i and let $(\bar{a}_1, \dots, \bar{a}_p)$ denote all columns of (a_1, \dots, a_p) which are not prescribed to be zero.

Nonzero Column Degrees. For the moment, let us assume $p_i > 0$. Recall the quasi companion form (1.4.3) stated in Section 1.4.2. Also recall that this state space system is

minimal if and only if $\bar{\Gamma}_p > 0$ and the column end matrix of $a(z)$, $\left([a_{p_1}]_{\cdot,1}, \dots, [a_{p_n}]_{\cdot,n}\right)$, is non-singular which is generic in $\Theta_{(p_1, \dots, p_n)}$.

Analogously to Section 3.2 we can write down the state space representation for $(Y_t)_{t \in 2\mathbb{Z}} = \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}_{t \in 2\mathbb{Z}}$

$$(3.3.1) \quad \begin{aligned} \bar{x}_{t+1} &= \bar{\mathcal{A}}^2 \bar{x}_{t-1} + \begin{pmatrix} \bar{\mathcal{B}} & \bar{\mathcal{A}}\bar{\mathcal{B}} \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} \\ Y_t &= \begin{pmatrix} \begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}}^2 \\ \begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}} \end{pmatrix} \bar{x}_{t-1} + \begin{pmatrix} b & \bar{a}_1 b \\ 0 & b \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} \end{aligned}$$

Clearly, the system (3.3.1) is minimal if and only if (1.4.3) is minimal. For the blocked observed process $(\tilde{y}_t)_{t \in 2\mathbb{Z}} = \begin{pmatrix} y_t \\ y_{t-1}^f \end{pmatrix}_{t \in 2\mathbb{Z}}$, we obtain the following state space representation

$$(3.3.2) \quad \begin{aligned} \bar{x}_{t+1} &= \bar{\mathcal{A}}^2 \bar{x}_{t-1} + \begin{pmatrix} \bar{\mathcal{B}} & \bar{\mathcal{A}}\bar{\mathcal{B}} \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} \\ \tilde{y}_t &= \begin{pmatrix} \begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}}^2 \\ \begin{pmatrix} I_{n_f} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}} \end{pmatrix} \bar{x}_{t-1} + \begin{pmatrix} b & \bar{a}_1 b \\ 0 & \begin{pmatrix} I_{n_f} & 0 \end{pmatrix} b \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} \end{aligned}$$

which is not miniphase.

We want to show that generically a transfer function corresponding to system (3.3.2) has the same McMillan degree as a causal, stable and miniphase transfer function for $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$.

THEOREM 3.3.1. *For $((\bar{a}_1, \dots, \bar{a}_p), \Sigma) \in \Theta_{(p_1, \dots, p_n)}$, $p_i > 0$, where the column end matrix of $a(z)$ is non-singular and $\bar{\Gamma}_p > 0$ and if for eigenvalues of $\bar{\mathcal{A}}$ such that $\lambda_i \neq \lambda_j$ it follows that $\lambda_i^2 \neq \lambda_j^2$ holds, the McMillan degree of a causal and miniphase spectral factor $k(z^2)$ of $f_{\tilde{y}}(z^2)$ is equal to $\sum_i p_i$.*

PROOF. The proof is completely analogous to the proof of Theorem 3.2.2.

Note that since $p_i > 0$

$$\begin{aligned} \gamma(k) &= \begin{pmatrix} I_n & 0 \end{pmatrix} \mathbb{E} x_{t+k} x_t^T \begin{pmatrix} I_n \\ 0 \end{pmatrix} = \begin{pmatrix} I_n & 0 \end{pmatrix} \mathcal{A}^k \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} I_n & 0 \end{pmatrix} \mathbb{E} \bar{x}_{t+k} \bar{x}_t^T \begin{pmatrix} I_n \\ 0 \end{pmatrix} = \begin{pmatrix} I_n & 0 \end{pmatrix} \bar{\mathcal{A}}^k \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix}. \end{aligned}$$

For ease of notation we let $m = \sum_i p_i$. Thus let us look at the Hankel matrix $\mathcal{H}_m^\gamma =$

$$\begin{array}{c}
\left(\begin{array}{cc|cc|c}
(I_n & 0) \bar{\mathcal{A}}^2 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^3 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^4 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^5 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots \\
(I_{n_f} & 0) \bar{\mathcal{A}} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^2 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^3 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^4 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots \\
\hline
(I_n & 0) \bar{\mathcal{A}}^4 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^5 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^6 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^7 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \\
(I_{n_f} & 0) \bar{\mathcal{A}}^3 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^4 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^5 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^6 \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \\
\hline
\vdots & \vdots & \vdots & \vdots & \\
\hline
(I_n & 0) \bar{\mathcal{A}}^{2m} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^{2m+1} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^{2m+2} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^{2m+3} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots \\
(I_{n_f} & 0) \bar{\mathcal{A}}^{2m-1} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^{2m} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^{2m+1} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^{2m+2} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} & \dots
\end{array} \right) \\
\begin{array}{c}
\dots \\
\dots \\
\hline
\dots \\
\dots \\
\hline
\dots \\
\dots
\end{array} \left(\begin{array}{cc|c}
(I_n & 0) \bar{\mathcal{A}}^{2m} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^{2m+1} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} \\
(I_{n_f} & 0) \bar{\mathcal{A}}^{2m-1} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^{2m} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} \\
\hline
(I_n & 0) \bar{\mathcal{A}}^{2m+2} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^{2m+3} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} \\
(I_{n_f} & 0) \bar{\mathcal{A}}^{2m+1} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^{2m+2} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} \\
\hline
\vdots & \vdots \\
\hline
(I_n & 0) \bar{\mathcal{A}}^{4m-2} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^{4m-1} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix} \\
(I_{n_f} & 0) \bar{\mathcal{A}}^{4m-3} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_{n_f} & 0) \bar{\mathcal{A}}^{4m-2} \bar{\Gamma}_p \begin{pmatrix} I_{n_f} \\ 0 \end{pmatrix}
\end{array} \right)
\end{array}$$

We look at the following submatrix of \mathcal{H}_m^γ :

$$\begin{aligned}
& \left(\begin{array}{cccc}
(I_n & 0) \bar{\mathcal{A}}^2 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^4 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & (I_n & 0) \bar{\mathcal{A}}^{2m} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \\
(I_n & 0) \bar{\mathcal{A}}^4 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^6 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & (I_n & 0) \bar{\mathcal{A}}^{2m+2} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \\
(I_n & 0) \bar{\mathcal{A}}^6 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^8 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & \\
\vdots & \vdots & & \\
(I_n & 0) \bar{\mathcal{A}}^{2m} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n & 0) \bar{\mathcal{A}}^{2m+2} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \dots & (I_n & 0) \bar{\mathcal{A}}^{4m-2} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix}
\end{array} \right) \\
& = \begin{pmatrix} \begin{pmatrix} I_n & 0 \\ I_n & 0 \\ I_n & 0 \end{pmatrix} \bar{\mathcal{A}}^2 \\ \vdots \\ \begin{pmatrix} I_n & 0 \end{pmatrix} \bar{\mathcal{A}}^{2m} \end{pmatrix} \left(\bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \quad \bar{\mathcal{A}}^2 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \quad \bar{\mathcal{A}}^4 \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \quad \dots \quad \bar{\mathcal{A}}^{2m-2} \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \right) = \mathcal{OC} = \mathcal{H}
\end{aligned}$$

Since we assumed that the eigenvalues are nonzero, $\lambda_i \neq 0$, and for eigenvalues $\lambda_i \neq \lambda_j$ of \mathcal{A} $\lambda_i^2 \neq \lambda_j^2$ holds, it is easy to see from Lemma 3.2.1 that q_i is an eigenvector of $\bar{\mathcal{A}}^2$ if and only if q_i is an eigenvector of $\bar{\mathcal{A}}$.

Observe that \mathcal{O} is always of full column rank if the column end matrix of $a(z)$ is non-singular. Again we use the PBH Test, Theorem 1.1.3, and the fact that for any right eigenvector q'_i of $\bar{\mathcal{A}}$ the first n components are not all equal to zero, as shown in Remark 1.4.7:

$$\begin{pmatrix} (\bar{\mathcal{A}}^2 - \lambda_i^2 I) \\ (I_n \quad 0) \bar{\mathcal{A}}^2 \end{pmatrix} q_i = \begin{pmatrix} 0 \\ [\lambda_i^2 q_i]_{1:n} \neq 0 \end{pmatrix}$$

Also observe that \mathcal{C} is always of full column rank if the column end matrix of $a(z)$ is non-singular and $\bar{\Gamma}_p > 0$: Again we are using the PBH Test, Theorem 1.1.3:

We have to test for all left eigenvectors q_i of $\bar{\mathcal{A}}^2$ or equivalently $\bar{\mathcal{A}}$ that

$$q'_i \begin{pmatrix} (\bar{\mathcal{A}}^2 - \lambda_i^2 I), \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0, q'_i \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \end{pmatrix} \neq 0$$

Thus if q_i is orthogonal to $\bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix}$ also

$$q'_i \bar{\mathcal{A}}^j \bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} = 0, \forall j \in \mathbb{N},$$

holds which implies $q'_i \bar{\Gamma}_p = 0$ which is in contradiction to $\bar{\Gamma}_p > 0$.

Therefore no eigenvector q_i of $\bar{\mathcal{A}}^2$ is orthogonal to $\bar{\Gamma}_p \begin{pmatrix} I_n \\ 0 \end{pmatrix}$.

Now, according to Theorem 2.3.2 in Hannan and Deistler [2012] \mathcal{H} has rank $\sum_i p_i$. \square

For obtaining the high frequency parameters $(\bar{a}_1, \dots, \bar{a}_p)$, we end up with a very similar procedure as in the case that only p is prescribed:

We need the following assumptions:

- The column end matrix of $a(z)$ is non-singular.
- $\bar{\Gamma}_p > 0$.
- For eigenvalues $\lambda_i \neq \lambda_j$ of $\bar{\mathcal{A}}$, $\lambda_i^2 \neq \lambda_j^2$ holds.
- The pair $\left(\begin{pmatrix} I_{n_f} & 0 & \dots & 0 \end{pmatrix}, \bar{\mathcal{A}} \right)$ is observable (which is generic, see Anderson et al. [2012b] and Remark 1.4.7).

Note that the subset defined by the assumptions listed above is generic in $\Theta_{(p_1, \dots, p_n)}$.

Realizing a stable and miniphase spectral factor of the population spectral density for the observed blocked process $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ gives us matrices \bar{A}_b and \bar{C}_b which are similar to $\bar{\mathcal{A}}^2$ and $\left(\begin{pmatrix} I_n & 0 & \dots & 0 \\ I_{n_f} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}}\right)$, see Theorem 3.2.3:

$$\bar{A}_b = T^{-1} \bar{\mathcal{A}}^2 T$$

and

$$\bar{C}_b = \left(\begin{pmatrix} I_n & 0 & \dots & 0 \\ I_{n_f} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}}\right) T.$$

Exactly as in the regular case, we can reconstruct the root $\bar{A} = T^{-1} \bar{\mathcal{A}} T$ of \bar{A}_b since we assume that for eigenvalues of $\bar{\mathcal{A}}$ such that $\lambda_i \neq \lambda_j$ it follows that $\lambda_i^2 \neq \lambda_j^2$ holds and $\left(\begin{pmatrix} I_{n_f} & 0 \end{pmatrix}, \bar{\mathcal{A}}\right)$ is observable.

Let us partition the $\sum_i p_i \times \sum_i p_i$ matrix T as

$$T = \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_p \end{pmatrix}$$

where the number of rows of T_i corresponds to the number of columns in \bar{a}_i .

As in Section 3.2 we can get T_1 , the first n rows of T , via

$$\bar{C}_b \bar{A}_b^{-1} = \left(\begin{pmatrix} I_n & 0 & \dots & 0 \\ I_{n_f} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}}\right) T T^{-1} \bar{\mathcal{A}}^{-2} T = \begin{pmatrix} T_1 \\ * \end{pmatrix}$$

Let $S = \text{diag}(S_i)$ as in Remark 1.4.7 and note that $p_i > 0$ implies $S_1 = I_n$. Then

$$\bar{A} = S A S' = \begin{pmatrix} \bar{a}_1 & \bar{a}_2 & \dots & \bar{a}_{p-2} & \bar{a}_{p-1} & \bar{a}_p \\ S_2 & 0 & & & & 0 \\ 0 & S_3 S_2' & & & & \\ & & \ddots & & & \\ & & & S_{p-1} S_{p-2}' & 0 & \\ 0 & & & 0 & S_p S_{p-1}' & 0 \end{pmatrix}.$$

Thus we can reconstruct T and subsequently \bar{A} from:

$$\begin{aligned}
(3.3.3) \quad \bar{a}_1 T_1 + \bar{a}_2 T_2 + \dots + \bar{a}_p T_p &= T_1 \bar{A} \\
S_2 T_1 &= T_2 \bar{A} \\
S_3 S_2' T_2 &= T_3 \bar{A} \\
&\vdots \\
S_p S_{p-1}' T_{p-1} &= T_p \bar{A}
\end{aligned}$$

We obtain the desired companion form $\bar{\mathcal{A}} = T \bar{A} T^{-1}$ where the free system parameters are in the first n rows and are uniquely determined in $\Theta_{(p_1, \dots, p_n)}$.

Zero Column Degrees. For the case that there is an i such that $p_i = 0$, let us assume that $\bar{\Gamma}_p > 0$ and the column end matrix of $S_1 \bar{a}(z)$ is non-singular. These conditions imply minimality of (1.4.3), see Remark 1.4.7.

The state space systems of $(Y_t)_{t \in 2\mathbb{Z}} = \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}_{t \in 2\mathbb{Z}}$ and $(\tilde{y}_t)_{t \in 2\mathbb{Z}} = \begin{pmatrix} y_t \\ y_{t-1}^f \end{pmatrix}_{t \in 2\mathbb{Z}}$ are slightly different from the case of nonzero column degrees, namely

$$\begin{aligned}
(3.3.4) \quad \bar{x}_{t+1} &= \bar{\mathcal{A}}^2 \bar{x}_{t-1} + \begin{pmatrix} \bar{\mathcal{B}} & \bar{\mathcal{A}} \bar{\mathcal{B}} \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} \\
Y_t &= \begin{pmatrix} (\bar{a}_1, \dots, \bar{a}_p) \bar{\mathcal{A}} \\ (\bar{a}_1, \dots, \bar{a}_p) \end{pmatrix} \bar{x}_{t-1} + \begin{pmatrix} b & \bar{a}_1 b \\ 0 & b \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}
\end{aligned}$$

and

$$\begin{aligned}
(3.3.5) \quad \bar{x}_{t+1} &= \bar{\mathcal{A}}^2 \bar{x}_{t-1} + \begin{pmatrix} \bar{\mathcal{B}} & \bar{\mathcal{A}} \bar{\mathcal{B}} \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} \\
\tilde{y}_t &= \begin{pmatrix} (\bar{a}_1, \dots, \bar{a}_p) \bar{\mathcal{A}} \\ (I_{n_f} \ 0) (\bar{a}_1, \dots, \bar{a}_p) \end{pmatrix} \bar{x}_{t-1} + \begin{pmatrix} b & \bar{a}_1 b \\ 0 & (I_{n_f} \ 0) b \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}.
\end{aligned}$$

Note that if a column degree is zero, since $(\bar{a}_1, \dots, \bar{a}_p) = (a_1, \dots, a_p) S'$ contains rows which are not contained in $\bar{\mathcal{A}} = S A S'$, $\begin{pmatrix} (\bar{a}_1, \dots, \bar{a}_p) \bar{\mathcal{A}} \\ (\bar{a}_1, \dots, \bar{a}_p) \end{pmatrix}$ and $\begin{pmatrix} (\bar{a}_1, \dots, \bar{a}_p) \bar{\mathcal{A}} \\ (I_{n_f} \ 0) (\bar{a}_1, \dots, \bar{a}_p) \end{pmatrix}$ only contain as a submatrix the matrices $\begin{pmatrix} (I_{n-s} \ 0 \ \dots \ 0) \bar{\mathcal{A}}^2 \\ (I_{n-s} \ 0 \ \dots \ 0) \bar{\mathcal{A}} \end{pmatrix}$ and $\begin{pmatrix} (I_{n-s} \ 0 \ \dots \ 0) \bar{\mathcal{A}}^2 \\ (I_{n_f-s_f} \ 0 \ \dots \ 0) \bar{\mathcal{A}} \end{pmatrix}$, respectively where s again is the number of i such that $p_i = 0$ and where s_f is the number of fast components such that $p_i = 0$.

Let the $(n - s)$ -dimensional process y_t^r contain all components of y_t with prescribed column degree $p_i > 0$, $y_t^r = S_1 y_t$.

We can obtain a quasi companion state space representation for $(y_t^r)_{t \in \mathbb{Z}}$ as

$$(3.3.6) \quad \begin{aligned} \bar{x}_{t+1} &= \bar{\mathcal{A}}\bar{x}_t + \bar{\mathcal{B}}^r \bar{\varepsilon}_t \\ y_t^r &= S_1 (\bar{a}_1, \dots, \bar{a}_p) \bar{x}_t + \bar{b} \bar{\varepsilon}_t \end{aligned}$$

where $\bar{b} \bar{\varepsilon}_t = S_1 b \varepsilon_t$ with $\bar{b} \bar{b}' = S_1 b b' S_1$, where \bar{b} has minimal column degree, r say, and $\Sigma_{\bar{\varepsilon}} = I_r$, and where $(\bar{\mathcal{B}}^r)' = \begin{pmatrix} \bar{b}' & 0 \end{pmatrix}$. As y_t^r clearly is an AR process and \bar{b} is of full column rank, (3.3.6) is miniphase.

Note that (3.3.6) is minimal if and only if $\bar{\Gamma}_p > 0$ and the column end matrix of $S_1 \bar{a}(z)$ is non-singular.

A stable and miniphase state space system for $(Y_t^r)_{t \in 2\mathbb{Z}} = \begin{pmatrix} y_t^r \\ y_{t-1}^r \end{pmatrix}_{t \in 2\mathbb{Z}}$ is given by

$$(3.3.7) \quad \begin{aligned} \bar{x}_{t+1} &= \bar{\mathcal{A}}^2 \bar{x}_{t-1} + (\bar{\mathcal{B}}^r, \bar{\mathcal{A}} \bar{\mathcal{B}}^r) \begin{pmatrix} \bar{\varepsilon}_t \\ \bar{\varepsilon}_{t-1} \end{pmatrix} \\ Y_t^r &= \begin{pmatrix} \begin{pmatrix} I_{n-s} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}}^2 \\ \begin{pmatrix} I_{n-s} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}} \end{pmatrix} \bar{x}_{t-1} + \begin{pmatrix} \bar{b} & \bar{a}_1 \bar{b} \\ 0 & \bar{b} \end{pmatrix} \begin{pmatrix} \bar{\varepsilon}_t \\ \bar{\varepsilon}_{t-1} \end{pmatrix} \end{aligned}$$

and a stable but not miniphase state space system for $(\tilde{y}_t^r)_{t \in 2\mathbb{Z}} = \begin{pmatrix} y_t^r \\ y_{t-1}^{rf} \end{pmatrix}_{t \in 2\mathbb{Z}}$ is given by

$$(3.3.8) \quad \begin{aligned} \bar{x}_{t+1} &= \bar{\mathcal{A}}^2 \bar{x}_{t-1} + (\bar{\mathcal{B}}^r, \bar{\mathcal{A}} \bar{\mathcal{B}}^r) \begin{pmatrix} \bar{\varepsilon}_t \\ \bar{\varepsilon}_{t-1} \end{pmatrix} \\ \tilde{y}_t^r &= \begin{pmatrix} \begin{pmatrix} I_{n-s} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}}^2 \\ \begin{pmatrix} I_{n_f-s_f} & 0 & \dots & 0 \end{pmatrix} \bar{\mathcal{A}} \end{pmatrix} \bar{x}_{t-1} + \begin{pmatrix} \bar{b} & \bar{a}_1 \bar{b} \\ 0 & \begin{pmatrix} I_{n_f-s_f} & 0 \end{pmatrix} \bar{b} \end{pmatrix} \begin{pmatrix} \bar{\varepsilon}_t \\ \bar{\varepsilon}_{t-1} \end{pmatrix}. \end{aligned}$$

Both systems (1.4.3) and (3.3.6) are stable and miniphase. Thus for both $(y_t)_{t \in \mathbb{Z}}$ and $(y_t^r)_{t \in \mathbb{Z}}$ the McMillan degree of a causal and miniphase transfer function is $\sum_{i=1}^n p_i$. As is easily seen, therefore also for both $(Y_t)_{t \in 2\mathbb{Z}}$ and $(Y_t^r)_{t \in 2\mathbb{Z}}$ the McMillan degree of a causal and miniphase transfer function is $\sum_{i=1}^n p_i$. Since $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ and $(\tilde{y}_t^r)_{t \in 2\mathbb{Z}}$ are subprocesses of $(Y_t)_{t \in 2\mathbb{Z}}$ (and $(Y_t^r)_{t \in 2\mathbb{Z}}$, respectively), the McMillan degrees of $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ and $(\tilde{y}_t^r)_{t \in 2\mathbb{Z}}$ smaller or equal to $\sum_{i=1}^n p_i$ which can easily be seen from the proof of Theorem 3.2.2 or also 3.3.1 since a Hankel matrix of covariances of a process e.g. $(Y_t)_{t \in 2\mathbb{Z}}$, contains the Hankel matrix of covariances of any subprocess, e.g. $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$. Under the assumptions that $\bar{\Gamma}_p > 0$ and the column end matrix of $S_1 \bar{a}(z)$ is non-singular and for eigenvalues $\lambda_i \neq \lambda_j$ of $\bar{\mathcal{A}}$ $\lambda_i^2 \neq \lambda_j^2$ holds, using Theorem 3.3.1

we obtain that a causal miniphase transfer function for $(\tilde{y}_t^r)_{t \in 2\mathbb{Z}}$ has McMillan degree $\sum_{i=1}^n p_i$ and since $(\tilde{y}_t^r)_{t \in 2\mathbb{Z}}$ is a subprocess of $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ also a causal miniphase transfer function for $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$ has McMillan degree $\sum_{i=1}^n p_i$.

Thus Theorem 3.2.3 tells us that realizing a stable and miniphase spectral factor of the population spectral density for the observed blocked process $(\tilde{y}_t)_{t \in 2\mathbb{Z}}$, e.g. by the subspace algorithm proposed in Section 1.5.2, gives us matrices \bar{A}_b and \bar{C}_b such that

$$\begin{aligned}\bar{A}_b &= T^{-1}\bar{\mathcal{A}}^2T \\ \bar{C}_b &= \begin{pmatrix} (\bar{a}_1, \dots, \bar{a}_p)\bar{\mathcal{A}} \\ \left(\begin{matrix} I_{n_f} & 0 \end{matrix} \right) (\bar{a}_1, \dots, \bar{a}_p) \end{pmatrix} T\end{aligned}$$

for a suitable non-singular matrix T .

\bar{C}_b contains as a submatrix

$$\begin{pmatrix} \left(\begin{matrix} I_{n-s} & 0 & \dots & 0 \end{matrix} \right) \bar{\mathcal{A}}^2 \\ \left(\begin{matrix} I_{n_f-s_f} & 0 & \dots & 0 \end{matrix} \right) \bar{\mathcal{A}} \end{pmatrix} T$$

As before, we can reconstruct the root $\bar{A} = T^{-1}\bar{\mathcal{A}}T$ of \bar{A}_b if we assume that for eigenvalues of $\bar{\mathcal{A}}$ such that $\lambda_i \neq \lambda_j$ it follows that $\lambda_i^2 \neq \lambda_j^2$ holds and $\left(\left(\begin{matrix} I_{n_f-s_f} & 0 & \dots & 0 \end{matrix} \right), \bar{\mathcal{A}} \right)$ is observable.

Let T be partitioned as in the case of nonzero prescribed column degrees. We can compute

$$\begin{pmatrix} \left(\begin{matrix} I_{n-s} & 0 & \dots & 0 \end{matrix} \right) \bar{\mathcal{A}}^2 \\ \left(\begin{matrix} I_{n_f-s_f} & 0 & \dots & 0 \end{matrix} \right) \bar{\mathcal{A}} \end{pmatrix} T T^{-1} \bar{\mathcal{A}}^{-2} T = (T_1)$$

From equation (3.3.3) we obtain the remaining rows of T . Subsequently from the now known matrices \bar{C}_b , $T^{-1}\bar{\mathcal{A}}^{-1}T$ and T^{-1} we can compute the system parameters $(\bar{a}_1, \dots, \bar{a}_p)$:

$$\begin{aligned}\bar{C}_b T^{-1} \bar{\mathcal{A}}^{-1} T \quad T^{-1} &= \begin{pmatrix} (\bar{a}_1, \dots, \bar{a}_p)\bar{\mathcal{A}} \\ \left(\begin{matrix} I_{n_f} & 0 \end{matrix} \right) (\bar{a}_1, \dots, \bar{a}_p) \end{pmatrix} T T^{-1} \bar{\mathcal{A}}^{-1} T T^{-1} \\ &= \left((\bar{a}_1, \dots, \bar{a}_p) \right).\end{aligned}$$

Thus, for prescribed column degrees we have proved a very similar result to the case that only the degree p of $a(z)$ is given:

THEOREM 3.3.2. *For $((\bar{a}_1, \dots, \bar{a}_p), \Sigma) \in \Theta_{(p_1, \dots, p_n)}$ where the column end matrix of $S_1 a(z)$ is non-singular, $\bar{\Gamma}_p > 0$, if for eigenvalues of $\bar{\mathcal{A}}$ such that $\lambda_i \neq \lambda_j$ it follows that $\lambda_i^2 \neq \lambda_j^2$ holds, $n_f - s_f > 0$, and the pair $\left(\left(\begin{matrix} I_{n_f-s_f} & 0 & \dots & 0 \end{matrix} \right), \bar{\mathcal{A}} \right)$ is observable the system*

parameters $(\bar{a}_1, \dots, \bar{a}_p)$ are uniquely determined from those population second moments which can be observed in principle.

As in the case that only p is specified, given the system parameters, the noise parameters can be determined as in Subsection 2.3.

3.4. Generic Identifiability for AR(p) Systems with General Sampling Frequency N

In this section we are shortly describing how to extend the results of Section 3.2 to the case that the sampling frequency of the slow components is $N > 2$. Again we assume that $\Gamma_p > 0$ and that a_p is non-singular.

The blocked process is $(Y_t)_{t \in N\mathbb{Z}} = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-N+1} \end{pmatrix}_{t \in N\mathbb{Z}}$. Analogously to Section 3.2 the state

spaces for $t \in N\mathbb{Z}$ of $(y_t)_{t \in \mathbb{Z}}$ and $(Y_t)_{t \in N\mathbb{Z}}$ are the same and both processes are AR processes with minimal state dimension np . A minimal, stable and miniphase state space representation for $(Y_t)_{t \in N\mathbb{Z}}$ is

$$(3.4.1) \quad x_{t+1} = \underbrace{\mathcal{A}^N}_{A_b} x_{t-N+1} + \underbrace{(\mathcal{B}, \mathcal{A}\mathcal{B}, \dots, \mathcal{A}^{N-1}\mathcal{B})}_{B_b} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-N+1} \end{pmatrix}$$

$$(3.4.2) \quad Y_t = \begin{pmatrix} \left(\begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} \mathcal{A}^N \right) \\ \left(\begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} \mathcal{A}^{N-1} \right) \\ \vdots \\ \left(\begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} \mathcal{A} \right) \end{pmatrix} x_{t-N+1} + \begin{pmatrix} \left(\begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} (\mathcal{B}, \mathcal{A}\mathcal{B}, \dots, \mathcal{A}^{N-1}\mathcal{B}) \right) \\ \left(\begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} (0, \mathcal{B}, \mathcal{A}\mathcal{B}, \dots, \mathcal{A}^{N-2}\mathcal{B}) \right) \\ \vdots \\ \left(\begin{pmatrix} I_n & 0 & \dots & 0 \end{pmatrix} (0, \dots, 0, \mathcal{B}) \right) \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-N+1} \end{pmatrix}.$$

Now the blocked observed process is $(\tilde{y}_t)_{t \in \mathbb{N}\mathbb{Z}} = \begin{pmatrix} y_t \\ y_{t-1}^f \\ \vdots \\ y_{t-N+1}^f \end{pmatrix}_{t \in \mathbb{N}\mathbb{Z}}$. A state space representation is

$$(3.4.3) \quad x_{t+1} = A_b x_{t-N+1} + B_b \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-N+1} \end{pmatrix}$$

$$(3.4.4) \quad \tilde{y}_t = \underbrace{\begin{pmatrix} \left(I_n \ 0 \ \cdots \ 0 \right) \mathcal{A}^N \\ \left(I_{n_f} \ 0 \ \cdots \ 0 \right) \mathcal{A}^{N-1} \\ \vdots \\ \left(I_{n_f} \ 0 \ \cdots \ 0 \right) \mathcal{A} \end{pmatrix}}_{C_b} x_{t-N+1} + \underbrace{\begin{pmatrix} \left(I_n \ 0 \ \cdots \ 0 \right) (\mathcal{B}, \mathcal{A}\mathcal{B}, \dots, \mathcal{A}^{N-1}\mathcal{B}) \\ \left(I_{n_f} \ 0 \ \cdots \ 0 \right) (0, \mathcal{B}, \mathcal{A}\mathcal{B}, \dots, \mathcal{A}^{N-2}\mathcal{B}) \\ \vdots \\ \left(I_{n_f} \ 0 \ \cdots \ 0 \right) (0, \dots, 0, \mathcal{B}) \end{pmatrix}}_{D_b} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-N+1} \end{pmatrix}.$$

The spectral factor of the spectral density $f_{\tilde{y}}(z^N)$ of $(\tilde{y}_t)_{t \in \mathbb{N}\mathbb{Z}}$ corresponding to (3.4.3), (3.4.4) is causal and stable but not miniphase. Analogously to Section 3.2 we want to show that a causal, stable and miniphase spectral factor $k(z^N) = \left(\bar{C}_b \left(I_{np} (z^N)^{-1} - \bar{A}_b \right)^{-1} \bar{B}_b + \bar{D}_b \right)$ of $f_{\tilde{y}}(z^N)$ has (generically) McMillan degree np . Then we have from Theorem 3.2.3 that

$$\begin{aligned} \bar{A}_b &= T^{-1} A_b T \\ \bar{C}_b &= C_b T \end{aligned}$$

Note that the next theorem here states a slightly weaker result as Theorem 3.2.2 since we assume that \mathcal{A} is diagonalizable.

THEOREM 3.4.1. *For $((a_1, \dots, a_p), \Sigma_\nu) \in \Theta$, if a_p non-singular, $\Gamma_p > 0$, \mathcal{A} diagonalizable, and if for eigenvalues of \mathcal{A} such that $\lambda_i \neq \lambda_j$ it follows that $\lambda_i^N \neq \lambda_j^N$ holds, the McMillan degree of a causal and miniphase spectral factor $k(z^N)$ of $f_{\tilde{y}}(z^N)$ is equal to np .*

PROOF. The proof is analogous to the proof of Theorem 3.2.2. Again we consider the Hankel matrix of covariances

$$\mathcal{H}_{np}^\gamma = \mathbb{E} \begin{pmatrix} \tilde{y}_{t+N} \\ \vdots \\ \tilde{y}_{t+npN} \end{pmatrix} \begin{pmatrix} \tilde{y}'_t & \cdots & \tilde{y}'_{t-(np-1)N} \end{pmatrix} \in \mathbb{R}^{(n+(N-1)n_f)np \times (n+(N-1)n_f)np}.$$

This matrix contains as a submatrix

$$\begin{aligned} & \begin{pmatrix} (I_n \ 0) \mathcal{A}^N \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n \ 0) \mathcal{A}^{2N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \cdots & (I_n \ 0) \mathcal{A}^{npN} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \\ (I_n \ 0) \mathcal{A}^{2N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n \ 0) \mathcal{A}^{3N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \cdots & (I_n \ 0) \mathcal{A}^{(np+1)N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \\ (I_n \ 0) \mathcal{A}^{3N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n \ 0) \mathcal{A}^{4N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \cdots & \\ \vdots & \vdots & & \\ (I_n \ 0) \mathcal{A}^{npN} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & (I_n \ 0) \mathcal{A}^{(np+1)N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \cdots & (I_n \ 0) \mathcal{A}^{(2np-1)N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \end{pmatrix} \\ & = \begin{pmatrix} (I_n \ 0) \mathcal{A}^N \\ (I_n \ 0) \mathcal{A}^{2N} \\ (I_n \ 0) \mathcal{A}^{3N} \\ \vdots \\ (I_n \ 0) \mathcal{A}^{npN} \end{pmatrix} \begin{pmatrix} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \mathcal{A}^N \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \mathcal{A}^{2N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} & \cdots & \mathcal{A}^{(np-1)N} \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \end{pmatrix} = \mathcal{O}\mathcal{C} = \mathcal{H} \end{aligned}$$

Since we assumed that \mathcal{A} is diagonalizable and the eigenvalues are nonzero, $\lambda_i \neq 0$, and for eigenvalues $\lambda_i \neq \lambda_j$ of \mathcal{A} $\lambda_i^N \neq \lambda_j^N$ holds, it is easy to see that q_i is an eigenvector of \mathcal{A}^N if and only if q_i is an eigenvector of \mathcal{A} . Then completely analogously to the proof of Theorem 3.2.2 we can show that \mathcal{O} and \mathcal{C} have full rank.

To show that \mathcal{O} is of full column rank we again use the PBH Test, Theorem 1.1.3, and the fact that for any right eigenvector q_i of \mathcal{A} the first n components are not all equal to zero, as shown in Anderson et al. [2012b], Lemma 2:

$$\begin{pmatrix} (\mathcal{A}^N - \lambda_i^N I) \\ (I_n \ 0) \mathcal{A}^N \end{pmatrix} q_i = \begin{pmatrix} 0 \\ [\lambda_i^N q_i]_{1:n} \neq 0 \end{pmatrix}$$

Again using the PBH Test, we show that \mathcal{C} is of full row rank if and $\Gamma_p > 0$:

We have to test for all left eigenvectors q_i of \mathcal{A}^N or equivalently \mathcal{A} that

$$q'_i \left((\mathcal{A}^N - \lambda_i^N I), \Gamma_p \begin{pmatrix} I_n \\ 0 \end{pmatrix} \right) = \left(0, q'_i \begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix} \right) \neq 0$$

Thus if q_i is orthogonal to $\begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix}$ also

$$q'_i \mathcal{A}^j \begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix} = 0, \forall j \in \mathbb{N},$$

holds which implies $q'_i \Gamma_p = 0$ which is in contradiction to $\Gamma_p > 0$.

Therefore for all eigenvectors q_i of \mathcal{A}^N

$$\left(0, q'_i \begin{pmatrix} \gamma(0) \\ \vdots \\ \gamma(1-p) \end{pmatrix} \right) \neq 0.$$

Now, according to Theorem 2.3.2 in Hannan and Deistler [2012] \mathcal{H} has rank np . \square

The set of θ such that \mathcal{A} is diagonalizable is generic. This result is implied by the following lemma:

LEMMA 3.4.2. *The set of θ such that the eigenvalues of \mathcal{A} are simple is generic in Θ .*

PROOF. The eigenvalues of \mathcal{A} are the inverse of the zeros of $\det a(z)$. Thus we can prove the lemma by showing that generically the zeros of $\det a(z)$ are simple. Note that a_p is non-singular on a generic subset of Θ . On this generic subset the degree of $\det a(z)$ is np .

Clearly, z_i is a zero of $\det a(z)$ with multiplicity larger than one if and only if z_i is also a zero of the derivative of $\det a(z)$. Therefore we use the Sylvester matrix of $\det a(z)$ and its derivative, see e.g. Kailath [1980] p 142. The nonzero elements of the Sylvester matrix are the coefficients of these two polynomials. The determinant of this matrix, the so-called Sylvester's resultant, is zero if and only if there is a common zero of $\det a(z)$ and its derivative. If we consider $a_i, i = 1, \dots, p$ diagonal, it is easy to find a special choice of (a_1, \dots, a_p) such that the resultant is nonzero. Since the resultant is polynomial in (a_1, \dots, a_p) it is generically nonzero and the lemma is proved. \square

The procedure for obtaining the high frequency parameters (a_1, \dots, a_p) :

We need the following assumptions:

- a_p is non-singular.
- $\Gamma_p > 0$.
- \mathcal{A} is diagonalizable.
- For eigenvalues $\lambda_i \neq \lambda_j$ of \mathcal{A} , $\lambda_i^N \neq \lambda_j^N$ holds.
- The pair $\left(\begin{pmatrix} I_{n_f} & 0 & \dots & 0 \end{pmatrix}, \mathcal{A} \right)$ is observable (which is generic, see Anderson et al. [2012b]).

Note that the subset defined by the assumptions listed above is generic in Θ .

Realizing a stable and miniphase spectral factor of the population spectral density $f_{\tilde{y}}(z^N)$ for the observed blocked process $(\tilde{y}_t)_{t \in N\mathbb{Z}}$ gives us matrices \bar{A}_b and \bar{C}_b which are similar to A_b and C_b according to Theorems 3.4.1 and 3.2.3.

We are left with the task to find the root $\bar{\mathcal{A}} = T^{-1}\mathcal{A}T$ of the matrix $\bar{A}_b = \bar{\mathcal{A}}^N$ and to find the transformation T corresponding to basis change to get $\mathcal{A} = T\bar{\mathcal{A}}T^{-1}$.

For the generic case that \mathcal{A} is diagonalizable, the eigendecomposition of \mathcal{A} is $\mathcal{A} = Q\Lambda Q^{-1}$ where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{np})$, and $Q = (q_1, \dots, q_{np})$ where q_i are the eigenvectors. Now it immediately follows that $\bar{\mathcal{A}} = T^{-1}Q\Lambda Q^{-1}T$ and $\bar{A}_b = T^{-1}Q\Lambda^N Q^{-1}T$, which is the eigendecomposition of \bar{A}_b . Note that $T^{-1}Q$ and Λ^N can be determined from \bar{A}_b . Now we have

$$\begin{aligned}
 (3.4.5) \quad \bar{C}_b T^{-1}Q &= \begin{pmatrix} \left(\begin{matrix} I_n & 0 & 0 & \dots & 0 \end{matrix} \right) \mathcal{A}^N \\ \vdots \\ \left(\begin{matrix} I_{n_f} & 0 & 0 & \dots & 0 \end{matrix} \right) \mathcal{A}^2 \\ \left(\begin{matrix} I_{n_f} & 0 & 0 & \dots & 0 \end{matrix} \right) \mathcal{A} \end{pmatrix} T T^{-1}Q \\
 &= \begin{pmatrix} \left(\begin{matrix} I_n & 0 & 0 & \dots & 0 \end{matrix} \right) \mathcal{A}^N Q \\ \vdots \\ \left(\begin{matrix} I_{n_f} & 0 & 0 & \dots & 0 \end{matrix} \right) \mathcal{A}^2 Q \\ \left(\begin{matrix} I_{n_f} & 0 & 0 & \dots & 0 \end{matrix} \right) \mathcal{A} Q \end{pmatrix}.
 \end{aligned}$$

Since Q is a matrix of eigenvectors for both \mathcal{A} and \mathcal{A}^2 , by looking at the submatrix

$$\begin{pmatrix} \left(\begin{matrix} I_{n_f} & 0 & \dots & 0 \end{matrix} \right) \mathcal{A}^2 Q \\ \left(\begin{matrix} I_{n_f} & 0 & \dots & 0 \end{matrix} \right) \mathcal{A} Q \end{pmatrix}$$

of (3.4.5) we can determine the eigenvalues λ_i of \mathcal{A} analogously to Section 3.2.

$\bar{\mathcal{A}}$ can be determined as follows:

$$(3.4.6) \quad \bar{\mathcal{A}} = T^{-1}Q\Lambda Q^{-1}T = T^{-1}\mathcal{A}T$$

We partition T as before: $T' = (T'_1 \ T'_2 \ \dots \ T'_p)'$, $T_i \in \mathbb{R}^{n \times np}$. Using the fact that we know T_1 from

$$(3.4.7) \quad \bar{C}_b \bar{A}_b^{-1} = \begin{pmatrix} \left(\begin{array}{cccc} I_n & 0 & 0 & \dots & 0 \end{array} \right) \mathcal{A}^N \\ \vdots \\ \left(\begin{array}{cccc} I_{n_f} & 0 & 0 & \dots & 0 \end{array} \right) \mathcal{A}^2 \\ \left(\begin{array}{cccc} I_{n_f} & 0 & 0 & \dots & 0 \end{array} \right) \mathcal{A} \end{pmatrix} T T^{-1} \mathcal{A}^{-N} T \\ = \begin{pmatrix} T_1 \\ \vdots \\ * \end{pmatrix}$$

we can calculate the remaining T_i as in Section 3.2 from $T_i = T_1 \bar{\mathcal{A}}^{-i+1}$, $i = 2, \dots, p$. Finally we obtain the desired companion form $\mathcal{A} = T \bar{\mathcal{A}} T^{-1}$ where the free system parameters are in the first n rows and are uniquely determined.

Thus we have shown:

THEOREM 3.4.3. *Under the assumptions of Theorem 3.4.1 and the additional assumptions that that the pair $\left(\begin{pmatrix} I_{n_f} & 0 & \dots & 0 \end{pmatrix}, \mathcal{A} \right)$ is observable the system parameters (A_1, \dots, A_p) are uniquely determined from those population second moments which can be observed in principle.*

Given the system parameters, the noise parameters can be determined as in Subsection 2.3.

Part 3

GDFMs Single and Mixed Frequency

Generalized Linear Dynamic Factor Models (GDFMs)

In the last part of the thesis we consider generalized linear dynamic factor models (GDFMs) which were a great incentive for us to closely investigate singular AR models. Singular AR models appear as models for the latent variables as well as for the static factor which will be explained below.

Factor models are a useful instrument when dealing with high dimensional data. In modeling high dimensional time series we usually face the problem that the number of parameters of a model used to describe the series is proportional to the squared cross-sectional dimension whereas the data only grows linearly in the cross-sectional dimension (and of course linearly in the time dimension). Factor models overcome this problem by condensing not only information which is contained in the time dimension but also information contained in the cross-sectional dimension and thus being able to reduce the number of parameters such that they depend only linearly on the cross-sectional dimension. As a trade off assumptions on the dependence in the cross-sectional dimension have to be made, which will ensure some sort of co-movement of the components of the time series. The time series can be explained by common factors plus some noise.

As classical factor models for many applications turn out to pose too strict assumptions on the structure of the noise, i.e. a diagonal covariance matrix, the GDFMs, which are a generalization of linear dynamic factor models, see e.g. Geweke [1977], Sargent and Sims [1977], Scherrer and Deistler [1998], and generalized static factor models, see e.g. Chamberlain and Rothschild [1983], Chamberlain [1983], prove to be more apt in many situations as they allow for cross-sectional dependence of the noise components, i.e. a non-diagonal covariance matrix.

GDFMs have been introduced by two different groups, a European group around Forni, Lippi, Hallin and Reichlin, see Forni et al. [2000], Forni and Lippi [2001], Forni et al. [2004, 2005], and a U.S. American group, Stock and Watson, see Stock and Watson [2002a,b, 2005].

In this thesis we consider the GDFMs introduced by Forni, Lippi, Hallin and Reichlin. In this chapter we give a short presentation of GDFMs as given in Deistler et al. [2010].

4.1. GDFMs: Definitions and Assumptions

In this section we introduce a set of assumptions defining the generalized linear dynamic factor models we are considering in this thesis.

Let us denote the cross-sectional dimension by n and the time dimension by T .

The fundamental idea of a factor model is to decompose the observations z_t^n into the so-called latent variables or common component \hat{z}_t^n and the idiosyncratic noise or idiosyncratic component ξ_t^n

$$z_t^n = \hat{z}_t^n + \xi_t^n$$

where the latent variables contain the co-movement, the similarities of all the observations and are “strongly dependent” and the idiosyncratic noise contains the information which is specific for a single (or for GDFMs only a few) time series and are “weakly dependent”. (We will define the terms weak and strong dependence below.) Then we want to write the latent variables as $\hat{z}_t^n = \Lambda y_t$ where $\Lambda \in \mathbb{R}^{n \times r}$ is the so-called static factor loading matrix and y_t are the static factors with $r \ll n$. The appeal of the factor model is clearly that modeling the factors instead of the high dimensional latent variables is much more convenient.

Of course assumptions have to be made to guarantee that the observed process (z_t^n) allows for a (unique) GDFM representation.

We assume that (\hat{z}_t^n) and (ξ_t^n) are wide sense stationary with zero mean

$$\mathbb{E}\hat{z}_t^n = \mathbb{E}\xi_t^n = 0, \forall t \in \mathbb{Z},$$

both have absolutely summable covariances in order to ensure the existence of the spectral densities $f_{\hat{z}}^n(\lambda)$ and $f_{\xi}^n(\lambda)$, $\lambda \in [-\pi, \pi]$, and are mutually orthogonal

$$\mathbb{E}\hat{z}_t^n (\xi_s^n)' = 0, \forall s, t \in \mathbb{Z},$$

and thus

$$f_z^n(\lambda) = f_{\hat{z}}^n(\lambda) + f_{\xi}^n(\lambda)$$

holds.

To allow for a more general covariance structure of the idiosyncratic noise ξ_t^n (than a classical factor model) and still have some sort of factor model representation we have to let T as well as n tend to infinity. Thus we are looking at a sequence of GDFMs where the index is the cross-sectional dimension n . Therefore the following assumption on our process of observations is useful:

ASSUMPTION 4.1.1. *The doubly indexed sequence $(z_t^n)_{t \in \mathbb{Z}}^{n \in \mathbb{N}}$ corresponds to a nested sequence of models in the sense that the one dimensional components \hat{z}_{it} and ξ_{it} of \hat{z}_t^n and ξ_t^n do not depend on n for $i \leq n$.*

ASSUMPTION 4.1.2. *The spectral density $f_{\hat{z}}^n$ is rational and there is a n_0 such that $\forall n \geq n_0$ $f_{\hat{z}}^n$ has constant rank $q < n$ on $[-\pi, \pi]$.*

ASSUMPTION 4.1.3. *There is a n_0 such that $\forall n \geq n_0$ the rank q of $f_{\hat{z}}^n$ is independent of n .*

The last two assumptions assure that $\forall n \geq n_0$ the latent variables can be modeled by a singular ARMA model which is strictly miniphase with q dimensional innovations from a certain n_0 onwards.

ASSUMPTION 4.1.4. *The dimension, m say, of a minimal state space realization of a stable and miniphase spectral factor of $f_{\hat{z}}^n$ is independent of n from a certain n_0 onwards.*

This assumption constrains the dynamics of the model, in fact it constrains the McMillan degree of a causal, stable and miniphase transfer function for the latent variables from a certain n_0 onwards.

Weak and strong dependence: Let us define $\mu_j(\lambda)$, $\lambda \in [-\pi, \pi]$, as the function associating λ with the j th eigenvalue of the spectral density f at λ the, see e.g. Brillinger [1981] Chapter 9 or Forni et al. [2000]. The function μ_j is called the j th dynamic eigenvalue of f .

ASSUMPTION 4.1.5 (weak dependence). *The largest dynamic eigenvalue of f_{ξ}^n is uniformly bounded in λ and n .*

Note that this assumption allows for serial dependence of (ξ_t^n) as well as for some dependence the cross-section.

ASSUMPTION 4.1.6 (strong dependence). *The largest dynamic q eigenvalues of $f_{\hat{z}}^n$ diverge to infinity for all frequencies as $n \rightarrow \infty$.*

The last two assumptions are crucial in uniquely decomposing the observations z_t^n into the latent variables \hat{z}_t^n and the idiosyncratic noise ξ_t^n . Since we allow for some cross-sectional dependence in ξ_t^n , \hat{z}_t^n and ξ_t^n are not uniquely identifiable from z_t^n for any fixed n . As mentioned above, we need $n \rightarrow \infty$ for a unique decomposition, see Forni and Lippi [2001].

In Forni and Lippi [2001] the authors have also proved the following:

LEMMA 4.1.7. *Let $\mu_{ni}^z(\lambda)$ denote the i th dynamic eigenvalue of the spectral density $f_z^n(\lambda)$ of process $(z_t^n)_{t \in \mathbb{Z}}$. Then given i for $n > i$ the function $\mu_{ni}^z(\lambda)$ is non-decreasing in n , i.e. $\forall \lambda$ $\mu_{ni}^z(\lambda) \leq \mu_{(n+1)i}^z(\lambda)$.*

This gives rise to the definition of the function $\mu_i^z(\lambda) = \sup_{n \in \mathbb{N}} \mu_{ni}^z(\lambda)$.

THEOREM 4.1.8 (Unique Representation Theorem). *The doubly indexed process $(z_t^n)_{t \in \mathbb{Z}}^{n \in \mathbb{N}}$ has a generalized linear dynamic factor model representation if and only if there is a integer q such that*

- $\mu_q^z(\lambda)$ diverges to infinity a.e. in $[-\pi, \pi]$ as $n \rightarrow \infty$ and
- $\mu_{q+1}^z(\lambda)$ is essentially bounded as $n \rightarrow \infty$.

Note that this theorem gives us necessary and sufficient conditions on the nested sequence of the the spectral density of the observations $(z_t^n)_{t \in \mathbb{Z}}$ such that the doubly indexed process $(z_t^n)_{t \in \mathbb{Z}}^{n \in \mathbb{N}}$ permits a representation as a generalized dynamic factor models. As we will see later, the integer q will be the dimension of the so-called minimal dynamic factor.

For ease of notation, from now on we omit the index n .

4.2. Minimal Static Factors of the Latent Variables

As already mentioned, we are interested in modeling the n -dimensional latent variables by a process of much smaller dimension, the (minimal) static factor process which we define in the following:

DEFINITION 4.2.1. A process $(y_t)_{t \in \mathbb{Z}}$ is called a *static factor* of the n -dimensional process $(\hat{z}_t)_{t \in \mathbb{Z}}$ if it has dimension smaller than n and there is a constant matrix Λ such that $\hat{z}_t = \Lambda y_t, \forall t \in \mathbb{Z}$. A *minimal* static factor is a static factor of least possible dimension.

From Assumption 4.1.2 it follows that there is a minimal, stable and miniphase state space system with the latent variables as outputs:

$$(4.2.1) \quad \begin{aligned} x_{t+1} &= Fx_t + G\varepsilon_{t+1} \\ \hat{z}_t &= Hx_t \end{aligned}$$

Note that the system (4.2.1) is a different state space system compared to the systems considered in Chapter 1 as a minimal state here is a basis of the space spanned by the projections of the components of $y^+(t-1) = (y'_t, y'_{t+1}, \dots)'$ onto the space of present and past $\mathbb{H}_y^-(t) = \overline{\text{span}} \{y_t^1, \dots, y_t^n, y_{t-1}^1, \dots, y_{t-1}^n, \dots\}$ and thus we have no error term in the second equation of (4.2.1). We call this an overlapping state space system.

Obviously, the state x_t is a static factor of \hat{z}_t but not necessarily a minimal one. The next Lemma shows how to determine the dimension of a minimal static factor.

LEMMA 4.2.2. *Let $(\hat{z}_t)_{t \in \mathbb{Z}}$ be a stationary n -dimensional process. Then the dimension of a minimal static factor is the rank $r \leq n$ of the zero-lag variance covariance matrix of $(\hat{z}_t)_{t \in \mathbb{Z}}$, $\mathbb{E}\hat{z}_t\hat{z}_t'$.*

PROOF. Suppose y_t is a static factor with $\hat{z}_t = \Lambda y_t$. Then $\mathbb{E}\hat{z}_t\hat{z}_t' = \Lambda\mathbb{E}y_t y_t'\Lambda'$ and it follows that there can be no static factor of dimension less than the rank r of $\mathbb{E}\hat{z}_t\hat{z}_t'$. Let S be a selector matrix selecting the first linearly independent components of \hat{z}_t then $y_t = S\hat{z}_t$ clearly is a r -dimensional static factor. \square

Let us call the space spanned by the one dimensional components of the minimal static factor y_t the *factor space*. As is easily seen, the factor space coincides with the space spanned by the one dimensional components of \hat{z}_t . Obviously, any basis of the factor space constitutes a minimal static factor. Thus without any further restrictions on the factor model, only the factor space, not the minimal static factor is determined. Any minimal static factors can be obtained by premultiplying the minimal static factor y_t by a non-singular constant matrix R : $\hat{z}_t = \Lambda R^{-1}Ry_t, \forall t \in \mathbb{Z}$.

Static Principal Components. Since we do not observe the minimal static factor directly, we need a procedure for consistently estimating minimal static factors from the observations. In Stock and Watson [2002a] the authors use static principal component analysis for estimation. Thus we give a short description of static principal component analysis, see e.g. Jolliffe [2002]. The main idea of principal component analysis is to compress information contained in a n -dimensional process $(z_t)_{t \in \mathbb{Z}}$. We want to find a r -dimensional process, in fact a linear transformation $O_1'z_t$ of z_t such that

$$\min \mathbb{E} (z_t - O_1'z_t)' (z_t - O_1'z_t)$$

subject to $O_1'O_1 = I_r$. As it turns out the solution of the minimization problem can be achieved via eigenvalue decomposition:

$$\mathbb{E}z_t z_t' = O\Lambda O' = O_1\Lambda_1 O_1' + O_2\Lambda_2 O_2'$$

where Λ_1 contains the r largest eigenvalues of $\mathbb{E}z_t z_t'$ with corresponding (orthonormal) eigenvectors O_1 and Λ_2 contains the $n - r$ smallest eigenvalues of $\mathbb{E}z_t z_t'$ with corresponding (orthonormal) eigenvectors O_2 . Thus the r first principal components $O_1'z_t$ explain as much of the variance covariance matrix of $\mathbb{E}z_t z_t'$ as possible. When dealing with data $\mathbb{E}z_t z_t'$ has to be estimated by $\frac{1}{T} \sum_{t=1}^T z_t z_t'$.

In Stock and Watson [2002a] it is shown that under certain assumptions the space spanned by the first r static principal components of the observations converges for $n, T \rightarrow \infty$ to the space spanned by the static factor. Therefore the first r static principal components of the observations are consistent estimators for a minimal static factor.

The assumptions under which the Stock and Watson [2002a] show consistency are the following. First we state the assumptions on the minimal static factors and the factor loading matrix:

ASSUMPTION 4.2.3. $\frac{1}{n}\Lambda'\Lambda \rightarrow I_r$

ASSUMPTION 4.2.4. $\mathbb{E}y_t y_t' = \gamma(0) > 0$ is diagonal with elements $\gamma_{ii} > \gamma_{jj} > 0$ for $i < j$.

These two assumptions on the factor loading matrix Λ and the minimal static factor y_t are made to determine a normalization up to sign change of the minimal static factor. We already mentioned that any minimal static factor can be obtained by premultiplying the minimal static factor by a non-singular constant matrix R , $\hat{z}_t = \Lambda R^{-1} R y_t, \forall t \in \mathbb{Z}$. The first assumption restricts R to be a constant orthogonal matrix. The second assumption restricts R to be diagonal. Thus we have $r_{ii} = \pm 1, r_{ij} = 0, i \neq j$.

Furthermore these two assumptions guarantee that each of the r components of the minimal factor contributes substantially to the observations z_t .

Further assumption on the factor loading matrix are:

ASSUMPTION 4.2.5. *There exists a bound $\bar{\lambda}$ such that for the elements λ_{ij} of the loading matrix Λ $|\lambda_{ij}| \leq \bar{\lambda} < \infty$ holds.*

ASSUMPTION 4.2.6. $\frac{1}{T} \sum_{t=1}^T y_t y_t' \xrightarrow{P} \gamma(0)$, i.e. ergodicity of the minimal static factor.

The assumptions on the idiosyncratic noise are:

ASSUMPTION 4.2.7. *Let $\gamma_{n,t}^\xi(h) = \frac{1}{n} \mathbb{E} \xi_t' \xi_{t-h}$ then $\lim_{n \rightarrow \infty} \sup_t \sum_{h=-\infty}^{\infty} |\gamma_{n,t}^\xi(h)| < \infty$.*

This assumption allows for some limited serial correlation of the idiosyncratic noise.

Let $\xi_t^{(i)}$ denote the i th component of ξ_t .

ASSUMPTION 4.2.8. $\lim_{n \rightarrow \infty} \sup_t \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left| \mathbb{E} \xi_t^{(i)} \xi_t^{(j)} \right| < \infty$.

This assumption allows the idiosyncratic noise to be weakly correlated in the cross-section.

The last two assumptions are closely related to the assumption of weak dependence of the idiosyncratic noise ξ_t which also allows for some serial and cross-sectional dependence.

ASSUMPTION 4.2.9. $\lim_{n \rightarrow \infty} \sup_{s,t} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left| \mathbb{E} \xi_s^{(i)} \xi_t^{(i)} \xi_s^{(j)} \xi_t^{(j)} \right| < \infty$.

This last assumption limits the fourth moments of the idiosyncratic noise.

4.3. Modeling the Minimal Static Factors

In this section we want to model the dynamics of the latent variables and of the minimal static factors.

Because of Assumptions 4.1.2 and 4.1.3 we know that we can find a tall rational $n \times q$ transfer function $w(z)$, a spectral factor of the spectral density of the latent variables $(\hat{z}_t)_{t \in \mathbb{Z}}$, with no zeros and poles inside and on the unit circle, see Theorem 1.3.1. Therefore we can find a singular ARMA model for the latent variables $(\hat{z}_t)_{t \in \mathbb{Z}}$ with q -dimensional innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$. Note that the q -dimensional innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ are minimal dynamic factors generating the latent variables:

DEFINITION 4.3.1. A process $(\varepsilon_t)_{t \in \mathbb{Z}}$ is called a *dynamic factor* of the n -dimensional process $(\hat{z}_t)_{t \in \mathbb{Z}}$ if it has dimension smaller than n and there is a causal, stable and strictly miniphase transfer function $w(z)$ such that $\hat{z}_t = w(z)\varepsilon_t, \forall t \in \mathbb{Z}$. A *minimal dynamic factors* is a dynamic factor of least possible dimension.

REMARK 4.3.2. Note that it is easy to see that the minimal dimension of a dynamic factor is $q = \text{rk} f_{\hat{z}}$. As already mentioned, the innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ are a minimal dynamic factor. All other minimal dynamic factors can be obtained as $u(z)\varepsilon_t$ where $u(z)$ is a $(q \times q)$ causal, stable and strictly miniphase transfer function.

As is easy to see for a minimal static factor y_t of $(\hat{z}_t)_{t \in \mathbb{Z}}$ there is a matrix Λ of full column rank such that $\hat{z}_t = \Lambda y_t$. Thus we have that $y_t = \Lambda^- \hat{z}_t$ with generalized inverse Λ^- . Therefore a state space system for $(y_t)_{t \in \mathbb{Z}}$ can be obtained from (4.2.1) as

$$(4.3.1) \quad \begin{aligned} x_{t+1} &= Fx_t + G\varepsilon_{t+1} \\ y_t &= Lx_t \end{aligned}$$

where $L = \Lambda^- H$. Obviously, this state space system is stable and controllable. Observability can be seen as follows: Since $\hat{z}_t = \Lambda y_t$ holds we have that $Hx_t = \Lambda \Lambda^- Hx_t$ and thus $H \underbrace{\mathbb{E}x_t x_t'}_{>0} = \Lambda \Lambda^- H \mathbb{E}x_t x_t'$ implies $H = \Lambda \Lambda^- H$. Thus the observability matrices of (4.2.1) and (4.3.1) are related by

$$\begin{pmatrix} H \\ HF \\ \vdots \\ HF^{m-1} \end{pmatrix} = \begin{pmatrix} \Lambda & & & \\ & \Lambda & & \\ & & \ddots & \\ & & & \Lambda \end{pmatrix} \begin{pmatrix} L \\ LF \\ \vdots \\ LF^{m-1} \end{pmatrix}.$$

Clearly, if the observability matrix $\begin{pmatrix} H' & F'H' & \dots & (F^{m-1})'H' \end{pmatrix}'$ has full column rank then also the observability matrix $\begin{pmatrix} L' & F'L' & \dots & (F^{m-1})'L' \end{pmatrix}'$ has full column rank. Since the first matrix on the right hand side of the equation above has full column rank, full column rank of $\begin{pmatrix} L' & F'L' & \dots & (F^{m-1})'L' \end{pmatrix}'$ implies full column rank of $\begin{pmatrix} H' & F'H' & \dots & (F^{m-1})'H' \end{pmatrix}'$. With the same reasoning we can show that the system matrices of (4.2.1) and (4.3.1) have the same zeros since they are related by

$$\begin{pmatrix} I - Fz & -G \\ H & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \Lambda \end{pmatrix} \begin{pmatrix} I - Fz & -G \\ L & 0 \end{pmatrix}.$$

Note that here we encounter a slightly different form of system matrix due to the fact that (4.2.1) and (4.3.1) are overlapping state space systems. Nevertheless, as is shown in Lemma 9.2.7 in Filler [2010] the zeros of these system matrices are the zeros of the corresponding transfer functions. Therefore (4.2.1) is miniphase or zeroless if and only if (4.3.1) is miniphase or zeroless, respectively.

Thus we have shown that the transfer function $k(z) = L(I - Fz)^{-1}G$ corresponding to the system (4.3.1) is causal, stable and miniphase if and only if the transfer function $w(z) = H(I - Fz)^{-1}G$ corresponding to the system (4.2.1) is causal, stable and miniphase. Furthermore $k(z)$ can be realized by an AR system if and only if $w(z)$ can be realized by an AR system. Since $k(z)$ is a $r \times q$ matrix independent of n this transfer function is much more convenient to realize than $w(z)$.

Let us now consider tall transfer functions. In Anderson and Deistler [2008], Filler [2010], Anderson et al. [2012c] it is proved that a tall transfer function with minimal (non-overlapping) state space realization $(A, B, C, D) \in (\mathbb{R}^{m \times m}, \mathbb{R}^{m \times q}, \mathbb{R}^{r \times m}, \mathbb{R}^{r \times q})$ is generically zeroless in $(\mathbb{R}^{m \times m}, \mathbb{R}^{m \times q}, \mathbb{R}^{r \times m}, \mathbb{R}^{r \times q})$. This is proved for $k(z^{-1}) = C(I_m z^{-1} - A)^{-1}B + D$ in the forward shift z^{-1} . We, however, consider a slightly different theorem as we consider $k(z)$ in the backward shift z . Also, the proof is a slight modification of the proof given in Anderson and Deistler [2008] as we have to adapt it for the backward shift and we include some considerations on the geometric multiplicity of zeros which were adopted from Anderson et al. [2012c].

THEOREM 4.3.3. *Let the $r \times q$ rational transfer function $k(z)$ be tall, i.e. $r > q$, with minimal state space realization (A, B, C, D) then for a generic set of parameter matrices $(A, B, C, D) \in (\mathbb{R}^{m \times m}, \mathbb{R}^{m \times q}, \mathbb{R}^{r \times m}, \mathbb{R}^{r \times q})$ $k(z)$ has no (finite) zeros.*

PROOF. In Lemma 1.5.1 we showed that the zeros of $k(z)$ are the same as the zeros of the $(m+r) \times (m+q)$ system matrix

$$M(z) = \begin{pmatrix} I_m - Az & B \\ -Cz & D \end{pmatrix}.$$

Generically $M(0) = \begin{pmatrix} I_m & B \\ 0 & D \end{pmatrix}$ has full column rank $(m+q)$ since D is generically of full column rank. Thus generically $z = 0$ is not a zero of $M(z)$ and the normal rank of $M(z)$ is $(m+q)$. A zero of $M(z)$ must be a zero of every $(m+q) \times (m+q)$ minor of $M(z)$. Since $M(z)$ is (generically) of normal rank $(m+q)$ there is a $(m+q) \times (m+q)$ minor of $M(z)$ which has full rank for almost all $z \in \mathbb{C}$.

Let us denote the first $(m+q) \times (m+q)$ minor by

$$\tilde{M}(z) = \begin{pmatrix} I_m - Az & B \\ -\tilde{C}z & \tilde{D} \end{pmatrix}$$

where \tilde{C} and \tilde{D} are the first q rows of C and D , respectively. As a square matrix, the zeros of $\tilde{M}(z)$ are the zeros of $\det \tilde{M}(z)$. We will show that generically the zeros of $\det \tilde{M}(z)$ are distinct:

We again use the following well known result that a scalar function $f : \mathbb{R}^{(m+r)(m+q)} \rightarrow \mathbb{R}$ which is polynomial in $\text{vec}(A, B, C, D) \in \mathbb{R}^{(m+r)(m+q)}$ is generically nonzero, see e.g. Lee and Markus [1967], Wonham [1985], Bochnak et al. [1998].

Since \tilde{D} is generically non-singular, using the Schur complement we have that generically

$$\det \tilde{M}(z) = \det \tilde{D} \det (I_m - Az + BD^{-1}Cz) = \det \tilde{D} \det (I_m - z(A - BD^{-1}C)).$$

Thus we see that the maximal degree of $\det \tilde{M}(z)$ is smaller or equal to m . Consider $B = 0$ and a diagonal matrix A then it is evident that the maximal degree of $\det \tilde{M}(z)$ is equal to m . Since the coefficients of $\det \tilde{M}(z)$ rationally depend on the parameter matrices (A, B, C, D) , generically the maximal degree of $\det \tilde{M}(z)$ will be m .

It is easy to see that a zero of $\det \tilde{M}(z)$ has multiplicity larger than one if and only if it is a zero of $\det \tilde{M}(z)$ and the derivative of $\det \tilde{M}(z)$ simultaneously. Thus we can use the Sylvester matrix of $\det \tilde{M}(z)$ and its derivative, see e.g. Kailath [1980] p 142, which is a matrix in the coefficients of these two polynomials. The determinant of this matrix is called Sylvester's resultant and it is zero if and only if there is a common factor, i. e. a common zero, of $\det \tilde{M}(z)$ and its derivative. If we again consider the special case $B = 0$ and diagonal matrix A now with distinct diagonal entries, we know that the resultant is not equal to zero

and therefore is nonzero on a generic set of parameters (A, B, C, D) . Thus generically the zeros of $\tilde{M}(z)$ have multiplicity one.

Let us consider the Smith McMillan form of the polynomial matrix

$$\tilde{M}(z) = u \begin{pmatrix} \phi_1 & & \\ & \ddots & \\ & & \phi_{m+q} \end{pmatrix} v.$$

Since generically the zeros are distinct, we have $\phi_1 = \cdots = \phi_{m+q-1} = 1$ and $c\phi_{m+q} = \det \tilde{M}(z)$. Thus generically the dimension of the kernel of any zero of $\tilde{M}(z)$ is one-dimensional.

Let $z_0 \neq 0$ be a zero of $\tilde{M}(z)$ then the kernel $\tilde{M}(z_0)$ is (generically) one-dimensional. Now consider the $(q+1)$ st row of $\begin{pmatrix} -\tilde{C}z_0 & \tilde{D} \end{pmatrix}$. Generically this row will not be orthogonal to the kernel of $\tilde{M}(z_0)$ and thus z_0 is not a zero of $M(z)$. Generically this row will also not be orthogonal to any other kernel corresponding of the other (finitely many) zeros of $\tilde{M}(z)$ and thus the theorem is proved. \square

Note that the set of stable and strictly miniphase state space systems is an open subset of $(\mathbb{R}^{m \times m}, \mathbb{R}^{m \times q}, \mathbb{R}^{r \times m}, \mathbb{R}^{r \times q})$ and thus we also have generic zerolessness of $k(z)$ in this subset.

As we already mentioned, GDFMs were a motivation for our close look at singular AR and ARMA models: The last theorem implies that if $r > q$, generically the minimal static factors can be modeled by a singular AR system. We already described an estimation procedure in Section 1.4.2. For the non-generic case that the minimal static factors have to be modeled by a singular ARMA or state space system, we already considered an estimation procedure in 1.5.2.

Mixed Frequency GDFMs

5.1. Mixed Frequency Static Factors

In this chapter we are considering the case that we have mixed frequency observations of the process $(z_t)_{t \in \mathbb{Z}}$ which permits a GDFM representation. We assume that we observe part of the one dimensional components of $(z_t)_{t \in \mathbb{Z}}$ at every time point, $(z_t^f)_{t \in \mathbb{Z}}$ say, and part of the components at every second time point, $(z_t^s)_{t \in 2\mathbb{Z}}$ say. Further we assume that both the dimension n_f of the fast components $(z_t^f)_{t \in \mathbb{Z}}$ and the dimension n_s of the slow components $(z_t^s)_{t \in 2\mathbb{Z}}$ tend to infinity.

Our goal is to determine a mixed frequency minimal static factor $(y_t) = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix}$ for which we can determine $(y_t^f)_{t \in \mathbb{Z}}$ and $(y_t^s)_{t \in 2\mathbb{Z}}$. More precisely, we want to determine a minimal static factor with the maximum number of fast components $(y_t^f)_{t \in \mathbb{Z}}$. This minimal static factor will have at least one fast component. Thus we can (generically) determine a high frequency AR model for this minimal static factor $(y_t)_{t \in \mathbb{Z}}$ from those second moments which are observed in principle using the techniques described in the second part of the thesis.

Note that here we have shown an interface of our analysis of mixed frequency data and modeling of high dimensional time series.

The procedure for determining the minimal static factor we are suggesting is a form of generalization of the approach in Hallin and Liska [2007] where the authors assume that the process $(z_t)_{t \in \mathbb{Z}}$ consists of two subpanels at the same sampling frequency. We, however, consider the case that the observations of the two blocks are available at different sampling frequencies, i.e. the subpanels are exactly the fast components $(z_t^f)_{t \in \mathbb{Z}}$ and the slow components $(z_t^s)_{t \in 2\mathbb{Z}}$.

In Hallin and Liska [2007] the authors prove that if $(z_t)_{t \in \mathbb{Z}}$ has a GDFM representation also $(z_t^f)_{t \in \mathbb{Z}}$ and $(z_t^s)_{t \in \mathbb{Z}}$ permit GDFM representations. The factor model decomposition of $(z_t^f)_{t \in \mathbb{Z}}$ and $(z_t^s)_{t \in \mathbb{Z}}$ is then closely analyzed:

If we - for the sake of notational simplicity - set aside the nestedness assumption for a moment, we can write the factor model decomposition of $(z_t)_{t \in \mathbb{Z}}$ as

$$z_t = \begin{pmatrix} z_t^f \\ z_t^s \end{pmatrix} = \begin{pmatrix} \hat{z}_t^f \\ \hat{z}_t^s \end{pmatrix} + \begin{pmatrix} \xi_t^f \\ \xi_t^s \end{pmatrix}, t \in \mathbb{Z}.$$

The individual factor model decompositions of $(z_t^f)_{t \in \mathbb{Z}}$ and $(z_t^s)_{t \in \mathbb{Z}}$ are

$$z_t^f = \chi_t^f + \zeta_t^f, t \in \mathbb{Z}$$

and

$$z_t^s = \chi_t^s + \zeta_t^s, t \in \mathbb{Z}$$

where χ_t^f and χ_t^s are the latent variables and ζ_t^f and ζ_t^s is the idiosyncratic noise respectively.

For $i = f, s$, let \mathbb{H}_i^C be the Hilbert space spanned by the one dimensional components of χ_t^i , \mathbb{H}_z^C the space spanned by the one dimensional components of \hat{z}_t and \mathbb{H}_z^I the space spanned by the one dimensional components of ξ_t . Then we can decompose the Hilbert space spanned by the one dimensional components of z_t^f into four orthogonal spaces

$$(\mathbb{H}_f^C \cap \mathbb{H}_s^C) + \mathbb{H}_f^C \setminus (\mathbb{H}_f^C \cap \mathbb{H}_s^C) + \mathbb{H}_z^C \setminus \mathbb{H}_f^C + \mathbb{H}_z^I$$

where $\mathbb{A} \setminus \mathbb{B}$ denotes the orthocomplement of \mathbb{B} in \mathbb{A} . Thus the two different factor model representations can be used to further decompose both z_t^f into four mutually orthogonal components.

$$z_t^f = \underbrace{\phi_t^f + \psi_t^f}_{\chi_t^f} + \underbrace{\omega_t^f + \xi_t^f}_{\zeta_t^f}, t \in \mathbb{Z}$$

where ϕ_t^f is called strongly common as $\phi_t^f \in (\mathbb{H}_f^C \cap \mathbb{H}_s^C)$, ψ_t^f is called weakly common as $\psi_t^f \in \mathbb{H}_f^C \setminus (\mathbb{H}_f^C \cap \mathbb{H}_s^C)$, ω_t^f is called weakly idiosyncratic as $\omega_t^f \in \mathbb{H}_z^C \setminus \mathbb{H}_f^C$, and ξ_t^f is called strongly idiosyncratic as $\xi_t^f \in \mathbb{H}_z^I$. Completely analogously we can decompose z_t^s into

$$z_t^s = \underbrace{\phi_t^s + \psi_t^s}_{\chi_t^s} + \underbrace{\omega_t^s + \xi_t^s}_{\zeta_t^s}, t \in \mathbb{Z}.$$

Note that ω_t^f is weakly idiosyncratic which means that for the individual factor model of $\left(z_t^f\right)_{t \in \mathbb{Z}}$ ω_t^f is part of the idiosyncratic noise, but for the full factor model of $(z_t)_{t \in \mathbb{Z}}$ ω_t^f is part of the common component, the latent variables. Thus $\omega_t^f \in \mathbb{H}_s^C$ holds. An analogous statement holds for ω_t^s .

Therefore we can justify the following procedure for our mixed frequency GDFM setting:

We treat the fast components $\left(z_t^f\right)_{t \in \mathbb{Z}}$ and the slow components $\left(z_t^s\right)_{t \in 2\mathbb{Z}}$ separately as two different blocks. Let Assumptions 4.2.3-4.2.9 hold for $\left(z_t^f\right)_{t \in \mathbb{Z}}$ and $\left(z_t^s\right)_{t \in 2\mathbb{Z}}$. Let r_f be the minimal number of static factors of a factor model decomposition of $\left(z_t^f\right)_{t \in \mathbb{Z}}$ and r_s the minimal number of static factors of a factor model decomposition of $\left(z_t^s\right)_{t \in 2\mathbb{Z}}$. We compute the r_f and r_s first static principal components of z_t^f for $t \in \mathbb{Z}$ and z_t^s for $t \in 2\mathbb{Z}$ respectively. We call these first fast and slow static factors $\left(y_t^f\right)$ and $\left(\tilde{y}_t^s\right)$. Then, we consider the covariance matrix of the stacked vector $\begin{pmatrix} y_t^f \\ \tilde{y}_t^s \end{pmatrix}$ and select the first basis in terms of rows for the row space of this covariance matrix. Then, we cancel those elements in \tilde{y}_t^s which do not correspond to a basis row in the covariance matrix described above. It is easy to see that the static factor $y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix}$ obtained by this procedure is minimal and contains the maximum number of fast components.

For n_f and n_s going to infinity, this leads to a consistent estimation procedure for a version of the true static factors.

A drawback of this procedure is the following: The fast static factor y_t^f is a minimal static factor of the fast subpanel and only explains χ_t^f which is the sum of the strongly common component ϕ_t^f and weakly common component ψ_t^f . The weakly idiosyncratic components ω_t^f is explained by y_t^s . If ω_t^f is nonzero, the latent variables of the full panel \hat{z}_t^f are not a linear transformation of y_t^f alone but of y_t^f and y_t^s . The factor loading matrix Λ in $\hat{z}_t = \Lambda y_t$ does not have the structure $\Lambda = \begin{pmatrix} \Lambda^f & 0_{n_f \times r_s} \\ \Lambda^{sf} & \Lambda^s \end{pmatrix}$ and thus we cannot determine all components of the fast latent variables \hat{z}_t^f for all $t \in \mathbb{Z}$. However, using the techniques of the second part of the thesis, we can (generically uniquely) determine the system and noise parameters of an AR model for the minimal static factor. Consequently, we can interpolate (with error) the missing slow components of $(y_t)_{t \in \mathbb{Z}}$ for instance using the Kalman smoother for missing observations, see Jones [1980]. This interpolated minimal static factor can then be used to interpolate (with error) the latent variables \hat{z}_t for all $t \in \mathbb{Z}$.

5.2. Simulation Example

In this section we illustrate the procedure described in the last section.

For $n = T = 1000$, we simulate observations of a GDFM

$$z_t = \Lambda y_t + \xi_t$$

with a two dimensional minimal static factor y_t . This minimal static factor y_t is the solution of a regular AR(1) system with parameters

$$\left(a_1 = \begin{pmatrix} -0.04765167 & 0.58589510 \\ 1.13259476 & 0.04362198 \end{pmatrix}, b = I_2 \right).$$

We partition the observations z_t such that $n_f = n_s = 500$. We choose a factor loading matrix

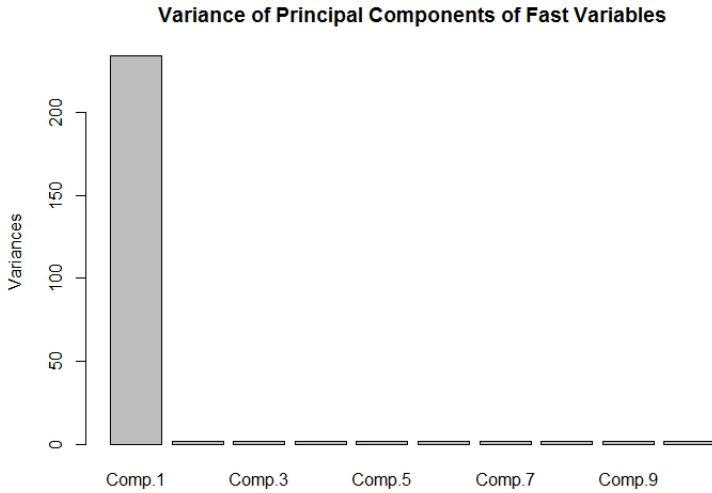
$$\Lambda = \begin{pmatrix} \Lambda^f & \Lambda^{fs} \\ \Lambda^{sf} & \Lambda^s \end{pmatrix} \text{ such that we have}$$

- nonzero entries of Λ^f and Λ^{sf} and thus $\phi_t^f \neq 0$ and $\phi_t^s \neq 0$, and
- nonzero entries of Λ^s and zeros for all but 5 entries of Λ^{fs} and thus $\psi_t^s \neq 0$ and $\omega_t^f \neq 0$.

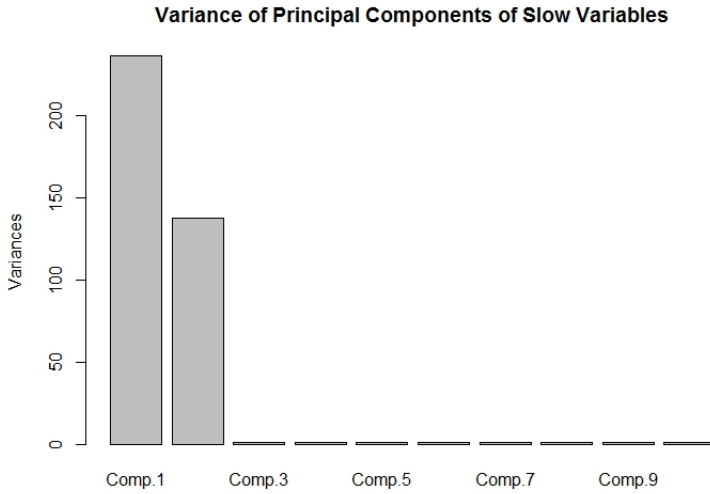
The idiosyncratic errors are chosen uncorrelated in the cross-section.

We want to estimate the mixed frequency minimal static factor via the procedure described in the last section:

Therefore we compute the static principal components \hat{y}_t^f of z_t^f , $t = 1, \dots, 1000$. The plot of the variances of the first principal components suggests one fast static factor.



For z_{2t}^s , $t = 1, \dots, 500$, we find two dominant slow static factors \hat{y}_t^s .



We compute the eigenvalues of the zero lag covariance of the (scaled) fast factor and the first slow factor $\begin{pmatrix} \hat{y}_t^f \\ \hat{y}_t^{s(1)} \end{pmatrix}$: 1.029 and 0.966. The components are linearly independent. The eigenvalues of the zero lag covariances of all the (scaled) estimated factors $\begin{pmatrix} \hat{y}_t^f \\ \hat{y}_t^s \end{pmatrix}$ are 1.994, 0.998 and 0.001. Thus we choose $\hat{y}_t = \begin{pmatrix} \hat{y}_t^f \\ \hat{y}_t^{s(1)} \end{pmatrix}$ as the minimal static factor. Thus we estimated $r = 2$ correctly.

To determine whether the space spanned by the estimated minimal static factor \hat{y}_t is a good estimate of the space spanned by the true minimal static factor y_t we regress \hat{y}_t onto y_t . We compute the coefficient of determination R^2 as a measure of the goodness of fit. The closer R^2 is to 1 the better is the static factor space estimated. For the fast static factor, we have $R_f^2 = 0.997$ and for the slow static factor we have $R_s^2 = 0.998$. Thus it seems the (mixed frequency) estimator \hat{y}_t of the minimal static factor is a good choice.

Bibliography

- H. Akaike. Stochastic Theory of Minimal Realization. *IEEE Transactions on Automatic Control*, 19(6):667–674, 1974.
- T. Amemiya and R. Y. Wu. The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association*, 67(339):628–632, 1972.
- B. D. O. Anderson and M. Deistler. Properties of Zero-free Transfer Function Matrices. *SICE Journal of Control, Measurement and System Integration*, 1(4):284–292, 2008.
- B. D. O. Anderson, M. Deistler, W. Chen, and A. Filler. Autoregressive models of singular spectral matrices. *Automatica*, 48:2843–2849, 2012a. doi: 10.1109/CDC.2009.5399891.
- B. D. O. Anderson, M. Deistler, E. Felsenstein, B. Funovits, P. Zadrozny, M. Eichler, W. Chen, and M. Zamani. Identifiability of regular and singular multivariate autoregressive models. *Proceedings of the 51th IEEE Conference on Decision and Control*, 2012b.
- B. D. O. Anderson, M. Deistler, and A. Filler. Generic properties of system zeros. Unpublished Working Paper, 2012c.
- B. D. O. Anderson, M. Deistler, E. Felsenstein, B. Funovits, Kölbl L., and M. Zamani. Multivariate ar systems and mixed frequency data: Identifiability and estimation. *Submitted to Econometric Theory*, February 2014.
- S. Bittanti, P. Colaneri, and G. De Nicolao. The Difference Periodic Riccati Equation for the Periodic Prediction Problem. *IEEE Transactions on Automatic Control*, 33(8):706–712, 1988.
- J. Bochnak, M. Coste, and M. Roy. *Real Algebraic Geometry*. Springer, 1998.
- D. R. Brillinger. *Time Series: Data Analysis and Theory*. Holden Day, San Francisco, 1981.
- P. J. Brockwell and R. A. Davis. *Time Series: Data Theory and Methods*. Springer, New York, 1987.
- G. Chamberlain. Funds, Factors and Diversification in Arbitrage Pricing Models. *Econometrica*, 51(5):1305–1324, 1983.
- G. Chamberlain and M. Rothschild. Arbitrage, Factor Structure and Meanvariance Analysis on Large Asset Markets. *Econometrica*, 51(5):1281–1304, 1983.

- B. Chen and P. A. Zdrozny. An Extended Yule-Walker Method for Estimating a Vector Autoregressive Model with Mixed-Frequency. *Advances in Econometrics*, 13:47–73, 1998.
- Tongwen Chen and Bruce Francis. *Optimal Sampled-Data Control Systems*. Springer Verlag, 1995.
- W. Chen, B. D. O. Anderson, M. Deistler, and A. Filler. Solutions of Yule-Walker equations for Singular AR processes. *Journal of Time Series Analysis*, 32:531–538, 2011. doi: 10.1111/j.1467-9892.2010.00711.x.
- M. Deistler. z-transform and identification of linear econometric models with autocorrelated errors. *Metrika*, 22:13–25, 1975.
- M. Deistler. *A Birds Eye View on System Identification*, volume 364 of *Lecture Notes in Control and Information Sciences*, pages 59–71. Springer Berlin/Heidelberg, 2007. ISBN 978-3-540-73569-4.
- M. Deistler, K. Peternell, and W. Scherrer. Consistency and relative efficiency of subspace methods. *Automatica*, 31(12):1865–1875, 1995.
- M. Deistler, B. D. O. Anderson, A. Filler, Ch. Zinner, and W. Chen. Generalized linear dynamic factor models - an approach via singular autoregressions. *European Journal of Control*, 16(3):211–224, 2010. doi: 10.3166/ejc.16.211-224.
- M. Deistler, A. Filler, and B. Funovits. AR Systems and AR Processes: The Singular Case. *Communications in Information and Systems*, 11(3):225–236, 2011.
- E. Felsenstein. Yule-Walker Gleichungen für singuläre AR Systeme. Master’s thesis, Vienna University of Technology, 2010.
- A. Filler. *Generalized Dynamic Factor Models - Structure Theory and Estimation for Single Frequency and Mixed Frequency Data*. PhD thesis, Vienna University of Technology, 2010.
- G. D. Forney jr. Minimal Bases of Rational Vector Spaces, with Applications to Multivariable Linear Systems. *SIAM Journal of Control*, 13(3):493–520, 1975.
- M. Forni and M. Lippi. The generalized dynamic factor model: representation theory. *Econometric Theory*, 17:1113–1141, 2001.
- M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The Generalized Dynamic Factor Model: Identification and Estimation. *The Review of Economics and Statistics*, 82(4):540–554, 2000.
- M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: Consistency and rates. *Journal of Econometrics*, 119:231–255, 2004.
- M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting. *Journal of the American Statistical Association*, 100:830–840, 2005.

- J. Geweke. *Latent Variables in Socio-Economic Models*, Ed.: Aigner, D. J., chapter 19, The Dynamic Factor Analysis of Economic Time Series Models, pages 365–383. North-Holland Amsterdam, 1977.
- E. Ghysels. Macroeconomics and the reality of mixed frequency data. *Working Paper, Department of Economics, University of North Carolina*, 2012.
- E. Ghysels, P. Santa-Clara, and R. Valkanov. Predicting Volatility: Getting the Most out of Return Data Sampled at Different Frequencies. *Journal of Econometrics*, 131(1-2):59–95, 2006.
- M. Hallin and R. Liska. Determining the number of factors in the general dynamic factor model. *J. of the American Statistical Association*, 102:603–617, 2007.
- E. J. Hannan. *Multiple Time Series*. Wiley, 1970.
- E. J. Hannan and M. Deistler. *The Statistical Theory of Linear Systems*. SIAM Classics in Applied Mathematics, Philadelphia, 2012.
- B. L. Ho and R. E. Kalman. Effective construction of linear state-variable models from input/output functions. *Regelungstechnik*, 14:545–548, 1966.
- Y. Inouye. Modeling of Multichannel Time Series and Extrapolation of Matrix-Valued Autocorrelation Sequences. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(1):45–55, 1983.
- I. T. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics, 2002.
- R. H. Jones. Maximum likelihood fitting of arma models to time series with missing observations. *Technometrics*, 22:389–395, 1980.
- T. Kailath. *Linear Systems*. Prentice Hall, Englewood Cliffs, N.J., 1980.
- R. E. Kalman. Mathematical description of linear dynamical systems. *SIAM J. Control*, 1:152–192, 1963.
- R. E. Kalman. Irreducible realizations and the degree of a rational matrix. *Journal of the Society for Industrial and Applied Mathematics*, 13(2):520–544, Jun 1965.
- R. E. Kalman. Algebraic geometric description of the class of linear systems of constant dimension. In *8th Annual Princeton Conference on Information Sciences and Systems, NJ*, 1974.
- W.E. Larimore. System identification, reduced-order filtering and modeling via canonical variate analysis. *American Control Conference*, pages 445–451, 1983.
- E. B. Lee and L. Markus. *Foundations of Optimal Control Theory*. John Wiley & Sons, New York, London, Sydney, 1967.
- G. Reinsel. *Elements of Multivariate Time Series Analysis*. Springer Series in Statistics, 1993.
- Y. A. Rozanov. *Stationary Random Processes*. Holden-Day, San Francisco, 1967.

- T. J. Sargent and C. A. Sims. Business cycle modeling without pretending to have too much a priori economic theory. *Working Paper 55, Federal Reserve Bank of Minneapolis*, 1977.
- W. Scherrer and M. Deistler. A Structure Theory for Linear Dynamic Errors-in-Variable Models. *SIAM Journal on Control and Optimization*, 36(6):2148–2175, 1998.
- M. S. Srivastava and D. von Rosen. Regression models with unknown singular covariance matrix. *Linear Algebra and its Applications*, 354:255–273, 2002.
- J. H. Stock and M. W. Watson. Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association*, 97:1167–79, 2002a.
- J. H. Stock and M. W. Watson. Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics*, 20:147–62, 2002b.
- J. H. Stock and M. W. Watson. An empirical comparison of methods for forecasting using many predictors. *Mimeo, Princeton University*, 2005.
- G. C. Tiao. Asymptotic Behaviour of Temporal Aggregates of Time Series. *Biometrika*, 59(3):525–531, 1972.
- P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Pearson Education India, 1993.
- W. M. Wonham. *Linear Multivariable Control - A Geometric Approach*. Springer-Verlag New York, 1985.
- P. A. Zdrozny. Forecasting US GNP at monthly intervals with an estimated bivariate time series model. *Federal Reserve Bank of Atlanta Economic Review*, 75:2–15, 1990.
- M. Zamani. *Modeling Multivariable Time Series Using Regular and Singular Autoregressions*. PhD thesis, The Australian National University, 2014.

Curriculum Vitae

Elisabeth Felsenstein

Date of Birth 07.02.1986
Place of Birth Ried im Innkreis
Citizenship Austrian

Education

11/2013 Research visit at the RSISE, ANU Canberra, Australia
11/2012 Research visit at the RSISE, ANU Canberra, Australia
11/2011 Research visit at the RSISE, ANU Canberra, Australia
Since 11/2010 Ph.D student in Technical Mathematics at the Vienna University of Technology
10/2005-10/2010 Dipl.-Ing. (M.Sc.) with distinction in Technical Mathematics, with focus on Mathematics in Economics (Vienna University of Technology), Master Thesis: Yule Walker Equations for Singular AR Systems, supervised by Em. O. Univ. Prof. Dipl.-Ing. Dr.techn. Manfred Deistler
01/2009-05/2009 Erasmus exchange semester at the KTH Stockholm, Sweden
10/2004-06/2005 Enrolled at the University of Vienna in Molecular Biology
06/2004 Matura (equivalent to high school diploma) with distinction at Akademisches Gymnasium Wien

Professional Career

Since 11/2010 Project assistant (FWF) at the Institute of Mathematical Methods in Economics, Vienna University of Technology
08/2008-09/2008 Internship at the AWS (Austria Wirtschaftsservice) in Risk Management