49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Multivariable Frequency Domain Identification using IV-based Linear Regression

Rogier S. Blom and Paul M. J. Van den Hof

Abstract— **Identification of output error models from frequency domain data generally results in a non-convex optimization problem. A well-known method to approach the output error minimum by iterative linear regression steps was formulated by Sanathanan and Koerner. A disadvantage of this approach is that in general convergence of the iterations only implies optimality under restrictive conditions. In the literature, an alternative iterative linear regression procedure is available, which ensures optimality upon convergence, also in case of undermodeling. This algorithm is known for time-domain identification as the Simplified Refined Instrumental Variable method (SRIV), and was recently formulated for frequency domain identification of SISO output error models. Here we generalize this formulation to MIMO identification of models in matrix fraction description. The effectiveness of the approach is demonstrated by its application to estimation of a parametric model of the multivariable dynamics of a spindle with Active Magnetic Bearings.**

I. Introduction

Identification of multivariable parametric models from frequency response function (FRF) data is applied in numerous application areas. Frequently, the emphasis is on identification of a parametric model of the plant dynamics. When the objective is to find a model in a fractional representation with a parametrized numerator that minimizes a quadratic output error (OE) criterion, in general a nonconvex optimization problem is obtained. Gradient-based optimization can be employed to solve such problems, however the computational complexity has stimulated many authors to look for simpler alternatives.

One approach is to replace the nonconvex optimization by a sequence of linear regression steps. A classical method to achieve this when estimating SISO models from FRF data is due to Sanathanan and Koerner [1]. This method forms the basis for several MIMO identification methods. These MIMO methods differ in the model structure that is used to parametrize the multivariable system and the ability to incorporate frequency dependent weighting to improve the estimate. In the approach by Bayard [2] a model representation with a matrix numerator polynomial and a scalar denominator polynomial is used. Here particular attention is given to the computational aspects given the sparsity of the resulting linear regression steps. While in the approach by Bayard only scalar frequency weighting can be applied, the approach by Verboven et al. [3] further extends and improves the method of Bayard for multivariable frequency dependent weighting, with applications in modal analysis. The approach by de Callafon et al. [4] uses a more flexible model representation by definition of model sets in left and right polynomial matrix fraction description (MFD). Multivariable frequency dependent weighting can be incorporated, where solutions for both input/output weighting and element-wise weighting are formulated. Also Gaikwad and Rivera [5] give an extension of the approach by Bayard for models in MFD and specifically utilize the possibility to perform pre and post weighting to identify a control-relevant model. Also non-iterative methods exist that estimate state-space models from FRF data using subspace algorithms [6], although these algorithms do not guarantee a cost criterion is minimized.

An important and well-known limitation of the algorithm proposed by Sanathanan and Koerner is that in general convergence of the iterations does not imply that the resulting parameter estimate minimizes the OE cost criterion, inevitably leading to a bias in the estimated model.

In the literature, an alternative iterative linear regression procedure is available, that is based on an Instrumental Variable (IV) approach. This algorithm is known for time-domain identification of OE models as the Simplified Refined Instrumental Variable method (SRIV) [7], [8], and was recently formulated for frequency domain identification of SISO OE models [9]. This method has the property that upon convergence of the iterations a stationary point of the cost function is reached, also in the case the system is not in the model set. In this paper, we generalize this formulation to MIMO identification of discrete-time and continuous-time models in matrix fraction description. Iterative linear regression algorithms are derived for the case pre and post, or element-wise multivariable frequency weighting of the OE is applied.

After introducing the identification setting in section II and discussing the model sets in section III, we give the linear regression algorithms for identification of OE models in matrix fraction description (section IV). The algorithms have been applied for identification of the dynamics of an Active Magnetic Bearing spindle, of which the results are discussed in section V.

II. Identification setting

The central objective in this paper is to find an LTI model P of a multivariable system with m inputs and p

This work was supported by MicroNed and the Delft Center for Mechatronics and Microsystems

R.S. Blom is with Precision and Microsystems Engineering and the Delft Center for System and Control, Faculty of 3mE, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands r.s.blom@tudelft.nl

P.M.J. Van den Hof is with the Delft Center for System and Control, Faculty of 3mE, Delft University of Technology, Mekelweg 2, 2628 CD, Delft, The Netherlands p.m.j.vandenhof@tudelft.nl

978-1-4244-7744-9/10/$26.00 ©2010 IEEE

outputs using measured data. For this, the set of N noisy multivariable frequency response function observations $\mathcal{G}$ is available, which is defined as

$$\mathcal{G} := \{G(\omega_k) | G(\omega_k) \in \mathbb{C}^{p\times m}, k = 1 \ldots N\}. \quad (1)$$

Here the model P is parametrized by either a left or right polynomial MFD depending on the real valued parameter vector θ. Further details on the parametrization are discussed in the next section. To describe the discrepancy between the data $\mathcal{G}$ and the model $P(\theta)$, we use the output error

$$E(\omega_k, \theta) = G(\omega_k) - P(\xi(\omega_k), \theta). \quad (2)$$

Here $\xi(\omega_k)$ is used to denote the frequency dependency of P, where $\xi(\omega_k) = i\omega_k$ or $\xi(\omega_k) = e^{i\omega_k}$ when P represents a continuous-time or discrete-time system respectively. With this, the aim is to solve the identification problem $\hat{\theta} = \arg\min_\theta V(\theta)$ with cost function

$$V(\theta) = \sum_{k=1}^{N} \|E(\omega_k, \theta)\|_F^2 . \quad (3)$$

Application of frequency dependent weighting of the OE can be used to obtain estimates with lower variance when the frequency response data have varying covariance, or to estimate control-relevant models. Cost functions with weighted errors can be obtained by substituting for $E(\omega_k, \theta)$ in (3) either the *input/output weighted OE*

$$E_{i/o}(\omega_k, \theta) = W_o(\omega_k)[G(\omega_k) - P(\xi(\omega_k), \theta)]W_i(\omega_k), \quad (4)$$

with $W_o(\omega_k) \in \mathbb{R}^{p\times p}$ and $W_i(\omega_k) \in \mathbb{R}^{m\times m}$, or the *Schur-weighted OE*

$$E_s(\omega_k, \theta) = W_s(\omega_k). * [G(\omega_k) - P(\xi(\omega_k), \theta)], \quad (5)$$

with $W_s(\omega_k) \in \mathbb{R}^{p\times m}$, and where $.*$ is used to denote the Schur matrix product (i.e. element-wise multiplication).

III. Model structure

A. Definition of the model set

The MIMO model is represented in a left or right matrix fraction description:

$$P(\xi, \theta) = B(\xi, \theta)A^{-1}(\xi, \theta) \quad \text{(R-MFD)}$$
$$P(\xi, \theta) = A^{-1}(\xi, \theta)B(\xi, \theta) \quad \text{(L-MFD)}$$

with $B(\xi, \theta) = B_{n_b}\xi^{n_b} + B_{n_b-1}\xi^{n_b-1} + \cdots + B_0$ and $A(\xi, \theta) = \xi^{n_a} + A_{n_a-1}\xi^{n_a-1} + \cdots + A_0$, where $B_i \in \mathbb{R}^{p\times m}$, $i = 0 \ldots n_b$ and $A_i \in \mathbb{R}^{m\times m}$ (for models in R-MFD), or $A_i \in \mathbb{R}^{p\times p}$ (for models in L-MFD), for $i = 0 \ldots n_a - 1$. The parameter vector θ is constructed by accumulating all elements of the matrices A_i, $i = 0 \ldots n_a - 1$ and B_i, $i = 0 \ldots n_b$. As will become clear in the sequel, it is convenient to choose an ordering of these elements that depends on the choice for a left or right matrix fraction description of the model set. To avoid unnecessary notational complexity, it is assumed here that the order of the polynomials of all elements of A are n_a and those of B are all n_b. However, the approach presented in this paper is equally suited for model sets where the matrix fractions have elements with varying polynomial orders. We will return to this in section IV-F.

IV. An IV-based iterative method to solve a multivariable OE identification problem

A. A criterion for optimality

The identification problem that was posed in section II, results in a nonconvex optimization problem. In this section we give an iterative linear regression algorithm to solve this optimization problem for both selected model sets, having the property that convergence implies (local) optimality. This algorithm is a MIMO extension of the frequency domain formulation of the SRIV method as given by Van den Hof and Douma in [9]. We first give the result for the unweighted OE cost function. In section IV-D we will extend this for the weighted OE cost functions.

Similar as with the SRIV method, the starting notion is that for all $\hat{\theta}$ that locally minimize $V(\theta)$, $\frac{\partial}{\partial\theta}V(\theta)\big|_{\theta=\hat{\theta}} = 0$. From the definition of $V(\theta)$ in (3), it follows that

$$\frac{\partial}{\partial\theta}V(\theta) = \frac{\partial}{\partial\theta}\sum_{k=1}^{N} \text{vec}\left[E(\omega_k, \theta)\right]^H \text{vec}\left[E(\omega_k, \theta)\right]$$
$$= \sum_{k=1}^{N} -2Re\left\{\text{vec}\left[E(\omega_k, \theta)\right]^H M_k(\theta)\right\}$$

where $(\cdot)^H$ denotes the complex conjugate transpose, $\text{vec}(\cdot)$ the vectorization operator, and $M_k(\theta) = \frac{\partial}{\partial\theta}\text{vec}[P(\xi(\omega_k), \theta)]$. Hence, for all $\hat{\theta}$ for which $\frac{\partial}{\partial\theta}V(\hat{\theta}) = 0$, the following equality holds:

$$\sum_{k=1}^{N} Re\left\{M_k^H(\hat{\theta})\text{vec}[E(\omega_k, \hat{\theta})]\right\} = 0. \quad (6)$$

B. Iterative procedure for models in L-MFD

In this section we will use (6) to arrive at an iterative linear regression algorithm to estimate $\hat{\theta}$ for models in L-MFD. For that, we will rewrite (6) in a regression format. Let us therefore introduce the notation

$$\Theta = \left[\begin{array}{cccccc} A_{n_a-1} & \ldots & A_0 & B_{n_b} & \ldots & B_0 \end{array}\right]$$
$$\theta = \text{vec}(\Theta). \quad (7)$$

We give the following two propositions:

Proposition 4.1: With $E(\xi(\omega_k), \theta)$ as defined in (2), where the model is represented in L-MFD, and with θ as defined in (7), the following identity holds

$$\text{vec}\left[E(\omega_k, \theta)\right] = Y_k(\theta) - X_k(\theta)\theta \quad (8)$$

with

$$Y_k(\theta) = \left[I \otimes A^{-1}(\xi(\omega_k), \theta)\right] \text{vec}[\xi(\omega_k)^{n_a} G(\omega_k)]$$
$$X_k(\theta) = \left[\Omega^T(\omega_k) \otimes A^{-1}(\xi(\omega_k), \theta)\right]$$

where $\otimes$ is the Kronecker product, and

$$\Omega(\omega_k) = \begin{bmatrix} -\xi(\omega_k)^{n_a-1}G(\omega_k) \\ \vdots \\ -\xi(\omega_k)^0 G(\omega_k) \\ \xi(\omega_k)^{n_b} I_{m\times m} \\ \vdots \\ \xi(\omega_k)^0 I_{m\times m} \end{bmatrix}.$$

Proof: For the given model parametrization, and using the given definition of $\Omega(\omega_k)$, observe that we can express $E(\xi(\omega_k),\theta)$ as

$$\begin{aligned}E(\xi(\omega_k),\theta) &= G(\omega_k) - A^{-1}(\xi(\omega_k),\theta)B(\xi(\omega_k),\theta)\\ &= A^{-1}(\xi(\omega_k),\theta)\left[A(\xi(\omega_k),\theta)G(\omega_k) - B(\xi(\omega_k),\theta)\right]\\ &= A^{-1}(\xi(\omega_k),\theta)\left[\xi(\omega_k)^{n_a}G(\omega_k) - \Theta\Omega(\omega_k)\right].\end{aligned}$$

To proceed, we need the following two identities:

$$\mathrm{vec}(AB) = (I\otimes A)\mathrm{vec}(B) = (B^T\otimes I)\mathrm{vec}(A) \tag{9a}$$
$$(A\otimes B)(C\otimes D) = AB\otimes DB. \tag{9b}$$

Using (9a) we can write

$$\mathrm{vec}\left[E(\omega_k,\theta)\right] = \left[I\otimes A^{-1}(\xi(\omega_k),\theta)\right]\cdot \cdot\,\mathrm{vec}\left[\xi(\omega_k)^{n_a}G(\omega_k) - \Theta\Omega(\omega_k)\right]$$

which by applying (9a) and subsequently (9b), we can rewrite to

$$\begin{aligned}\mathrm{vec}\left[E(\omega_k,\theta)\right] &= \left[I\otimes A^{-1}(\xi(\omega_k),\theta)\right]\cdot\\ &\cdot\left(\mathrm{vec}[\xi(\omega_k)^{n_a}G(\omega_k)] - (\Omega^T(\omega_k)\otimes I)\mathrm{vec}(\Theta)\right)\\ &= \left[I\otimes A^{-1}(\xi(\omega_k),\theta)\right]\mathrm{vec}[\xi(\omega_k)^{n_a}G(\omega_k)] +\\ &\quad - \left[\Omega^T(\omega_k)\otimes A^{-1}(\xi(\omega_k),\theta)\right]\theta\end{aligned}$$

which is the claimed result. ■

Proposition 4.2: For models parametrized in L-MFD,

$$M_k(\theta) = \Phi_k(\theta)^T\otimes A^{-1}(\xi(\omega_k),\theta) \tag{10}$$

with

$$\Phi_k(\theta) = \begin{bmatrix} -\xi(\omega_k)^{n_a-1}P(\xi(\omega_k),\theta)\\ \vdots\\ -\xi(\omega_k)^0P(\xi(\omega_k),\theta)\\ \xi(\omega_k)^{n_b}I_{m\times m}\\ \vdots\\ \xi(\omega_k)^0 I_{m\times m}\end{bmatrix}.$$

Proof: See appendix VIII-A. ■

With the results in equations (8) and (10), we can recast (6) into

$$\sum_{k=1}^{N} Re\left\{M_k^H(\hat\theta)(Y_k(\hat\theta) - X_k(\hat\theta)\hat\theta)\right\} = 0 \tag{11}$$

or equivalently

$$\sum_{k=1}^{N}\left[\begin{array}{cc} Re\{M_k^T(\hat\theta)\} & Im\{M_k^T(\hat\theta)\}\end{array}\right]\cdot \cdot\left(\begin{bmatrix} Re\{Y_k(\hat\theta)\}\\ Im\{Y_k(\hat\theta)\}\end{bmatrix} - \begin{bmatrix} Re\{X_k(\hat\theta)\}\\ Im\{X_k(\hat\theta)\}\end{bmatrix}\hat\theta\right) = 0.$$

With the notation

$$\mathbf{M}^T(\theta) := \left[\begin{array}{ccc} Re\{M_1^T(\theta)\} & Im\{M_1^T(\theta)\} & \dots\\ \dots & Re\{M_N^T(\theta)\} & Im\{M_N^T(\theta)\}\end{array}\right]$$

$$\mathbf{X}(\theta) := \begin{bmatrix} Re\{X_1(\theta)\}\\ Im\{X_1(\theta)\}\\ \vdots\\ Re\{X_N(\theta)\}\\ Im\{X_N(\theta)\}\end{bmatrix},\quad \mathbf{Y}(\theta) := \begin{bmatrix} Re\{Y_1(\theta)\}\\ Im\{Y_1(\theta)\}\\ \vdots\\ Re\{Y_N(\theta)\}\\ Im\{Y_N(\theta)\}\end{bmatrix}$$

it follows that the solution of (6) is characterized by

$$\mathbf{M}^T(\hat\theta)(\mathbf{Y}(\hat\theta) - \mathbf{X}(\hat\theta)\hat\theta) = 0. \tag{12}$$

From this, a natural iterative identification algorithm follows:

$$\mathbf{M}^T(\hat\theta_{j-1})(\mathbf{Y}(\hat\theta_{j-1}) - \mathbf{X}(\hat\theta_{j-1})\hat\theta_j) = 0 \tag{13}$$

with solution

$$\hat\theta_j = \left[\mathbf{M}^T(\hat\theta_{j-1})\mathbf{X}(\hat\theta_{j-1})\right]^{-1}\mathbf{M}^T(\hat\theta_{j-1})\mathbf{Y}(\hat\theta_{j-1}) \tag{14}$$

When this algorithm converges, necessarily $V'(\hat\theta) = 0$, ensuring that $\hat\theta$ is a stationary point of the cost function. Observe that (14) has the structure of an IV estimator. Also, note that replacing $\mathbf{M}$ by $\mathbf{X}$ would give the Sanathanan-Koerner iteration for the given model set.

C. Iterative procedure for models in R-MFD

Analogous to the analysis in the previous section, we will now use equation (6) to derive an iterative linear regression algorithm for systems in R-MFD . To that end, we introduce a different notation for Θ, i.e.

$$\Theta^T = \left[\begin{array}{cccccc} A_{n_a-1}^T & \dots & A_0^T & B_{n_b}^T & \dots & B_0^T\end{array}\right]$$
$$\theta = \mathrm{vec}(\Theta). \tag{15}$$

With this, we give the following two propositions. The proofs of these are similar to those of the corresponding propositions for the L-MFD case, and are omited for reasons of space.

Proposition 4.3: With $E(\xi(\omega_k),\theta)$ as defined in (2), where the model is represented in R-MFD, and with θ as in defined (15), the following identity holds

$$\mathrm{vec}\left[E(\omega_k,\theta)\right] = Y_k(\theta) - X_k(\theta)\theta \tag{16}$$

with

$$\begin{aligned}Y_k(\theta) &= \left[A^{-T}(\xi(\omega_k),\theta)\otimes I\right]\mathrm{vec}\left(\xi(\omega_k)^{n_a}G(\omega_k)\right)\\ X_k(\theta) &= \left[A^{-T}(\xi(\omega_k),\theta)\otimes\Omega(\omega_k)\right]\end{aligned}$$

where

$$\Omega(\omega_k) = \left[\begin{array}{ccc} -\xi(\omega_k)^{n_a-1}G(\omega_k) & \dots & -\xi(\omega_k)^0G(\omega_k)\\ \xi(\omega_k)^{n_b}I_{p\times p} & \dots & \xi(\omega_k)^0 I_{p\times p}\end{array}\right].$$

Proposition 4.4: For models parametrized in R-MFD,

$$M_k(\theta) = A^{-T}(\xi(\omega_k),\theta)\otimes\Phi_k(\theta) \tag{17}$$

with

$$\Phi_k(\theta) = \left[\begin{array}{ccc} -\xi(\omega_k)^{n_a-1}P(\xi(\omega_k),\theta) & \dots & -\xi(\omega_k)^0P(\xi(\omega_k),\theta)\\ \xi(\omega_k)^{n_b}I_{p\times p} & \dots & \xi(\omega_k)^0 I_{p\times p}\end{array}\right].$$

With (16) and (17), we can rewrite (6) also for models in R-MFD to (11), albeit with different definitions of the matrices Y_k, X_k and M_k. Hence, by applying these definitions, a similar iterative algorithm can be followed as the one derived for systems in L-MFD in the previous section.

D. Minimization of weighted OE cost criteria

The algorithms in the previous sections were derived for the cost function based on the unweighted error. Here we will show how these results can be generalized for the case input/output weighting or Schur weighting is applied.

1) Input-output weighting: Observe that with the input/output weighted error $E_{i/o}(\omega_k,\theta)$ as defined in (4), setting the first derivative of the cost function to zero yields the equality

$$\sum_{k=1}^{N} Re\left\{M_{i/o,k}^{H}(\theta)\text{vec}\left[W_o(\omega_k)E(\omega_k,\theta)W_i(\omega)\right]\right\}=0 \tag{18}$$

where $M_{i/o,k}(\theta)=\frac{\partial}{\partial\theta}\text{vec}[W_o(\omega_k)P(\xi(\omega_k),\theta)W_i(\omega_k)]$. Using the identity $\text{vec}(ABC)=(C^T\otimes A)\text{vec}(B)$, we derive that

$$\begin{aligned}&\text{vec}[W_o(\omega_k)E(\omega_k,\theta)W_i(\omega_k)]\\&\qquad\qquad=[W_i^T(\omega_k)\otimes W_o(\omega_k)]\text{vec}[E(\omega_k,\theta)]\end{aligned}$$

and

$$\begin{aligned}M_{i/o,k}(\theta)&=\frac{\partial}{\partial\theta}[W_i^T(\omega_k)\otimes W_o(\omega_k)]\text{vec}[P(\xi(\omega_k),\theta)]\\&=[W_i^T(\omega_k)\otimes W_o(\omega_k)]M_k(\theta).\end{aligned}$$

By substituting these identities in (18), and using the expressions derived for $\text{vec}[E(\omega_k,\theta)]$ and $M_k(\omega_k,\theta)$, iterative algorithms to minimize the input/output weighted cost can be obtained in the same fashion as derived above.

2) Schur weighting: Similarly as for input/output weighting, note that with the Schur-weighted error $E_s(\omega_k,\theta)$ as defined in (5), $\frac{\partial}{\partial\theta}V(\theta)=0$ implies

$$\sum_{k=1}^{N} Re\left\{M_{s,k}^{H}(\theta)\text{vec}\left[W_s(\omega_k).*E(\omega_k,\theta)\right]\right\}=0 \tag{19}$$

where $M_{s,k}(\theta)=\frac{\partial}{\partial\theta}\text{vec}[W_s(\omega_k).*P(\xi(\omega_k),\theta)]$. We derive that

$$\text{vec}[W_s(\omega_k).*E(\omega_k,\theta)]=\text{vec}[W_s(\omega_k)].*\text{vec}[E(\omega_k,\theta)]$$

and

$$\begin{aligned}M_{s,k}(\theta)&=\frac{\partial}{\partial\theta}\text{vec}[W_s(\omega_k)].*\text{vec}[P(\xi(\omega_k),\theta)]\\&=\text{vec}[W_s(\omega_k)].*M_k(\theta).\end{aligned}$$

Again, substitution of these identities in (19), in conjuction with the derived expressions for $\text{vec}[E(\omega_k,\theta)]$ and $M_k(\omega_k,\theta)$, allows to derive iterative algorithms that minimize the Schur weighted cost upon convergence.

E. Estimation of common denominator models

The algorithm that is described in this paper, ensures that converging iterations imply an optimal estimate of the parameters in a matrix fraction representation is obtained. Here we will demonstrate that this property can also be obtained for model representations with a common denominator.

For this, let the model set be defined by $P(\xi,\theta)=B(\xi,\theta)A^{-1}(\xi,\theta)$, where $B(\xi,\theta)$ is as defined before, and $A(\xi,\theta)=I\cdot a(\xi,\theta)$ with $a(\xi,\theta)$ a scalar polynomial. Estimation of the parameters in this represention can be reformulated to estimation of the parameters in a representation that matches the fully parametrized matrix fraction representation of section III. For this we will show that there exists $G_v(\omega_k)$, $B_v(\xi,\theta)$ and $A_v(\xi,\theta)$, with $B_v(\xi,\theta)$ and $A_v(\xi,\theta)$ fully parametrized, such that

$$\|E(\omega_k,\theta)\|_F^2=\|E_v(\omega_k,\theta)\|_F^2 \tag{20}$$

where $E_v(\omega_k,\theta)=G_v(\omega_k)-B_v(\xi(\omega_k),\theta)A_v^{-1}(\xi(\omega_k),\theta)$. Indeed note that for models with a common denominator, we can write

$$\begin{aligned}&\|E(\omega_k,\theta)\|_F^2\\&=\left|\text{vec}[G(\omega_k)]-\text{vec}[B(\xi(\omega_k),\theta)]\cdot a^{-1}(\xi(\omega_k),\theta)\right|^2.\end{aligned}$$

With this it follows that if $G_v(\omega_k)=\text{vec}[G(\omega_k)]$, $B_v(\xi,\theta)=B_{v,n_b}\xi^{n_b}+\cdots+B_{v,0}$, where $B_{v,i}\in\mathbb{R}^{p\cdot m\times 1}$, $i=0\ldots n_b$ and $B_{v,i}=\text{vec}(B_i)$, and $A_v(\xi,\theta)=a(\xi,\theta)$, identity (20) will hold. Once having obtained estimates $\hat{B}_{v,i}$, we can directly construct estimates $\hat{B}_i$ using the identity $B_{v,i}=\text{vec}(B_i)$.

F. Extension for model sets with non-full parametrization

Until now, it was assumed that polynomial matrices $A(\xi,\theta)$ and $B(\xi,\theta)$ are fully parametrized. However, it is straightforward to deal with model descriptions for non-full parametrizations. Observe that following the approach in the previous sections for a non-full parameterization would result in a parameter vector θ with one or more zero elements. Deletion of these elements from θ, as well as deletion of the corresponding columns from the matrices M_k and X_k, yields the desired result.

V. Results

The algorithm has been applied for estimation of a parametric model of a micro-milling spindle with Active Magnetic Bearings (more details on the application and the setting can be found in e.g. [10], [11]). The radial dynamics of an AMB spindle constitute a 4×4 MIMO system, where the inputs represent the currents through the electromagnetic coils and the outputs the displacement of the rotor shaft at the location of the bearings. Here the objective is to find an accurate model of the resonant behavior of the spindle.

Noisy FRF data at a frequency grid consisting of $N=312$ points ranging from 1.2 to 4.4 kHz was available. A model set in R-MFD with $n_a=5$ and $n_b=4$ was selected, resulting in a total number of 128 parameters to be estimated. Schur weighting was applied, where for $W_s(\omega_k)$ the inverse of the estimate of the standard deviation of the FRF was used. The iterations were initiated by a least squares estimate. After 7 iterations the algorithm converged, with final cost of 2.94. For comparison, also a model was estimated using SK iterations. These converged after 27 iterations, yielding a final cost of 4.50. We note that during the SK iterations, some of the intermediate estimates had a lower cost than the final estimate (the minimum obtained cost was 4.29). In contrast to this, the final estimate obtained with the IV-based method had the least cost. These results confirm that the IV-based method outperforms the method of Sanathanan and Koerner.

In figure 1 the results with the IV-based method are depicted. The frequency reponse of the estimated model shows very high correspondence to the dataset. The dynamics including the ill-damped resonances are estimated correctly. Moreover, the applied weighting effectively avoids modeling errors due to large variance errors in the FRF data. Here such variance errors are particularly present around the harmonics of the rotational frequency of the spindle ($80,000$ rpm).

VI. Conclusions

Iterative linear regression algorithms are given for estimation of OE models in left or right matrix fraction description from frequency response data. These algorithms are extensions of the SISO IV-based linear regression algorithm, which has the property that convergence implies a stationary point of the cost function is reached. This property, in combination with the freedom in the definition of the model set and the possibility to incorporate pre, post or element-wise multivariable frequency weighting, makes this an attractive approach for MIMO frequency domain identification of OE models. It is not claimed that the method discussed here can outperform gradient-based optimization methods. However, it appears to be a favorable alternative for the classically applied SK-iterations. Application of the approach to estimation of a model of a spindle with Active Magnetic Bearings demonstrates this.

VII. Acknowledgement

The authors thankfully acknowledge Prof. R.A. de Callafon of the Department of Mechanical and Aerospace Engineering, University of California, San Diego, U.S.A. for making available the LTSFIT toolbox, allowing implementation of the algorithm described in this paper.

VIII. Proofs

In this appendix we give the proof of proposition 4.2. We will first give the following lemma.

Lemma 8.1: Suppose $A(\theta)$ is a complex square matrix, depending on the complex vector θ, which takes values in an open subset $V \subseteq \mathbb{C}^l$. Furthermore, suppose that $A(\theta)$ is invertible and analytical for all $\theta \in V$. Then for $\theta \in V$

$$\frac{d}{d\theta}\mathrm{vec}[A^{-1}(\theta)] = [-A^{-T}(\theta) \otimes A^{-1}(\theta)]\frac{d}{d\theta}\mathrm{vec}[A(\theta)]. \tag{21}$$

Proof: Let θ_i be the i^{th} element of θ. Using the product rule for matrix differentation, we derive that

$$\frac{\partial}{\partial\theta_i}(A(\theta)\cdot A^{-1}(\theta)) = \frac{\partial A(\theta)}{\partial\theta_i}A^{-1}(\theta) + A(\theta)\frac{\partial A^{-1}(\theta)}{\partial\theta_i} = 0$$

from which immediately follows that

$$\frac{\partial A^{-1}(\theta)}{\partial\theta_i} = -A^{-1}(\theta)\frac{\partial A(\theta)}{\partial\theta_i}A^{-1}(\theta).$$

Using the identity $\mathrm{vec}(ABC) = (C^T \otimes A)\mathrm{vec}(B)$, we infer

$$\frac{\partial}{\partial\theta_i}\mathrm{vec}[A^{-1}(\theta)] = -[A^{-T}(\theta) \otimes A^{-1}(\theta)]\mathrm{vec}[\frac{\partial A(\theta)}{\partial\theta_i}],$$

implying

$$\frac{d}{d\theta}\mathrm{vec}[A^{-1}(\theta)] = -[A^{-T}(\theta) \otimes A^{-1}(\theta)]\frac{d}{d\theta}\mathrm{vec}[A(\theta)].$$

This proofs the claim. ■

A. Proof of proposition 4.2

We will first introduce the notation $\theta^T = \begin{bmatrix} \theta_A^T & \theta_B^T \end{bmatrix}$ where θ_A only contains the parameters used to define $A(\xi,\theta)$ and θ_B those to define $B(\xi,\theta)$. For brevity, we will drop the dependency of P, B and A on $\xi(\omega_k)$ from here on. Observe that with this, we can write

$$M_k(\theta) = \begin{bmatrix} \frac{\partial}{\partial\theta_A}\mathrm{vec}[P(\theta)] & \frac{\partial}{\partial\theta_B}\mathrm{vec}[P(\theta)] \end{bmatrix}$$

Using the identities (9a) and (9b), we rewrite this to

$$M_k(\theta) = \Big[\left[B^T(\theta)\otimes I\right]\frac{\partial}{\partial\theta_A}\mathrm{vec}[A^{-1}(\theta)] \quad \left[I\otimes A^{-1}(\theta)\right]\frac{\partial}{\partial\theta_B}\mathrm{vec}[B(\theta)]\Big]. \tag{22}$$

We will derive expressions for $\frac{\partial}{\partial\theta_A}\mathrm{vec}[A^{-1}(\theta)]$ and $\frac{\partial}{\partial\theta_B}\mathrm{vec}[B(\theta)]$. Application of lemma 8.1 yields

$$\begin{aligned}
\frac{\partial}{\partial\theta_A}\mathrm{vec}[A^{-1}(\theta)] &= \left[-A^{-T}(\theta)\otimes A^{-1}(\theta)\right]\frac{\partial}{\partial\theta_A}\mathrm{vec}[A(\theta)] \\
&= \left[-A^{-T}(\theta)\otimes A^{-1}(\theta)\right]\cdot \\
&\quad \cdot \begin{bmatrix} I_{pp\times pp}\xi(\omega_k)^{n_a-1} & \dots & I_{pp\times pp}\xi(\omega_k)^0 \end{bmatrix} \\
&= \left[-A^{-T}(\theta)\otimes A^{-1}(\theta)\right]\cdot \\
&\quad \cdot \left(\begin{bmatrix} I_{p\times p}\xi(\omega_k)^{n_a-1} & \dots & I_{p\times p}\xi(\omega_k)^0 \end{bmatrix}\otimes I_{p\times p}\right).
\end{aligned}$$

Furthermore, we derive that

$$\begin{aligned}
\frac{\partial}{\partial\theta_B}\mathrm{vec}[B(\theta] &= \begin{bmatrix} I_{mp\times mp}\xi(\omega_k)^{n_b} & \dots & I_{mp\times mp}\xi(\omega_k)^0 \end{bmatrix} \\
&= \begin{bmatrix} I_{m\times m}\xi(\omega_k)^{n_b} & \dots & I_{m\times m}\xi(\omega_k)^0 \end{bmatrix}\otimes I_{p\times p}.
\end{aligned}$$

With this, we express the first element of $M_k(\theta)$ in (22) as

$$\begin{aligned}
&\left[-B^T(\theta)\otimes I\right]\left[-A^{-T}(\theta)\otimes A^{-1}(\theta)\right]\cdot \\
&\quad \cdot\left(\begin{bmatrix} I_{p\times p}(\xi(\omega_k)^{n_a-1} & \dots & I_{p\times p}\xi(\omega_k)^0 \end{bmatrix}\otimes I_{p\times p}\right) \\
&= \begin{bmatrix} -P^T(\theta)\xi(\omega_k)^{n_a-1} & \dots & -P^T(\theta)\xi(\omega_k)^0 \end{bmatrix}\otimes A^{-1}(\theta),
\end{aligned}$$

and the second element of $M_k(\theta)$ in (22) as

$$\begin{aligned}
&\left[I\otimes A^{-1}(\theta)\right]\cdot \\
&\quad \cdot\left(\begin{bmatrix} I_{m\times m}\xi(\omega_k)^{n_b} & \dots & I_{m\times m}\xi(\omega_k)^0 \end{bmatrix}\otimes I_{p\times p}\right) \\
&= \begin{bmatrix} I_{m\times m}\xi(\omega_k)^{n_b} & \dots & I_{m\times m}\xi(\omega_k)^0 \end{bmatrix}\otimes A^{-1}(\theta).
\end{aligned}$$

Combining these results, and using the identity $\begin{bmatrix} A_1\otimes B & A_2\otimes B \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \end{bmatrix}\otimes B$, we infer

$$M_k(\theta) = \Phi_k(\theta)^T \otimes A^{-1}(\theta), \tag{23}$$

which proves the claim.

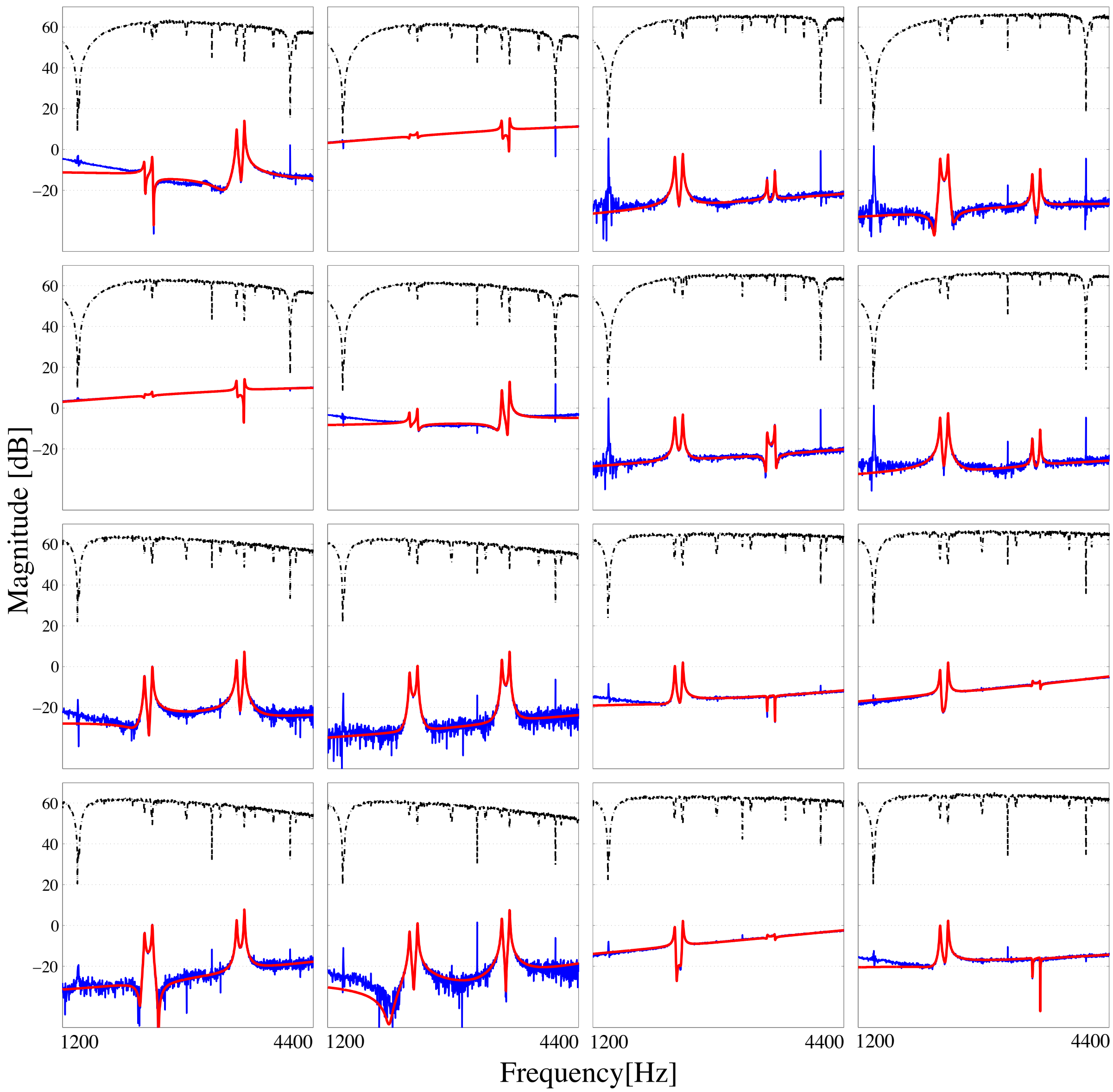

Fig. 1. Results with the AMB spindle system: amplitude plot of the FRF data $G(\omega_k)$ (blue), the Schur weighting $W_s(\omega_k)$ (black dash-dotted) and the estimated model $P(\xi(\omega_k), \hat{\theta})$ (red).

REFERENCES

[1] S. G. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *IEEE Trans. Autom. Control*, pp. 56–58, 1963.

[2] D. S. Bayard, "High-order multivariable transfer function curve fitting: Algorithms, sparse matrix methods and experimental results," *Automatica*, vol. 30, no. 9, pp. 1439–1444, 1994.

[3] P. Verboven, P. Guillaume, and B. Cauberghe, "Multivariable frequency–response curve fitting with application to modal parameter estimation," *Automatica*, vol. 41, no. 10, pp. 1773–1782, 2005.

[4] R. A. de Callafon, D. de Roover, and P. M. J. Van den Hof, "Multivariable least squares frequency domain identification using polynomial matrix fraction descriptions," *Proc. 35th IEEE Conference on Decision and Control, Kobe, Japan*, vol. 2, pp. 2030–2035, 1996.

[5] S. Gaikwad and D. Rivera, "Multivariable frequency-response curve fitting with application to control-relevant parameter estimation," *Automatica*, vol. 33, no. 6, pp. 1169–1174, 1997.

[6] T. McKelvey, H. Akcay, and L. Ljung, "Subspace-based multivariable system identification from frequency response data," *IEEE Trans. Autom. Control*, vol. 41, no. 7, pp. 960–979, 1996.

[7] P. Young, "Some observations on instrumental variable methods of time-series analysis." *Int. J. Control*, vol. 23, no. 5, pp. 593–612, 1976.

[8] P. Young, H. Garnier, and M. Gilson, *Identification of Continuous-time Models from Sampled Data*. London, UK: Springer, 2008, ch. Refined instrumental variable identification of continuous-time hybrid Box-Jenkins models, pp. 91–131.

[9] P. M. J. Van den Hof and S. G. Douma, "An IV-based iterative linear regression algorithm with optimal output error properties," *DCSC Technical report nr. 09-018, http://www.dcsc.tudelft.nl/~pvandenhof/Reportfiles/StM2OE_Report.pdf, Delft University of Technology*, 2008.

[10] R. S. Blom, A. M. Hoogstrate, P. M. J. Van den Hof, and H. H. Langen, "Indirect force measurement of micro-milling with an active magnetic bearing spindle," *Proc. 10th CIRP international workshop on Modeling of Machining Operations, Reggio Calabria, Italy, August 27-28*, pp. 525–532, 2007.

[11] R. S. Blom and P. M. J. Van den Hof, "Estimating cutting forces in micromilling by input estimation from closed-loop data," *Proc. 17th IFAC WC, Seoul, Korea, July 6–11*, pp. 468–473, 2008.