# The Lindquist Symposium in Systems Theory

Manfred Deistler (TUW),
B.D.O. Anderson (ANU) and A. Braumann (TU Braunschweig)

KTH Stockholm, Nov 24-25, 2017

TECHNISCHE UNIVERSITÄT WIEN Vienna University of Technology

Australian National University

# Contributors

B.D.O. Anderson

A. Braumann

# Main References to our work

- Anderson & Deistler (2008)
- Deistler et al. (2010)
- Deistler et al. (2011)
- Anderson et al. (2012)
- Anderson et al. (2016a)
- Anderson et al. (2016b)
- Koelbl et al. (2016)
- Filler (2010)
- Felsenstein (2014)
- Koelbl (2015)
- Deistler et al. (2017)
- Koelbl & Deistler (2016)
- Braumann et al. (2016)

# Outline

# Model Classes and Observation Schemes

## VAR Systems:

$$a(z)y_t = v_t, \quad t \in \mathbb{Z} \quad \text{(high frequency)} \tag{1}$$

z... backward shift as well as a complex variable

$$a(z) = I_n - A_1 z - \cdots - A_p z^p, \quad A_i \in \mathbb{R}^{n \times n}$$

## Assumptions

- $\det(a(z)) \neq 0, \forall |z| \leq 1$ (stability)
- $(v_t)$ white-noise, $\Sigma = \mathbb{E}\left(v_t v_t^T\right) > 0$. We write $v_t = b\varepsilon_t$, $(\varepsilon_t)$ white-noise with $\mathbb{E}\left(\varepsilon_t \varepsilon_t^T\right) = I_n$.
- $A_p$ is nonsingular

# Companion Form

$$\underbrace{\begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}}_{=x_{t+1}} = \underbrace{\begin{pmatrix} A_1 & \cdots & A_{p-1} & A_p \\ I_n & & & \\ & \ddots & & \\ & & I_n & 0 \end{pmatrix}}_{\mathscr{A}} \underbrace{\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}}_{=x_t} + \underbrace{\begin{pmatrix} b \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\mathscr{B}} \varepsilon_t$$

or in short, written as a state space system

$$x_{t+1} = \mathscr{A} x_t + \mathscr{B} \varepsilon_t \tag{2}$$

$$y_t = \underbrace{(A_1, \ldots, A_p)}_{=\mathscr{C}} x_t + \varepsilon_t \tag{3}$$

$$\Gamma_p = \mathbb{E} x_t x_t^T$$

fulfills the Lyapunov equation

$\Gamma_p = \mathscr{A} \Gamma_p \mathscr{A}^T + \mathscr{B}\mathscr{B}^T = \sum_{j=0}^{\infty} \mathscr{A}^j \mathscr{B}\mathscr{B}^T \left(\mathscr{A}^T\right)^j$

# Companion Form

An alternative state space system is of the form

$$x_{t+1} = \mathscr{A} x_t + \mathscr{B} \varepsilon_t \tag{4}$$

$$y_t = \underbrace{(I_n, 0, \dots, 0)}_{=\mathscr{C}} x_{t+1} \tag{5}$$

In the following we use both representations.

Note that a state space system $(\mathscr{A}, \mathscr{B}, \mathscr{C})$ is minimal if and only if it is controllable and observable. Since for $\Sigma > 0$ always $\Gamma_p > 0$ holds, the system is always controllable. It is observable and thus minimal, if and only if $A_p$ is nonsingular.

## Other model classes considered here

- VARMA systems
- Singular VAR and VARMA systems
- Generalized linear dynamic factor models (GDFM's)

# VAR Case

**Parameter space for high frequency AR model (fixed $p$)**

$$\Theta = \underbrace{\left\{(A_1, \ldots, A_p) \in \mathbb{R}^{n \times np} \middle| \det\left(a(z)\right) \neq 0, |z| \leq 1\right\}}_{=S} \times \left\{\Sigma \middle| \Sigma = \Sigma^T,\ \Sigma > 0\right\}$$

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

### Transfer Function

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

### Steady state solution

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

### Spectral Density

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

# Observation Schemes

"Stock" data:

$$y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} \quad \begin{array}{l} \text{of dimension } (n_f \times 1) \\ \text{of dimension } (n_s \times 1) \end{array}$$

where $y_t^f,\ t \in \mathbb{Z}$, and $y_t^s,\ t \in N\mathbb{Z}$, $N = 2, 3, \ldots,$

**Population second moments which can be directly observed**

$$\gamma^{ff}(h) = \mathbb{E}\left(y_{t+h}^f (y_t^f)^T\right), \quad h \in \mathbb{Z}$$

$$\gamma^{sf}(h) = \mathbb{E}\left(y_{t+h}^s (y_t^f)^T\right), \quad h \in \mathbb{Z}$$

$$\gamma^{ss}(h) = \mathbb{E}\left(y_{t+h}^s (y_t^s)^T\right), \quad h \in N\mathbb{Z}$$

# Flow Data and More General Aggregation Schemes

## More general linear aggregation schemes

$$y_t = \begin{pmatrix} y_t^f \\ w_t \end{pmatrix} \quad \begin{array}{l} \text{of dimension } (n_f \times 1) \\ \text{of dimension } (n_s \times 1) \end{array}$$

where $w_t = k_0 y_t^s + \cdots + k_{N-1} y_{t-N+1}^s$, $k_0$ non-singular, $t \in N\mathbb{Z}$

## Flow data

Here $k_0 = \cdots = k_{N-1} = I_{n_s}$

# Outline

# g-Identifiability

**Identifiability**

Can the system and noise parameters $\theta \in \Theta$ of the underlying system be uniquely determined from those second moments which can be directly observed?

**Genericity**

A property is said to hold generically if it holds on a set containing an open and dense subset of the parameter space.

In many cases we cannot show identifiability, but only generic identifiability, which we call g-identifiability.

# g-Identifiability

## Complement of a Proper Algebraic Variety

Our results are even stronger as we show identifiability on the complement of a proper algebraic variety.

## Well-Posedness

In addition to identifiability, we explicitly describe the mapping attaching the parameters to the second moments which can be directly observed as well as the continuity of this mapping.

As a consequence of well-posedness, consistent estimators of the second moments, which can be directly observed, give consistent estimators for the high-frequency parameters in this case.

# Outline

# General Remarks

- Mixed frequency (MF) data are quite common, in particular in high dimensional time series.
- Aims:
  - Estimation of high frequency (VAR) models
  - Forecasting, Nowcasting, Interpolation
- Approaches:
  - Use only data at the lowest frequency: information loss
  - Interpolate the missing observations and estimate using high frequency "data".
  - Estimate a continuous-time model from low frequency data by "inverting" the aliasing formula (Phillips 1973, Hansen and Sargent 1983). No use of cross correlations.

# Relation to Continuous-Time Systems

## Idea

If all data, when sampled at the slowest rate, determine the parameters of "the corresponding" continuous-time system, then we have solved our problem.

## Aliasing

$$f^{(\Delta)}(\lambda) = \sum_{j=-\infty}^{\infty} f\left(\lambda + \frac{2\pi j}{\Delta}\right)$$

$f^{(\Delta)}$: spectral density of discrete-time process, sampling interval $\Delta$
$f$: spectral density of continuous-time process

Without further restrictions, the mapping $f \rightarrow f^{\Delta}$ is not injective.
Discrete-to-discrete analogon: finite sum

# Relation to Continuous-Time Systems

## Band limited processes

Nyquist: Sample at a rate exceeding twice the bandwidth, then the errors are “small”.

# Further Approaches

- MIDAS regression (Ghysels et al. 2006, 2007)
- Kalman filtering (Zadrozny 1990)

## Our approach:

- Direct estimation of the parameters of the high frequency (VAR) model from MF data.
  - First step (Anderson et al. 2012, Anderson et al. 2016): Identifiability: "constructive", well-posedness, consistent estimators provided.
  - Once the VAR parameters are given, all second moments of the output process are obtained and all linear least squares approximations (forecasting, nowcasting, interpolation) can be performed, e.g. via the Kalman filter.

# Outline

# AR case, $n = 2$, $p = 1$

$$y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} = \begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \end{pmatrix} \begin{pmatrix} y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} + \begin{pmatrix} \varepsilon_t^f \\ \varepsilon_t^s \end{pmatrix}, \; t \in \mathbb{Z}$$

$$y_t^f = \begin{pmatrix} a_{ff} & a_{fs} \end{pmatrix} \begin{pmatrix} y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} + \varepsilon_t^f \tag{6}$$

Here all moments needed for projecting $y_t^f$ on $y_{t-1}^f$ and $y_{t-1}^s$ and thus $a_{ff}$, $a_{fs}$ and $\sigma_f^2 = \left(\mathbb{E}\varepsilon_t^f\right)^2$ are uniquely determined.

# AR case, cont.

$$y_t^s = \begin{pmatrix} a_{sf} & a_{ss} \end{pmatrix} \begin{pmatrix} y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} + \varepsilon_t^s, \qquad \mathbb{E}\varepsilon_t \varepsilon_t^T = \begin{pmatrix} \sigma_f^2 & 0 \\ 0 & \sigma_s^2 \end{pmatrix}$$

The autocovariance $\mathbb{E} y_t^s y_{t-1}^s$ is not available. We replace $y_{t-1}^s$ by $a_{sf} y_{t-2}^f + a_{ss} y_{t-2}^s + \varepsilon_{t-2}^s$ and project $\begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix}$ onto $\begin{pmatrix} y_{t-1}^f \\ y_{t-2}^f \\ y_{t-2}^s \end{pmatrix}$:

$$\begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} = \begin{pmatrix} a_{ff} & a_{fs} a_{sf} & a_{fs} a_{ss} \\ a_{sf} & a_{ss} a_{sf} & a_{ss}^2 \end{pmatrix} \underbrace{\begin{pmatrix} y_{t-1}^f \\ y_{t-2}^f \\ y_{t-2}^s \end{pmatrix}}_{=(*)} + \begin{pmatrix} \bar{\varepsilon}_t^f \\ \bar{\varepsilon}_t^s \end{pmatrix} \qquad (7)$$

where the components of $(*)$ are linearly independent and $\begin{pmatrix} \bar{\varepsilon}_t^f \\ \bar{\varepsilon}_t^s \end{pmatrix} = \begin{pmatrix} a_{fs} \\ a_{ss} \end{pmatrix} \varepsilon_{t-1}^s + \begin{pmatrix} \varepsilon_t^f \\ \varepsilon_t^s \end{pmatrix}$. The parameters $a_{ff}$, $a_{fs}$ and $\sigma_f^2$ are unique from (6). If $a_{fs} \neq 0$ or $a_{sf} \neq 0$ or $a_{ss} = 0$, then the system is identified from (7).

## MA case, $n = 2, q = 1$

$$y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} = I \begin{pmatrix} \varepsilon_t^f \\ \varepsilon_t^s \end{pmatrix} + \underbrace{\begin{pmatrix} b_{ff} & b_{fs} \\ b_{sf} & b_{ss} \end{pmatrix}}_{=B} \begin{pmatrix} \varepsilon_{t-1}^f \\ \varepsilon_{t-1}^s \end{pmatrix}$$

Miniphase condition: $\det \underbrace{(I - Bz)}_{=b(z)} \neq 0, \ |z| \leq 1$

Parameter space:

$$\Theta = \underbrace{\left\{ B \in \mathbb{R}^{2\times 2} |\det(b(z)) \neq 0, |z| \leq 1 \right\}}_{=S} \times \{\Sigma | \Sigma > 0\}$$

$S$ is open in $\mathbb{R}^{2\times 2}$

$B \overset{i_P}{\to} b(e^{-i\lambda})$, $||b(e^{-i\lambda})||_{\text{sup}} = \sup_{\lambda \in [-\pi,\pi]} ||b(e^{-i\lambda})||_{\max}$, $i_P$ is a homeomorphism

# MA case, cont.

Autocovariances: $(\underbrace{\mathbb{E}y_0y_0^T}_{\gamma(0)}, \underbrace{\mathbb{E}y_1y_0^T}_{\gamma(1)})$, $\gamma(j) = 0, \quad |j| \geq 2$

Spectral density:
$f(\lambda) = (2\pi)^{-1} \sum_{j=-1}^{1} \gamma(j)e^{-i\lambda j} = b(e^{-i\lambda})\Sigma b(e^{-i\lambda})^*$

Let $\Gamma = \left\{(\mathrm{vech}\gamma(0)^T, \mathrm{vech}\gamma(1)^T, f(\lambda) > 0, \lambda \in [-\pi, \pi]\right\} \in \mathbb{R}^{3+4}$ is open

We have identifiability for the high frequency case and in addition $\Theta \overset{i}{\longleftrightarrow} \Gamma$ is a homeomorphism. Thus, if $\gamma_{ss}(1)$ is not directly observed, $i^{-1}\left(\gamma(0), \begin{pmatrix} \gamma_{ff}(1) & \gamma_{fs}(1) \\ \gamma_{sf}(1) & * \end{pmatrix}\right)$ is a one-dimensional equivalance class in $\Theta$.

# Outline

# g-Identifiability for the VAR Case, Stock Variables

## Identifiability of system parameters $(A_1, \ldots, A_p)$

Right-multiplying fast, lagged variables, taking expectations

$$\underbrace{\mathbb{E}\left[y_t\left((y^f_{t-1})^T,(y^f_{t-2})^T,\ldots\right)\right]}_{=Z_1^\infty} = (A_1,\ldots,A_p)\underbrace{\mathbb{E}\left[\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p}\end{pmatrix}\left((y^f_{t-1})^T,(y^f_{t-2})^T,\ldots\right)\right]}_{=Z^\infty}$$

contains only second moments which can be estimated from mixed frequency data. However, not all second moments which can be directly observed are used.

# The extended Yule-Walker equations

## Using Cayley-Hamilton

Define

$$Z = \mathbb{E}\left[\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}\left((y^f_{t-1})^T, \dots, (y^f_{t-np})^T\right)\right] \in \mathbb{R}^{np \times n_f np}.$$

$$= \left(\mathscr{K}, \mathscr{A}\mathscr{K}, \mathscr{A}^2\mathscr{K}, \dots, \mathscr{A}^{np-1}\mathscr{K}\right), \quad \mathscr{K} = \mathbb{E}\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}(y^f_{t-1})^T$$

By Cayley-Hamilton $rk(Z^\infty) = rk(Z)$.

## Extended Yule-Walker equations

$$\mathbb{E}\left[y_t\left((y^f_{t-1})^T, \dots (y^f_{t-np})^T\right)\right] = (A_1, \dots, A_p)Z$$

# g-Identifiability of System Parameters

**Theorem**

*The matrix $Z = \left(\mathscr{K}, \mathscr{A}\mathscr{K}, \mathscr{A}^2\mathscr{K}, \ldots, \mathscr{A}^{np-1}\mathscr{K}\right)$ has full row rank on a generic subset of the parameter space. Therefore the system parameters are identifiable on this set.*

- Use rational structure of $vec(\Gamma_p) = \left(I_{(np)^2} - (\mathscr{A} \otimes \mathscr{A})\right)^{-1} vec(\mathscr{B}\mathscr{B}^T)$.
- There is a particular $((A_1, \ldots, A_p), b)$, w.l.o.g. $b$ can be chosen as an $n$-vector, in the parameter space $\Theta$ such that $\det(Z) \neq 0$ holds. Consider

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

- Thus $\det(Z) \neq 0$ holds on the complement of a proper algebraic subset of $\Theta$.

# Proof continued

- Note, that an algebraic variety is the set of common zeros of a finite number of polynomials.
- It is called proper if it is not equal to the embedding Euclidean space.
- Since the set of system parameters $S$ is open in this embedding Euclidean space, the result follows.

# g-Identifiability of Noise Parameters

For given $(A_1 \ldots, A_p)$,
$\Sigma$ is generically obtained from a vectorization of the equations

$$\Gamma_p = \mathscr{A}\Gamma_p\mathscr{A}^T + \mathscr{H}^T\Sigma\mathscr{H}$$
$$\gamma(0) = E\left(y_t y_t^T\right) = \mathscr{H}\Gamma_p\mathscr{H}^T,$$

where

$$\mathscr{H} = (I_n, 0_{n\times n}, \ldots, 0_{n\times n})$$

which leads to

$$vec\,(\gamma(0)) = (\mathscr{H}\otimes\mathscr{H})\left(I_{(np)^2} - \mathscr{A}\otimes\mathscr{A}\right)^{-1}(\mathscr{H}^T\otimes\mathscr{H}^T)vec\,(\Sigma)\,.$$

# Theorem and Idea of Proof

**Theorem**

The matrix $(\mathscr{H} \otimes \mathscr{H})\left(I_{(np)^2} - \mathscr{A} \otimes \mathscr{A}\right)^{-1}(\mathscr{H}^T \otimes \mathscr{H}^T)$ is generically non-singular. *Therefore $\Sigma$ is unique generically.*

- Find $(A_1, \dots, A_p)$ such that $(\mathscr{H} \otimes \mathscr{H})\left(I_{(np)^2} - \mathscr{A} \otimes \mathscr{A}\right)^{-1}(\mathscr{H}^T \otimes \mathscr{H}^T)$ is non-singular.
- In an open neighborhood of this $(A_1, \dots, A_p)$ non-singularity holds, so we have a nonempty intersection with the generic set of identifiable system parameters.
- The determinant is a rational function in the free entries $(A_1, \dots, A_p)$, thus non-singularity holds for a generic set in the parameter space.

**$rk(Z) = np$ not necessary for identifiability**

as Extended Yule Walker equations do not use the full information contained in the second moments which can be observed in principle.

We have not been able to precisely describe the set of high frequency parameters where $Z$ has full row rank.

# Outline

# Blocking

An alternative way to retrieve the parameters of the high-frequency system from the second moments which can be directly observed:
This approach (Anderson et al. 2016):

- provides additional insights
- all moments which can be directly observed are used
- leads to alternative estimation procedures

Consider the case $N = 2$ and the processes

$$\tilde{y}_t = \begin{pmatrix} y_t \\ y^f_{t-1} \end{pmatrix}, t \in 2\mathbb{Z} \qquad Y_t = \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}, t \in 2\mathbb{Z}$$

Then $(y_t | \, t \in \mathbb{Z})$ and $(Y_t | \, t \in 2\mathbb{Z})$ are both AR with minimal state dimension $np$.
Note that $(\tilde{y}_t | \, t \in 2\mathbb{Z})$ is not necessarily $AR$, but we stay in the class of $ARMA$ processes.

# Blocking

Consider the spectral density of $(\tilde{y}_t | \, t \in 2\mathbb{Z})$. This spectrum can be factorized as

$$f_{\tilde{y}}(z^2) = k(z^2)k(z^2)^* \tag{8}$$

where $k(z^2)$ is a stable and miniphase spectral factor (with a minimum number of columns).

Let $\left(\bar{A}_b, \bar{B}_b, \bar{C}_b, \bar{D}_b\right)$ denote a minimal state space system realizing such a spectral factor, i.e.

$$k(z^2) = \bar{C}_b \left( I \left(z^2\right)^{-1} - \bar{A}_b \right)^{-1} \bar{B}_b + \bar{D}_b \tag{9}$$

**McMillan degree of a transfer function**

State dimension of a minimal state space system.

**Theorem:**

For $((A_1, \ldots, A_p), \Sigma_v) \in \Theta$, if $A_p$ is non-singular, $\Gamma_p > 0$, and if for eigenvalues of $\mathscr{A}$, $\lambda_i \neq \lambda_j$ implies $\lambda_i^2 \neq \lambda_j^2$, the McMillan degree of a causal and miniphase spectral factor $k\left(z^2\right)$ of $f_{\tilde{y}}\left(z^2\right)$ is equal to $np$.

This result is quite remarkable, as it tells us that generically, the McMillan degree does not drop if we omit the slow, odd variables from $Y_t = \begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix}$.

The proof is quite tedious and shows that the Hankel matrix of covariance matrices

$$\mathbb{E} \begin{pmatrix} \tilde{y}_{t+2} \\ \vdots \\ \tilde{y}_{t+2np} \end{pmatrix} \begin{pmatrix} \tilde{y}_t^T & \cdots & \tilde{y}_{t-2(np-1)}^T \end{pmatrix}$$

has rank $np$ under the assumptions of the theorem.

# Blocking

On the other hand we obtain from (2), (3)

$$x_{t+1} = \underbrace{\mathscr{A}^2}_{A_b} x_{t-1} + \underbrace{(\mathscr{B}, \mathscr{A}\mathscr{B})}_{B_b} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} \tag{10}$$

$$\tilde{y}_t = \underbrace{\begin{pmatrix} I_n & 0 & 0 & \cdots & 0 \\ 0 & (I_{n_f}, 0) & 0 & \cdots & 0 \end{pmatrix} \mathscr{A}^2}_{C_b} x_{t-1} + \underbrace{\begin{pmatrix} b & A_1 b \\ 0 & (I_{n_f}, 0)\, b \end{pmatrix}}_{D_b} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}$$

As is easy to see, the driving noise of this system are not innovations, the system has a fat transfer function and will not be miniphase.

# Blocking

Nevertheless, one can show:

**Theorem:**

Under the assumptions of the theorem above $(A_b, C_b)$ and $(\bar{A}_b, \bar{C}_b)$ are related by

$$\bar{A}_b = T^{-1} A_b T$$
$$\bar{C}_b = C_b T$$

for a suitably chosen non-singular $np \times np$ matrix $T$.

Note that $(\bar{A}_b, \bar{C}_b)$ can be obtained from the spectral density $f_{\tilde{y}}(z^2)$ of the observed process $(\tilde{y}_t | \, t \in 2\mathbb{Z})$ uniquely up to basis change.

# Blocking

Using the structure of $C_b = \begin{pmatrix} I_n & 0 & 0 & \cdots & 0 \\ 0 & (I_{n_f}, 0) & 0 & \cdots & 0 \end{pmatrix} \mathscr{A}^2$ and

$\mathscr{A} = \begin{pmatrix} A_1 & \cdots & A_{p-1} & A_p \\ I_n & & & \\ & \ddots & & \\ & & I_n & 0 \end{pmatrix}$, we obtain $T$ and thus the system parameters in a unique way:

**Theorem:**

Under the assumptions of previous theorem and the additional assumptions that the matrix

$$\begin{pmatrix} \begin{pmatrix} I_{n_f} & 0 & \cdots & 0 \end{pmatrix} \\ \begin{pmatrix} I_{n_f} & 0 & \cdots & 0 \end{pmatrix} \mathscr{A} \\ \vdots \\ \begin{pmatrix} I_{n_f} & 0 & \cdots & 0 \end{pmatrix} \mathscr{A}^{np} \end{pmatrix}$$

has rank $np$ and that all eigenvalues of $\mathscr{A}$ are distinct, the system parameters $(A_1, \ldots, A_p)$ are uniquely determined from those population second moments which are directly observed.

# Outline

# g-Identifiability for the Case of Flow Data

(Anderson et al. 2016)

**$(w_t)_{t\in\mathbb{Z}}$ is a linear transformation of $(y^s_t)_{t\in\mathbb{Z}}$**

- Time Domain: linear transformation
  $w_t = (1+z+z^2+\cdots+z^{N-1})y^s_t,\ t\in\mathbb{Z}$
- Frequency Domain:
  $f_{ww}(\lambda) =$
  $\left(1+e^{i\lambda}+\cdots+\left(e^{i\lambda}\right)^{N-1}\right) f_{y^s y^s}(\lambda) \left(1+e^{-i\lambda}+\cdots+\left(e^{-i\lambda}\right)^{N-1}\right)$
  - Inverse transfer function
    $\left(1+e^{i\lambda}+\left(e^{i\lambda}\right)^2+\cdots+\left(e^{i\lambda}\right)^{N-1}\right)^{-1} I_{n_s} \in \mathscr{L}_2(f_{ww}d\lambda)$
  - Isomorphism between frequency and time domain

# Recovering $f_{y^s y^f}$

- Cross spectral densities between $(y_t^s)_{t\in\mathbb{Z}}$ and $\left(y_t^f\right)_{t\in\mathbb{Z}}$

$$f_{y^s y^f}(\lambda) = \left(1 + e^{i\lambda} + \left(e^{i\lambda}\right)^2 + \cdots + \left(e^{i\lambda}\right)^{N-1}\right)^{-1} f_{wy^f}(\lambda)$$

- One to one relationship of the spectral density and the covariance data: covariances $\gamma^{sf}(h)$ can be recovered from $\gamma_{wy^f}(h)$ which are observed.

## Remarks

- Analysis presented can easily be extended to general linear aggregation schemes:

$$w_t = k_0 y_t^s + k_1 y_{t-1}^s + \ldots + k_{N-1} y_{t-N+1}^s, \quad k_0 \,\text{non-singular}$$

# Outline

# ARMA Systems

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

$$b(z) = \sum_{j=0}^{q} B_j z^j, \quad B_0 = I$$

Assumptions:

- $\det(a(z)) \neq 0$, $|z| \leq 1$, $\mathrm{rk}(A_p) = n$
- $\det(b(z)) \neq 0$, $|z| \leq 1$, $\mathrm{rk}(B_q) = n$
- $(a(z), b(z))$ is left coprime

Parameter space:

$$\Theta_I = \{(A_1, \dots, A_p, B_1, \dots, B_q) \,|\, \text{Assumptions from above hold}\} \times \left\{\Sigma \,|\, \Sigma > 0, \Sigma^T = \Sigma\right\}$$

# ARMA Systems - Main results

Main results:

- for $p \geq q$: g-identifiability, Anderson et al. (2016), JoE
- for $p < q$: non-identifiability; equivalence classes are (topological) manifolds of dimension larger then or equal to one, Deistler et al. (2016), EcoSta
- Simplest Case: MA case; Here the not directly observed covariances of the outputs can be "freely varied" and this corresponds to equivalance classes in the parameter space

# Outline

# Estimation: VAR Case

Estimation VAR Case:
Koelbl et al. (2016), Koelbl (2015), Koelbl and Deistler (2016)

- Quick and Dirty: Extended YW, IVL
  initial estimators
- MLE and EM algorithm

Estimation GDFM's
Braumann, Fresoli, Deistler (2016)

Evaluation of these estimators by bootstrap methods is ongoing (starting) work

# Extended Yule-Walker Estimator

- We replace the population second moments by their sample counterparts:

$$\begin{aligned}\hat{\gamma}^{ff}(h) &= \frac{1}{T}\sum_{t=1}^{T-h} y_{t+h}^{f}\left(y_t^{f}\right)^T \\ \hat{\gamma}^{sf}(h) &= \frac{1}{T/2}\sum_{t=t_1}^{t_2} y_{2t}^{s}\left(y_{2t-h}^{f}\right)^T, \quad T \text{ sample size}\end{aligned}$$

- If $Z$ has full row rank, its estimator, $\hat{Z}$ say, will be of full row rank, from a certain $T_0$ onwards, too. We may now define the XYW estimators as

$$\left(\hat{A}_1, \ldots, \hat{A}_p\right) = \hat{Z}_1 \hat{Z}^{\dagger}$$

where $\hat{Z}^{\dagger} = \hat{Z}^T\left(\hat{Z}\hat{Z}^T\right)^{-1}$ is the Moore-Penrose pseudo inverse of $\hat{Z}$ (see Chen & Zadrozny (1998)).
- Note that not all second moments which can be directly observed are used in this estimator.

# Maximum Likelihood Estimation

$$\tilde{y}_t = \begin{pmatrix} y_t \\ y^f_{t-1} \end{pmatrix}, \; t \in 2\mathbb{Z}$$

$$\tilde{Y}_{obs} = \begin{pmatrix} \tilde{y}_2^T & \tilde{y}_4^T & \cdots & \tilde{y}_T^T \end{pmatrix}^T$$

$$\theta = \left(\mathrm{vec}(A)^T, \mathrm{vech}(\Sigma)^T\right)^T, \quad A = (A_1, \ldots, A_p)$$

$$\begin{aligned} \tilde{L}_T(\theta) &= \log(\det(\tilde{\Gamma}_T(\theta))) + \tilde{Y}_{obs}^T (\tilde{\Gamma}_T(\theta))^{-1} \tilde{Y}_{obs} \\ [\tilde{\Gamma}_T(\theta)]_{i,j} &= \begin{pmatrix} \gamma^{ff}(i-j) & \gamma^{fs}(i-j) & \gamma^{ff}(i-1-j) \\ \gamma^{sf}(i-j) & \gamma^{ss}(i-j) & \gamma^{fs}(i-1-j) \\ \gamma^{ff}(i-j+1) & \gamma^{sf}(i-j+1) & \gamma^{ff}(i-j) \end{pmatrix} \; i,j \in 2\mathbb{N} \end{aligned}$$

- Problems:
  - size of the matrix $\tilde{\Gamma}_T(\theta)$
  - no explicit formula for the score function

# EM Algorithm

The idea of the algorithm is to find $\hat{\theta} = \arg\min_{\theta \in \Theta} \tilde{L}_T(\theta)$ iteratively as follows. Consider the negative log-likelihood of the complete (i.e. high-frequency) data $Y_T = \left(y_1^T, y_2^T, \dots, y_T^T\right)^T$:

$$L_T(\theta) = \log\det(V_1) + x_1^T V_1^{-1} x_1 + T \log\det(\Sigma) + \sum_{t=1}^{T} (y_t - Ax_t)^T \Sigma^{-1} (y_t - Ax_t)$$

where we assume for the initial values $x_1 = (y_{-p+1}^T, \dots, y_0^T)^T \sim \mathcal{N}(0, V_1)$.

The $k+1$-th iteration consists of two steps:

1. E-step: For given parameters $\theta^{(k)}$, compute the conditional expectation $\mathbb{E}_{\theta^{(k)}}\left(L_T(\theta) \,|\, \tilde{Y}_{obs}\right)$, where $\tilde{Y}_{obs} = \begin{pmatrix} \tilde{y}_2^T & \tilde{y}_4^T & \cdots & \tilde{y}_T^T \end{pmatrix}^T$
2. M-step: Obtain $\theta^{(k+1)}$ by optimizing the conditional expectation $\mathbb{E}_{\theta^{(k)}}\left(L_T(\theta) \,|\, \tilde{Y}_{obs}\right)$ with respect to $\theta$.

# EM Algorithm - E-step

$$\mathbb{E}_{\theta^{(k)}}\left(L_T(\theta)\,|\,\tilde{Y}_{obs}\right) = \log\det(V_1) + \operatorname{trace}\left(V_1^{-1}\mathbb{E}_{\theta^{(k)}}\left(x_1 x_1^T\,|\,\tilde{Y}_{obs}\right)\right) +$$

$$T\log\det(\Sigma) + \operatorname{trace}(\Sigma^{-1}\underbrace{\sum_{t=1}^{T}\mathbb{E}_{\theta^{(k)}}\left((y_t - Ax_t)(y_t - Ax_t)^T\,|\,\tilde{Y}_{obs}\right)}_{=(*)})$$

Let $x_{t|T} = \mathbb{E}_{\theta^{(k)}}\left(x_t|\,\tilde{Y}_{obs}\right)$ and consider term (*) using $y_t = \underbrace{(I_n, 0, \ldots, 0)}_{\mathscr{C}} x_{t+1}$:

$$(*) = \mathscr{C}\left(\sum_{t=1}^{T} x_{t+1|T}x_{t+1|T}^T + P_{t+1|T}\right)\mathscr{C}^T + A\left(\sum_{t=1}^{T} x_{t|T}x_{t|T}^T + P_{t|T}\right)A^T$$

$$-\mathscr{C}\left(\sum_{t=1}^{T} x_{t+1|T}x_{t|T}^T + P_{t+1,t|T}\right)A^T - A\left(\sum_{t=1}^{T} x_{t|T}x_{t+1|T}^T + P_{t,t+1|T}\right)\mathscr{C}^T$$

where $P_{t,t-j|T} = \mathbb{E}\left(x_t - x_{t|T}\right)\left(x_{t-j} - x_{t-j|T}\right)^T$.

# EM Algorithm - E-step cont.

We compute the conditional expectations and its errors,

- $x_{t|T} = \mathbb{E}_{\theta^{(k)}}\left(x_t | \tilde{Y}_{obs}\right)$
- $P_{t,t-j|T} = \mathbb{E}\left(x_t - x_{t|T}\right)\left(x_{t-j} - x_{t-j|T}\right)^T$

through the following time variable system

$$x_{t+1} = \mathscr{A} x_t + \mathscr{B}\varepsilon_t$$
$$y_t^\times = C_t x_{t+1}$$

where $x_t = (y_{t-1}^T, \ldots, y_{t-p}^T)$ and $C_{2t} = (I_n 0 \cdots 0)$, $C_{2t-1} = (I_{n_f} 0 \cdots 0)$. The vector $y_t^\times$ contains only observable components.

Since $x_t = (y_{t-1}^T, \ldots, y_{t-p}^T)$, calculating $x_{t|T}$ means interpolating the missing $y_t^s$.

# EM Algorithm - M-step

Obtain $A^{(k+1)}, \Sigma^{(k+1)}$ by solving $\frac{\partial \mathbb{E}_{\theta^{(k)}}\left(L_T(\theta) \mid \tilde{Y}_{obs}\right)}{\partial X} = 0$, $X = A, \Sigma$:

$$
\begin{aligned}
A^{(k+1)} &= \mathscr{C} S_{10} S_{00}^{-1} \\
\Sigma^{(k+1)} &= T^{-1}\left(\mathscr{C} S_{11} \mathscr{C}^T - \mathscr{C} S_{10} S_{00}^{-1} S_{10}^T \mathscr{C}^T\right) \\
S_{ii} &= \sum_{t=1}^{T}\left(x_{t+j|T} x_{t+j|T}^T + P_{t+j|T}\right), j = 0, 1 \\
S_{10} &= \sum_{t=1}^{T}\left(x_{t+1|T} x_{t|T}^T + P_{t+1,t|T}\right)
\end{aligned}
$$

The matrix $\mathscr{C} = (I_n, 0, \ldots, 0)$ selects the first $n \times np$ block row.

Note that this is a version of the EM algorithm for state space models, first proposed in Shumway & Stoffer (1982).

# Asymptotic Distribution of the XYW Estimators

- The derivation of the asymptotic distribution of the XYW estimator is tedious (no analytic formula for asym. variance)
- First, we consider Bartlett's formula for the mixed-frequency estimators of the autocovariance function:

$$\sqrt{T}\left(\mathrm{vec}\begin{pmatrix}\hat{\gamma}^{ff}(h)\\ \hat{\gamma}^{sf}(h)\end{pmatrix} - \mathrm{vec}\begin{pmatrix}\gamma^{ff}(h)\\ \gamma^{sf}(h)\end{pmatrix}\right)_{h=1-p}^{np} \xrightarrow{d} \mathcal{N}\left(0, \Sigma_\gamma\right)$$

- In the next step we consider the linearization of the mapping attaching the parameters to the second moments at the "true" point:

$$\mathrm{vec}\left(\hat{A}_{\mathrm{XYW}} - A\right) = \hat{Z}^\dagger \mathrm{vec}\left(\hat{Z}_1 - A\hat{Z}\right) = \hat{Z}^\dagger L_A \left(\mathrm{vec}\begin{pmatrix}\hat{\gamma}^{ff}(h)\\ \hat{\gamma}^{sf}(h)\end{pmatrix} - \mathrm{vec}\begin{pmatrix}\gamma^{ff}(h)\\ \gamma^{sf}(h)\end{pmatrix}\right)_{h=1-}^{np}$$

- Finally, we can conclude that

$$\sqrt{T}\left(\mathrm{vec}\left(\hat{A}_{\mathrm{XYW}}\right) - \mathrm{vec}(A)\right) \xrightarrow{d} \mathcal{N}\left(0, Z^\dagger L_A \Sigma_\gamma L_A^T \left(Z^\dagger\right)^T\right)$$

# Compare Properties

Asymptotic and finite sample properties of

- HF Maximum Likelihood, which is asymptotically equivalent to the (HF) Yule-Walker estimator
- HF XYW: Loss of uniqueness and thus of consistency if $Z$ is of rank smaller than $np$. In general, efficiency loss compared to HF Yule-Walker estimator, in particular if $Z$ is "almost" rank deficient
- MF Maximum Likelihood: asymptotically efficient for MF data. EM algorithm
- MF XYW: two sources of efficiency loss relative to HF maximum likelihood: caused by MF data and relative to the MF maximum likelihood estimator

# Loss of Identifiability

**Theorem**

Assume that $p = 1$, $n_f = n_s = 1$, $\Sigma = I_2$. The system parameters $\begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \end{pmatrix}$ are not identifiable if and only if they satisfy the equations

$$a_{fs} = 0,\ a_{sf} = 0,\ a_{ss} \neq 0.$$

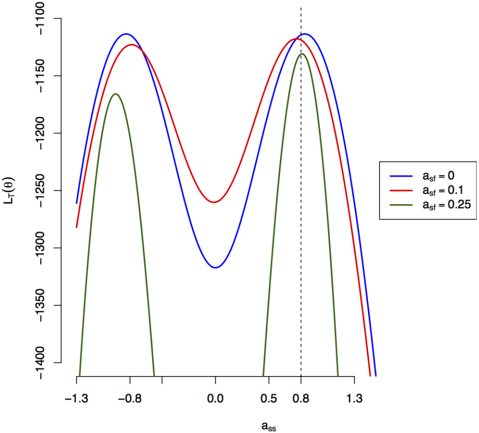
**Example: AR(1)**

$$y_t = \begin{pmatrix} 0.9 & 0 \\ a_{sf} & 0.8 \end{pmatrix} y_{t-1} + \varepsilon_t$$

$$(\varepsilon_t) \sim WN(0, I_2),\ a_{sf} \in \{0, 0.1, 0.25\}$$

# Example: AR(1)

Evaluate the log-likelihood $-L_T(\theta)$ fixed at the true parameters $a_{ff}, a_{fs}, a_{sf}, \Sigma$.



# Finite Sample Examples

Model “The Good”:

$$y_t = \begin{pmatrix} 0.9556 & 0.8611 \\ -0.6914 & 0.2174 \end{pmatrix} y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, I_2)$$

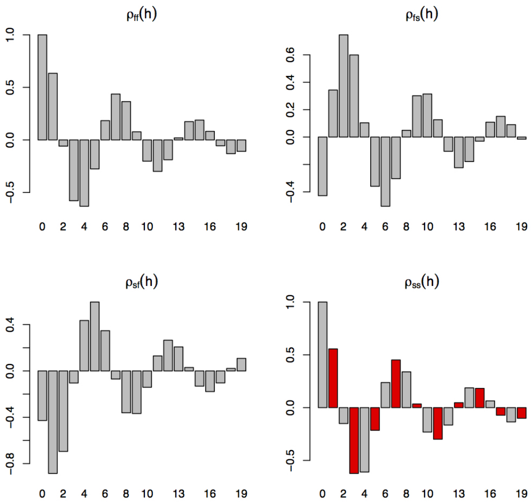
$$z_{0,1} = 0.7303 \pm 0.8437i$$

Model “The Bad”:

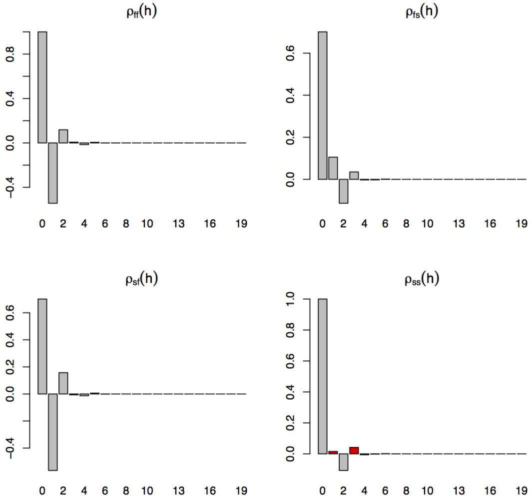
$$y_t = \begin{pmatrix} -1.2141 & 1.1514 \\ -0.9419 & 0.8101 \end{pmatrix} y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, I_2)$$

$$z_{0,1} = -2 \pm 2.4294i$$

# Correlations - Model "The Good"



# Correlations - Model "The Bad"



# Finite Sample - MSE $\left(\hat{\theta}\right)$

$$\theta = \left(\text{vec}\left(A_1\right)^T, \text{vech}\left(\Sigma\right)^T\right)^T,$$
$$\text{MSE}\left(\hat{\theta}\right) = \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{7}\left(\theta_i - \hat{\theta}_i^j\right)^2, \; m = 10^3$$

| | | Model "The Good" | | Model "The Bad" | |
|---|---|---|---|---|---|
| | | $T = 500$ | | $T = 500$ | |
| | Estimator | absolute | relative | absolute | relative |
| HF | YW | 0.0067 | 1 | 0.0092 | 1 |
| HF | XYW | 0.0111 | 1.66 | 0.2255 | 24.51 |
| MF | MLE-EM | 0.0101 | 1.51 | 0.0451 | 4.90 |
| MF | XYW | 0.0783 | 11.69 | 1.3226 | 143.76 |

# Finite Sample - One-step-ahead Forecasting Errors

- Comparing LF-MF-HF

| | | Model 1 | | Model 2 | |
|---|---|---|---|---|---|
| | Estimators | Absolute | Relative | Absolute | Relative |
| LF | YW | 3.6071 | 1 | 2.8594 | 1 |
| MF | MLE-EM | 2.3709 | 0.66 | 2.8574 | 0.99 |
| | XYW | 2.3883 | 0.66 | 34.1492 | 11.94 |
| HF | YW | 1.9948 | 0.55 | 1.9981 | 0.70 |

# Outline

# Modeling High Dimensional Time Series

"Structured big data"

Curse of dimensionality vs. gaining additional information from additional time series

Model Classes: Sparsity (either a-priori knowledge or model selection)

- Dynamic Factor Models
- Dynamic PCA
- Structural Models
- Graphical Time Series Models

# Areas of Application

Areas of Application:

- System Identification: Oversensoring
- System Identification: Internet of Things
- Macroeconometric Modeling: Disaggregated Data
- High Density EEG's

# GDFM's

$$y_t = \hat{y}_t + u_t, \qquad t \in \mathbb{Z}, \quad \text{High-frequency system}$$

$y_t$ ... $n$-dimensional observations (e.g. n=150)
$\hat{y}_t$ ... latent variables, strongly dependent
$u_t$ ... noise, weakly dependent (weakly idiosyncratic)
$(\hat{y}_t) \perp (u_t)$

$$\hat{y}_t = Lz_t, \qquad L \in \mathbb{R}^{n \times s}$$

$z_t$ ... $s$-dimensional static factors, $s < n$

$$a(z)z_t = b\varepsilon_t, \qquad b \in \mathbb{R}^{s \times q}$$

$\varepsilon_t$ ... $q$-dimensional dynamic factors, $q \leq s$
For $q < s$ (singular case), the AR case is generic, Anderson et al. (2016b)

# GDFM's

Separate denoising (i.e. estimating latent variables or static factors from observations) from MF observations to obtain a MF AR(MA) model for static factors

Aim: Obtain a minimal static factor with a maximum number of high frequency variables.
Stock variables:

- Estimate factor $z_t^f$ from $y_t^f$
- Estimate factor $z_t^s$ from $y_t^s$
- For the estimate $\hat{z}_t$ eliminate linear dependent components of second step.

The HF AR model estimated this way then may be used for "structured" interpolation.
Felsenstein (2014), Braumann, Fresoli, Deistler (2016)

Denoising in other cases is still an open problem

# GDFM's

Integrated denoising via a state space formulation of the GDFM (Mariano & Murasawa (2003); Bańbura & Modugno (2014)). Assume that

$$
\begin{aligned}
y_t &= Lz_t + u_t \\
z_t &= A_1 z_{t-1} + \cdots + A_4 z_{t-4} + \varepsilon_t \\
u_{it} &= \alpha_i u_{it-1} + \eta_{it}
\end{aligned}
$$

where $|\alpha_i| < 1$ and $\det\left(I_s - A_1 z - \cdots - A_4 z^4\right) \neq 0$, $\forall |z| \leq 1$ and $(\varepsilon_t) \sim \mathrm{WN}(0, I_s)$. In addition $(u_t) \sim \mathrm{WN}(0, R)$.

# GDFM's

State space formulation of the GDFM:

$$y_t = (L, 0, \ldots, 0, I_n, 0) \begin{pmatrix} z_t \\ z_{t-1} \\ z_{t-2} \\ z_{t-3} \\ u_t \end{pmatrix}$$

$$\begin{pmatrix} z_t \\ z_{t-1} \\ z_{t-2} \\ z_{t-3} \\ u_t \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & A_3 & A_4 & 0 \\ I_n & & & & \\ & I_n & & & \\ & & I_n & & \\ & & & & \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_N \end{pmatrix} \end{pmatrix} \begin{pmatrix} z_{t-1} \\ z_{t-2} \\ z_{t-3} \\ z_{t-4} \\ u_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \\ \eta_t \end{pmatrix}$$

The parameters can be estimated e.g. by the EM algorithm.

# Conclusions

- For the VAR case as well as for the VARMA case where the MA order is less than or equal to the AR order we have generic identifiability from MF data.
- Estimation in the VAR case: Fast methods: XYW and IVL, consistent but not asymptotically efficient. Initial estimators for MLE and EM
- VMA case as well as the VARMA case where the MA order exceeds the AR order: Genuine nonidentifiablity.
- Future work: Evaluation of estimators by bootstrap methods

## Questions?

Thank you!

# Bibliography I

Anderson, B. D. O. & Deistler, M. (2008). Properties of Zero-free Transfer Function Matrices. *SICE Journal of Control, Measurement, and System Integration*, 1(1), 1–9.

Anderson, B. D. O., Deistler, M., Felsenstein, E., Funovits, B., Koelbl, L., & Zamani, M. (2016a). Multivariate AR Systems and Mixed Frequency Data: g-Identifiability and Estimation. *Econometric Theory*, 32(4), 793–826.

Anderson, B. D. O., Deistler, M., Felsenstein, E., Funovits, B., Zadrozny, P. A., Eichler, M., Chen, W., & Zamani, M. (2012). Identifiability of Regular and Singular Multivariate Autoregressive Models. *Proceedings of the 51th IEEE Conference on Decision and Control (CDC)*, 184-189.

# Bibliography II

Anderson, B. D. O., Deistler, M., Felsenstein, E., & Koelbl, L. (2016b). The Structure of Multivariate AR and ARMA Systems: Regular and Singular Systems; the Single and the Mixed Frequency Case. *Journal of Econometrics*, 192(2), 366–373.

Bańbura, M. & Modugno, M. (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, 29(1), 133–160.

Braumann, A., Fresoli, D., & Deistler, M. (2016). Generalized linear dynamic factor models - estimation and forecasting in the single- and mixed-frequency case. Working Paper.

Chen, B. & Zadrozny, P. A. (1998). An Extended Yule-Walker Method for Estimating a Vector Autoregressive Model with Mixed-Frequency Data. *Advances in Econometrics*, 13, 47–73.

# Bibliography III

Deistler, M., Anderson, B. D. O., Filler, A., Zinner, C., & Chen, W. (2010). Generalized Linear Dynamic Factor Models - An Approach via Singular Autoregressions. *European Journal of Control*, 16(3), 211–224.

Deistler, M., Filler, A., & Funovits, B. (2011). AR Systems and AR Processes: The Singular Case. *Communications in Information and Systems*, 11(3), 225–236.

Deistler, M., Koelbl, L., & Anderson, B. (2017). Non-Identifiability of VMA and VARMA Systems in the Mixed Frequency Case. *Econometrics and Statistics*.

Felsenstein, E. (2014). *Regular and Singular AR and ARMA Models: The Single and The Mixed Frequency Case*. PhD thesis, Vienna University of Technology.

# Bibliography IV

Filler, A. (2010). *Generalized Dynamic Factor Models - Structure Theory and Estimation for Single Frequency and Mixed Frequency Data*. PhD thesis, Vienna University of Technology.

Koelbl, L. (2015). *VAR Systems: g-Identifiability and Asymptotic Properties of Parameter Estimates for the Mixed-Frequency Case*. PhD thesis, Vienna University of Technology.

Koelbl, L., Braumann, A., Felsenstein, E., & Deistler, M. (2016). Estimation of VAR Systems from Mixed-Frequency Data: The Stock and the Flow Case. *Advances in Econometrics*, 35, 43–73.

Koelbl, L. & Deistler, M. (2016). A new approach for estimating var systems in the mixed-frequency case. Submitted to Statistical Papers, Springer.

# Bibliography V

Mariano, R. S. & Murasawa, Y. (2003). A New Coincident Index of Business Cycles based on Monthly and Quarterly Series. *Journal of Applied Econometrics*, 18(4), 427–443.

Shumway, R. & Stoffer, D. (1982). An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm. *Journal of Time Series Analysis*, 3(4), 253–264.