Using Orthogonal Rational Functions for System Identification

Patrick Van gucht, Adhemar Bultheel

Abstract—Recently, there has been a growing interest in the use of orthogonal rational functions (ORF) in system identification. There are many advantages over more classical techniques. Probably due to a known explicit expression for the basis functions when the orthogonality weight is uniformly equal to 1 (the so called Malmquist basis), the attention has been on the development of methods using this basis. However, for some discrete identification problems, this is not the appropriate basis to use: not from a numerical viewpoint and not from the viewpoint of fast recursive updating procedures. In this note we give an algorithm based on general ORF to overcome these problems.

Keywords—orthogonal rational functions, system identification, least squares

I. INTRODUCTION

The use of orthogonal rational functions (ORF) in system identification, has been advertised by many authors in the recent literature. The advantages of representing a transfer function as a linear combination of ORF are numerous. However, in our opinion, some aspect has not been given proper attention. It concerns the choice of the appropriate inner product for the orthogonality. The ORF that are used in the literature are usually orthogonal with respect to a uniform measure on the complex unit circle or the real line.

For example, in the polynomial case, the polynomials orthogonal with respect to the uniform Lebesgue measure are simply the monomials $\{z^k : k = 0, 1, \dots\}$ (Fourier basis), while for a general measure it results in more general Szegő polynomials. The computation of the monomials is trivial since they are explicitly known. The general orthogonal polynomials with respect to an arbitrary measure are however also simple to compute. Given the moments of the orthogonality measure (e.g., the autocorrelation coefficients of the impulse response), they can be efficiently computed by the Levinson or Schur algorithm. The use of the Fourier basis is inappropriate in such cases. It does not allow a recursive computation (updating an AR model of degree n to degree $n + 1$ requires the recomputation of all the coefficients with respect to the Fourier basis) and it may result in a very bad numerical conditioning of the problem, which is manifested by very big rounding errors that may overrule the exact results completely.

Exactly the same situation occurs in the case of ORF. If the Lebesgue measure is used, an explicit form for the ORF is known. This is what Walsh has called the Malmquist basis (see below). The advantage of knowing the basis

functions explicitly is completely overshadowed by the disadvantages from a computational and numerical viewpoint. Not only is the recursive computation of the model not nested with respect to the degree, but there can be a considerable loss of accuracy due to rounding errors so that results become totally worthless for certain problem settings like for example frequency data in a narrow band and/or a high degree model.

In this note we want to show that by choosing an appropriate weight function in discrete time frequency domain identification, the ORF can still be computed in an efficient way, at the same time improving the numerical conditioning to the best possible one.

Of course, given the same poles of the system and given the same objective function, the optimal model that we compute will theoretically be exactly the same as the one computed by any other method. It may also be that with respect to other optimality criteria, our computed solution is worse than the one computed by another method. We do not want to enter a polemic of what an optimality criterion should be, or whether computations should be time or frequency domain based. Our main point is that there is no reason why, from a numerical point of view, one should restrict the discussion to the Malmquist basis. Although the theoretical analysis of the more general basis can be considerably more complicated, the numerics are not, so that if another orthogonality weight is more appropriate, then it should be used. We illustrate our point with a very simple case of discrete frequency domain identification, but the idea is applicable in many other situations as well.

We will consider a discrete time system ($t \in \{0, \pm 1, \pm 2, \dots\}$) with input $\{u(t)\} \in \ell^2$ and output $\{y(t)\} \in \ell^2$ and with transfer function G . The z -transform is denoted by capitals and it is defined using the $1/z$ convention. For example the z -transform of the input is $U(z) = \sum_k u(k)z^{-k}$. We shall estimate G as $\hat{G}_n = \sum_{k=0}^n \theta_k \phi_k$ where the $\{\theta_k\}$ are the expansion coefficients to be determined and the $\{\phi_k\}$ form an ORF basis. We want to find $\{\theta_k\}_{k=0}^n$ from frequency domain data. Thus we suppose that $U(z)$ and $Y(z)$, or equivalently $G(z)$, are known for a certain set of measurement points $\{z_k = e^{i\omega_k}\}$ on the unit circle.

Note that because of our z -transform convention, the transfer function $G(z) = Y(z)/U(z)$ is a function that is naturally expressed in terms of the variable z^{-1} rather than in the variable z . The most simple orthogonal functions are the negative powers of z (polynomials in z^{-1}). These are orthogonal with respect to the inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} f(e^{i\omega}) \overline{g(e^{i\omega})} d\lambda$, with $d\lambda = \frac{1}{2\pi} d\omega$ the (normalized) Lebesgue measure on the unit circle.

The simplest form of non-polynomial ORF are the *Laguerre functions* obtained by orthogonalisation of the basis $\{\zeta^k : k = 0, 1, 2, \dots\}$ with $\zeta = \zeta(z) = (z - a)/(1 - az)$, used by Lee [1] as early as 1925, but in the context of identification only in the 80's by Wahlberg [2], [3] who also introduced the two-parameter *Kautz models* where not one, but a pair of complex conjugate poles is used. In the 90's Van den Hof, Heuberger et al. introduced a finite num-

Patrick Van gucht and Adhemar Bultheel are with the Department of Computer Science, K.U.Leuven, Belgium. Their work is partially supported by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Sciences, Technology and Culture. The scientific responsibility rests with the authors. E-mail of corresponding author: Patrick.Vangucht@cs.kuleuven.ac.be

ber of poles that are repeated cyclically [4], [5], [6]. They also give a state space description. Shortly after, Ninness, Bokor and co-workers in [7], [8], [9] use an arbitrary sequence of poles, which results in the Malmquist basis [10, pp. 224]. For a state space description see [11].

The orthogonality considered in all these papers is with respect to the Lebesgue measure. A theory of ORF with respect to a more general inner product of the form $\langle f, g \rangle_\mu = \int f(z) \overline{g(z)} d\mu$ where μ is a finite positive measure on the unit circle has been given in [12].

II. SPACES OF RATIONAL FUNCTIONS

The unit circle in $\mathbb{C}$, its exterior and interior are denoted by $\mathbb{T}, \mathbb{E}, \mathbb{D} = \{z \in \mathbb{C} : |z| =, >, < 1\}$.

Let $\underline{\alpha} = \{\alpha_1, \alpha_2, \dots\} \subset \mathbb{D}$ be an arbitrary sequence of preselected poles. The space of models of degree n at most is denoted as

$$\mathcal{L}_n = \left\{ \frac{p_n(z)}{\pi_n(z)} : p_n \in \Pi_n, \pi_n(z) = \prod_{k=1}^n (z - \alpha_k) \right\},$$

where Π_n denotes the space of polynomials of degree less than or equal to n .

Let μ be a positive measure on $\mathbb{T}$. An inner product is defined by

$$\langle f, g \rangle_\mu = \int f(e^{i\omega}) \overline{g(e^{i\omega})} d\mu, \quad f, g \in L^2(\mu). \quad (1)$$

The orthonormal basis functions $\{\phi_0, \phi_1, \dots\}$ for this inner product satisfy $\langle \phi_k, \phi_l \rangle_\mu = \delta_{kl}$, with δ_{kl} the Kronecker delta, and are such that $\mathcal{L}_n = \text{span}\{\phi_0, \phi_1, \dots, \phi_n\}$, $n = 0, 1, \dots$. They may e.g. be obtained by a Gram-Schmidt procedure applied to the basis of Blaschke products $B_n(z) = \prod_{l=1}^n \zeta_l(z)$, $n = 1, 2, \dots$ ($B_0 = 1$) with Blaschke factors $\zeta_l(z) = (1 - \overline{\alpha_l}z)/(z - \alpha_l)$. Note that if all $\alpha_k = 0$, then the space $\mathcal{L}_n$ reduces to Π_n , the space of all polynomials in z^{-1} of degree less than or equal to n .

For convergence reasons it is important that the ultimate space $\mathcal{L} = \cup_{n \geq 0} \mathcal{L}_n$ is dense in $\mathcal{H}^2(\mathbb{E})$, the Hardy space of all functions, which are square integrable on $\mathbb{T}$ and analytic in $\mathbb{E}$. This is true provided the Blaschke condition holds.

Theorem 1: The space $\mathcal{L}$ is dense in $\mathcal{H}^2(\mathbb{E})$, iff the Blaschke condition $\sum_{k=1}^\infty (1 - |\alpha_k|) = \infty$ is satisfied. For a proof when the poles are all in $\mathbb{E}$, see [12]. The adaptation to the present situation is trivial.

Note however, that density is not guaranteed for the norm $\|\cdot\|_\mu$ where μ is a more general (non-negative) measure. Therefore one needs additionally that the Szegő condition for the measure is satisfied: $\log \mu'$ should be Lebesgue integrable over the unit circle. This is not satisfied for example if μ is only supported on part of the unit circle. We can not expect any convergence where the weight is zero.

The Malmquist basis is obtained when $\mu = \lambda$, the normalized Lebesgue measure [10, pp. 224].

Theorem 2 (Malmquist basis) If μ is the normalized Lebesgue measure, then the orthonormal basis is given by

$$\phi_0 = 1, \quad \phi_k(z) = \frac{\sqrt{1 - |\alpha_k|^2} z}{z - \alpha_k} B_{k-1}(z), \quad k = 1, 2, \dots$$

This basis is used e.g., in [7] and several other places. For a state space based proof see [11].

III. PROBLEM SETTING

Let $G \in \mathcal{H}^2(\mathbb{E})$ be the transfer function of a system so that in the z -transform domain $Y(z) = G(z)U(z)$. We also suppose that the input system is causal: $U(z) = \sum_{k \geq 0} u(k)z^{-k} \in \mathcal{H}^2(\mathbb{E})$. We want to identify the system given frequency domain data.

If the Blaschke condition is satisfied, then the transfer function G has a unique series expansion

$$G(z) = \sum_{k=0}^{\infty} \theta_k \phi_k(z), \quad z \in \mathbb{E}, \quad (2)$$

converging in $\mathbb{E}$ and whose non-tangential limit for z approaching $\mathbb{T}$ exists a.e. For convergence results of these general rational Fourier series, see [13].

The model poles are of course a set of nonlinear parameters that should be chosen in an optimal way. This is a highly nonlinear and nontrivial problem, that we do not discuss in this paper. However, when they are considered as the parameters of a constrained (to enforce e.g., stability or complex conjugate pairs) nonlinear least squares problem, the coefficients are linear parameters that should be computed in each iteration as a function of the nonlinear parameters in a linear least squares subproblem. The latter subproblem is the one considered in this paper.

It is a well known fact that if the ϕ_k are orthogonal, then the least squares approximant of a prescribed degree is given by a truncation of the series expansion (2) after the term of degree n . Thus we are trying to identify $\underline{\theta} = [\theta_0, \dots, \theta_n]^T$ in the model

$$\hat{G}_n(z, \underline{\theta}) = \sum_{k=0}^n \theta_k \phi_k(z). \quad (3)$$

Note that the approximation $\hat{G}_n$ of the transfer function G is proper, but not strictly proper. A strictly proper transfer function can be modelled by redefining $\mathcal{L}_n$ as the span of the Blaschke products $B_k(z)/z$ and some other small changes to the theory, but we are not going into detail here.

Given the experimental frequency data $\{G(z_k) = Y(z_k)/U(z_k)\}_{k=1}^N$, we want to minimize $\sum_{k=1}^N |E_n(z_k, \underline{\theta})|^2$ where

$$E_n(z, \underline{\theta}) = Y(z) - \hat{Y}(z, \underline{\theta}) = Y(z) - \sum_{k=0}^n \theta_k \phi_k(z) U(z). \quad (4)$$

The z_k are measurement points on the unit circle.

IV. THE MEASURE μ

For simplicity of the presentation, we assume that we want to minimize the unweighted least squares norm $\|\mathbf{E}\| = (\mathbf{E}^H \mathbf{E})^{1/2}$ of the vector $\mathbf{E} = [E_n(z_1, \underline{\theta}), \dots, E_n(z_N, \underline{\theta})]^T$, $z_k = e^{i\omega_k}$, but everything goes trough if we use a norm

e.g., based on the covariance matrix of the measurement errors if this would be available.

That this norm is not weighted may falsely suggest that the Malmquist basis of Theorem 2 (orthogonal with respect to a uniform measure) is the appropriate basis for solving the unweighted least squares problem. However

Theorem 3: The appropriate ORF basis for minimizing $\|\mathbf{E}\|$ with $\mathbf{E}$ as given above should be orthogonal to a discrete measure taking the value $|U(z_k)|^2$ at the points $z_k \in \mathbb{T}$.

Proof: Recall the definition of $\mathbf{E}$ and define $\mathbf{G} = [G(z_1), \dots, G(z_N)]^T$ and $\mathbf{Y} = [Y(z_1), \dots, Y(z_N)]^T$, etc. Then

$$\begin{aligned} \|\mathbf{E}\|^2 &= \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 = \sum_{k=1}^N \left| \left(G(z_k) - \hat{G}_n(z_k, \underline{\theta}) \right) U(z_k) \right|^2 \\ &= \|G - \hat{G}_n(\cdot, \underline{\theta})\|_\mu^2, \end{aligned}$$

where μ is as claimed. $\blacksquare$

Note that the function $|U(z)|^2$, with $z \in \mathbb{T}$ is nothing else but the spectrum of the input signal. Thus if the spectrum of the input signal is flat, then the weighting will not be of great importance, but if it is not flat, then one can expect numerical problems.

Example 1: Choose as basis the Laguerre functions with pole $\alpha = 0.5$. For the input $u(t) = 1$, $t = 1, \dots, 20$ and frequency data equidistant in the interval $[-\pi/6, \pi/6]$, the least squares problem for an approximation of order 10, has a condition number of order $O(10^9)$. This means that we loose approximately 9 decimal digits in the computation of $\underline{\theta}$. This simple example has a very oscillating spectrum, which results in the numerical instability.

Thus using an orthogonal basis with respect to the Lebesgue measure does not solve the numerical problems. A data-dependent measure μ , should be used instead. Before we formulate our identification algorithm, we first recall a recursion formula for the computation of the ORF ϕ_k for an arbitrary measure on $\mathbb{T}$.

V. ORTHOGONAL RATIONAL FUNCTIONS, ANALYTIC OUTSIDE THE UNIT DISC

The basic theory of ORF with prescribed poles in $\mathbb{E}$ can be found in [12]. It generalizes the results for Szegő-Levinson polynomials in a natural way. For our purposes we only need the recurrence relation. The translation from the z notation to the $1/z$ notation is trivial. For details see [14].

Theorem 4: Let ϕ_n be the orthonormal ORF with respect to a positive measure μ on $\mathbb{T}$. Define $\phi_n^*(z) = B_n(z)\overline{\phi_n(1/\bar{z})}$, then

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = e_n \frac{z - \alpha_{n-1}}{z - \alpha_n} \begin{bmatrix} 1 & \overline{L_n} \\ L_n & 1 \end{bmatrix} \times \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix},$$

where

$$L_n = - \frac{\left\langle \phi_k, \frac{1 - \overline{\alpha_{n-1}}z}{z - \alpha_n} \phi_{n-1} \right\rangle_\mu}{\left\langle \phi_k, \frac{z - \alpha_{n-1}}{z - \alpha_n} \phi_{n-1}^* \right\rangle_\mu}, \quad \forall k \leq n-1, \quad (5)$$

$$e_n = \left(\frac{1 - |\alpha_n|^2}{1 - |\alpha_{n-1}|^2} \frac{1}{1 - |L_n|^2} \right)^{1/2}. \quad (6)$$

The initial conditions are $\phi_0(z) \equiv \langle 1, 1 \rangle_\mu^{-1/2} \equiv \phi_0^*(z)$.

VI. THE IDENTIFICATION ALGORITHM

Suppose we are given the frequency response of the input and output $\{U(e^{i\omega_k}), Y(e^{i\omega_k})\}_{k=1}^N$ in the frequencies $\omega_k \in [0, 2\pi]$.

We denote $w_k = |U(e^{i\omega_k})|^2$. Furthermore, we define the inner product

$$\langle f, g \rangle_w = \sum_{k=1}^N w_k f(e^{i\omega_k}) \overline{g(e^{i\omega_k})}, \quad f, g \in L^2(w). \quad (7)$$

Recall

$$\begin{aligned} \hat{Y}(e^{i\omega_k}, \underline{\theta}) &= \hat{G}_n(e^{i\omega_k}, \underline{\theta}) U(e^{i\omega_k}) = \sum_{l=0}^n \theta_l \phi_l(e^{i\omega_k}) U(e^{i\omega_k}) \\ &= \sum_{l=0}^n \theta_l \psi_l(e^{i\omega_k}), \quad \text{with } \psi_l(e^{i\omega_k}) = \phi_l(e^{i\omega_k}) U(e^{i\omega_k}). \end{aligned}$$

We introduce the following notation

$$\begin{aligned} \underline{\psi}_l &= [\psi_l(e^{i\omega_1}), \psi_l(e^{i\omega_2}), \dots, \psi_l(e^{i\omega_N})]^T, \\ \underline{\Psi} &= [\underline{\psi}_0 \ \underline{\psi}_1 \ \dots \ \underline{\psi}_n]. \end{aligned}$$

Then $\mathbf{E} = \mathbf{Y} - \underline{\Psi}\underline{\theta}$. The identification problem is to determine $\underline{\theta}$ in a least square sense. By the normal equations $\underline{\theta}^{LS} = (\underline{\Psi}^H \underline{\Psi})^{-1} \underline{\Psi}^H \mathbf{Y}$ but since the ϕ_n are orthogonal with respect to the weights $\{w_k\}$, the matrix $\underline{\Psi}$ is orthogonal so that becomes $\underline{\theta}^{LS} = \underline{\Psi}^H \mathbf{Y}$ or $\theta_k^{LS} = \langle G, \phi_k \rangle_\mu$, which gives us a powerful interactive identification tool.

For frequency domain identification using the Malmquist basis and cyclically repeated poles see e.g. [15]. This basis is however not orthogonal with respect to the discrete inner product of the space in which the norm is taken.

Remark: Similar techniques can be used in the time domain. If the model is $y(t) = \sum_k \phi_k(q)u(t) + d(t)$ where q is the shift operator: $qu(t) = u(t-1)$ and $d(t)$ is zero mean unit white noise, then an estimator is obtained from data in the points $t = t_1, t_2, \dots, t_N$ by minimizing $E_n(t, \underline{\theta}) = y(t) - \sum_{k=0}^n \theta_k \psi_k(t)$ with $\psi_k(t) = \phi_k(q)u(t)$ in least squares norm. Thus replacing the frequency domain quantities by their time domain equivalent, basically the same problems occur with the same type of solution using ORF. However, since now orthogonality is with respect to summation over the t_k which are all on the real line, we need the three-term recurrence relation for the ORF which generalizes the classical three-term recurrence relation for orthogonal polynomials on the real line. For details see chapter 11 of [12].

VII. TWO NUMERICAL EXAMPLES

A. An artificial example

We consider the following transfer function (see [6])

$$G(z) = \frac{.25299z^{-1} - .9724z^{-2} + 1.4282z^{-3} - .9493z^{-4} + .241z^{-5}}{1 - 4.15z^{-1} + 6.8831z^{-2} - 5.6871z^{-3} + 2.3333z^{-4} - .3787z^{-5}},$$

which has poles $.95 \pm .2i$, $.85 \pm .09i$ and $.55$.

As input, we consider a zero-mean unit-variance white noise signal of length $N = 1200$ and frequency data equidistant in $[-\pi/6, \pi/6]$. For the model of degree 20 with poles $.9 \pm .3i$ and $.7 \pm .2i$ repeated 5 times as in [6], the condition number of Ψ is $O(10^{17})$ for the Lebesgue measure and $1 + O(10^{-12})$ for the ORF. The relative error $\|Y - \hat{Y}\|/\|Y\|$ for the ORF turns out to be 63% better!

B. A robotarm

Suppose we are given measured frequency data of the transfer function $\{G(e^{i\omega_k})\}$ and the variances $\{\sigma_k\}$ on these measurements.

We are interested in an approximation $\hat{G}_n$ of G that minimizes the following cost function

$$\sum_{k=1}^N \left| \frac{G(e^{i\omega_k}) - \hat{G}_n(e^{i\omega_k})}{\sigma_k} \right|^2.$$

We can consider σ_k^{-1} as the input and $G(e^{i\omega_k})/\sigma_k$ as the output in our algorithm and search for the least squares solution $\hat{G}_n$.

We have experimental data from a robot arm problem. The length of the signal is $N = 100$. The transfer function and the variances are shown in Figure 1. Note that due to the great uncertainty around the peaks, every approximation will have trouble matching these peaks.

We do not know the exact location of the poles, but the problem has been a test case for several methods which lead to an approximate location of the optimal poles near $-0.0174 \pm 0.9318i$, $0.8768 \pm 0.4763i$ and $0.9733 \pm 0.2265i$.

We use the poles $.9 \pm .3i$ and $.5 \pm .4i$, repeated 7 times, to obtain an approximation of degree 28. The condition number of the system matrix in the Malmquist case is $O(10^{17})$ meaning that it is numerically rank deficient, while in our approach the condition number is almost perfect: $1 + O(10^{-12})$.

The relative error $\|Y - \hat{Y}\|/\|Y\|$ is 0.0835 in the Malmquist case and 0.0644 in the ORF case, which is an improvement of 30%.

In Figure 1 we have plotted the transfer function and the approximations. As expected, the two peaks are not very well approximated.

VIII. CONCLUSIONS

We have given an identification algorithm for the frequency domain. The main advantage of this approach, where we incorporate a weight which depends on the input signal, is that the least squares problem becomes well conditioned and is straightforward. This results in faster algorithms, which are numerically stable.

Fig. 1. Plot of the simulated system G (dotted line), the approximating system found by the Malmquist basis (dashed line) and the one found by our approach (solid line). The lowest graph is a rescaled version of the variance on the measured frequency data.

REFERENCES

- [1] Y. W. Lee, "Synthesis of electrical networks by means of the Fourier transforms of Laguerre functions," *J. Math. Phys.*, vol. 11, pp. 83-113, 1933.
- [2] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans. Automat. Control*, vol. AC-36, no. 5, pp. 551-562, 1991.
- [3] B. Wahlberg, "System identification using Kautz models," *IEEE Trans. Automat. Control*, vol. AC-39, no. 6, pp. 1276-1282, 1994.
- [4] P.S.C. Heuberger, P.M.J. Van den Hof, and O. Bosgra, "A generalized orthogonal basis for linear dynamical systems," *IEEE Trans. Automat. Control*, vol. 40, pp. 451-465, 1995.
- [5] Z. Szabó, J. Bokor, and F. Schipp, "Identification of rational approximate models in H_∞ using generalized orthogonal basis," *IEEE Trans. Automat. Control*, vol. 44, no. 1, pp. 153-158, 1999.
- [6] P.M.J. Van den Hof, P.S.C. Heuberger, and J. Bokor, "System identification with generalized orthogonal basis functions," *Automatica*, vol. 31, pp. 1821-1834, 1995.
- [7] B. Ninness and F. Gustafsson, "A unified construction of orthogonal bases for system identification," *IEEE Trans. Automat. Control*, vol. 42, pp. 515-522, 1997.
- [8] H. Akçay and B. Ninness, "Rational basis functions for robust identification from frequency and time domain measurements," *Automatica*, vol. 34, no. 9, pp. 1101-1117, 1998.
- [9] B. Ninness, "Frequency domain estimation using orthonormal bases," in *Proceedings of the 13th IFAC World Congress, San Francisco 1996, Volume 8*, 1996, pp. 381-386.
- [10] J.L. Walsh, *Interpolation and approximation*, vol. 20 of *Amer. Math. Soc. Colloq. Publ.*, Amer. Math. Soc., Providence, Rhode Island, 3rd edition, 1960, First edition 1935.
- [11] P. Van gucht and A. Bultheel, "State-space realization and orthogonal rational functions," Technical Report TW292, Department of Computer Science, K.U. Leuven, September 1999.
- [12] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad, *Orthogonal rational functions*, vol. 5 of *Cambridge Monographs on Applied and Computational Mathematics*, Cambridge University Press, 1999.
- [13] A. Bultheel and P. Van gucht, "Boundary asymptotics for orthogonal rational functions on the unit circle," *Acta Applicandae Mathematicae*, vol. 61, pp. 333-349, 2000.
- [14] P. Van gucht and A. Bultheel, "Computing orthogonal rational functions analytic outside the unit disc," Technical Report TW312, Department of Computer Science, K.U. Leuven, September 2000.
- [15] D.K. de Vries and P. van den Hof, "Frequency domain identification with generalized orthonormal basis functions," *IEEE Trans. Automat. Control*, vol. AC-43, no. 5, pp. 656-669, 1994.