Numerically robust frequency domain identification of multivariable systems

R. Pintelon*, Y. Rolain*, A. Bultheel[#] and M. Van Barel[#]
*Vrije Universiteit Brussel, dept. ELEC, Pleinlaan 2, 1050 Brussels, Belgium, Rik.Pintelon@vub.ac.be
[#]Katholieke Universiteit Leuven, dept. CS, Celestijnenlaan 200A, 3001 Heverlee, Belgium

Abstract

Frequency domain methods are known to suffer from a poor numerical conditioning when the frequency span and/or the model order become large (more than 2 decades and an order more than 20). This numerical problem ruins both the modeling performance and the model order selection capability. This paper presents a modeling approach that leads to (almost) optimal numerical conditioning of the normal equations (condition number = 1). The key idea is to expand the numerator and denominator of the transfer function model in a vector orthogonal basis. The theory is illustrated on a real measurement example.

1 Introduction

Several attempts have been made in the past to circumvent numerical degeneracy in frequency domain system identification. These range from an appropriate scaling of frequencies [1], to the decomposition of the numerator and denominator of the model equation in a separate basis of polynomials orthogonal w.r.t. the inner product defined by the normal equations of the estimator [1], [2]. Even if these approaches increase the limit on bandwidth and complexity, they do not remove numerical problems totally, neither do they extrapolate gracefully to multiple input multiple output (MIMO) systems. The approach proposed here consists in expanding the numerator and denominator polynomials in a vector orthogonal polynomial basis [3], [4]. It solves the numerical conditioning problem perfectly (condition number = 1) for all frequency domain identification methods whose cost function can be reduced to a weighted linear least squares problem or a weighted (generalised) total least squares problem. For the nonlinear least squares and the maximum likelihood estimators [1] we use the vector polynomial basis of the starting values (weighted linear least squares or weighted generalised total least squares) in each step of the Newton-Gauss iteration. This implies that the condition number will be somewhat larger than 1.

In [5] the vector orthogonal polynomial modeling approach is applied to single input, single output (SISO) systems. The main contribution of this paper is the extension of the method to multivariable systems.

The outline of the paper is as follows. First, it is shown that a transfer function model can be parametrized in vector polynomials (Section 2). Next, Section 3 tackles the problem of the construction of a vector orthogonal polynomial basis that leads to the (almost) best conditioned normal equations. Finally, the theory is illustrated on a multivariable (two input, two output) modal analysis problem (see Section 4). Although the theory is elaborated for continuous-time systems (s-domain) the presented results also apply to discrete-time modeling (z-domain).

2 Parametrization transfer function model in vector polynomials

Consider the transfer function model $G(s, \theta)$ of a SISO continuous-time system

$$G(s, \theta) = \frac{B(s, \theta)}{A(s, \theta)} = \frac{\sum_{r=0}^{n_b} b_r s^r}{\sum_{r=0}^{n_a} a_r s^r} \qquad (1)$$

with s the Laplace variable, $A(s, \theta)$ and $B(s, \theta)$ the denominator and numerator polynomials of order n_a and n_b respectively, and $\theta = [a_0 \ldots a_{n_a} b_0 \ldots b_{n_b}]^T$ the model parameters. Stacking $A(s)$ and $B(s)$ on top of each other defines a 2 by 1 vector polynomial $Q(s)$

$$Q(s) = \begin{bmatrix} A(s) \\ B(s) \end{bmatrix} = \begin{bmatrix} \sum_{r=0}^{n_a} a_r s^r \\ \sum_{r=0}^{n_b} b_r s^r \end{bmatrix} \qquad (2)$$

A (non-orthogonal) basis for the vector polynomial $Q(s)$ is, for example,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} s \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} s^{n_a} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ s \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ s^{n_b} \end{bmatrix} \qquad (3)$$

Using the basis (3), $Q(s)$ can be written as

$$Q(s) = \begin{bmatrix} A(s) \\ B(s) \end{bmatrix} = \sum_{r=0}^{n_a + n_b + 1} \lambda_r \begin{bmatrix} q_r(s) \\ p_r(s) \end{bmatrix} \qquad (4)$$

with $\lambda_r = a_r$ for $r = 0, 1, \ldots, n_a$, $\lambda_r = b_{r-n_a-1}$ for $r = n_a+1, \ldots, n_a+n_b+1$, and $[q_r(s)\ p_r(s)]^T$ the rth vector of the basis. Eq. (4) allows to parametrize transfer function model (1) as

$$G(s, \theta) = \frac{B(s, \theta)}{A(s, \theta)} = \frac{\sum_{r=0}^{n_a+n_b+1} \lambda_r p_r(s)}{\sum_{r=0}^{n_a+n_b+1} \lambda_r q_r(s)} \quad (5)$$

where $\theta = [\lambda_0\ \lambda_1 \ldots \lambda_{n_a+n_b+1}]^T$. In Section 3 it is explained how an orthogonal basis can be constructed for $Q(s)$, that minimizes the condition number of the normal equations.

Exactly the same reasoning can be applied to the n_y by n_u transfer function matrix $G(s, \theta)$ of a multivariable system with n_u inputs and n_y outputs. Indeed, consider, for example, a left matrix fraction description (LMFD)

$$G(s, \theta) = A^{-1}(s, \theta)B(s, \theta) \quad (6)$$

where $A(s, \theta)$ and $B(s, \theta)$ are n_y by n_y and n_y by n_u matrix polynomials respectively

$$A(s, \theta) = \sum_{r=0}^{n_a} A_r s^r \text{ and } B(s, \theta) = \sum_{r=0}^{n_b} B_r s^r \quad (7)$$

with $A_r \in \mathbb{R}^{n_y \times n_y}$ and $B_r \in \mathbb{R}^{n_y \times n_u}$ (see [7]). Following the lines of the SISO case, a vector polynomial basis can be constructed for the $(n_y^2 + n_u n_y)$ by 1 vector polynomial

$$Q(s) = \begin{bmatrix} \text{vec}(A(s)) \\ \text{vec}(B(s)) \end{bmatrix} \quad (8)$$

where the operator vec( ) stacks the columns of a matrix on top of each other, and the transfer function $G(s, \theta)$ (6) can be parametrized in that basis.

3 Construction of a vector orthogonal polynomial basis

For simplicity the idea is illustrated first on a weighted linear least squares estimator for SISO systems. Next, it is shown how the idea can be extended to weighted nonlinear least squares estimators for multivariable systems.

3.1 Weighted linear least squares

The weighted linear least squares estimator of the SISO model (1) minimizes

$$\sum_{k=1}^{F} w^2(k) \left| e(s_k, \theta) \right|^2 \quad (9)$$

w.r.t. θ, where

$$e(s_k, \theta) = A(s_k, \theta)Y(k) - B(s_k, \theta)U(k) \quad (10)$$

with $w(k)$ a real weighting function, $U(k)$ and $Y(k)$ the input and output DFT spectra, and F the number of frequencies (see [1]). Using

$$e(s_k, \theta) = \begin{bmatrix} Y(k) & -U(k) \end{bmatrix} \begin{bmatrix} A(s_k, \theta) \\ B(s_k, \theta) \end{bmatrix} \quad (11)$$

cost function (9) can be written as

$$\sum_{k=1}^{F} Q^H(s_k, \theta) W(k) Q(s_k, \theta) \quad (12)$$

where Q is defined in (4) and with

$$W(k) = w^2(k) \begin{bmatrix} |Y(k)|^2 & -\overline{Y}(k)U(k) \\ -Y(k)\overline{U}(k) & |U(k)|^2 \end{bmatrix} \quad (13)$$

a positive (semi-)definite matrix. Clearly (12) defines an inner product for the real 2 by 1 vector polynomials $P(s)$ and $R(s)$

$$\langle P(s), R(s) \rangle = \text{Re}\left(\sum_{k=1}^{F} R^H(s_k) W(k) P(s_k)\right) \quad (14)$$

where Re( ) denotes the real part of. Using the inner product one can calculate an orthogonal basis

$$Q_r(s) = \begin{bmatrix} q_r(s) \\ p_r(s) \end{bmatrix}, \ r = 0, 1, \ldots, n_a + n_b + 1 \quad (15)$$

for the 2 by 1 vector polynomials $Q(s)$ (4) via a Gram-Schmidt procedure [8]. A numerical stable and time efficient implementation of this procedure can be found in [3], [4].

Expanding $Q(s)$ (4) in the vector orthogonal basis (15)

$$\begin{bmatrix} A(s, \theta) \\ B(s, \theta) \end{bmatrix} = \sum_{r=0}^{n_a+n_b+1} \lambda_r Q_r(s) = \sum_{r=0}^{n_a+n_b+1} \lambda_r \begin{bmatrix} q_r(s) \\ p_r(s) \end{bmatrix} \quad (16)$$

the $[k, r]$th entry of the Jacobian matrix of the weighted residual $w(k)e(s_k, \theta)$ (10) equals

$$J_{[k,r]} = w(k)\frac{\partial e(s_k, \theta)}{\partial \lambda_r} = w(k) \begin{bmatrix} Y(k) & -U(k) \end{bmatrix} Q_r(s_k) \quad (17)$$

(use eq. (11) and (16)). Using (17) it can easily be verified that the $[n, m]$th entry of the matrix of the normal equation is given by

$$\text{Re}((J^H J)_{[n,m]}) = \langle Q_m(s), Q_n(s) \rangle = \delta_{nm} \quad (18)$$

with δ_{nm} the Kronecker delta. Eq. (18) proves that the normal equations are best conditioned (condition number = 1).

Using (14), (16) and (18), it is easy to verify that cost function (9) can be written as

$$\sum_{r=0}^{n_a+n_b+1} \lambda_r^2 \quad (19)$$

Finally, (19) should be minimized w.r.t. to $\theta = [\lambda_0 \lambda_1 \ldots]^T$. To avoid the trivial solution $\theta = 0$, the parameter vector must be constrained. For example, if $n_b < n_a$ and a solution where $A(s, \theta)$ has strict order n_a is sought ($\lambda_{n_a+n_b+1} = 1$), then

$$\lambda_0 = \lambda_1 = \ldots = \lambda_{n_a+n_b} = 0$$

minimizes (19) and the solution is given by $Q_{n_a+n_b+1}(s)$. Due to the orthogonality, the other vector polynomials $Q_r(s)$, $r = 0, 1, \ldots, n_a+n_b$, represent the lower order weighted linear least squares solutions.

Note that the orthogonalization procedure can also be applied to iteratively weighted linear least squares estimators like, for example, the method of Sanathanan and Koerner [6], where the weighting function $w(k)$ depends on a initial (previous) guess $\theta^{(i-1)}$ of the model parameters $w(k) = w(k, \theta^{(i-1)})$, giving

$$\sum_{k=1}^{F} w^2(k, \theta^{(i-1)}) \left| e(s_k, \theta^{(i)}) \right|^2 \qquad (20)$$

3.2 Weighted nonlinear least squares

Assuming that n_u experiments with different periodic excitation signals have been performed, the maximum likelihood (ML) estimator of the transfer function matrix (6) minimizes

$$\sum_{k=1}^{F} \mathrm{vec}^H(E(s_k, \theta)) C_{\mathrm{vec}E}^{-1}(s_k, \theta) \mathrm{vec}(E(s_k, \theta)) \qquad (21)$$

w.r.t. θ, with superscript H the hermitian transpose (transpose + complex conjugate), and where

$$\begin{aligned} E(s, \theta) &= A(s, \theta)Y(k) - B(s, \theta)U(k) \\ C_{\mathrm{vec}E}(s_k, \theta) &= \mathrm{Cov}(\mathrm{vec}(E(s_k, \theta))) \end{aligned} \qquad (22)$$

with $\mathrm{Cov}(\ )$ the covariance matrix (see [1]). $Y(k)$ and $U(k)$ are n_y by n_u and n_u by n_u matrices containing the output and input DFT spectra of n_u (independent) experiments (each column corresponds to an experiment with one particular choice of the n_u excitation signals, see also Section 4) at frequency k. Using

$$\begin{aligned} \mathrm{vec}(E) &= \left(\begin{bmatrix} Y \\ -U \end{bmatrix}^T \otimes I_{n_y} \right) \begin{bmatrix} \mathrm{vec}(A) \\ \mathrm{vec}(B) \end{bmatrix} \\ \mathrm{Cov}(\mathrm{vec}(E)) &= (I_{n_u} \otimes A) C_{\mathrm{vec}Y} (I_{n_u} \otimes A^H) \\ &\quad + (I_{n_u} \otimes B) C_{\mathrm{vec}U} (I_{n_u} \otimes B^H) \\ &\quad - 2\mathrm{herm}((I_{n_u} \otimes A) C_{\mathrm{vec}Y \mathrm{vec}U} (I_{n_u} \otimes B^H)) \end{aligned} \qquad (23)$$

with $\otimes$ the Kronecker product (see [9]), I_n the n by n identity matrix, $\mathrm{herm}(x) = (x + x^H)/2$, $C_{\mathrm{vec}x} = \mathrm{Cov}(\mathrm{vec}(x))$, $C_{\mathrm{vec}x\mathrm{vec}y} = \mathrm{Cov}(\mathrm{vec}(x), \mathrm{vec}(y))$ the ML cost function (21) can be written as

$$\sum_{k=1}^{F} Q^H(s_k, \theta) W(k, \theta) Q(s_k, \theta) \qquad (24)$$

with Q defined in (8) and

$$W = \left(\begin{bmatrix} \overline{Y} \\ -\overline{U} \end{bmatrix} \otimes I_{n_y} \right) C_{\mathrm{vec}E}^{-1} \left(\begin{bmatrix} Y \\ -U \end{bmatrix}^T \otimes I_{n_y} \right) \qquad (25)$$

where the overline denotes the complex conjugate. Cost function (24) is of the same form as (12) and, therefore, also defines an inner product. Since the weighting matrix $W(k, \theta)$ depends on the actual value of the model parameters θ, the vector orthogonal polynomial basis must be updated in each iteration step of the Newton-Gauss minimization procedure. The transformation of the model parameters from the old orthogonal basis (step i) to the new orthogonal basis (step $i+1$) is, however, numerically ill conditioned. Therefore, instead of updating the orthogonal basis we use the vector orthogonal basis of the starting values in each step of the Newton-Gauss iteration. As a consequence the condition number of the normal equations (iterative Newton-Gauss equations) will not be one, and will grow during the iteration. Practical experience has shown that when the initial guess is good enough , the condition number of the maximum likelihood solution will not grow above 10.

As starting values the iterative quadratic maximum likelihood (IQML) can be used. The idea is to evaluate $C_{\mathrm{vec}E}^{-1}(s_k, \theta)$ in an initial guess $\theta^{(0)}$ of the model parameters, reducing (21) to the following weighted linear least squares problem

$$\sum_{k=1}^{F} \mathrm{vec}^H(E(s_k, \theta)) C_{\mathrm{vec}E}^{-1}(s_k, \theta^{(0)}) \mathrm{vec}(E(s_k, \theta)) \qquad (26)$$

The linear least squares estimator which minimizes

$$\sum_{k=1}^{F} \mathrm{vec}^H(E(s_k, \theta)) \mathrm{vec}(E(s_k, \theta)) \qquad (27)$$

w.r.t. θ is taken as initial guess $\theta^{(0)}$. The improved estimate $\theta^{(1)}$ (minimizer of (26)) can be used to calculate an improved weighting $C_{\mathrm{vec}E}^{-1}(s_k, \theta^{(1)})$ and so on and so forth leading to the following iteratively weighted linear least squares estimator

$$\sum_{k=1}^{F} \mathrm{vec}^H(E(s_k, \theta^{(i)})) C_{\mathrm{vec}E}^{-1}(s_k, \theta^{(i-1)}) \mathrm{vec}(E(s_k, \theta^{(i)}))$$

It defines an inner product for the real $(n_y^2 + n_u n_y)$ by 1 vector polynomials $P(s)$ and $R(s)$

$$\langle P(s), R(s) \rangle = \mathrm{Re}\left(\sum_{k=1}^{F} R^H(s_k) W(k, \theta^{(i-1)}) P(s_k) \right) \qquad (28)$$

Finally, the vector orthogonal polynomial basis calculated with inner product (28) is used in the Newton-Gauss minimization of the ML estimator (21).

Another possibility for generating high quality starting values (and the corresponding vector orthogonal polynomial basis) are the iteratively weighted generalized total least squares algorithms [10].

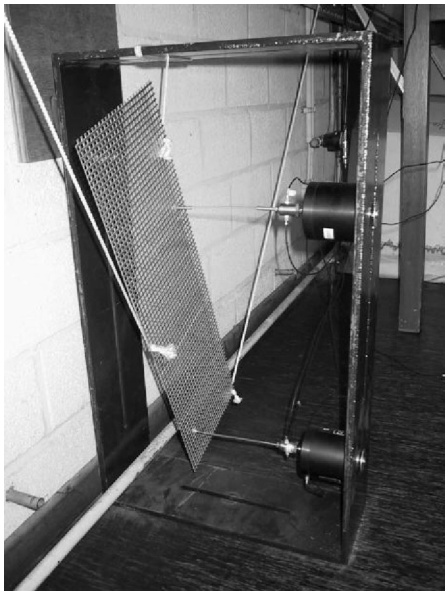

Figure 1: Experimental set up: the aluminium plate hung by three nylon threads is excited by two mini-shakers via plexi-glass stinger rods.

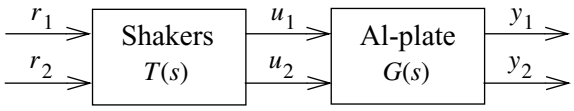

Figure 2: Block diagram of the measurement set up shown in Figure 1: r_1, r_2 are the excitation signals stored in the arbitrary function generators, u_1, u_2 the forces applied by the shakers to the Al-plate, and y_1, y_2 the observed accelerations of the Al-plate. $G(s)$ and $T(s)$ are the 2 by 2 transfer function matrices of the Al-plate and shakers respectively. The rms values of the measured forces and accelerations is about 70 mV .

Similar to the SISO case the parameter vector θ should be constrained to avoid the trivial solution $\theta = 0$. For example, one can choose the solution where the diagonal entries of $A(s, \theta)$ have strict order n_a and the non-diagonal entries have strict order less than n_a, which corresponds to the case $A_{n_a} = I_{n_y}$ in (7). This solution is obtained by setting appropriate lambda values to one and zero.

4 Measurement example

4.1 Measurement set up

An aluminium plate (185mm x 63mm x 1.5mm) hung by three nylon threads is excited by two mini-shakers (B&K 4810) via plexi-glass stinger rods (see Figures 1 and 2). The forces (u_1, u_2 in Figure 2) and accelerations (y_1, y_2 in Figure 2) at the excitations points are measured with impedance heads (B&K 8001). These signals are amplified (charge amplifier B&K 2635) and buffered ($Z_{\text{in}} > 5\ \text{M}\Omega$, $Z_{\text{out}} = 50\ \Omega$) before being applied to the acquisition channels (HPE 1430A, $Z_{\text{in}} = 50\ \Omega$) of the VXI measurement device. The periodic excitation signals (r_1, r_2 in Figure 2) are generated by two arbitrary function generators (HPE 1445A, $Z_{\text{out}} = 50\ \Omega$) at a sampling frequency $f_s = 10\ \text{MHz}/2^{12} \approx 2.44\ \text{kHz}$. The output of the arbitrary function generators is lowpass filtered (7th order inverse Chebyshev filter with a cut off frequency of 1 kHz) before being applied to the mini-shaker. To reduce the effect of the inductive impedance of the shaker on the generator unit, an $18\ \Omega/5\ \text{W}$ resistance in put is series with its input.

4.2 Measurement results

For each MIMO measurement two random phase multisines (= sum of harmonically related sinewaves with deterministic amplitude spectrum and random phase spectrum) with flat amplitude spectrum, interleaved frequency grids, and $N = 4096$ points per period are used as excitation signals. Expressed in harmonic numbers w.r.t. the frequency resolution $f_0 = f_s/N \approx 0.6\ \text{Hz}$ the interleaved frequency grids are

$$\begin{array}{ll} \text{grid 1} & 119{:}2{:}1637 \\ \text{grid 2} & 120{:}2{:}1638 \end{array} \tag{29}$$

which corresponds to $F = 1520$ frequencies in a frequency band of about $[71\ \text{Hz}, 976\ \text{Hz}]$.

Twenty five MIMO experiments with different realisations of the random phase multisines are performed. Each MIMO experiment consists of $n_u = 2$ MIMO measurements. Hence, four different random phase multisines r_{11} (frequency grid 1), r_{21} (frequency grid 2), r_{12} (frequency grid 2), and r_{22} (frequency grid 1) are calculated for each MIMO experiment. In the jth measurement ($j = 1, 2$) of each experiment signals r_{1j} and r_{2j} are applied to shakers 1 and 2 respectively. Hundred consecutive periods of the steady state response are observed, which allows to calculate the sample means and sample covariance matrices of the nth MIMO experiment (see [1], [11])

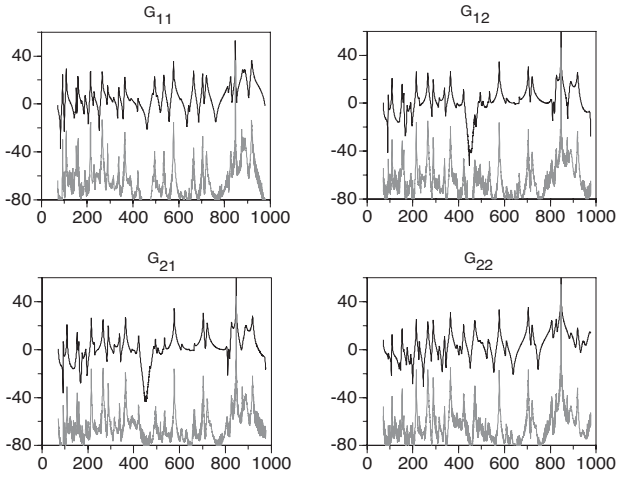

Figure 3: Measured frequency response matrix of one MIMO experiment: magnitude (black lines) and standard deviation (gray lines) in dB as a function of the frequency in Hz.

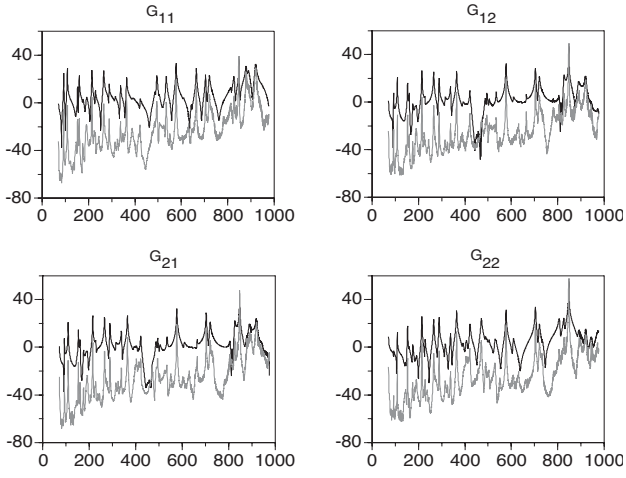

Figure 4: Measured frequency response matrix (average over all MIMO experiments): magnitude (black lines) and standard deviation (gray lines) in dB as a function of the frequency in Hz.

$$\hat{X}^{[n]}(k) = \begin{bmatrix} \hat{X}_{11}^{[n]}(k) & \hat{X}_{12}^{[n]}(k) \\ \hat{X}_{21}^{[n]}(k) & \hat{X}_{22}^{[n]}(k) \end{bmatrix}$$

$$\hat{X}_{ij}^{[n]}(k) = \frac{1}{M}\sum_{m=1}^{M} X_{ij}^{[n,m]}(k) \tag{30}$$

$$M\hat{C}_{XZ}^{[n]}(k) = \frac{1}{M-1}\sum_{m=1}^{M} \Delta X_{[:,j]}^{[n,m]}(\Delta Z_{[:,j]}^{[n,m]})^H$$

$$\Delta X_{[:,j]}^{[n,m]} = X_{[:,j]}^{[n,m]}(k) - \hat{X}_{[:,j]}(k)$$

with $X, Z = Y$ and/or U, $\hat{C}_{XX} = \hat{C}_X$, $M = 100$, and where, for example, $U_{12}^{[n,m]}(k)$ is the DFT spectrum of the mth period of the first input in the second measurement of the nth MIMO experiment. Since the $n_u = 2$ measurements are independent, the covariance matrices in (21) can be written as

$$\hat{C}_{\mathrm{vec}U} = I_{n_u} \otimes \hat{C}_U \tag{31}$$

and similarly for $\hat{C}_{\mathrm{vec}Y}$ and $\hat{C}_{\mathrm{vec}Y\mathrm{vec}U}$. Figure 3 shows the measured frequency response matrix

$$\hat{G}^{[n]}(j\omega_k) = \hat{Y}^{[n]}(k)(\hat{U}^{[n]}(k))^{-1} \tag{32}$$

together with its standard deviation calculated via

$$\hat{C}_{\mathrm{vec}\hat{G}^{[n]}} \approx C^{[n]} \otimes \hat{C}_Y^{[n]} + C^{[n]} \otimes (\hat{G}^{[n]}\hat{C}_U^{[n]}(\hat{G}^{[n]})^H) - 2\mathrm{herm}(C^{[n]} \otimes (\hat{C}_{YU}^{[n]}(\hat{G}^{[n]})^H)) \tag{33}$$

$$C^{[n]} = (\overline{\hat{U}^{[n]}(\hat{U}^{[n]})^H})^{-1}$$

(see Appendix I).

Referring the observed input/output DFT spectra to the exactly known reference signals r_1, r_2 (see Figure 2),

$$\begin{aligned} U_R^{[n]}(k) &= \hat{U}^{[n]}(k)(R^{[n]}(k))^{-1} \\ Y_R^{[n]}(k) &= \hat{Y}^{[n]}(k)(R^{[n]}(k))^{-1} \end{aligned} \tag{34}$$

where $\hat{U}^{[n]}(k)$, $\hat{Y}^{[n]}(k)$ and $R^{[n]}(k)$ are the mean (over the hundred periods) 2 by 2 input, output, and reference DFT spectra of the nth MIMO experiment (see eq. (30)), it is also possible to calculate the sample means and sample covariance matrices of the measurements over the twenty five MIMO experiments

$$\hat{X}(k) = \begin{bmatrix} \hat{X}_{11}(k) & \hat{X}_{12}(k) \\ \hat{X}_{21}(k) & \hat{X}_{22}(k) \end{bmatrix}$$

$$\hat{X}_{ij}(k) = \frac{1}{N}\sum_{n=1}^{N} \hat{X}_{ij}^{[n]}(k)$$

$$N\hat{C}_{xz}(k) = \frac{1}{N-1}\sum_{n=1}^{N} \Delta x^{[n]}(\Delta z^{[n]})^H \tag{35}$$

$$\Delta x^{[n]} = x^{[n]}(k) - \hat{x}(k)$$

$$x = \mathrm{vec}(X)$$

with $X, Z = Y_R$ and/or U_R, $\hat{C}_{xx} = \hat{C}_x$, $N = 25$, and where, for example, $\hat{Y}_{R22}^{[n]}(k)$ is the mean (over the hundred periods) DFT spectrum of the second output in the second measurement of the nth MIMO experiment Contrary to $\hat{U}^{[n]}(k)$ and $\hat{Y}^{[n]}(k)$, the columns of $\hat{U}_R(k)$ and $\hat{Y}_R(k)$ are not independently distributed, and, therefore, $\hat{C}_{\mathrm{vec}\hat{U}_R}(k)$, $\hat{C}_{\mathrm{vec}\hat{Y}_R}(k)$ and $\hat{C}_{\mathrm{vec}\hat{Y}_R\mathrm{vec}\hat{U}_R}(k)$ cannot be simplified as in (31). Figure 4 shows the measured FRM

$$\hat{G}(j\omega_k) = \hat{Y}_R(k)\hat{U}_R^{-1}(k) \tag{36}$$

together with its standard deviation calculated via

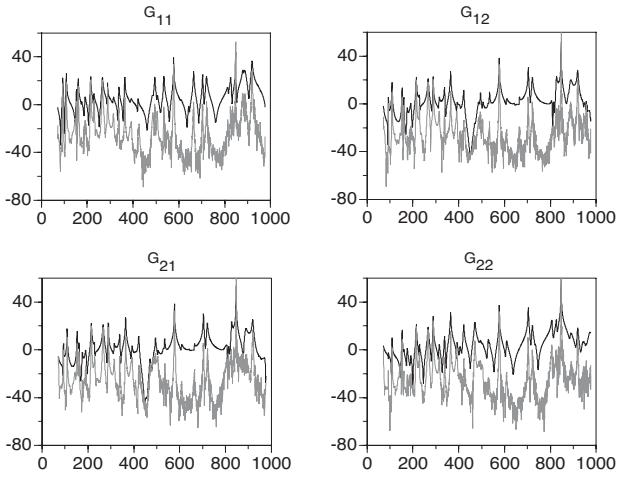

Figure 5: Estimated frequency response matrix (FRM) of the MIMO experiment shown in Figure 3: magnitude FRM (black lines) and difference between measured and modelled FRM (gray lines) in dB as a function of the frequency in Hz.

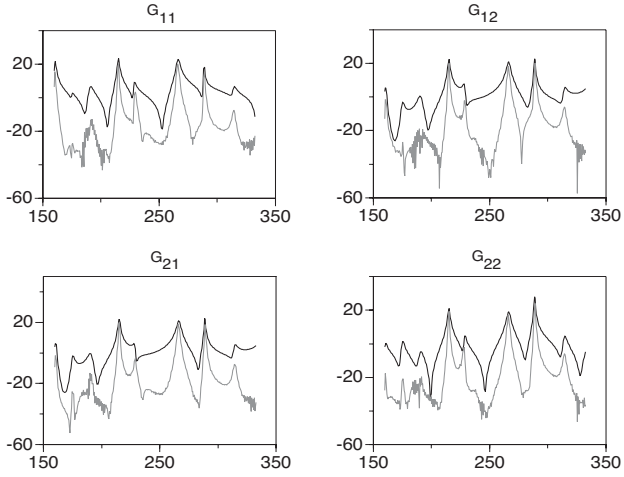

Figure 6: Zoom of Figure 5 in the band 150 Hz to 350 Hz: magnitude FRM (black lines) and difference between measured and modelled FRM (gray lines) in dB as a function of the frequency in Hz.

$$\hat{C}_{\mathrm{vec}\hat{G}} \approx (\hat{U}_R^{-T} \otimes I_{n_y}) C_{\mathrm{vec}Y_R} (\overline{\hat{U}_R^{-1}} \otimes I_{n_y}) + (\hat{U}_R^{-T} \otimes \hat{G}) C_{\mathrm{vec}U_R} (\overline{\hat{U}_R^{-1}} \otimes \hat{G}^H) - 2\mathrm{herm}((\hat{U}_R^{-T} \otimes I_{n_y}) C_{\mathrm{vec}Y_R \mathrm{vec}U_R} (\overline{\hat{U}_R^{-1}} \otimes \hat{G}^H)) \quad (37)$$

(see Appendix I).

Comparing the uncertainty of Figures 3 and 4 reveals that the uncertainty of the FRM estimate over all MIMO experiments is significantly larger than that of one MIMO experiment. The apparent contradiction (the standard deviation should decrease with the square root of the number of experiments) is explained by the fact that the sample covariance matrix (33), calculated over the hundred periods of one MIMO experiment, contains the contribution of the disturbing noise only, while the sample covariance matrix (37), calculated over the twenty five MIMO experiments with different realisations of the random phase multisines, contains the contribution of the disturbing noise as well as the stochastic nonlinear distortions [1], [12].

4.3 Identification results

The iterative quadratic maximum likelihood (IQML) estimator (see Section 3.2) is used to estimate transfer function model (6) with $n_a = n_b = 62$ for the measurements shown in Figure 3. The orthogonal basis contains $(n_u^2 + n_u n_y)(n_a + 1) = 504$ vector polynomials of size $n_u^2 + n_u n_y = 8$ by 1. In each step of the iterative algorithm (see eq. (26) and further) the weighted linear least squares solution is found as the sum of the two 8 by 1 vector orthogonal polynomials of guaranteed order 62 corresponding to the main diagonal entries of the matrix polynomial $A(s, \theta)$. The results after 55 iterations are shown in Figures 5 and 6. It follows that the estimated model follows quite accurately the measured frequency response matrix (FRM). The deviation between measured and modelled FRM in the band 400 Hz to 976 Hz is due to the nonlinear distortions (compare Figures 5 and 4). This is not the case in the band 71 Hz to 400 Hz, where the deviation between the measurements and the model is especially apparent at the resonance frequencies (see Figure 6). Note, however, that the IQML solution is an excellent starting value for the maximum likelihood estimator (21).

5 Conclusion

A numerically robust parametrization of transfer function models for multivariable systems has been presented. Although the approach has been elaborated for left matrix fraction descriptions (LMFD), it can also be applied to common denominator and right matrix fraction descriptions.

For a left matrix fraction description, the ML estimator (21) uses, in general, sample noise covariance matrices of size $n_y n_u$ by $n_y n_u$ ($C_{\mathrm{vec}Y}$), n_u^2 by n_u^2 ($C_{\mathrm{vec}U}$), and $n_y n_u$ by n_u^2 ($C_{\mathrm{vec}Y\mathrm{vec}U}$); and requires, in general, $(n_a + 1)n_y^2 + (n_b + 1)n_u n_y$ vector polynomials of size $n_y^2 + n_u n_y$ by 1. Therefore, the proposed method is suitable for moderate values of n_u and n_y only. For large values of n_y (typically 20 and more) the techniques presented in [13] and [14] are recommended.

6 Acknowledgement

This work is supported by the Fund for Scientific Research (FWO-Vlaanderen), the Flemish Government (GOA-IMMI) and the Belgian Government (IUAP V-22).

7 Appendix I

For easy of notation the frequency arguments are dropped. The noisy input/output DFT spectra U, Y resulting from n_u independent MIMO experiments are related to the true values U_0, Y_0 by

$$\begin{aligned} Y &= Y_0 + N_Y \\ U &= U_0 + N_U \end{aligned} \tag{38}$$

Using the first order Taylor series expansion

$$(U_0 + N_U)^{-1} \approx U_0^{-1} - U_0^{-1} N_U U_0^{-1}$$

we get

$$\begin{aligned} G &= YU^{-1} \\ &= (Y_0 + N_Y)(U_0 + N_U)^{-1} \\ &\approx G_0 + N_Y U_0^{-1} - G_0 N_U U_0^{-1} \end{aligned} \tag{39}$$

Using

$$\mathrm{vec}(G_0 N_U U_0^{-1}) = (U_0^{-T} \otimes G_0)\mathrm{vec}(N_U)$$

$$\mathrm{vec}(N_Y U_0^{-1}) = \mathrm{vec}(I_{n_y} N_Y U_0^{-1}) = (U_0^{-T} \otimes I_{n_y})\mathrm{vec}(N_Y)$$

(see [9]), it can easily be verified that

$$\begin{aligned} C_{\mathrm{vec}G} &= (U_0^{-T} \otimes I_{n_y}) C_{\mathrm{vec}Y} (\overline{U}_0^{-1} \otimes I_{n_y}) \\ &\quad + (U_0^{-T} \otimes G_0) C_{\mathrm{vec}U} (\overline{U}_0^{-1} \otimes G_0^H) \\ &\quad - 2\mathrm{herm}((U_0^{-T} \otimes I_{n_y}) C_{\mathrm{vec}Y\,\mathrm{vec}U} (\overline{U}_0^{-1} \otimes G_0^H)) \end{aligned} \tag{40}$$

If the columns of U and Y are independently distributed, then, using (31) and $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ (see [9]), expression (40) can be simplified as

$$\begin{aligned} C_{\mathrm{vec}G} &\approx C \otimes C_Y + C \otimes (G_0 C_U G_0^H) \\ &\quad - 2\mathrm{herm}(C \otimes (C_{YU} G_0^H)) \\ C &= (\overline{U_0 U_0^H})^{-1} \end{aligned} \tag{41}$$

Replacing the true values U_0 and G_0 by the measured values, and the exact noise covariance matrices by the sample noise covariance matrices finally gives (33) and (37).

8 References

[1] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach,* IEEE Press, Piscataway, 2001.

[2] Y. Rolain, R. Pintelon, K. Q. Xu, and H. Vold (1995). Best conditioned parametric identification of transfer function models in the frequency domain. *IEEE Trans. Autom. Contr.,* vol. 40, no. 11, pp. 1954-1960.

[3] Van Barel, M. and A. Bultheel (1992). A parallel algorithm for discrete least squares rational approximation. *Numer. Math.,* vol. 63, pp. 99-121.

[4] Van Barel, M. and A. Bultheel (1994). Discrete linearized least-squares rational approximation on the unit circle. *Journal of Computational and Applied Mathematics,* vol. 50, pp. 545-563.

[5] A. Bultheel, M. Van Barel, Y. Rolain and J. Schoukens (2001). Robust rational approximation for identification. *Proceedings of the 40th IEEE Conference on Decision and Control,* Orlando, Florida (USA), pp. 4770-4775.

[6] Sanathanan, C. K. and J. Koerner (1963). Transfer function synthesis as a ratio of two complex polynomials. *IEEE Trans. on Automatic Control,* vol. AC-8, pp. 56-58.

[7] Kailath, T. (1980). *Linear Systems.* Prentice Hall, Englewood Cliffs.

[8] Golub, G. H. and C. F. Van Loan (1996). *Matrix Computations (3rd ed.).* John Hopkins University Press, Baltimore.

[9] Brewer, J. W. (1978). Kronecker products and matrix calculus in system theory. *IEEE Trans. on Circuits and Systems,* vol. 25, pp. 772-781.

[10] Pintelon, R., P. Guillaume, G. Vandersteen and Y. Rolain (1998). Analyses, development and applications of TLS algorithms in frequency-domain system identification. *SIAM J. Matrix Anal. Appl.,* vol. 19, no 4, pp. 983-1004.

[11] Schoukens, J., R. Pintelon, G. Vandersteen and P. Guillaume (1997). Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica,* vol. 33, no 6, pp. 1073-1086.

[12] Schoukens, J., T. Dobrowiecki, and R. Pintelon (1998). Identification of linear systems in the presence of nonlinear distortions. A frequency domain approach. *IEEE Trans. on Automatic Control*, vol. 43, no 2, pp. 176-190.

[13] Guillaume P., P. Verboven and S. Vanlanduit (1998). Frequency-Domain Maximum Likelihood Estimation of Modal Parameters with Confidence Intervals. *Proceedings of the International Conference on Noise and Vibration Engineering (ISMA-23),* Leuven (Belgium), Sept. 16-18, pp. 359-366.

[14] Van Der Auweraer H., P. Guillaume, P. Verboven and S. Vanlanduit (2001). Application of a fast-stabilizing frequency domain parameter identification method. *ASME Journal of Dynamic Systems, Measurement, and Control,* vol. 123, no. 4, pp. 651-658.