Robust rational approximation for identification.

A. Bultheel*, M. Van Barel*, Y. Rolain[#] and J. Schoukens[#].

* Department of Computer Science, K.U.Leuven, Belgium, [#]Department ELEC, VUB, Brussels, Belgium

Abstract - Using vector orthogonal polynomials as basis functions for the maximum-likelihood (ML) frequency domain identification of the rational form of a linear time invariant system is shown to circumvent all the well known numerical conditioning problems. For identification of very high order systems (e.g. 100/100), systems that operate over a wide frequency band, or even in the presence of over- and undermodelling, condition numbers of less than 10 are reported on real measurements and simulation.

I. INTRODUCTION.

Frequency domain identification of linear time-invariant systems has regained some interest during the last years ([2],[5],[10]). The main driving force behind this revival is the use of periodic excitation signals. Periodic excitation signals not only allow to calculate reliable and very cheap non-parametric frequency response function (frf) estimations by the division of the measured spectra, they do also yield a reliable measure of the stochastic uncertainty on these measured spectra at the cost of a low number (>4) of repeated synchronised measurements([6]).

The knowledge of the frequency response function and its uncertainty allow visual interpretation of the measured system, the estimated model and the match between them. Selection of a simplified model, based on application dependent criteria, is possible through the comparison of the simplified model with the measured non-parametric frf. Based on the measured frf variance, model simplification is statistically sound up to a specified significance level, whenever the simpler model remains within the frf uncertainty bound. Over-simplification of the model can be detected by comparison of the model cost function with its known expected value. Combination of all these indicators criterion leads to fully automatic model order selection methods ([7],[8]).

On the dark side, frequency domain methods suffer from a poor numerical conditioning when the frequency span and/or the model order become large (more than 2 decades frequency span and orders of more than 20 in the laplace domain, narrow banded (a few %) systems of high order (>30) in the discrete-time domain). This numerical problem ruins both the modelling performance and the model order selection capability.

Several attempts have been made in the past to circumvent numerical degeneracy. These range from an appropriate scaling of the frequencies ([5]) to the decomposition of the numerator and denominator of the model equation in a separate basis of polynomials that are orthogonal on the inner product defined by the normal equations of the estimator([4]). 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 solves the numerical conditioning issue perfectly (condition number = 1) for all the frequency domain methods whose cost function can be reduced to a linear weighted least squares problem or a weighted (generalised) total least squares problem. As such, it opens the road to automatic model estimation procedures in the frequency domain for both SISO and MIMO linear time-invariant systems.

II. THE MODEL

To simplify notation, the reasoning will be first performed for the single input single output (SISO) case and will be extended to multiple input multiple output (MIMO) case afterwards. Consider a linear, time-invariant SISO system as in fig. 1.

Fig. 1. The system and noise model

Here, the input/output spectra are labeled $U(\Omega_k) \in C$ and $Y(\Omega_k) \in C$, the exact (unknown) excitation spectra are denoted by a subscript 0, the noise perturbations are labeled N , with a subscript to indicate the noise source. The $C^{2 \times 2}$ covariance $\Sigma(\Omega_k) = \text{cov}([N_U^T(\Omega_k) N_Y^T(\Omega_k)]^T)$ is considered to be known at all measurement frequencies. In practice, it can easily be measured using a low number of repeated synchronised measurements. The system model is $R(\Omega_k, \theta) = A^{-1}(\Omega_k, \theta)B(\Omega_k, \theta)$ with $A(\Omega_k, \theta)$, $B(\Omega_k, \theta)$ polynomials of degree n_a and n_b respectively. The independent variable Ω equals the time lag z for discrete time and the Laplace variable s for continuous time systems. The model equation for the plant is

$$Y_0(f_k) = A^{-1}(f_k, \theta)B(f_k, \theta)U_0(f_k) \quad (1)$$

III. THE ESTIMATOR

The estimators that are considered here all boil down to special cases of a nonlinear least squares estimator. They optimize a cost function of the following form: cost in de vorm van een norm

$$V_{NLS} = \sum_{k=1}^F W^2(f_k, \theta) \begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix}^H \begin{bmatrix} |Y_k|^2 & -\bar{Y}_k U_k \\ -Y_k \bar{U}_k & |U_k|^2 \end{bmatrix} \begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix} \quad (2)$$

where the superscript bar denotes a complex conjugate, and a superscript H a complex conjugate transpose. The parameter dependent weighting in this estimation makes the estimator a nonlinear least squares estimator. One well known example of such an estimator is the maximum likelihood (ML) estimator, which weights measurement data according to measurement

quality. The weighting W_{ML}^2 is thereto set equal to the inverse of the variance of the equation error (the matrix expression in (2)).

$$W_{ML}^{-2}(f_k, \theta) = \sigma_{YU}^2(k) |A(f_k, \theta)|^2 + \sigma_U^2(k) |B(f_k, \theta)|^2 - \dots \quad (3)$$

$$2\text{Re}(\sigma_{YU}^2(k) A(f_k, \theta) \bar{B}(f_k, \theta))$$

where σ are the (co)variances of the noise sources in subscript. The main disadvantage of such a setup is that a nonlinear optimisation results, that has to be solved by quite expensive iterative methods. Because the problem is not convex, it is usually not guaranteed that an optimum is reached, unless a good starting value is known. The way out is to find a good approximation (that could be used as an initial condition to this nonlinear problem) by solving a linearized problem.

IV. THE LINEARISED PROBLEM

Several possible approaches exist to perform the linearisation. All these methods rely on the approximation of the weight W by some expression that is independent on the actual parameter values.

To start the process, a parameter independent weight is required. The most easy way is to set $W(f_k, \theta) = 1$, but this results in a very noise sensitive solution. More sophisticated approximations use the measurement noise information to get a closer match to W_{ML} ([9]).

Once an initial guess of the parameters is known, a refinement step can be started, where the weighting is evaluated in the known parameters θ to determine the new parameter set θ . The cost function then becomes:

$$V_{ls} = \sum_{k=1}^F W^2(f_k, \tilde{\theta}) \begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix}^H \begin{bmatrix} |Y_k|^2 & -\bar{Y}_k U_k \\ -Y_k \bar{U}_k & |U_k|^2 \end{bmatrix} \begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix} \quad (4)$$

$$= \sum_{k=1}^F \left(\begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix}^H W_k^H \begin{bmatrix} \bar{Y}_k \\ -\bar{U}_k \end{bmatrix} \right) \left(\begin{bmatrix} Y_k & -U_k \end{bmatrix} W_k \begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix} \right)$$

An alternative set of methods replaces the cost function (4) by the generalised total least squares cost function,

$$V_{gtls} = \frac{\sum_{k=1}^F \begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix}^H \begin{bmatrix} |Y_k|^2 & -\bar{Y}_k U_k \\ -Y_k \bar{U}_k & |U_k|^2 \end{bmatrix} \begin{bmatrix} A(f_k, \theta) \\ B(f_k, \theta) \end{bmatrix}}{\sum_{k=1}^F W^2(f_k, \tilde{\theta})} \quad (5)$$

The global minimum of this cost function can be calculated directly through a generalised singular value decomposition.

To rule out the trivial solution $\theta = 0$, which is of course not acceptable, an additional constraint is required. For the LS cost function, one of the parameters will be fixed to a non-zero value. For the gtls case, a unity norm constraint is imposed for the parameter vector θ .

It is classical to start with a unity weighting and then to solve a series of weighted linear estimation problems, whose weighting depends in some way on the parameter values of the previous iteration. This iterative ‘refinement’ process will then eventually converge to a close approximation of the ML estimates ([2]). However, if the problem is too complex, the estimation obtained in the first step of the process (unity weighting) may be inadequate to start the process successfully. In these cases, the parameter independent approximation of the ML weighting has been successfully used in practice([9]).

V. VECTOR ORTHOGONAL POLYNOMIALS

Providing a sensible weighting W is a necessary but not a sufficient condition to obtain high quality estimates. All the methods cited above, even the linearised ones, still suffer from the poor numerical conditioning of the Jacobian matrix of the cost function towards the parameters. To enhance numerical conditioning, a number of methods have been proposed. The most classical one is to scale the frequencies by

$$f_{sc} = (\min(f_k) + \max(f_k))/2 \quad (6)$$

This scaling will be used throughout the paper. However, this does not solve the conditioning issue.

A. Least Squares Cost Function

For the moment, only the least squares cost functions as defined by (4) are considered. If the cost function is rewritten as $V = e^H e, e \in C^{F \times 1}$, numerically reliable estimates are obtained if the set of normal equations

$$(J^H J)^{-1} \Delta \theta = -J^H e \quad (7)$$

can be solved with an appropriate numerical precision. Here, the element k, r in the Jacobian matrix J is

$$J_{[k, r]} = \partial e_k / \partial \theta_{[r]} = \begin{cases} -U_k W_k \Omega_k^r & r \leq n_b + 1 \\ Y_k W_k \Omega_k^{r - n_b - 1} & r > n_b + 1 \end{cases} \quad (8)$$

with e_k the k th element of the vector. The classical method to solve such polynomial least squares problems is to use orthogonal polynomials. Using orthogonal polynomials has several advantages over other methods. Not the least important one being the numerical stability that can be guaranteed. Moreover, orthogonal polynomials satisfy some recurrence relations and this allows for a recursive solution. That is, to compute the solution for degree n , one has to compute the successive solutions for all the lower degrees and increasing the degree by one, i.e., $n \rightarrow n + 1$ is obtained by an update of the previous solution. Such an update is simple and cheap so that this recursion also gives a fast algorithm. This is another advantage over other straightforward solutions. The simplest and best known recurrences for orthogonal polynomials are those that result from orthogonality with respect to a weight on the real line or with respect to a weight on the unit circle.

In our case the inner product is with respect to a discrete measure.

In a first attempt, a separate Forsythe polynomial basis was selected for the numerator B and the denominator A of the rational form. Despite its optimality for this dual basis approach, the method failed to obtain optimal conditioning ($\kappa = 1$) for the system (7).

In this paper, a basis of vector orthogonal polynomials is constructed to orthogonalise the numerator and the denominator of the rational form together. If the last element of the parameter vector θ , $\theta_{[n_a + n_b + 2]}$ is fixed to 1, the denominator A is monic. The matrix J is then of full rank $n_a + n_b + 1$ under standard identifiability conditions and the cost function (4) defines an inner product of vector polynomials

$$\langle p(\Omega), q(\Omega) \rangle_w = \text{Re}(W^2(\Omega_k, \tilde{\theta}) p^H(\Omega_k) \begin{bmatrix} |Y_k|^2 & -\bar{Y}_k U_k \\ -Y_k \bar{U}_k & |U_k|^2 \end{bmatrix} q(\Omega_k)) \quad (9)$$

where $p(\Omega)$ and $q(\Omega)$ are 2 by 1 vector polynomials of order smaller or equal to $n = \max(n_a, n_b) + 1$.

In this setting, one needs a polynomial vector $p \in \mathbb{P}_n^{2 \times 1}$ that has minimal cost V_{ls} and is of strict degree n . The problem has been considered in previous publications in [11],[12],[13],[14] in all the gradations of its generality.

The general idea behind the method proposed here is to construct a solution of the form

$$p(\Omega) = \begin{bmatrix} B(\Omega) \\ A(\Omega) \end{bmatrix} = \sum_{k=0}^n \varphi_k(\Omega) \theta_k \quad \theta_k = \begin{bmatrix} \theta_{[2k+1]} \\ \theta_{[2k+2]} \end{bmatrix} \in \mathbb{C}^{2 \times 1} \quad (10)$$

where $\varphi_k \in \mathbb{P}_k^{2 \times 2} \setminus \mathbb{P}_{k-1}^{2 \times 2}$ are 2×2 polynomial matrices of strict degree k that are orthonormal in the sense that

$$\langle \varphi_k(\Omega), \varphi_l(\Omega) \rangle_w = \delta_{kl} I_2. \quad (11)$$

The least squares equation error $e = 0$ now becomes

$$\text{diag}(\tilde{w}_1, \dots, \tilde{w}_F) \begin{bmatrix} \varphi_0(\Omega_1) & \dots & \varphi_n(\Omega_1) \\ \dots & \dots & \dots \\ \varphi_0(\Omega_F) & \dots & \varphi_n(\Omega_F) \end{bmatrix} \begin{bmatrix} \theta_0 \\ \dots \\ \theta_{2n+2} \end{bmatrix} = 0 \quad (12)$$

with $\tilde{w}_k = \begin{bmatrix} -U_k W_k & Y_k W_k \end{bmatrix}$ or $\tilde{W} \Phi \theta = 0$. Taking into account the orthogonality relation (11), the cost function to be minimised becomes

$$V_{ls} = \theta^H \Phi^H \tilde{W}^H \tilde{W} \Phi \theta = \sum_{k=0}^{2n+2} |\theta_k|^2 \quad (13)$$

Taking into account that A was assumed to be monic above, it is clear that $\theta_1 = \dots = \theta_{2n} = 0$ to minimise the cost function, and the cost function boils down to

$$V_{ls} = |\hat{\theta}_{2n+1}|^2 + |\hat{\theta}_{2n+2}|^2 \quad (14)$$

and thus

$$\hat{p}(\Omega) = \begin{bmatrix} \hat{B}(\Omega) \\ \hat{A}(\Omega) \end{bmatrix} = \varphi_n(\Omega) \begin{bmatrix} \hat{\theta}_{2n+1} \\ \hat{\theta}_{2n+2} \end{bmatrix} \quad (15)$$

To deal with the degree condition, assume that the leading coefficient of $\varphi_n(\Omega)$ is upper triangular

$$\varphi_n(\Omega) = \begin{bmatrix} \alpha_n & \gamma_n \\ 0 & \beta_n \end{bmatrix} \Omega^n + \dots \quad (16)$$

with $\alpha_n > 0$ and $\beta_n > 0$. Such a normalisation is always possible if this is a nonsingular recursion. Now it is clear how to fix the only remaining freedom: $\theta_{2n+1} = 0$. To obtain a monic polynomial, $\theta_{2n+2} = 1/\beta_n$. Thus

$$\hat{p}(\Omega) = \begin{bmatrix} \hat{B}(\Omega) \\ \hat{A}(\Omega) \end{bmatrix} = \varphi_n(\Omega) \begin{bmatrix} 0 \\ 1/\beta_n \end{bmatrix}. \quad (17)$$

Note that other normalisation conditions are possible. In fact, an arbitrary degree structure can be imposed. For a full discussion, see [13].

B. Generalised Total Least Squares Cost Function

For this type of cost function, a direct solution without any matrix operation is less straightforward to obtain.

To obtain the estimates, the generalised singular value decomposition of the matrix pair J, C_J is to be calculated. Here, J is defined as in (8) and C_J is the column covariance matrix of J , which can be evaluated since the experimental variances have also been measured.

If an orthogonal vector polynomial basis is constructed in exactly the same way as above, this basis can be introduced in the Jacobian matrix J , and will also result in an orthogonalisation. This can be accounted for during the evaluation of the generalised singular value decomposition, which becomes fast to evaluate and numerically robust.

C. Application to iterative weight refinement

This orthogonalisation procedure can now be used for all the least squares and total least squares based methods. As was discussed in section IV, it is common to construct a series of 'linearised' cost functions that come close to the maximum likelihood cost function. To apply the orthogonalisation here is straightforward, as the weight calculated in the old vector polynomial basis can be used to evaluate the new vector polynomial basis.

D. Maximum Likelihood Cost Function

Since this is a nonlinear least squares estimator, the weighting function W now becomes parameter dependent. Hence, orthogonalisation becomes impossible. To get around this, one could evaluate a new vector polynomial basis in each iteration step during ML optimisation to maintain the orthogonality of the matrix J . However, since the iterative procedure calculates a parameter increment in the new basis while the parameters are only known in the old one, the parameter vector must be transformed to the new basis before the addition can take place. Unfortunately, there is no numerically stable

way to calculate this transformation. This makes the approach inadequate. The solution is to fix the polynomial basis during the whole maximum likelihood optimisation. Practical experience shows that the starting values for the ML optimisation are usually good enough to keep the condition number after the ML optimisation below 10.

VI. POLES AND ZEROS

The pair estimated parameters θ and vector polynomial basis $\phi_k(\Omega)$ determines the estimated polynomials totally. For simulation purposes, this representation is adequate, but the poles and/or zeros of the polynomials are not immediately accessible. Fortunately, they can be calculated using the recurrence relation of the polynomials and the coefficient values. Since the vector recurrence relation complicates this process, a workaround has been used in this paper. A separate Forsythe orthogonal representation is estimated for B and A . This noiseless fitting of a polynomial of known order can easily be accomplished with epsilon small errors. The roots of this representation are then evaluated using the modified companion method used in [4].

VII. EXTENSION TO MIMO

All the reasoning and the equations can immediately be extended to MISO systems (n_i inputs), where $B(\Omega)$ becomes a vector of polynomials instead of scalar polynomial. In this case, the expression for maximum likelihood weighting $W_{ML}(f_k, \theta)$ is still a real scalar, and hence the cost function can be written in the form of (4). Using vector polynomials of size $n_i + 1$ instead of 2 for the SISO case will hence result in a linearised cost function that comes very close to the full ML cost.

The same holds for SIMO systems (n_o outputs), but now $A(\Omega)$ becomes a vector of polynomials. The new problem now can be decomposed in n_o independent scalar problems, that can be solved in one estimation stacking the equations of the subproblems in the Jacobian matrix. Again, the linearised ML cost can be rewritten in the form of (4), and vector polynomials of size $n_o + 1$ clear the job.

For the full MIMO case, the extension is again straightforward if the full problem can be decomposed in n_o MISO problems. To allow this, the covariance matrix of the output part of the spectrum has to be diagonal, meaning that the output noise sources on the system outputs have to be statistically independent. If this is the case, the orthogonalisation as proposed will also work in the MIMO case.

VIII. EXPERIMENTAL RESULTS

To show that the proposed method leads to practically applicable results, it has been decided to apply it on real measured data of a real-world system. The systems are selected in such a way that the most common types of practical non-ideal behaviour are present in the examples. The first example is a complex mechanical vibrating system with a very high number of vibration modes. It shows the ability of the method to cope with wideband, very high order systems.

The second example, the radial response of a CD head, illustrates the robustness of the method against lousy model order

selection. This property can then be exploited to obtain an automatic model order selection procedure.

A. A very high order measured vibrating mechanical structure in the Laplace domain.

A composed steel beam structure (Fig. 2) is excited by a mini-shaker. The force and the acceleration are measured by an impedance head mounted on the shaker using a stinger rod. Signals are acquired by an FFT-based dynamic signal analyser (HP3562A). The excitation signal is a periodic chirp between 0 and 5000 Hz with a frequency resolution of 6.25 Hz. Fig. 4 shows the magnitude of the FRF obtained by the division of the output and input spectra. These spectra will then be used for the estimation.

Fig. 2. Used structure for the modal analysis measurement.

The rational transfer function is modelled between 493 Hz and 5 kHz at $F=722$ equidistant frequency lines. 32 consecutive synchronised measurements allow to obtain the mean value of the measured spectral lines and their variance as a function of frequency. The model order is chosen to be order numerator / order denominator = $n/d = 120/120$.

The estimation is performed in two steps. First, a weighted generalised total least squares (GTLS) estimator with a parameter-independent weighting is used to get an initial guess for the parameter vector. Next, the bootstrapped total least squares (BTLS) estimator is used for 10 iterations to refine the initial guess. A new vector orthogonal polynomial basis is calculated for each step in this process. Finally, the maximum likelihood ELiS estimator is used to tweak the results using the vector orthogonal basis that was calculated for the last BTLS step. Since orthogonality cannot be imposed during the iteration of the ML estimator, the condition number of the Jacobian matrix is evaluated at the end of the ML iteration to evaluate the degradation of the conditioning.

Modelled and measured magnitude of the transfer function are shown in fig. 4. Even if both functions are present there, differences can hardly be found. The phase of both quantities is shown in fig. 3. To show that there is actually a difference between model and measurement, the magnitude of the complex difference between model and measurement is shown together with the magnitude of the transfer functions in fig. 5. In general, the ratio between transfer function and residue is everywhere below the -50dB . The phase difference between model and measurement as shown in fig. 3 is everywhere smaller than 1 degree.

The estimated model can predict the linear dynamics of the system with a high accuracy. The model combines pole/zero pairs with high quality factors, that result in sharp resonances, with low quality factor roots to describe slower variations in

the transfer function. The latter can not be determined at all if the frequency characteristic is sliced into smaller bands as proposed by generic modal analysis approaches.

The value of the ML cost function is 642 while the expected value ([5]) is $F - (n + d + 1)/2 = 601.5$. The estimated cost falls within the uncertainty bound of the theoretically predicted one (50% for 32 measurements). This means that the residue between model and measurement can be explained by the noise variation only. The (linear) condition number of the Jacobian matrix degrades from 1 to 1.4 during the ML estimation. This is due to the fact that the orthogonal basis is not updated during the ML iteration. This is only a marginal degradation, and shows that the BTLS estimation, that was obtained in the last iteration of this estimator, matches the ML solution very closely. Calculation of the poles/zeros of the obtained model show a stable, minimum-in-phase behaviour. Based on the cost function and the excellent conditioning, one can conclude that the model captures the dynamics of the system adequately. The ML residual shown in fig. 6, contains a residual correlation as a function of the frequency. Maximal residue magnitudes are mainly attained at the sharp resonances of the transfer function. This points to the presence of small non-linearities at these frequencies. Practical experience learns that this often happens for modal analysis measurements with sharp resonances.

B. Modelling a CD-radial positioner.

The device under test is a Philips CD320/00G compact disk drive modified to get access to the control loop of the radial servo system. All measurements were done in closed loop operation at the start of track 1. The maximal experiment time was 26.2 s. The closed loop system is excited by a HPE1445A VXI arbitrary waveform generator operating in the ZOH-mode without reconstruction filter. The sampling frequency was set to $f_s = 1 \times 10^7 / 2^{10} \approx 9765.6 \text{ Hz}$.

The system response is measured by HPE1430A VXI ADC cards with anti-alias protection on and sampling frequency equal to f_s . High impedance buffers ($1 \text{ M}\Omega$, 10 pF) with a gain of about 0.5 were used to isolate the plant from the 50Ω input impedance of the acquisition frontend.

The excitation signal is a special odd multisine ([10]) with $F = 154$ spectral components located at frequencies $k f_0$, $k = 1, 3, 9, 11, 17, 19, \dots, 611$. The crest factor of the signals was compressed to about 1.55 starting from random initialized phases. The RMS value of the applied excitation signals was about 1 Volt.

The system is modelled in the Laplace domain. Several models are extracted. The degree of the transfer function model is varied from $2/2$ to $6/6$ in increments of 1 order and from $8/8$ to $20/20$ with increments of 2.

To obtain the final ML estimates, an initial GTLS estimation and a series of 10 BTLS estimations are performed as in the previous example.

Results of the series of estimations are summarised in fig. 7. Clearly, the cost function obtained ranging from $50e3 \rightarrow 500$ is too large when compared to the expected value of the cost ranging from $151.5 \rightarrow 133.5$, even for the most complex model. An additional test on the whiteness of residual shows that for model orders larger than $8/8$, there are no linear

unmodelled dynamics left, and that the increased value of the cost is caused by small nonlinear distortions.

Note that, even under severe overmodelling and undermodelling, the (linear) condition number of the ML estimator remains very close to the optimal value of 1 that is obtained as a result for the BTLS initial step. The numerical properties of the estimator are hereby shown to be robust to both overmodelling and undermodelling, and this opens the door to sensible and easy to interpret model order selection procedures.

To show that the modelling as obtained for the different orders is indeed sensible, three estimated models are shown. The first model of order $3/3$ is shown in fig. 8. It results in a severe undermodelling of the measured characteristics, especially at low frequencies where the magnitude of the complex model error is much larger than the value of the measurement. A sensible model is obtained at the order $5/5$ (fig. 9). The low frequency dynamics are now nicely captured, the model error is about 20 dB under the model value over the whole band. When the model order is further increased to $18/18$ (fig. 10), the overall level of the error decreases further, especially at the center of the band where the remaining bumps in the characteristic are now captured. However, the model contains 6 unstable poles, that are almost perfectly cancelled by very close zeros. The closed loop system resulting from this open loop model is still perfectly stable.

Based on this example, one can conclude that the proposed method is robust to over- and undermodelling problems. As such, the tool may be handy to the community of identification users, that have no extensive identification knowledge but require sensible models for use in their own expertise fields. Even if the initial model order is poorly selected, the method will come up with a sensible model that comes close to the behaviour of the system to be modelled.

IX. CONCLUSION

A method is proposed and experimentally verified to calculate the frequency domain maximum likelihood estimates of linear systems with almost optimal numerical conditioning. Even for very wide frequency bands, very high model order and in the presence of over- and undermodelling, or a combination of these factors, the method performs extremely well. As such, it enables the realisation of automatic model order selection procedures. On top of that, these methods extend gracefully to MISO and MIMO systems.

X. ACKNOWLEDGEMENT

The work of all authors has been sponsored by the Belgian programme on interuniversity poles of attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The work of Y. Rolain and J. Schoukens is also sponsored by the VUB research Council, GOA-IMMI II project.

XI. REFERENCES

- [1] R. Pintelon and J. Schoukens, 'Robust Identification of Transfer Functions in the s- and z- Domains', *IEEE Trans. Instrumentation and Measurement*, Vol. IM-39, No. 4, pp. 565-573, 1990.

- [2] R. Pintelon, P. Guillaume, Y. Rolain, J. Schoukens and H. Van hamme, 'Parameter Identification of Transfer Functions in the Frequency Domain, A Survey', *IEEE Trans. Autom. Contr.*, Vol. AC-39, No. 11, 1994.
- [3] R. Pintelon, P. Guillaume, Y. Rolain and F. Verbeyst, 'Identification of Linear Systems Captured in a Feedback Loop', *IEEE Trans. Instrumentation and Measurement*, Vol. IM-41, No. 6, pp. 747-754, 1992.
- [4] Y. Rolain, R. Pintelon, K. Q. Xu and H. Vold, 'On the Use of Orthogonal Polynomials in High Order Frequency Domain System Identification and its Application to Modal Parameter Estimation', *Proceedings of 33rd IEEE Conference on Decision and Control, Orlando (USA), December 14-16, 1994*, pp. 3365-3373.
- [5] J. Schoukens and R. Pintelon, 'Identification of Linear Systems: A Practical Guideline to Accurate Modelling', Pergamon Press, London (UK), 1991.
- [6] J. Schoukens, G. Vandersteen, R. Pintelon and P. Guillaume, '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, 1997.
- [7] R. Pintelon, J. Schoukens and G. Vandersteen, 'Model Selection through a Statistical Analysis of the Global Minimum of a Weighted Non-Linear Least Squares Cost Function', *IEEE Trans. Sign. Proc.*, Vol. SP-45, no. 3, pp. 686-693, 1997.
- [8] J. Schoukens, T. Dobrowiecki and R. Pintelon, 'Parametric Identification of Linear Systems in the Presence of Nonlinear Distortions. A Frequency Domain Approach', *IEEE Trans. Autom. Contr.*, Vol. AC-43, no. 2, pp. 176-190, 1998.
- [9] Y. Rolain and R. Pintelon, 'Generating Robust Starting Values for Frequency Domain Transfer Function Estimation', *Automatica*, Vol. 35, no. 5, pp. 965-972, 1999.
- [10] Pintelon R. and J. Schoukens: 'System Identification: A Frequency Domain Approach', IEEE Press, Piscataway, 2001
- [11] M. Van Barel and A. Bultheel, 'Discrete linearized least squares approximation on the unit circle', *J. Comput. Appl. Math.*, vol. 50, pp. 545-563, 1994.
- [12] M. Van Barel and A. Bultheel, 'A parallel algorithm for discrete least squares rational approximation', *Numer. Math.*, vol. 63, pp. 99-121, 1992.
- [13] A. Bultheel and M. Van Barel, 'Vector orthogonal polynomials and least squares approximation', *SIAM J. Matrix Anal. Appl.*, vol. 16, no. 3, pp. 863-885, 1995.
- [14] M. Van Barel and A. Bultheel, 'Orthogonal polynomial vectors and least squares approximation for a discrete inner product', *Electron. Trans. Numer. Anal.*, vol. 3, pp. 1-23, 1995.

Fig. 3. Phase of the meas. (dashed) and estim. (full) frf

Fig. 4. Magn. of the meas. (dashed) and estim. (full) frf

Fig. 5. Magn. estim. (full), meas. (dashed) frf & magn. of cmplx error (dotted)

Fig. 6. Magn. of the ML eqn. error

Fig. 7. ML cost (crosses) and condition nr. (triangles) vs. increasing model complexity

Fig. 8. Estim. model 3/3, Model (full), meas. (dotted) & magn. cmplx error (dash-dot)

Fig. 9. Estim. model order 5/5, (Model full, meas. dotted & magn. of cmplx error dash-dot)

Fig. 10. Estim. model order 18/18, overmodelling, (Model full, meas. dotted, magn. cmplx error dash-dot)