PARAMETRIC IDENTIFICATION OF FREQUENCY DOMAIN SYSTEMS USING ORTHONORMAL RATIONAL BASES

Dirk Deschrijver [*,1] **Tom Dhaene** [**]

[*] *Dept. Math. and Comp. Sc, University of Antwerp, dirk.deschrijver@ua.ac.be*
[**] *Dept. Math. and Comp. Sc, University of Antwerp, tom.dhaene@ua.ac.be*

Abstract: This paper presents a new robust algorithm for the identification of linear time-invariant systems in the frequency domain. The method is computationally very efficient, and accurately synthesises transfer functions which require a large amount of poles. Some numerical difficulties are avoided by combining the use of a Sanathanan-Koerner iteration and orthonormal rational functions.

Keywords: Identification Algorithms, Parameter Estimation, Continuous-time systems, Estimation Algorithms, Frequency Responses

1. INTRODUCTION

Parametric identification of continuous-time linear time-invariant (LTI) systems is becoming increasingly important for accurate transfer function synthesis in multiple scientific communities such as e.g. Power Systems and Microwave Engineering. In this paper, a new identification method is proposed which reliably parameterizes the transfer function of highly dynamical systems without requiring any a-priori knowledge of the system poles. The identification algorithm minimizes a weighted linear cost function, by iteratively relocating a set of prescribed poles using a Sanathanan-Koerner iteration (Sanathanan and Koerner, 1963). The numerical conditioning of the system equations is improved by using Muntz-Laguerre orthonormal rational basis functions (Knockaert, 2001)(Akcay and Ninness, 1999).

The final transfer function is represented as an accurate state-space realization or pole-residue model. Such representations are often preferred over transfer functions which are expanded in an orthonormal basis, since they are very lenient towards the extraction of the poles and zeros. Also, the method is computationally more efficient than many existing approaches.

First, the new iterative method is placed in a broader context of system identification and it is related to some of the existing work. Afterwards, the technique is described in detail and the robustness is illustrated by a simulation-based example. Rather than divulging into theoretical aspects, this paper intends to give some considerations from a practical point of view.

2. IDENTIFICATION ALGORITHM

2.1 Goal

The major goal of system identification is to identify the mapping between the inputs and outputs of a complex system by an analytic model. For continuous-time LTI systems in the frequency domain, this reduces to finding a rational transfer function

[1] This work was supported by the Fund for Scientific Research Flanders (FWO-Vlaanderen)

$$R(s) = \frac{N(s)}{D(s)} = \frac{\sum_{n=0}^{N} N_n \phi_n(s)}{\sum_{d=0}^{D} D_d \phi_d(s)} \qquad s = i2\pi f \quad (1)$$

which approximates the spectral response of a system over some predefined frequency range of interest [f_{min},f_{max}]. The spectral behavior is characterized by a set of frequency-domain data samples $(s_k, H(s_k))$, $\forall k = 0, ..., K$, which can be obtained from observations, such as e.g. measurements or circuit simulations (Pintelon and Schoukens, 2001).

N_n and D_d are the real-valued system parameters which need to be estimated, and N and D represent the order of numerator and denominator respectively. In many situations, the amount of available data samples is quite numerous, so numerically stable fitting techniques are required which estimate the model coefficients in a least-squares sense (Golub and Loan, 1989).

2.2 Non-linearity of the estimator

Rational least-squares approximation is essentially a non-linear problem, and corresponds to minimizing the following cost function (Pintelon *et al.*, 1994)

$$\arg \min_{N_n, D_d} \sum_{k=0}^{K} \left| H(s_k) - \frac{N(s_k)}{D(s_k)} \right|^2 \qquad (2)$$

$$= \arg \min_{N_n, D_d} \sum_{k=0}^{K} \frac{1}{|D(s_k)|^2} |D(s_k)H(s_k) - N(s_k)|^2$$

Due to its non-linear nature, it becomes quite hard to estimate the system parameters in a fast and accurate way.

In many papers, e.g. (Wahlberg and Makila, 1996),this difficulty is avoided by assuming that a-priori knowledge about the poles is available. In this case, the non-linear problem reduces to a linear problem since the denominator parameters are assumed to be known. In practice, however, this situation is often not a realistic one.

Another possible option is the use of non-linear optimization techniques, such as Newton-Gauss type algorithms, in order to minimize (2). This approach is computationally not always efficient, and the solutions may converge to local minima, even when Levenberg-Marquardt algorithms are used to extend the region of convergence.

In (Akcay and Ninness, 1999), it was proposed to minimize a Kalman-linearized cost function which is non-quadratic in the system parameters (Kalman, 1958)(Levi, 1959)

$$\arg \min_{N_n, D_d} \sum_{k=0}^{K} |D(s_k)H(s_k) - N(s_k)|^2 \qquad (3)$$

This formulation basically reduces to (2), if the weighting factor $\frac{1}{|D(s_k)|^2}$ is set equal to one for all frequencies s_k. Clearly, this weighting will bias the fitted transfer function, and this often results in poor low-frequency fits, due to an undesired overemphasis of high-frequency errors.

In this paper, the use of a Sanathanan-Koerner (SK) iteration is advocated (Sanathanan and Koerner, 1963). First, an estimate of the poles is obtained by minimizing the Kalman-linearized cost function. Given this initial (iteration step 0) or previous (iteration step $t-1$) estimate of the poles, the model parameters of the next iteration step are calculated by minimizing the weighted linear cost function

$$\arg \min_{N_n^{(t)}, D_d^{(t)}} \left(\sum_{k=0}^{K} \frac{\left| D^{(t)}(s_k)H(s_k) - N^{(t)}(s_k) \right|^2}{\left| D^{(t-1)}(s_k) \right|^2} \right) (4)$$

By analyzing the gradients of the error criterion, it is straightforward to show that this method generates solutions that don't converge asymptotically to the solution of (2) either, even though the error criterion itself tends asymptotically to the fundamental least squares criterion (Whitfield, 1987). In practice, however, this approach often gives favorable results for sufficiently high signal-to-noise ratios and sufficiently small modeling errors.

The interested reader is hereby referred to an excellent survey of (Pintelon *et al.*, 1994) which analyses these and several other techniques in more detail.

2.3 Choice of basisfunctions

To solve the identification problem, equation (4) reduces naturally to a linear set of least-squares equations, which needs to be solved with sufficient accuracy

Suppose that $H = diag(H(s_0), ..., H(s_K))$, $w_k = [D^{(t-1)}(s_k)]^{-1}$, and $\Phi_{0:X}$ is defined as

$$\Phi_{0:X} = \begin{pmatrix} w_0\phi_0(s_0) & ... & w_0\phi_X(s_0) \\ ... & ... & ... \\ w_K\phi_0(s_K) & ... & w_K\phi_X(s_K) \end{pmatrix} \qquad (5)$$

Then the least-squares solution of $Ax = b$ can be calculated to estimate the parameter vector x, provided that A, x and b are defined as ($D_0 = 1$)

$$A = \begin{pmatrix} \Re e \begin{pmatrix} \Phi_{0:N} & -H\Phi_{1:D} \end{pmatrix} \\ \Im m \begin{pmatrix} \Phi_{0:N} & -H\Phi_{1:D} \end{pmatrix} \end{pmatrix} \qquad (6)$$

$$x = (N_0^{(t)} \ ... \ N_N^{(t)} \ D_1^{(t)} \ ... \ D_D^{(t)})^T \qquad (7)$$

$$b = \begin{pmatrix} \Re e(H\Phi_0) \\ \Im m(H\Phi_0) \end{pmatrix} \qquad (8)$$

Each equation is split in its real and imaginary part to enforce the poles and zeros to be real, or to occur in complex conjugate pairs (under the assumption that the basis functions $\phi(s)$ are real-valued as well). This ensures that the coefficients of the transfer function are real, and that no imaginary terms occur in the time-domain.

Now it's easy to estimate the system parameters by solving the normal equations

$$x = (A^T A)^{-1} A^T b \tag{9}$$

or e.g. by using a QR decomposition with column pivoting.

It becomes clear that the accuracy of the parameter vector x, and the numerical conditioning of this problem is highly dependent on the structure of the normal equations $A^T A$ (Golub and Loan, 1989). If the basisfunctions $\phi(s)$ are chosen to be a monomial power series basis (1,s,s^2,...), the matrix Φ will be a Vandermonde matrix which is notoriously ill-conditioned.

(Adcock and Potter, 1985) suggested the use of polynomials which are orthogonal with respect to a continuous inner product, such as Chebyshev polynomials, as basis functions. The large variation of the Chebyshev polynomials with increase in order makes it possible to downsize the effects of ill-conditioning.

On the other hand, (Richardson and Formenti, 1982) proposed the use of Forsythe polynomials which are orthonormal with respect to a discrete inner product, defined by the normal equations of the estimator. This implies that a different set of basis functions is used for numerator and denominator. (Rolain *et al.*, 1995) have shown that a basis transformation from the Forsythe polynomials to a different, arbitrary polynomial basis results in an inferior conditioning of $A^T A$. Hence, the Forsythe polynomial basis is optimal in a sense that there doesn't exist any other polynomial basis resulting in a better conditioned form of the normal equations.

Although polynomial bases are probably the most natural choice, it is well-known that rational basis functions have a lot of numerical advantages. Quite recently, (Gustavsen and Semlyen, 1999) proposed the use of partial fractions as basis functions for the numerator and denominator

$$R(s) = \frac{N(s)}{D(s)} = \frac{\sum_{p=1}^{P} \frac{c_p}{s+\bar{a}_p}}{1 + \sum_{p=1}^{P} \frac{\tilde{c}_p}{s+\bar{a}_p}} \tag{10}$$

provided that c_p represent the residues, and $-\bar{a}_p$ are a set of prescribed poles. The denominator has an additional basisfunction which equals the constant value 1. Its coefficient can be fixed to one, since numerator and denominator can be divided by the same constant value without loss of generality. Given the constraint that the poles of the numerator and denominator expression of (10) are the same, it's easy to see that these basis functions are complete, in a sense that they can approximate any strictly proper transfer function with distinct poles arbitrarily well. The parameter vector now consists of unknown residues

$$x = (c_1^{(t)} \ ... \ c_P^{(t)} \ \tilde{c}_1^{(t)} \ ... \ \tilde{c}_P^{(t)})^T \tag{11}$$

To approximate systems which require a proper or improper transfer function, an optional constant and linear term can be added to the numerator expression.

After simplification of the transfer function (10), the zeros of the denominator become the poles of the transfer function. These zeros can easily be calculated from the minimal state space realization of $D(s)$ as was shown in appendix B of (Gustavsen and Semlyen, 1999). Once the poles are known, the residues can be estimated as a linear problem, and the transfer function is realized as a partial fraction expansion.

2.4 Continuous-time Malmquist basis

Instead of using the partial fractions as rational basis functions, it was shown that orthonormal rational basis functions can lead to significant improvements in numerical conditioning (Ninness *et al.*, 2000)(Deschrijver and Dhaene, 2005).

$$R(s) = \frac{N(s)}{D(s)} = \frac{\sum_{p=1}^{P} c_p \phi_p(s)}{1 + \sum_{p=1}^{P} \tilde{c}_p \phi_p(s)} \tag{12}$$

A straightforward way to calculate an orthonormal basis, is to apply a Gram-Schmidt procedure on the partial fractions (Akcay and Ninness, 1999)(Oliveira e Silva, 1995). Hence, orthonormal rational functions $\phi_p(s)$ are obtained, which are in fact linear combinations of the partial fractions, of the form

$$\phi_p(s) = \frac{Q_p(s)}{\prod_{j=1}^{p}(s + a_j)} \tag{13}$$

for $p = 1, ..., P$ and $Q_p(s)$ an arbitrary polynomial of order $p-1$, such that $(1 \leq m, n \leq P)$.

$$\langle \phi_m(s), \phi_n(s) \rangle = \delta_{mn} \tag{14}$$

If the inner product is defined as

$$\langle \phi_m(s), \phi_n(s) \rangle = \frac{1}{2\pi i} \int_{i\mathbb{R}} \phi_m(s) \phi_n^*(s) ds \tag{15}$$

then the $Q_p(s)$ polynomial can be determined by imposing the orthonormality conditions on the

basis functions. The Gram-Schmidt procedure is done analytically and gives rise to the following closed form expression

$$\phi_p(s) = \kappa_p \sqrt{2\Re e(a_p)} \left(\prod_{j=1}^{p-1} \frac{s - a_j^*}{s + a_j} \right) \frac{1}{s + a_p} \quad (16)$$

provided that the poles are stable, i.e. located in the left half of the complex plane. κ is an arbitrary unimodular complex number. This basis originates from the discrete-time Takenaka-Malmquist basis (Takenaka, 1925)(Malmquist, 1926), and has later been transformed to the continuous time domain. It is a generalization of the Laguerre basis (Clement, 1982), where all poles $\{-a_p\}$ are the same real number, and the 2-parameter Kautz bases (Kautz, 1954) where all poles $\{-a_p, -a_{p+1}\}$ are the same complex conjugate pair with $-a_p^* = -a_{p+1}$. A theoretical analysis of these basis functions is well-described in literature. The interested reader is referred to (Heuberger *et al.*, 2005) which gives an excellent survey.

To make sure that the transfer function has real-valued coefficients, a linear combination of $\phi_p(s)$ and $\phi_{p+1}(s)$ is formed which can be made real-valued if the poles are real or occur in a complex conjugate pair. This way, two orthonormal functions of the following form are obtained

$$\phi_p(s) = \gamma_p \left(\prod_{j=1}^{p-1} \frac{s - a_j^*}{s + a_j} \right) \frac{s - x}{(s + a_p)(s + a_{p+1})} \quad (17)$$

$$\phi_{p+1}(s) = \gamma_{p+1} \left(\prod_{j=1}^{p-1} \frac{s - a_j^*}{s + a_j} \right) \frac{s - y}{(s + a_p)(s + a_{p+1})} \quad (18)$$

To satisfy the orthonormality, the indeterminates x and y are chosen to be $\sqrt{a_p a_{p+1}} = |a_p|$ and $-\sqrt{a_p a_{p+1}} = -|a_p|$ respectively, and γ_p and γ_{p+1} are set to $\sqrt{a_p + a_{p+1}} = \sqrt{2\Re e(a_p)}$. Note that this choice is not unique, and that other possibilities exist.

The minimal continuous-time LTI state-space realization

$$sX(s) = AX(s) + BU(s) \quad (19)$$
$$Y(s) = CX(s) + DU(s) \quad (20)$$

of the denominator

$$D(s) = 1 + \sum_{p=1}^{P} \tilde{c}_p \phi_p(s) \quad (21)$$

can then be calculated, by cascading the minimal state-space realization of smaller, first and second order sections (Gomez, 1998)

$$\frac{s - a_1^*}{s + a_1} \rightarrow \frac{s - a_2^*}{s + a_2} \rightarrow ... \rightarrow \frac{s - a_{P-1}^*}{s + a_{P-1}} \rightarrow \frac{1}{s + a_P} \quad (22)$$

The minimal state-space realization (A_p, B_p, C_p, D_p) of the all-pass function

$$\frac{Y_p(s)}{U_p(s)} = \frac{s - a_p^*}{s + a_p} \quad (23)$$

for $p = 1, ..., P-1$ is given as $A_p = -a_p,\ B_p = 1,\ C_p = 2\Re e(-a_p),\ D_p = 1$. The minimal state-space realization (A_p, B_p, C_p, D_p) of the low-pass function

$$\frac{Y_p(s)}{U_p(s)} = \frac{1}{s + a_p} \quad (24)$$

is given as $A_p = -a_p,\ B_p = 1,\ C_p = 1,\ D_p = 0$ for $p = P$. Then the minimal state-space realization of the compound system (22) is obtained as the cascade construction

$$A = \begin{bmatrix} A_1 & 0 & ... & 0 \\ B_2 C_1 & A_2 & ... & 0 \\ B_3 D_2 C_1 & B_3 C_2 & ... & 0 \\ B_4 D_3 D_2 C_1 & B_4 D_3 C_2 & ... & 0 \\ ... & ... & ... & ... \\ B_P D_{P-1} ... D_2 C_1 & B_P D_{P-1} ... D_3 C_2 & ... & A_P \end{bmatrix}$$

$$B = \begin{bmatrix} B_1 \\ B_2 D_1 \\ B_3 D_2 D_1 \\ B_4 D_3 D_2 D_1 \\ ... \\ B_P D_{P-1} ... D_1 \end{bmatrix}, C = \begin{bmatrix} D_P ... D_2 C_1 \\ D_P ... D_3 C_2 \\ ... \\ C_P \end{bmatrix}^T \quad (25)$$

$$D = D_P ... D_1$$

of the smaller state space models, with $y_p(t) = u_{p+1}(t)$.

The state matrix A and the input vector B are build such that the states contain exactly the unnormalized basis functions. The output vector C and scalar D are chosen to obtain the denominator expression (21), by compensating for the coefficients $\tilde{c}_p$ and normalization constant $\sqrt{2\Re e(a_p)}$ in the vector C, and setting the scalar D equal to the constant value 1. Hence, the following real-valued state space realization is obtained

$$A_{P \times P} = \begin{bmatrix} -a_1 & 0 & 0 & ... & 0 \\ 2\Re e(-a_1) & -a_2 & 0 & ... & 0 \\ 2\Re e(-a_1) & 2\Re e(-a_2) & -a_3 & ... & 0 \\ 2\Re e(-a_1) & 2\Re e(-a_2) & 2\Re e(-a_3) & ... & 0 \\ ... & ... & ... & ... & ... \\ 2\Re e(-a_1) & 2\Re e(-a_2) & 2\Re e(-a_3) & ... & -a_P \end{bmatrix}$$

$$B_{1 \times P} = \begin{bmatrix} 1 \\ 1 \\ ... \\ 1 \end{bmatrix},\ C_{P \times 1} = \begin{bmatrix} \tilde{c}_1 \sqrt{2\Re e(a_1)} \\ \tilde{c}_2 \sqrt{2\Re e(a_2)} \\ ... \\ \tilde{c}_P \sqrt{2\Re e(a_P)} \end{bmatrix}^T \quad (26)$$

$$D_{1 \times 1} = 1$$

provided that the poles $-a_p$ are real.

If $-a_p$ and $-a_{p+1}$ constitute a complex conjugate pair of poles (i.e. $-a_{p+1} = -a_p^*$), a real-valued state-space realization is obtained by replacing

$$\frac{s - a_p^*}{s + a_p} \rightarrow \frac{s - a_{p+1}^*}{s + a_{p+1}} \tag{27}$$

in the cascade scheme (22) by

$$\frac{(s - a_p^*)(s - a_{p+1}^*)}{(s + a_p)(s + a_{p+1})} = 1 + \frac{4\Re e(-a_p)s}{(s + a_p)(s + a_p^*)} \tag{28}$$

This corresponds to replacing

$$\begin{pmatrix} -a_p & 0 \\ 2\Re e(-a_p) & -a_{p+1} \end{pmatrix} \tag{29}$$

in the state matrix A, by

$$\begin{pmatrix} \Re e(-a_p) & \Re e(-a_p) - |a_p| \\ \Re e(-a_p) + |a_p| & \Re e(-a_p) \end{pmatrix} \tag{30}$$

The other state space matrices remain unchanged. This transformation makes the state-space realization of $D(s)$

$$D(s) = C(sI - A)^{-1}B + D \tag{31}$$

real-valued, such that the poles and zeros occur as complex conjugate pairs. When simplifying (12), the common poles of $N(s)$ and $D(s)$ cancel out and the zeros of the denominator become the identified poles of $R(s)$. From (31), they can be found by solving the eigenvalues of $A - BC$. By replacing the prescribed poles -a_p with the identified poles -$a_p^{(t)}$, this procedure can be repeated iteratively ($t = 1, ..., T$) until it converges to the minimum of the SK cost function. Unstable poles can be flipped into the left half plane before each iteration.

Based on the final converged poles -$a_p^{(T)}$, a linear identification problem is obtained. If the poles are stable, the residues can be estimated in the orthonormal basis

$$\arg\min_{C} \sum_{k=0}^{K} \left| H(s_k) - \left(\sum_{p=1}^{P} c_p \phi_p^{(T)}(s) \right) \right|^2 \tag{32}$$

or if unstable poles are allowed, one can resort to the partial fraction basis

$$\arg\min_{C} \sum_{k=0}^{K} \left| H(s_k) - \left(\sum_{p=1}^{P} \frac{c_p}{s_k + a_p^{(T)}} \right) \right|^2 \tag{33}$$

Here, $\phi^{(t)}(s)$ represent the orthonormal rational functions $\phi(s)$, based on the calculated poles of iteration t. Both model representations are easily realizable to state-space.

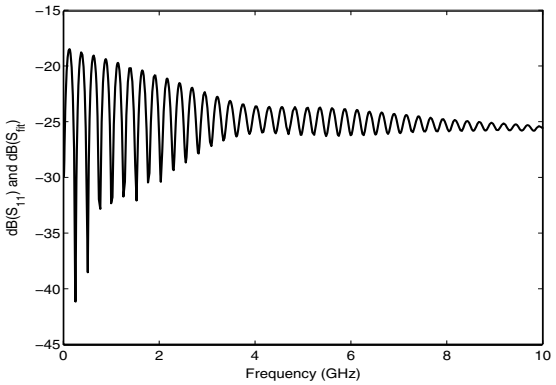

Fig. 1. Magnitude Lossy Coupled Lines (S11).

3. EXAMPLE

The reflection coefficient S_{11} of 2 symmetric coupled dispersive striplines (length = 13000 mil, width = 7.501 mil, spacing = 9.502 mil, thickness = 0.36 mil, conductivity = 5.8*10^7 S/m), laying in between 2 lossy substrate layers (substrate1 : height = 13.9 mil, ε_r = 4.2, tg δ = 0.024 & substrate2 : height = 5.24 mil, ε_r = 2.6, tg δ = 0.004), are modeled using the proposed technique. Figure 1 shows the magnitude of the spectral response over the frequency range of interest [50 Hz - 10 GHz]. First, a prescribed set of complex conjugate starting poles is chosen as was proposed by (Gustavsen and Semlyen, 1999)

$$-a_p = -\alpha + \beta i, -a_{p+1} = -\alpha - \beta i \tag{34}$$

$$\alpha = \beta/100 \tag{35}$$

with imaginary parts β covering the frequency range of interest. Other prescribed pole-location schemes are also possible, however they often require more iterations before the poles are converged. The weighted linear cost function (4) is solved using the orthonormal rational basis functions (16)(17)(18) and an estimate for the residues c_p and $\tilde{c}_p$ is obtained. Using the residues $\tilde{c}_p$ and the poles $-a_p$, the minimal state-space realization (A, B, C, D) of the denominator $D(s)$ (21) is calculated. From this state-space model, the poles of the transfer function are calculated by solving the eigenvalues of $A - BC$. These poles are chosen as new starting poles, and the method iterates until the poles are converged to their optimal location. Once the poles are known, the residues of the transfer function c_p can be estimated by solving (32) or (33).

In this example, the number of poles was set equal to 86, and the model is approximated in a least squares sense, using 4 SK-iterations. The final accuracy of the model $Error = \max(dB(|S_{11} - S_{fit}|))$ corresponds to -63 dB, which is quite close to the numerical noise level of the simulator (Agilent EEsof Comms EDA, n.d.). Table 1 shows the corresponding condition numbers of the fit.

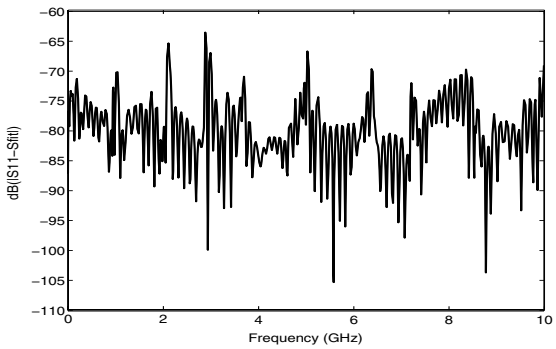

Fig. 2. Error fitting model.

Note that the condition numbers of the pole identification in the 3^{rd} and 4^{rd} iteration are similar, because they converge as the iterated poles converge to their optimal location.

Table 1. Condition number fit.

Nr. iter	Cond.Nr	
1	4.9439 x 10^{11}	Pole Ident
2	2.0255 x 10^{8}	Pole Ident
3	5.8277 x 10^{4}	Pole Ident
4	5.4260 x 10^{4}	Pole Ident
	2.2332 x 10^{3}	Res. Ident

In (Van Gucht and Bultheel, 2003), an orthonormal rational basis is proposed, which generalizes the 3-term recurrence relation for orthonormal polynomials. This approach minimizes the condition number of the system equations, by choosing an inner product which matches that of the cost function. It should be noted that the quality of a fit is not solely dependent on its condition number. In fact, the ill-conditioning is now hidden in the calculation of the recurrence coefficients, which deteriorates the fitting accuracy. Also from a computational perspective, the approach is more expensive.

4. CONCLUSION

This paper presents a technique for reliable identification of continuous-time LTI systems which require a large number of poles. It is numerically robust, computationally efficient, and doesn't require a-priori knowledge of the system. The transfer function representation is very lenient towards the accurate extraction of poles and zeros.

REFERENCES

Adcock, J. and R. Potter (1985). A frequency domain curve fitting algorithm with improved accuracy. *Proc. 3rd Int. Modal Analysis Conf.* **1**, 541–547.

Agilent EEsof Comms EDA (n.d.). ADS Momentum Software, Agilent Technologies, Santa Rosa, CA.

Akcay, H. and B. Ninness (1999). Orthonormal basis functions for modelling continuous-time systems. *Signal Processing* **77 (1)**, 261–274.

Clement, P.R. (1982). Laguerre functions in signal analysis and parameter identification. *J. of the Franklin Institute* **313**, 85–95.

Deschrijver, D. and T. Dhaene (2005). Broadband macromodelling of passive components using orthonormal vector fitting. *IEE Electronics Letters* **41 (6)**, 299–300.

Golub, G.H. and C.F. Van Loan (1989). *Matrix Computations, 2nd edition*. The John Hopkins University Press.

Gomez, J.C. (1998). Analysis of dynamic system identification using rational orthonormal bases. *Phd Thesis, University of Newcastle*.

Gustavsen, B. and A. Semlyen (1999). Rational approximation of frequency domain responses by Vector Fitting. *IEEE Trans. on Power Delivery* **14 (3)**, 1052–1061.

Heuberger, P.S.C., P.M.J. Van Den Hof and B. Wahlberg (2005). *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer-Verlag London.

Kalman, R.E. (1958). Design of a self optimizing control system. *Trans. ASME* **80**, 468–478.

Kautz, W.H. (1954). Transient synthesis in the time-domain. *IRE Trans. Circuit Theory* **1**, 29–39.

Knockaert, L. (2001). On orthonormal muntz-laguerre filters. *IEEE Trans. on Signal Processing* **49 (4)**, 790–793.

Levi, E.C. (1959). Complex curve fitting. *IEEE Trans. on Automatic Control* **AC-4**, 37–43.

Malmquist, F. (1926). Sur la dtermination d'une classe de fonctions analytiques par leurs valeurs dans un ensemble donn de points. *Compte Rendus de Sixieme Congres de Mathematiciens Scandinaves* pp. 253–259.

Ninness, B., S. Gibson and S.R. Weller (2000). Practical aspects of using orthonormal system parameterisations in estimation problems. *12th IFAC Symposium on System Identification*.

Oliveira e Silva, T. (1995). Rational orthonormal functions on the unit circle and on the imaginary axis, with applications in system identification.

Pintelon, R. and J. Schoukens (2001). *System Identification : A Frequency Domain Approach*. IEEE Press. Piscataway, NJ (USA).

Pintelon, R., P. Guillaume, Y. Rolain, J. Schoukens and H.V. Hamme (1994). Parametric identification of transfer functions in the frequency domain - a survey. *IEEE Trans. on Automatic Control* **39 (11)**, 2245–2260.

Richardson, M. and D.L. Formenti (1982). Parameter estimation from frequency response measurements using rational fraction polynomials. *Proc. 1st Int. Modal Analysis Conf.* pp. 167–182.

Rolain, Y., R. Pintelon, K.Q. Xu and H. Vold (1995). Best conditioned parametric identification of transfer function models in the frequency domain. *IEEE Trans. on Automatic Control* **AC-40 (11)**, 1954–1960.

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

Takenaka, S. (1925). On the orthogonal functions and a new formula of interpolation. *Japanese J. of Math.* pp. 129–145.

Van Gucht, P. and A. Bultheel (2003). Orthogonal rational functions for system identification : Numerical aspects. *IEEE Transactions on Automatic Control* **48 (4)**, 705–709.

Wahlberg, B. and P. Makila (1996). On approximation of stable linear dynamical systems using laguerre and kautz functions. *Automatica* **32**, 693–708.

Whitfield, A.H. (1987). Asymptotic behavior of transfer function synthesis methods. *Int. J. of Control* **45**, 1083–1092.