

Identifiability of regular and singular multivariate autoregressive models from mixed frequency data*

Brian D.O. Anderson¹, Manfred Deistler², Elisabeth Felsenstein³, Bernd Funovits⁴,
Peter Zdrozny⁵, Michael Eichler⁶, Weitian Chen⁷, Mohsen Zamani⁸

Abstract—This paper is concerned with identifiability of an underlying high frequency multivariate AR system from mixed frequency observations. Such problems arise for instance in economics when some variables are observed monthly whereas others are observed quarterly. If we have identifiability, the system and noise parameters and thus all second moments of the output process can be estimated consistently from mixed frequency data. Then linear least squares methods for forecasting and interpolating nonobserved output variables can be applied. Two ways for guaranteeing generic identifiability are discussed.

I. INTRODUCTION

In a number of applications, for a multivariate time series the component time series may be available only at different sampling frequencies. For such a situation, the term mixed frequency data is used. For instance, in economic applications some time series may be available monthly, e.g., unemployment data, whereas other time series are available only quarterly, e.g., GDP data. Another area of application is the analysis of environmental time series. We are interested in the underlying high frequency system, i.e. the system generating the data at the highest sampling frequency, or, to be more precise, in its system and noise parameters.

We restrict ourselves to the case where this high frequency

*Support by the FWF (Austrian Science Fund under contract P20833/N18) and the ARC (Australian Research Council under Discovery Project Grant DP1092571) is gratefully acknowledged.

¹Brian D.O. Anderson is with the Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, and Australia and Canberra Research Laboratory, National ICT Australia Ltd., e-mail: brian.anderson@anu.edu.au

²Manfred Deistler is with the Department of Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria, and Institute for Advanced Studies, Vienna, Austria, e-mail: deistler@tuwien.ac.at

³Elisabeth Felsenstein is with the Department of Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria, e-mail: elisabeth.felsenstein@tuwien.ac.at

⁴Bernd Funovits is with the Vienna Graduate School of Economics, University of Vienna, Austria and Department of Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria, e-mail: bernd.funovits@univie.ac.at

⁵Peter Zdrozny is with the Bureau of Labor Statistics, Washington D.C., e-mail: zdrozny.peter@bls.gov

⁶Michael Eichler is with the Department of Quantitative Economics, Faculty of Economics and Business Administration, Maastricht University, The Netherlands, e-mail: m.eichler@maastrichtuniversity.nl

⁷Weitian Chen is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada, e-mail: cwt6122@uwindsor.ca

⁸Mohsen Zamani is with the Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, e-mail: mohsen.zamani@anu.edu.au

system is a vector autoregression of order p , i.e.

$$y_t = \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} = \underbrace{\begin{pmatrix} a_{ff}(1) & a_{fs}(1) \\ a_{sf}(1) & a_{ss}(1) \end{pmatrix}}_{=A_1} \begin{pmatrix} y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} + \dots + \underbrace{\begin{pmatrix} a_{ff}(p) & a_{fs}(p) \\ a_{sf}(p) & a_{ss}(p) \end{pmatrix}}_{=A_p} \begin{pmatrix} y_{t-p}^f \\ y_{t-p}^s \end{pmatrix} + \underbrace{\begin{pmatrix} \nu_t^f \\ \nu_t^s \end{pmatrix}}_{=\nu_t}, \quad t \in \mathbb{Z}, \quad (1)$$

where $a_{ff}(i) \in \mathbb{R}^{n_f \times n_f}$, $a_{ss}(i) \in \mathbb{R}^{n_s \times n_s}$, $a_{fs}(i) \in \mathbb{R}^{n_f \times n_s}$, and where n_f is the number of the components observed at highest frequency, n_s is the number of components observed only for $t \in N\mathbb{Z}$, i.e., every N th time point, and $n = n_f + n_s$. Throughout we assume that the high frequency system (1) is stable, and that we restrict ourselves to the steady state and thus stationary solution. We consider the case where the innovation variance $\Sigma = \mathbb{E} \left(\begin{pmatrix} \nu_t^f \\ \nu_t^s \end{pmatrix} \begin{pmatrix} (\nu_t^f)^T & (\nu_t^s)^T \end{pmatrix} \right) = \begin{pmatrix} \Sigma_{ff} & \Sigma_{fs} \\ \Sigma_{sf} & \Sigma_{ss} \end{pmatrix}$, where $\Sigma_{ff} \in \mathbb{R}^{n_f \times n_f}$, $\Sigma_{ss} \in \mathbb{R}^{n_s \times n_s}$, and $\Sigma_{fs} \in \mathbb{R}^{n_f \times n_s}$, is regular as well as the case where this variance is singular. Then, the corresponding autoregressive systems are called regular and singular respectively. Singular autoregressive systems are important for generalized linear dynamic factor models (see [4], [3]). In the singular case, when Σ is of rank $q < n$, we can write Σ as $\Sigma = bb^T$ where b is an $(n \times q)$ matrix. Accordingly, $\nu_t = b\varepsilon_t$, where $\mathbb{E}(\varepsilon_t \varepsilon_t^T) = I_q$. For given Σ , b is unique up to postmultiplication by an orthogonal matrix and thus b can be made unique. Whereas in the regular case Σ is parameterized by its on and above diagonal elements, in the singular case b can be used for the parameterization of Σ , because its free parameters are more easily seen in this way.

System (1) can be written in block companion form as

$$\underbrace{\begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}}_{=x_{t+1}} = \underbrace{\begin{pmatrix} A_1 & \cdots & A_{p-1} & A_p \\ I_n & & & \\ & \ddots & & \\ & & I_n & 0 \end{pmatrix}}_{\mathcal{A}} \underbrace{\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}}_{=x_t} + \underbrace{\begin{pmatrix} b \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\mathcal{B}} \varepsilon_t. \quad (2)$$

The Lyapunov equation, where $\Gamma = \mathbb{E}(x_t x_t^T)$, for the system (2) is

$$\Gamma - \mathcal{A}\Gamma\mathcal{A}^T = \mathcal{B}\mathcal{B}^T. \quad (3)$$

The central problem considered in this paper is identi-

fiability, i.e. whether $\begin{pmatrix} a_{ff}(i) & a_{fs}(i) \\ a_{sf}(i) & a_{ss}(i) \end{pmatrix}$, $i \in \{1, \dots, p\}$ and Σ are uniquely identified from those population second moments which can be observed in principle, i.e. $\gamma^{ff}(h) = \mathbb{E}\left(y_{t+h}^f (y_t^f)^T\right)$, $h \in \mathbb{Z}$; $\gamma^{fs}(h) = \mathbb{E}\left(y_{t+h}^f (y_t^s)^T\right)$, $h \in \mathbb{Z}$; $\gamma^{ss}(h) = \mathbb{E}\left(y_{t+h}^s (y_t^s)^T\right)$, $h \in N\mathbb{Z}$. Note that if identifiability holds (and there is available an algorithm to compute the parameters from the observed second moments) we can reconstruct the missing moments $\gamma^{ss}(h) = \mathbb{E}\left(y_{t+h}^s (y_t^s)^T\right)$, $h \in N\mathbb{Z} - j$; $j = 1, \dots, N - 1$ and then linear least squares methods for forecasting and interpolating nonobserved output variables can be applied. In other words, this identifiability is an important step in getting consistent estimators of the system and noise parameters and thus of the missing second moments $\gamma^{ss}(h)$, $h \in N\mathbb{Z} - j$; $j = 1, \dots, N - 1$ based on the mixed frequency data available.

The paper consists of two parts. In the first part, we analyze (a modification of) the method proposed by Chen and Zadrozny ([2]) and we prove that it generically gives the parameters of the high frequency system uniquely. In the second part, i.e. section III, we consider a restrictive setting, i.e. regular AR(1) systems with $N = 2$, $n_f = n_s = 1$, and derive a sufficient and necessary condition for identifiability in this case. In particular, this shows that the conditions for identifiability formulated by [2] are only sufficient but not necessary. Moreover, these second moments can be used for forecasting and interpolation of missing observations.

II. EXTENDED YULE WALKER EQUATIONS

In this section we discuss the extended Yule Walker equations as proposed by [2]. A central result in this paper is that for generic AR(p) systems these equations have a unique solution. Thus, we generically have identifiability which in turn implies that the system and noise parameters can be estimated consistently under an assumption guaranteeing consistent estimation of those population second moments which can be observed in principle. The result given holds for regular as well as for singular AR systems.

A. Derivation of the extended Yule Walker equations for mixed frequency data

By postmultiplying equation (1) by y_{t-j}^T , $j > 0$ and forming expectations, we obtain the extended Yule Walker equations

$$\begin{pmatrix} \gamma^{ff}(1) & \gamma^{fs}(1) & \dots & \gamma^{ff}(p) & \gamma^{fs}(p) & \dots \\ \gamma^{sf}(1) & \gamma^{ss}(1) & \dots & \gamma^{sf}(p) & \gamma^{ss}(p) & \dots \end{pmatrix} = \begin{pmatrix} a_{ff}(1) & a_{fs}(1) & \dots & a_{ff}(p) & a_{fs}(p) \\ a_{sf}(1) & a_{ss}(1) & \dots & a_{sf}(p) & a_{ss}(p) \end{pmatrix} \times \begin{pmatrix} \gamma^{ff}(0) & \gamma^{fs}(0) & \dots & \gamma^{ff}(p-1) & \gamma^{fs}(p-1) & \dots \\ \gamma^{sf}(0) & \gamma^{ss}(0) & \dots & \gamma^{sf}(p-1) & \gamma^{ss}(p-1) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma^{ff}(1-p) & \gamma^{fs}(1-p) & \dots & \gamma^{ff}(0) & \gamma^{fs}(0) & \dots \\ \gamma^{sf}(1-p) & \gamma^{ss}(1-p) & \dots & \gamma^{sf}(0) & \gamma^{ss}(0) & \dots \end{pmatrix} \quad (4)$$

$$\text{Let } G = \mathbb{E}(x_t y_{t-1}^T) = \begin{pmatrix} \gamma^{ff}(0) & \gamma^{fs}(0) \\ \gamma^{sf}(0) & \gamma^{ss}(0) \\ \vdots & \vdots \\ \gamma^{ff}(1-p) & \gamma^{fs}(1-p) \\ \gamma^{sf}(1-p) & \gamma^{ss}(1-p) \end{pmatrix} = \Gamma \begin{pmatrix} I_n \\ \vdots \\ 0 \end{pmatrix}.$$

The block columns of the second matrix on the right hand side of (4) are of the form $\mathbb{E}(x_t y_{t-j-1}^T) = \mathbb{E}((\mathcal{A}x_{t-1} + \mathcal{B}\varepsilon_{t-1}) y_{t-j-1}^T) = \mathbb{E}(\mathcal{A}x_{t-1} y_{t-j-1}^T) = \mathcal{A}E(x_{t+j-1} y_{t-1}^T) = \dots = \mathcal{A}^j G$, $j \geq 0$. Thus this matrix can be written as $(G, \mathcal{A}G, \mathcal{A}^2 G, \dots)$. From the Cayley Hamilton theorem we see that the second matrix on the right hand side of (4) has full row rank if and only if the matrix consisting of the first np blocks has full row rank. We now suggest to use only those equations in (4) where the columns on the left hand side or the columns of the second matrix on the right hand side contain only second moments which can be observed. In other words we consider the equation system

$$\mathbb{E}y_t (y_{t-1}^f, \dots, y_{t-np}^f)^T = (A_1, \dots, A_p) \underbrace{\mathbb{E} \begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} \begin{pmatrix} (y_{t-1}^f)^T, \dots, (y_{t-np}^f)^T \end{pmatrix}}_{=Z}. \quad (5)$$

Note that [2] uses a larger subsystem of equations of (4) and thus, in particular, our identifiability result implies theirs.

As is easily seen, Z can be written as $(K, \mathcal{A}K, \mathcal{A}^2 K, \dots, \mathcal{A}^{np-1} K)$, where $K = \Gamma \begin{pmatrix} I_{n_f} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and therefore has the structure of a controllability matrix.

Clearly, the AR parameters A_1, \dots, A_p of the system (1) are identifiable if Z has full row rank np , or equivalently the pair (\mathcal{A}, K) is controllable. As will be shown below, this rank condition is not necessary for identifiability.

B. Generic identifiability of system parameters

Consider the set of all AR systems for given order p and given rank q of the innovation covariance matrix satisfying the stability condition. As easily can be seen, the parameter space for this set is an open subset of $\mathbb{R}^{pn^2+nq-\frac{q(q-1)}{2}}$ if we take into account the uniqueness of the transfer function only up to orthogonal postmultiplication. A property on this parameter space is said to hold generically if it holds on a superset of an open and dense subset of the parameter space. For simplicity of notation, we do not take into account the restriction arising from making the orthogonal postmultiplication unique and thus consider a parameter space $\Theta \subseteq \mathbb{R}^{pn^2+nq}$. In this case, the notions of genericity are the same in both spaces, that means a generic set in Θ intersected with the zero restrictions corresponding to normalizing the orthogonal postmultiplication gives a generic set in the restricted parameter space.

The next two theorems, which are the central results of this paper, show that the matrix Z in equation (5) is generically of full row rank and thus we have generic identifiability of the A_i . In addition, Σ is shown to be generically identifiable.

Note that this holds both for regular and singular AR systems, for all sampling frequency ratios N , and all $n^f \geq 1$.

Theorem 1: The matrix Z in equation (5) has full row rank $n \cdot p$ on a generic subset of the parameter space Θ .

The proof of the theorem, which uses system theoretic tools (see [5], [1]), is quite intricate and we start by proving three lemmas and a corollary.

Lemma 1: Let \mathcal{A} denote the block companion matrix defined above, and let $\bar{A}(z)$ denote the polynomial matrix $z^p I - A_1 z^{p-1} - A_2 z^{p-2} - \dots - A_p$. Suppose that $\alpha^T = (\alpha_1^T, \alpha_2^T, \dots, \alpha_p^T)$, where $\alpha \in \mathbb{R}^{np}$, $\alpha_i \in \mathbb{R}^n$. If α^T is a left eigenvector of \mathcal{A} corresponding to eigenvalue λ , then $\alpha_1^T \neq 0$ and lies in the left kernel of $\bar{A}(\lambda)$, i.e. $\alpha_1^T \bar{A}(\lambda) = 0$. Conversely, if $\alpha_1 \neq 0$ is such that $\alpha_1^T \bar{A}(\lambda) = 0$, and $\alpha_2^T = \alpha_1^T (\lambda I - A_1)$, $\alpha_3^T = \alpha_1^T (\lambda^2 I - \lambda A_1 - A_2)$, \dots , holds for some λ , then $\alpha^T = (\alpha_1^T, \alpha_2^T, \dots, \alpha_p^T)$ is a left eigenvector of \mathcal{A} corresponding to eigenvalue λ .

Proof: Suppose that α^T is a left eigenvector of \mathcal{A} corresponding to eigenvalue λ . Then at once

$$\begin{aligned} \alpha_1^T A_1 + \alpha_2^T &= \lambda \alpha_1^T \\ \alpha_1^T A_2 + \alpha_3^T &= \lambda \alpha_2^T \\ &\vdots \\ \alpha_1^T A_{p-1} + \alpha_p^T &= \lambda \alpha_{p-1}^T \\ \alpha_1^T A_p &= \lambda \alpha_p^T \end{aligned} \quad (6)$$

Evidently, $\alpha_1 \neq 0$, else these equations yield that α itself would be zero, contradicting the fact that α is an eigenvector. It is trivial to eliminate $\alpha_2, \alpha_3, \dots$ to obtain

$$\alpha_1^T (\lambda^p I - \lambda^{p-1} A_1 - \lambda^{p-2} A_2 - \dots - A_p) = 0 \quad (7)$$

Conversely, suppose $\alpha_1^T \bar{A}(\lambda) = 0$ with $\alpha_1 \neq 0$, and that $\alpha_2, \alpha_3, \dots$ are defined as in the lemma hypothesis. Equations (6) easily follow and then the eigenvalue property $\alpha^T \mathcal{A} = \lambda \alpha^T$ is immediate. ■

Corollary 1: With \mathcal{A}, \mathcal{B} as above, the pair $(\mathcal{A}, \mathcal{B})$ is controllable for a generic subset of the parameter space.

Proof: We argue first that generically, for each value of λ for which $\bar{A}(\lambda)$ is singular, the left kernel is one-dimensional:

Let $\alpha_1 \neq 0$ be such that $\alpha_1^T \bar{A}(\lambda) = 0$. Then we can find a nonsingular matrix T such that the first row of $T \bar{A}(\lambda)$ is zero. Thus,

$$0 = \alpha_1^T T^{-1} T \bar{A}(\lambda)$$

and thus the first element of $(\alpha_1^T T^{-1})$ can be chosen to be nonzero. Now, consider the polynomial matrix formed by the rows 2 till n of the polynomial matrix $T \bar{A}(z)$. By the result on generic zerolessness of tall rational transfer functions, see [1], the components 2 till n of the vector $(\alpha_1^T T^{-1})$ must be zero which gives the desired result. For any nonzero vector α , say, generating such a left kernel, there holds $\alpha^T \mathcal{B} = \alpha_1^T b \neq 0$, by virtue of the genericity, and this will hold for an arbitrary but finite number of α . Now the inequality above is easily seen to hold in an open subset of the parameter space: This is a consequence of the continuity of the mapping

attaching a suitably normalized eigenvector α of \mathcal{A} to \mathcal{A} and the continuity of the inner product $\alpha^T \mathcal{B}$. This means that controllability is established in an open subset of the parameter space. The density of this set is shown as follows: If $\alpha_1^T b = 0$, then there is a sequence $b_i \rightarrow b$ such that $\alpha_1^T b_i \neq 0$. ■

Lemma 2: Let \mathcal{A} denote the block companion matrix defined above, and let $\bar{A}(z)$ denote the polynomial matrix defined above. Suppose that c is an np -vector, with $c = (c_1^T, c_2^T, \dots, c_p^T)^T$ and each c_i an n -vector. Then if c is a right eigenvector of \mathcal{A} corresponding to eigenvalue $\lambda \neq 0$, then $c_1 \neq 0$ and lies in the kernel of $\bar{A}(\lambda)$. Conversely, if $c_1 \neq 0$ is such that $\bar{A}(\lambda)c_1 = 0$ with $\lambda \neq 0$ and $c_i = \lambda^{-i+1} c_1$, then $c = (c_1^T, c_2^T, \dots, c_p^T)^T$ is a right eigenvector of \mathcal{A} with eigenvalue λ .

Proof: Assume that $\mathcal{A}c = \lambda c$. Then it is easily seen that

$$\begin{aligned} \sum_{i=1}^p A_i c_i &= \lambda c_1 \\ c_1 &= \lambda c_2 \\ c_2 &= \lambda c_3 \\ &\vdots \\ c_{p-1} &= \lambda c_p \end{aligned}$$

The result is easily proved using these equations. ■

Denote by e_j the n -vector with 1 in the j -th entry and all other entries zero. Denote by \mathcal{E}_j the np -vector $\mathcal{E}_j = (e_j^T, 0, 0, \dots, 0)^T$. Then we have

Lemma 3: The pair $(\mathcal{A}, \mathcal{E}_j)$ is observable on a generic subset of the parameter space Θ .

Proof: It is well known that nonobservability of the pair $(\mathcal{A}, \mathcal{E}_j)$ is equivalent to the existence of a nonzero vector c for which, for some λ , there holds $\mathcal{A}c = \lambda c, \mathcal{E}_j^T c = 0$.

Suppose the conclusion of the lemma is false. Let c_1 denote the vector comprising the first n entries of c and let λ be the corresponding eigenvalue of \mathcal{A} which is nonzero because A_p can be assumed as nonsingular on a generic set. Then by the result of Lemma 2, there holds

$$\begin{aligned} \bar{A}(\lambda)c_1 &= 0 \\ e_j^T c_1 &= 0. \end{aligned}$$

Now observe that the second equation requires that the j -th entry of c_1 be zero. From the generic zerolessness of tall transfer functions, see [1], we conclude that the other $n-1$ entries of c_1 must be zero for a generic subset of the parameter space, i.e. that c_1 itself is zero. This is clearly false. Hence the conclusion of the lemma is established by contradiction. ■

Proof: of the main theorem. The following equations can be verified using equation (3)

$$(zI - \mathcal{A})\Gamma(z^{-1}I - \mathcal{A}^T) + \mathcal{A}\Gamma(z^{-1}I - \mathcal{A}^T) + (zI - \mathcal{A})\Gamma\mathcal{A}^T = \mathcal{B}\mathcal{B}^T$$

$$\begin{aligned} \Gamma + (zI - \mathcal{A})^{-1}\mathcal{A}\Gamma + \Gamma\mathcal{A}^T(z^{-1}I - \mathcal{A}^T)^{-1} &= \\ = (zI - \mathcal{A})^{-1}\mathcal{B}\mathcal{B}^T(z^{-1}I - \mathcal{A}^T)^{-1} & \end{aligned}$$

Premultiplying and postmultiplying by \mathcal{E}_1^T and \mathcal{E}_1 leads to

$$\mathcal{E}_1^T \Gamma \mathcal{E}_1 + \mathcal{E}_1^T (zI - \mathcal{A})^{-1} \mathcal{A} \Gamma \mathcal{E}_1 + \mathcal{E}_1^T \Gamma \mathcal{A}^T (z^{-1}I - \mathcal{A}^T)^{-1} \mathcal{E}_1 = \mathcal{E}_1^T (zI - \mathcal{A})^{-1} \mathcal{B} \mathcal{B}^T (z^{-1}I - \mathcal{A}^T)^{-1} \mathcal{E}_1 \quad (8)$$

Since by corollary 1 $(\mathcal{A}, \mathcal{B})$ is controllable on a generic subset of the parameter space, and by lemma 3, we have that $(\mathcal{A}, \mathcal{E}_1)$ is observable on a generic subset of the parameter space, and since the intersection of two generic sets is generic again, $(\mathcal{A}, \mathcal{B}, \mathcal{E}_1)$ is minimal and thus the McMillan degree of $\mathcal{E}_1^T (zI - \mathcal{A})^{-1} \mathcal{B}$ is np , the dimension of \mathcal{A} . It follows that the transfer function $\mathcal{B}^T (z^{-1}I - \mathcal{A}^T)^{-1} \mathcal{E}_1$ also has McMillan degree np , since transpose operations preserve McMillan degree, and replacements of a variable by a Mobius transformation (see [6]) of that variable preserve McMillan degree.

Further, by the stability assumption on the underlying AR system and because A_p is assumed as nonsingular, the McMillan degree of the product $\mathcal{E}_1^T (zI - \mathcal{A})^{-1} \mathcal{B} \mathcal{B}^T (z^{-1}I - \mathcal{A}^T)^{-1} \mathcal{E}_1$ will be $2np$, due to the absence of any pole-zero cancellations. Now the two nonconstant transfer functions in (8) on the left side necessarily have the same McMillan degree, one being obtainable from the other by transposition and Mobius transformation of the variable. Further, the nonconstant transfer functions on the left side of equation share no common poles, so that their sum has McMillan degree equal to the sum of the two McMillan degrees, or twice the McMillan degree of one of the transfer functions. Hence on the left side, we must have the McMillan degree of $\mathcal{E}_1^T (zI - \mathcal{A})^{-1} \mathcal{A} \Gamma \mathcal{E}_1$ equal to np , so that $\mathcal{A}, \mathcal{A} \Gamma \mathcal{E}_1$ is controllable. It follows trivially that $(\mathcal{A}, \Gamma \mathcal{E}_1)$ is controllable.

Up to now, the proof has only been given for the case $n_f = 1$. As is easily seen, the result is true a fortiori for $n_f > 1$. ■

The importance of theorem 1 is that if Z has full row rank, the mapping from the second moments which can be observed in principle to the parameters is continuous and thus consistent estimators of the corresponding population second moments give consistent estimators for the underlying high frequency parameters. This also holds for the case of generalized factor models where the static factors can be estimated by principal component analysis, see [3]. Of course, also the asymptotic covariances of the estimators are of interest. This however is left to future research.

C. Generic identifiability of the noise parameters

To show generic identifiability of the noise parameters Σ we proceed as follows. We commence from identifiable system parameters A_1, \dots, A_p . Rewriting equation (2) as

$$\underbrace{\begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}}_{=x_{t+1}} = \mathcal{A} \underbrace{\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix}}_{=x_t} + \underbrace{\begin{pmatrix} I_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=\mathcal{G}} \nu_t. \quad (9)$$

$$y_t = \underbrace{(I_n, 0, \dots, 0)}_{=\mathcal{H}} x_{t+1} \quad (10)$$

we obtain through vectorization of

$$\Gamma = \mathbb{E}(x_t x_t^T) = \mathcal{A} \Gamma \mathcal{A} + \mathcal{G} \Sigma \mathcal{G}$$

and

$$\gamma_0 = \mathbb{E}(y_t y_t^T) = \mathcal{H} \Gamma \mathcal{H}^T$$

that

$$\begin{aligned} \text{vec}(\Gamma) &= (\mathcal{A} \otimes \mathcal{A}) \text{vec}(\Gamma) + (\mathcal{G} \otimes \mathcal{G}) \text{vec}(\Sigma) \\ &= (I - \mathcal{A} \otimes \mathcal{A})^{-1} (\mathcal{G} \otimes \mathcal{G}) \text{vec}(\Sigma) \end{aligned} \quad (11)$$

and

$$\text{vec}(\gamma_0) = (\mathcal{H} \otimes \mathcal{H}) \text{vec}(\Gamma). \quad (12)$$

Therefore, we obtain that

$$\text{vec}(\gamma_0) = (\mathcal{H} \otimes \mathcal{H}) (I - \mathcal{A} \otimes \mathcal{A})^{-1} (\mathcal{G} \otimes \mathcal{G}) \text{vec}(\Sigma). \quad (13)$$

Note that the absolute value of all eigenvalues of \mathcal{A} is smaller than one by the stability assumption and therefore the same holds for the eigenvalues of $(\mathcal{A} \otimes \mathcal{A})$. This implies that $(I - \mathcal{A} \otimes \mathcal{A})$ is regular. For $A_1 = \dots = A_p = 0$, the matrix $(I - \mathcal{A} \otimes \mathcal{A})^{-1}$ is triangular with ones on its diagonal. Thus $(\mathcal{H} \otimes \mathcal{H}) (I - \mathcal{A} \otimes \mathcal{A})^{-1} (\mathcal{G} \otimes \mathcal{G})$ is a principal submatrix with the same property and is therefore nonsingular. This nonsingularity holds in an open neighborhood of $A_1 = \dots = A_p = 0$ and this neighborhood has a nonempty intersection with the generic set of identifiable system parameters as described in theorem 1. Now, there exists a point in this intersection for which the determinant of $(\mathcal{H} \otimes \mathcal{H}) (I - \mathcal{A} \otimes \mathcal{A})^{-1} (\mathcal{G} \otimes \mathcal{G})$ is unequal to zero. Since this determinant is a rational function in the free entries A_1, \dots, A_p the nonsingularity holds for a generic set in the parameter space. For the properties of the set of zeros of multivariate polynomials and thus rational functions see e.g. [7].

Clearly, if the matrix on the right hand side of equation (13) is nonsingular we have identifiability of Σ .

Thus we obtain the desired result:

Theorem 2: The noise parameters Σ are generically identifiable.

Note that, as immediate from the proof above, for generic values of A_1, \dots, A_p , Σ is always identifiable.

A consistent estimator in this context has been derived in [2].

III. SUBSTITUTION METHOD FOR THE REGULAR CASE WITH $p = 1$, AND $N = 2$

In this section, we discuss a procedure obtained by substitution which gives more detailed identifiability results. However, the derivation is only for the regular AR(1) case with $N = 2$ and $n_f = n_s = 1$.

A. Transformation of the high frequency AR(1) system into a system in observed variables

For ease of notation we write AR(1) systems as

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

The system (1) can be decomposed in two subsystems,

$$y_t^f = (a_{ff} \quad a_{fs}) \begin{pmatrix} y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} + \nu_t^f \quad (14)$$

$$y_t^s = (a_{sf} \quad a_{ss}) \begin{pmatrix} y_{t-1}^f \\ y_{t-1}^s \end{pmatrix} + \nu_t^s. \quad (15)$$

Note that in (14) all the moments needed for a projection of y_t^f on y_{t-1}^f and y_{t-1}^s are available and thus a_{ff} , a_{fs} , and Σ_{ff} are uniquely determined. On the other hand, for the second equation (15) the autocovariance $\gamma^{ss}(1) = \mathbb{E}(y_t^s (y_{t-1}^s)^T)$ is not available. For this reason, we replace y_{t-1}^s by $a_{sf}y_{t-2}^f + a_{ss}y_{t-2}^s + \nu_{t-1}^s$ which gives the projection of $\begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix}$ onto $\begin{pmatrix} y_{t-1}^f \\ y_{t-2}^f \\ y_{t-2}^s \end{pmatrix}$.

$$\begin{aligned} \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} &= \begin{pmatrix} a_{ff} & a_{fs}a_{sf} & a_{fs}a_{ss} \\ a_{sf} & a_{ss}a_{sf} & a_{ss}^2 \end{pmatrix} \begin{pmatrix} y_{t-1}^f \\ y_{t-2}^f \\ y_{t-2}^s \end{pmatrix} + \\ &+ \begin{pmatrix} a_{fs} \\ a_{ss} \end{pmatrix} \mathbb{E}(\nu_{t-1}^s | y_{t-1}^f, y_{t-2}^f, y_{t-2}^s) + \begin{pmatrix} \tilde{\nu}_t^f \\ \tilde{\nu}_t^s \end{pmatrix} = \\ &= \begin{pmatrix} a_{ff} & a_{fs}a_{sf} & a_{fs}a_{ss} \\ a_{sf} & a_{ss}a_{sf} & a_{ss}^2 \end{pmatrix} \begin{pmatrix} y_{t-1}^f \\ y_{t-2}^f \\ y_{t-2}^s \end{pmatrix} + \\ &+ \begin{pmatrix} a_{fs}\Sigma_{sf}\Sigma_{ff}^{-1} & -a_{fs}\Sigma_{sf}\Sigma_{ff}^{-1}a_{ff} & -a_{fs}\Sigma_{sf}\Sigma_{ff}^{-1}a_{fs} \\ a_{ss}\Sigma_{sf}\Sigma_{ff}^{-1} & -a_{ss}\Sigma_{sf}\Sigma_{ff}^{-1}a_{ff} & -a_{ss}\Sigma_{sf}\Sigma_{ff}^{-1}a_{fs} \end{pmatrix} \begin{pmatrix} y_{t-1}^f \\ y_{t-2}^f \\ y_{t-2}^s \end{pmatrix} \\ &+ \begin{pmatrix} \tilde{\nu}_t^f \\ \tilde{\nu}_t^s \end{pmatrix}. \quad (16) \end{aligned}$$

In equation (16), the error term is

$$\begin{pmatrix} \tilde{\nu}_t^f \\ \tilde{\nu}_t^s \end{pmatrix} = \begin{pmatrix} a_{fs} \\ a_{ss} \end{pmatrix} (\nu_{t-1}^s - \mathbb{E}(\nu_{t-1}^s | y_{t-1}^f, y_{t-2}^f, y_{t-2}^s)) + \begin{pmatrix} \nu_t^f \\ \nu_t^s \end{pmatrix}. \quad (17)$$

Note that the components of the regression vector $\begin{pmatrix} y_{t-1}^f \\ y_{t-2}^f \\ y_{t-2}^s \end{pmatrix}$ are linearly independent and thus the regression coefficients b_{ij} , $i = 1, 2$, $j = 1, 2, 3$, which are the block entries of the matrix multiplying the regression vector in (16) are uniquely determined. To sum up, and recalling that a_{ff} , a_{fs} , and Σ_{ff} can be assumed known based on use of (14), the problem of identifiability has now been reduced to the unique solvability of the following equation system in the variables a_{sf} , a_{ss} , Σ_{fs} , and Σ_{ss} :

$$\begin{aligned} b_{11} &= a_{ff} + a_{fs}\Sigma_{sf}\Sigma_{ff}^{-1} \\ b_{12} &= a_{fs}a_{sf} - a_{fs}\Sigma_{sf}\Sigma_{ff}^{-1}a_{ff} \\ b_{13} &= a_{fs}a_{ss} - a_{fs}\Sigma_{sf}\Sigma_{ff}^{-1}a_{fs} \\ b_{21} &= a_{sf} + a_{ss}\Sigma_{sf}\Sigma_{ff}^{-1} \\ b_{22} &= a_{ss}a_{sf} - a_{ss}\Sigma_{sf}\Sigma_{ff}^{-1}a_{ff} \\ b_{23} &= a_{ss}^2 - a_{ss}\Sigma_{sf}\Sigma_{ff}^{-1}a_{fs}. \end{aligned}$$

Finally we obtain for the covariances of the new error term

$$\begin{aligned} \tilde{\Sigma}_{ff} &= \Sigma_{ff} + a_{fs} \left(\Sigma_{ss} - \Sigma_{sf}\Sigma_{ff}^{-1}\Sigma_{fs} \right) a_{fs}^T \\ \tilde{\Sigma}_{ss} &= \Sigma_{ss} + a_{ss} \left(\Sigma_{ss} - \Sigma_{sf}\Sigma_{ff}^{-1}\Sigma_{fs} \right) a_{ss}^T \\ \tilde{\Sigma}_{fs} &= \Sigma_{fs} + a_{fs} \left(\Sigma_{ss} - \Sigma_{sf}\Sigma_{ff}^{-1}\Sigma_{fs} \right) a_{ss}^T. \end{aligned}$$

As can be seen, no further information for the underlying parameters can be obtained by further substitutions in this case.

B. Identifiability Analysis for the case $n_f = n_s = 1$

In this subsection, nonidentifiability is characterized by the existence of a static linear transformation which orthogonalizes the two components and leaves the second component as an AR(1) process (i.e. not white noise).

1) The case $\Sigma_{sf} = 0$:

Lemma 4: The system and noise parameters $\begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \end{pmatrix}$, Σ_{ff} , and Σ_{ss} are not identified if and only if $(a_{fs} = 0) \wedge (a_{sf} = 0) \wedge (a_{ss} \neq 0)$.

Proof:

a) If either $a_{fs} \neq 0$ or $a_{sf} \neq 0$ or $a_{ss} = 0$, then the system is identifiable: The regression coefficients of the two systems (14) and (16) are

$$\beta^{(1)} = \begin{pmatrix} a_{ff} & a_{fs} \end{pmatrix} \quad (18)$$

$$\beta^{(2)} = \begin{pmatrix} a_{ff} & a_{fs}a_{sf} & a_{fs}a_{ss} \\ a_{sf} & a_{ss}a_{sf} & a_{ss}^2 \end{pmatrix}. \quad (19)$$

Their respective error covariance matrices are Σ_{ff} and (for (16))

$$\begin{aligned} \mathbb{E} \left(\begin{pmatrix} \tilde{\nu}_t^f \\ \tilde{\nu}_t^s \end{pmatrix} \begin{pmatrix} (\tilde{\nu}_t^f)^T & (\tilde{\nu}_t^s)^T \end{pmatrix} \right) &= \begin{pmatrix} \tilde{\Sigma}_{ff} & \tilde{\Sigma}_{fs} \\ \tilde{\Sigma}_{sf} & \tilde{\Sigma}_{ss} \end{pmatrix} = \\ &= \begin{pmatrix} a_{fs} \\ a_{ss} \end{pmatrix} \Sigma_{ss} \begin{pmatrix} a_{fs}^T & a_{ss}^T \end{pmatrix} + \begin{pmatrix} \Sigma_{ff} & 0 \\ 0 & \Sigma_{ss} \end{pmatrix} \end{aligned}$$

which is obtained from $\begin{pmatrix} \tilde{\nu}_{1,t} \\ \tilde{\nu}_{2,t} \end{pmatrix} = \begin{pmatrix} a_{fs} \\ a_{ss} \end{pmatrix} \nu_{t-1}^s + \begin{pmatrix} \nu_t^f \\ \nu_t^s \end{pmatrix}$.

Thus, we obtain from equation (14) a_{ff} , a_{fs} and Σ_{ff} and from (16) a_{sf} and Σ_{ss} (from $\tilde{\Sigma}_{ss}$).

If $a_{fs} \neq 0$ or $a_{sf} \neq 0$ or $a_{ss} = 0$, then the system is identified from the coefficients $\beta_{13}^{(2)}$, $\beta_{22}^{(2)}$ and $\beta_{23}^{(2)}$ in equation (19).

b) If $(a_{fs} = 0) \wedge (a_{sf} = 0) \wedge (a_{ss} \neq 0)$, then the system is not identifiable: Obviously, then the process consists of two orthogonal components, and therefore only a_{ss}^2 can be obtained. ■

Remark 1: As is seen from the discussion above, in the nonidentifiable case the fast process contains no information about the parameters of the slow process, which is an AR(1) process on $t \in 2\mathbb{Z}$. Thus only a_{ss}^2 (and Σ_{ss}) can be obtained from the observable second moments of the observations. If $a_{ss}^2 \neq 0$ then $+\sqrt{a_{ss}^2}$ and $-\sqrt{a_{ss}^2}$ are possible solutions. Note that observable covariances corresponding to higher lags provide no additional information.

Remark 2: According to the lemma 4, we have identifiability if $a_{fs} \neq 0$ or $a_{sf} \neq 0$ or $a_{ss} = 0$. The matrix Z in equation (5) is rank deficient for $a_{ss} = 0$, $a_{sf} = 0$ even if $a_{fs} \neq 0$ which shows that the condition that $rk(Z) = np$ hold, is not necessary for identifiability. Whereas the equivalence classes of observational equivalent parameters in Remark 1 consist of two points the solution set of the extended Yule Walker equations consist of affine subspaces. This shows that the extended Yule Walker equations do not use the full information contained in the second moments which are in principle observed.

2) *General case (no assumption on Σ_{sf}):*

Theorem 3: 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 are on the manifold described by the equations

$$\begin{aligned} a_{fs} &= 0 \\ a_{sf} + \frac{\Sigma_{sf}}{\Sigma_{ff}}(a_{ss} - a_{ff}) + \left(\frac{\Sigma_{sf}}{\Sigma_{ff}}\right)^2 a_{fs} &= 0 \\ a_{ss} - a_{fs} \frac{\Sigma_{sf}}{\Sigma_{ff}} &\neq 0 \end{aligned}$$

whose complement is a superset of an open and dense set (with respect to the whole parameter space). In the nonidentifiable case, it is possible to orthogonalize the system by a linear static transformation.

Proof: The proof consists in tracing the case without restrictions on Σ_{sf} back to the case with $\Sigma_{sf} \neq 0$.

a) *Applying a static linear transformation on the system (1):* Applying the transformation T defined by

$$T = \begin{pmatrix} I & 0 \\ -\Sigma_{sf}\Sigma_{ff}^{-1} & I \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} I & 0 \\ \Sigma_{sf}\Sigma_{ff}^{-1} & I \end{pmatrix},$$

we obtain as new system matrix $\begin{pmatrix} \tilde{a}_{ff} & \tilde{a}_{fs} \\ \tilde{a}_{sf} & \tilde{a}_{ss} \end{pmatrix}$ the following:

$$\begin{aligned} T \begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \end{pmatrix} T^{-1} &= \begin{pmatrix} \tilde{a}_{ff} & \tilde{a}_{fs} \\ \tilde{a}_{sf} & \tilde{a}_{ss} \end{pmatrix} \\ &= \left(\begin{array}{c|c} a_{ff} + a_{fs}\Sigma_{sf}\Sigma_{ff}^{-1} & a_{fs} \\ \hline a_{sf} + \frac{\Sigma_{sf}}{\Sigma_{ff}}(a_{ss} - a_{ff}) + \left(\frac{\Sigma_{sf}}{\Sigma_{ff}}\right)^2 a_{fs} & a_{ss} - \Sigma_{sf}\Sigma_{ff}^{-1}a_{fs} \end{array} \right) \end{aligned}$$

Thus, the new system with diagonal error covariance matrix is

$$\begin{aligned} T \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} &= T \begin{pmatrix} a_{ff} & a_{fs} \\ a_{sf} & a_{ss} \end{pmatrix} T^{-1} T \begin{pmatrix} y_t^f \\ y_t^s \end{pmatrix} + T \begin{pmatrix} \nu_t^f \\ \nu_t^{f\perp} \end{pmatrix} \\ \begin{pmatrix} y_t^f \\ \tilde{y}_t^s \end{pmatrix} &= \begin{pmatrix} \tilde{a}_{ff} & \tilde{a}_{fs} \\ \tilde{a}_{sf} & \tilde{a}_{ss} \end{pmatrix} \begin{pmatrix} y_{t-1}^f \\ \tilde{y}_{t-1}^s \end{pmatrix} + \begin{pmatrix} \nu_t^f \\ \nu_t^{f\perp} \end{pmatrix}. \end{aligned}$$

b) *Using the new system for identifiability analysis.:* As in the diagonal, i.e. $\Sigma_{sf} = 0$, case, we obtain identifiability if and only if $(\tilde{a}_{fs} = 0) \wedge (\tilde{a}_{sf} = 0) \wedge (\tilde{a}_{ss} \neq 0)$. ■

IV. CONCLUSION

This paper has demonstrated that vector autoregressions are generically identifiable from covariance data in which significant information is missing, corresponding to the fact that some system outputs are only available every N -th time instants for some $N > 1$. Identifiability is however then a generic property. We have yet to determine what systems will be hard to identify, i.e. close to the nonidentifiable set in some metric.

REFERENCES

- [1] 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.
- [2] B. Chen and P. A. Zadrozny. An Extended Yule-Walker Method for Estimating a Vector Autoregressive Model with Mixed-Frequency. *Advances in Econometrics*, 13:47–73, 1998.
- [3] 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.
- [4] M. Forni and M. Lippi. The generalized dynamic factor model: representation theory. *Econometric Theory*, 17:1113–1141, 2001.
- [5] E. J. Hannan and M. Deistler. *The Statistical Theory of Linear Systems*. SIAM, Philadelphia, 2012.
- [6] 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.
- [7] I. Shafarevich. *Basic Algebraic Geometry 1*. Springer-Verlag, 1994.