16th IFAC Symposium on System Identification
The International Federation of Automatic Control
Brussels, Belgium. July 11-13, 2012

Numerically Reliable Frequency-Domain Estimation of Transfer Functions: A Computationally Efficient Methodology

Robbert van Herpen * **Tom Oomen** * **Okko Bosgra** *

* *Eindhoven University of Technology, The Netherlands, Control Systems Technology group, (R.M.A.v.Herpen@tue.nl)*

Abstract: Parametric identification of lightweight motion systems requires solving large weighted least-squares problems. The numerical conditioning of such problems, which determines the solution accuracy, crucially depends on the polynomial basis that is used to formulate the problem. The aim of this paper is to optimize numerical conditioning by constructing a polynomial basis that is orthonormal with respect to a data-dependent inner product. This basis is constructed in a computationally efficient way by exploiting underlying structure of the problem, related to polynomial recurrence relations. Through a confrontation with an industrial system with large dynamical complexity, numerical accuracy and efficiency of the method are confirmed.

1. INTRODUCTION

Parametric identification, see Pintelon et al. [1994] for a survey, typically involves solving a weighted least squares optimization problem. It is well-known that the accuracy of least-squares estimates heavily depends on the numerical conditioning of the underlying system of equations, see, *e.g.*, Golub and Van Loan [1989]. Identification problems where i) the considered frequency grid, ii) the order of the approximant, or iii) the number of system inputs/outputs becomes large, are notoriously ill-conditioned. This is also observed in Wills and Ninness [2008], where the numerical conditioning of the problem is enhanced by discarding the least significant part of the system of equations. An alternative approach to enhancement of numerical conditioning, which retains all available information on system behavior, is to reformulate the identification problem using a judiciously selected coordinate frame.

The polynomial basis in which the identification problem is posed crucially determines the numerical conditioning of the corresponding least-squares problem. Therefor, this basis should be selected with care, bearing in mind the problem-specific data. In Rolain et al. [1994], system identification using a polynomial basis that is orthonormal with respect to a data-dependent inner product is considered. Although enhanced numerical conditioning is observed indeed, the polynomial basis for the numerator and denumerator polynomials are selected individually. Instead, Van Barel and Bultheel [1994] propose an orthonormal *vector polynomial* basis, which yields optimal numerical conditioning of the weighted least squares problem. A computationally efficient way of computing this basis is given as well, based on work of Reichel et al. [1991].

Although data-dependent orthonormal vector polynomials have been studied for a generic class of multivariable systems, see Van Barel and Bultheel [1995], and their potential in parametric identification has been illustrated by means of an experimental study, see Bultheel et al. [2005], actual application of the method is far from trivial. Several key steps that are needed for efficient construction of the polynomial basis, in which underlying structure of the problem related to polynomial recurrence relations needs to be exploited, are obscure in existing literature. Nevertheless, *computational efficiency* is crucial for large problems. The main contribution of this paper is to provide a detailed discussion on how to efficiently construct a data-dependent orthonormal polynomial basis. This enables reliable and fast transfer function identification for highly complex systems.

This paper is organized as follows. In Sect. 2, parametric identification using a data-dependent orthonormal polynomial basis is discussed. In Sect. 3, polynomial recurrence relations are introduced. In Sect. 4, computation of the polynomial recurrence coefficients in matrix form is described. A detailed exposition of the efficient computation of these recurrence coefficients in given in Sect. 5, which forms the main contribution of this paper. Finally, Sect. 6 provides experimental results, whereas conclusion are drawn in Sect. 7.

2. PARAMETRIC MODELING OF SYSTEMS

In this section, least-squares (LS) approximation of the frequency response function (FRF) of a system P_o by a rational transfer function $\hat{P}(z)$ is considered. Throughout, P_o is assumed to be single input, single output (SISO).

Definition 1. Let $z_k = e^{j\theta_k}, k = 1, \dots, m$ be distinct nodes on the unit circle, ordered such that $0 < \theta_1 < \cdots < \theta_m < \pi$. Let $P_o(z_k)$ denote the FRF of the true system. Let $W(z_k)$ denote a given weight sequence. Weighted LS approximation of P_o by a transfer function $\hat{P}(z) = \frac{n(z)}{d(z)}$ is defined as:

$$\min_{\hat{P}} \left\| W \left(P_o - \hat{P}\right) \right\|_2^2 \quad := \tag{1}$$

$$\min_{\{n,d\}} \sum_{k=1}^{m} \left[\left(P_o(z_k) - \frac{n(z_k)}{d(z_k)}\right)^* W(z_k)^* W(z_k) \left(P_o(z_k) - \frac{n(z_k)}{d(z_k)}\right) \right].$$

Optimization problem (1) is nonlinear in the denumerator polynomial $d(z)$, hence, not trivial to solve. A frequently used approach to address this is to use SK iterations.

* This research is supported by ASML, Veldhoven, The Netherlands.

Algorithm 2. Sanathanan and Koerner [1963] Let the polynomials $\{n, d\}^{<k-1>}$ be determined in a previous step. Both polynomials are updated by solving the weighted LS problem:

$$\min_{\{n,d\}^{<k>}} \left\| \tilde{W} \; [P_o \;\; -I] \begin{bmatrix} d^{<k>} \\ n^{<k>} \end{bmatrix} \right\|_2^2, \tag{2}$$

where $\tilde{W} = W/d^{<k-1>}$. ∎

Observe that LS problem (2) is *linear* in the vector polynomial $\left[d^{<k>} \;\; n^{<k>}\right]^T$. Typically, the iterations are initialized by setting $d^{<0>} = 1$.

2.1 Selection of a polynomial basis

To actually solve (2), a polynomial basis is to be selected for $\left[d^{<k>} \;\; n^{<k>}\right]^T$. This step is crucial, since it determines the numerical conditioning of the resulting LS problem. It is well-known that certain standard bases, like, *e.g.*, the monomial basis, can lead to a LS problem that is extremely poorly conditioned, depending on the problem data at hand. In fact, the formulation of a well-conditioned LS problem hinges on the selection of a coordinate frame that explicitly accounts for the particular problem data. Herein, a data-dependent inner product plays an essential role.

Definition 3. Consider m nodes z_k on the unit circle, see Def. 1. Let $w_k = \tilde{W}(z_k)[-P_o(z_k) \;\; I] \in \mathbb{C}^{1\times 2}$ be corresponding weight tuples. For block polynomials $\phi(z), \psi(z) \in \mathbb{C}^{2\times 2}[z]$, the discrete inner product with respect to the given nodes and weights is defined as:

$$\langle \phi(z), \psi(z) \rangle := \sum_{k=1}^{m} \phi(z_k)^H \, w_k^H w_k \, \psi(z_k) \,. \tag{3}$$

By expressing the $\left[d^{<k>} \;\; n^{<k>}\right]^T$ in a polynomial basis that is orthonormal with respect to the inner product in Def. 3, a LS problem with optimal condition number is obtained.

Definition 4. Orthonormal block polynomials (OBPs) are defined as a set of block polynomials $\varphi_j(z) \in \mathbb{C}^{2\times 2}[z]$, $j = 0, \ldots, \ell$, that are orthonormal with respect to the data-dependent inner product in Def. 3, *i.e.*, $\langle \varphi_i, \varphi_j \rangle = \delta_{ij} I_2$. Here, block polynomial $\varphi_j(z)$ is of degree j.

Proposition 5. Bultheel and Van Barel [1995] Let the vector polynomial $\left[d^{<k>} \;\; n^{<k>}\right]^T$ be expressed in OBPs:

$$\begin{bmatrix} d^{<k>} \\ n^{<k>} \end{bmatrix} = \sum_{j=0}^{\ell} \varphi_j(z)\, \theta_j \,, \tag{4}$$

where $\theta_j \in \mathbb{C}^{2\times 1}$ are corresponding polynomial coefficient vectors. Then, the weighted LS problem (2) is cast into a linear system of equations $A\theta = b$ with $\mathrm{cond}(A) = 1$.

Proof: Substituting the OBP basis (4) in (2) yields:

$$\min_{\Theta_{\mathrm{OBP}}} \| A_{\mathrm{OBP}} \, \Theta_{\mathrm{OBP}} \; - \; b_{\mathrm{OBP}} \|_2^2 \tag{5}$$

where $\Theta_{\mathrm{OBP}} = \left[\theta_0^T \;\; \theta_1^T \;\; \ldots \;\; \theta_{\ell-1}^T\right]^T$, and:

$$A_{\mathrm{OBP}} = U \begin{bmatrix} \varphi_0(z_1) & \varphi_1(z_1) & \ldots & \varphi_{\ell-1}(z_1) \\ \varphi_0(z_2) & \varphi_1(z_2) & \ldots & \varphi_{\ell-1}(z_2) \\ \vdots & \vdots & & \vdots \\ \varphi_0(z_m) & \varphi_1(z_m) & \ldots & \varphi_{\ell-1}(z_m) \end{bmatrix}, \tag{6}$$

$$b_{\mathrm{OBP}} = U \; \left[(\varphi_\ell(z_1)\, \theta_\ell)^T \;\; \ldots \;\; (\varphi_\ell(z_m)\, \theta_\ell)^T \right]^T, \tag{7}$$

with

$$U := \mathrm{diag}(\tilde{W}(z_1)\,[-P_o(z_1) \;\; I], \ldots, \tilde{W}(z_m)\,[-P_o(z_m) \;\; I]). \tag{8}$$

By virtue of (3), $A_{\mathrm{OBP}}^T A_{\mathrm{OBP}} = I_{2n}$, *i.e.*, A_{OBP} is inner. Hence, $\mathrm{cond}(A_{\mathrm{OBP}}) = 1$. ∎

Remark: In order to avoid a trivial solution to (2), the highest degree coefficient θ_ℓ is selected such that the denumerator polynomial $d^{<k>}$ is monic.

Since the use of OBPs yields an optimal conditioning of the LS problem, the propagation of numerical round-off errors remains limited, *cf.* [Golub and Van Loan, 1989, Sect. 5.3.7]. Thus, a highly accurate solution to (2) is obtained, even when the number of data points m is large.

2.2 Estimation of real-rational transfer functions

In the modeling of physical systems, capturing a relation between measured *real* system inputs and outputs demands for a transfer function with real coefficients. Therefore, in (2) the polynomials $\{d^{<k>}, n^{<k>}\}$ should have real coefficients. In turn, this demands for basis polynomials $\varphi(z)$ with real coefficients. A well-known property of such polynomials is, Oppenheim and Schafer [1975]:

$$\varphi(z_k^*) = \varphi^*(z_k) \,. \tag{9}$$

The data should be consistent with this property, *i.e.*, the nodes z_k^*, $k = 1, \ldots, m$ with corresponding weight tuples w_k^* should be added to the problem. From (3), it follows that for real $\phi(z), \psi(z)$:

$$\sum_{k=1}^{m} \phi(z_k^*)^H \, (w_k^H w_k)^* \, \psi(z_k^*) \; = \; \langle \phi(z), \psi(z) \rangle^* \,. \tag{10}$$

This motivates the following data-dependent inner product for *real* polynomials.

Definition 6. Consider m nodes z_k with corresponding weights w_k as defined in Def. 3. For *real* block polynomials $\phi(z), \psi(z) \in \mathbb{R}^{2\times 2}[z]$, the discrete inner product with respect to the given nodes and weights is defined as:

$$\langle \phi(z), \psi(z) \rangle := 2\,\mathrm{Re} \left\{ \sum_{k=1}^{m} \phi(z_k)\, w_k^H w_k \, \psi(z_k) \right\} . \tag{11}$$

The remainder of this paper is concerned with the construction of a *data-dependent real OBP basis* for reliable weighted LS transfer function estimation using FRF data.

3. POLYNOMIAL RECURRENCE RELATIONS

The study of orthonormal polynomials has a long-standing history, which can be traced back to the pioneering work of, *e.g.*, Legendre, Chebyshev, Hermite, and Laguerre. It is well-known that classical polynomial bases all satisfy a three-term-recurrence relation. This recurrence relation enables efficient construction of new basis polynomials from given lower order polynomials. Such efficient construction is possible for OBPs with respect to the data-dependent inner product in Def. 6 as well, using *Szegö's recurrence relation*, Szegö [1939]. Generalizations for block-polynomials have been derived in, *e.g.*, Delsarte et al. [1978], Morf et al. [1978], and Youla and Kazanjian [1978]. The particular form in Prop. 7 connects to matrix-representations that are considered in subsequent sections.

Proposition 7. The OBPs $\varphi_j(z)$ in Def. 4 satisfy the three-term-recurrence relations:

$$\varphi_j(z) = (z\varphi_{j-1}(z) \; + \; \tilde{\varphi}_{j-1}(z)\, \Gamma_j) \; \Sigma_j^{-1} \,, \tag{12}$$

$$\tilde{\varphi}_j(z) = (z\varphi_{j-1}(z)\, \Gamma_j^T \; + \; \tilde{\varphi}_{j-1}(z)) \; \hat{\Sigma}_j^{-T} . \tag{13}$$

Here, $\Gamma_j, \Sigma_j, \hat{\Sigma}_j \in \mathbb{R}^{2\times 2}$ are predetermined recursion coefficients. The recursion is initialized with $\varphi_0 = \tilde{\varphi}_0 = \Sigma_0^{-1}$. ■

Note that $\varphi_j(z) \in \mathbb{R}^{2\times 2}[z]$, since the recursion coefficients are real blocks. These coefficients have a particular interpretation that is encountered frequently in literature on orthonormal polynomials.

Proposition 8. Van Barel and Bultheel [1994] The coefficients $\Gamma_j, \Sigma_j, \hat{\Gamma}_j, \hat{\Sigma}_j \in \mathbb{R}^{2\times 2}$ in Prop. 7 are *block Schur parameters*, for which holds:

$$\begin{bmatrix} -\Gamma_j & \hat{\Sigma}_j \\ \Sigma_j & \hat{\Gamma}_j \end{bmatrix}^T \begin{bmatrix} -\Gamma_j & \hat{\Sigma}_j \\ \Sigma_j & \hat{\Gamma}_j \end{bmatrix} = I_4 . \tag{14}$$

Moreover, Σ_j is constrained to be upper triangular with positive diagonal elements. ■

The fundamental indeterminate block-Schur parameter is Γ_j, as the remaining parameters can be derived from it. This is an essential observation that will be used in Sect. 5 to derive an algorithm for the fast reconstruction of the block Schur parameters from data.

Proposition 9. Let Γ_j be given. The upper triangular block Schur parameter Σ_j is defined uniquely by:

$$\Sigma_j = \mathrm{chol}(I_2 - \Gamma_j^T \Gamma_j) , \tag{15}$$

where chol(A) denotes the Cholesky factorization of A. The remaining Schur blocks $\hat{\Gamma}_j$ and $\hat{\Sigma}_j$, which are not unique, are selected in accordance with (14). *Any* singular value decomposition of $\begin{bmatrix} -\Gamma_j^T & \Sigma_j^T \end{bmatrix}$ yields the desired block Schur parameters, since:

$$USV^T := [\, I_2 \,] \; [\, I_2 \mid 0_2 \,] \left[\begin{array}{c|c} -\Gamma_j & \hat{\Sigma}_j \\ \Sigma_j & \hat{\Gamma}_j \end{array} \right]^T = \begin{bmatrix} -\Gamma_j^T & \Sigma_j^T \end{bmatrix} . \tag{16}$$

■

The gist of the construction of a *data-dependent* orthonormal basis is the efficient and reliable derivation of the block Schur parameters from the given data. In the next section, the problem of deriving block Schur parameters from a given set of nodes z_k with corresponding weights w_k is posed as an inverse eigenvalue problem.

4. COMPUTING BLOCK SCHUR PARAMETERS

Numerical algorithms for widespread orthonormal *vector* basis decompositions, including the eigenvalue decomposition, the singular value decomposition, and QR-factorization, rely on matrix zeroing operations, see, *e.g.*, Golub and Van Loan [1989]. In general, these algorithms produce very accurate results, since numerical round-off errors are not amplified in successive zeroing steps. In this section, it is discussed how a similar methodology can be used for the computation of block Schur parameters, *cf.* Def. 8, which are the fundamental indeterminates in the *polynomial* recurrence relations (12)–(13). Hereto, a matrix representation of these recurrence relations is formulated, in which a block-Hessenberg is central.

Definition 10. An upper 2 × 2 - block-Hessenberg matrix $H \in \mathbb{R}^{2m\times 2m}$ satisfies $H_{i,j} = 0 \;\forall\; i \geq j+3$. Moreover, the elements of the 2nd subdiagonal of H are constrained to be strictly positive, *i.e.*, $H_{i+2,\,i} > 0$ for $i = 1, \ldots, 2m-2$.

Now, a particular 2 × 2-block-Hessenberg matrix is constructed, which has the (conjugate) nodes specific to the approximation problem in Def. 1 as its eigenvalues.

Proposition 11. Reichel et al. [1991] Define the node matrix $Z \in \mathbb{C}^{2m\times 2m}$ and weight matrix $W \in \mathbb{C}^{2m\times 2}$ as (*cf.* Def. 3):

$$Z := \mathrm{diag}(z_1,\; z_1^*,\; z_2,\; z_2^*,\; \ldots\;,\; z_m,\; z_m^*) , \tag{17}$$

$$W := \begin{bmatrix} w_1^T \; w_1^H & w_2^T \; w_2^H & \ldots & w_m^T \; w_m^H \end{bmatrix}^T . \tag{18}$$

There exists a unitary matrix $Q \in \mathbb{C}^{2m\times 2m}$ such that:

$$\begin{bmatrix} I_2 & \\ & Q^H \end{bmatrix} \left[\begin{array}{c|c} 0_2 & W^T \\ \hline W & Z \end{array} \right] \begin{bmatrix} I_2 & \\ & Q \end{bmatrix} = \left[\begin{array}{c|c} 0_2 & \Sigma_0^T \;\; 0_{2,2m-2} \\ \hline \begin{array}{c} \Sigma_0 \\ 0_{2m-2,2} \end{array} & H \end{array} \right], \tag{19}$$

where $H \in \mathbb{R}^{2m\times 2m}$ is a 2 × 2 - block-Hessenberg matrix, see Def. 10, and $\Sigma_0 \in \mathbb{R}^{2\times 2}$ is upper triangular and has positive diagonal elements. ■

Note that, since all nodes z_k are taken on the unit circle, Z is unitary. Consequently, the block-Hessenberg matrix H is unitary as well. This enables a decomposition of H in elementary block-Givens-reflectors.

Proposition 12. Reichel et al. [1991] Every *unitary* 2 × 2-block-Hessenberg matrix $H \in \mathbb{R}^{2m\times 2m}$ can be written as:

$$H = G_1 \, G_2 \, \ldots \, G_m , \tag{20}$$

where the j^{th} unitary block-Givens-reflector G_j is defined:

$$G_j \;=\; \begin{bmatrix} I_{2(j-1)} & & & \\ & -\Gamma_j & \hat{\Sigma}_j & \\ & \Sigma_j & \hat{\Gamma}_j & \\ & & & I_{m-2(j+1)} \end{bmatrix} . \tag{21}$$

Here, $\Gamma_j, \Sigma_j, \hat{\Gamma}_j, \hat{\Sigma}_j \in \mathbb{R}^{2\times 2}$ are block Schur parameters, *cf.* Prop. 8. ■

By virtue of the special structure underlying the block-Hessenberg matrix, the block-columns of Q in (19) satisfy a three-term-recurrence relation similar to Prop. 7.

Definition 13. The unitary similarity transformation matrix $Q \in \mathbb{C}^{2m\times 2m}$ in Prop. 11 consists of m block-columns, *i.e.*:

$$Q := [\, Q_1 \mid Q_2 \mid \; \ldots \; \mid Q_m \,], \tag{22}$$

where $Q_j \in \mathbb{C}^{2m\times 2}$, $j = 1, \ldots, m$.

Proposition 14. Van Barel and Bultheel [1994], Bultheel and Van Barel [1995] Let Z, W be defined in (17)–(18). The block-columns Q_j in Def. 13 satisfy the three-term-recurrence relations:

$$Q_{j+1} = (Z \, Q_j \; + \, \tilde{Q}_j \, \Gamma_j) \; \Sigma_j^{-1} , \tag{23}$$

$$\tilde{Q}_{j+1} = (Z \, Q_j \, \Gamma_j^T \; + \, \tilde{Q}_j) \; \hat{\Sigma}_j^{-T} , \tag{24}$$

The recursion is initialized with $Q_1 = \tilde{Q}_1 = W \Sigma_0^{-1}$. ■

Observe that the same block Schur parameters have been used in Prop. 7 and Prop. 14. Indeed, an explicit connection exists between the OBPs $\varphi_j(z)$ in Def. 4 and the block-columns Q_j in Def. 13.

Proposition 15. Van Barel and Bultheel [1994] Define the weight matrix $D \in \mathbb{C}^{m\times 2m}$ as:

$$D := \mathrm{diag}(w_1, w_1^*, \ldots, w_m, w_m^*), \tag{25}$$

with w_k, $k = 1, \ldots, m$ given in Def. 3. Denote with $\Phi_j, \tilde{\Phi}_j \in \mathbb{C}^{2m\times 2}$ the block-columns obtained after evaluation of $\varphi(z), \tilde{\varphi}(z)$ at the considered nodes z_k, *i.e.*:

$$\Phi_j := [\, \varphi_j(z_1)^T,\; \varphi_j(z_1^*)^T,\; \ldots\;,\; \varphi_j(z_m)^T,\; \varphi_j(z_m^*)^T]^T, \tag{26}$$

$$\tilde{\Phi}_j := [\, \tilde{\varphi}_j(z_1)^T,\; \tilde{\varphi}_j(z_1^*)^T,\; \ldots\;,\; \tilde{\varphi}_j(z_m)^T,\; \tilde{\varphi}_j(z_m^*)^T]^T. \tag{27}$$

The block-columns Q_j in Def. 13 and the OBPs $\varphi_j(z)$ in Def. 4 are related as follows:

$$Q_{j+1} = D\,\Phi_j\,, \tag{28}$$

$$\tilde{Q}_{j+1} = D\,\tilde{\Phi}_j\,. \tag{29}$$

■

In conclusion, the block-Schur-parameters $\Gamma_j, \Sigma_j, \hat{\Gamma}_j, \hat{\Sigma}_j$ that constitute the block Hessenberg matrix H, see Prop. 12, are the recursion coefficients in Prop. 7. Hence, after solving the inverse eigenvalue problem in Prop. 11, the OBP basis can be constructed using Szegö's recurrence relations. In the next section, a fast and reliable algorithm is discussed that solves the inverse eigenvalue problem using *matrix zeroing operations.*

5. RELIABLE AND FAST COMPUTATION OF OBPS

It is straightforward to solve the inverse eigenvalue problem in Prop. 11 by means of matrix zeroing operations. In particular, the unitary matrix Q in (19) can be decomposed as a sequence of standard Householder reflections, see, *e.g.*, [Golub and Van Loan, 1989, Sect. 5.1]. Each Householder reflection transforms a subsequent column of the initial node-weight matrix into the desired block-Hessenberg form by enforcing zero elements.

The round-off properties associated with Householder reflections are very favorable, [Golub and Van Loan, 1989, Sect. 5.1.5, Sect.9.2]. Hence, the inverse eigenvalue problem in Prop. 11 is solved with high accuracy. However, $\mathcal{O}(m^2\ell)$ floating point operations (flops) are required, where m is the number of frequency nodes and ℓ the degree of the approximant. When considering a large number of frequencies, the involved computation time turns out to be unacceptably large.

As observed in, *e.g.*, Reichel et al. [1991], Van Barel and Bultheel [1994], Faßbender [1997], it is possible to solve the inverse eigenvalue problem in $\mathcal{O}(m\ell)$ flops. Such efficient construction of OBPs, which is crucial for large problems, can be accomplished by explicitly making use of the underlying structure of the Hessenberg matrix in Prop. 12. Although this concept is presented on an abstract level in existing literature, several algorithmic steps that are essential for actual implementation of the approach in system identification procedures are not available. One particularly important aspect that requires further attention is the construction of block polynomials in *real* coefficients. A main contribution of this paper is to provide a thorough exposition of a fast, reliable algorithm for construction of a real 2×2 OBP basis.

5.1 Enforcing the derivation of real block-Schur parameters

As motivated in Sect. 2.2. it is desired to built up an OBP basis in real coefficients. To this end, the recursion coefficients in Prop. 7 should be real-valued, which holds if $H \in \mathbb{R}^{2m\times 2m}$ indeed, *cf.* Prop. 11. To enforce the construction of a real block-Hessenberg matrix, the fact that complex-conjugate node and weight pairs have been introduced into the problem, *cf.* Sect. 2.2, is used explicitly.

Proposition 16. Let $z_k = e^{j\theta_k}$ and w_k be given in Def. 3. Let $\iota = \sqrt{-1}$ denote the imaginary unit. Define:

$$R_0 := I_m \otimes \frac{1}{\sqrt{2}}\begin{bmatrix}1 & \iota\\ 1 & -\iota\end{bmatrix}, \tag{30}$$

where $\otimes$ denotes the Kronecker product. The initial node-weight matrix in Prop. 11 is transformed into a real matrix under unitary similarity as follows:

$$\left[\begin{array}{c|c}0_2 & W_{\text{real}}^T\\ \hline W_{\text{real}} & Z_{\text{real}}\end{array}\right] = \begin{bmatrix}I_2 & \\ & R_0^H\end{bmatrix}\left[\begin{array}{c|c}0_2 & W^T\\ \hline W & Z\end{array}\right]\begin{bmatrix}I_2 & \\ & R_0\end{bmatrix}, \tag{31}$$

where $W_{\text{real}} \in \mathbb{R}^{2m\times 2}$ and $Z_{\text{real}} \in \mathbb{R}^{2m\times 2m}$ are given by:

$$W_{\text{real}} = \sqrt{2}\,\left[\text{Re}\{w_1^T\}, \text{Im}\{w_1^T\}, \ldots, \text{Re}\{w_m^T\}, \text{Im}\{w_m^T\}\right]^T,$$

$$Z_{\text{real}} = \begin{bmatrix}\begin{bmatrix}\cos(\theta_1) & -\sin(\theta_1)\\ \sin(\theta_1) & \cos(\theta_1)\end{bmatrix} & & \\ & \ddots & \\ & & \begin{bmatrix}\cos(\theta_m) & -\sin(\theta_m)\\ \sin(\theta_m) & \cos(\theta_m)\end{bmatrix}\end{bmatrix}.$$

■

Observe that Z_{real} contains the real part $\cos(\theta_k)$ and imaginary part $\sin(\theta_k)$ of the nodes z_k, $k = 1, \ldots, m$ only, *i.e.*, the matrix is defined by $2m$ real numbers. Hence, considering the conjugate nodes and weights in (17)–(18) does not lead to an increase of the problem complexity.

The initial *real* node-weight matrix (31) can be transformed into a real 2×2 block-Hessenberg matrix using *real* zeroing operations.

5.2 Extending a Hessenberg matrix with node-weight blocks

Instead of manipulating the entire matrix (31) all at once, 2×2 node-weight pairs $(Z_{\text{real},k}, W_{\text{real},k})$, *cf.* Prop. 16, are introduced into the problem one-by-one. Each time, the subset of node-weight blocks that has been considered thus far is transformed into a block-Hessenberg matrix of appropriate size. Algorithm 17 explains how an existing block-Hessenberg matrix can be updated after addition of a new 2×2 node-weight block-pair.

Algorithm 17. Let $T_j \in \mathbb{R}^{2j\times 2j}$ be a predetermined unitary matrix, which transforms the first j node-weight blocks into a real 2×2 block-Hessenberg matrix H_j under similarity. Determine a new unitary matrix $T_{\text{new}} \in \mathbb{R}^{2(j+1)\times 2(j+1)}$ that constitutes the transformation matrix:

$$T := \begin{bmatrix}I_2 & \\ & T_{j+1}\end{bmatrix} = \begin{bmatrix}I_2 & \\ & \begin{bmatrix}I_2 & \\ & T_j\end{bmatrix}\end{bmatrix}\begin{bmatrix}I_2 & \\ & T_{\text{new}}\end{bmatrix} \tag{32}$$

such that:

$$T^T\left[\begin{array}{c|c}0_2 & \left[W_{\text{real},1}^T \ \ldots \ W_{\text{real},j}^T \,\middle|\, W_{\text{real},j+1}^T\right]\\ \hline \left[\begin{array}{c}W_{\text{real},1}\\ \vdots\\ W_{\text{real},j}\\ \hline W_{\text{real},j+1}\end{array}\right] & \left[\begin{array}{ccc|c}Z_{\text{real},1} & & & \\ & \ddots & & \\ & & Z_{\text{real},j} & \\ \hline & & & Z_{\text{real},j+1}\end{array}\right]\end{array}\right]T =$$

$$\begin{bmatrix}I_2 & \\ & T_{\text{new}}^T\end{bmatrix}\left[\begin{array}{c|c}0_2 & W_{\text{real},j+1}^T \quad \Sigma_{0,j}^T \quad 0_{2,2j-2}\\ \hline \begin{array}{c}W_{\text{real},j+1}\\ \Sigma_{0,j}\\ 0_{2j-2,2}\end{array} & \begin{array}{cc}Z_{\text{real},j+1} & \\ & \begin{bmatrix}H_j\end{bmatrix}\end{array}\end{array}\right]\begin{bmatrix}I_2 & \\ & T_{\text{new}}\end{bmatrix}$$

$$= \left[\begin{array}{c|c}0_2 & \Sigma_{0,j+1}^T \quad 0_{2,2j}\\ \hline \begin{array}{c}\Sigma_{0,j+1}\\ 0_{2j,2}\end{array} & \begin{bmatrix}H_{j+1}\end{bmatrix}\end{array}\right]. \tag{33}$$

Here, $H_{j+1} \in \mathbb{R}^{2(j+1)\times 2(j+1)}$ is a block-Hessenberg matrix with eigenvalues $(z_1, z_1^*, \ldots, z_{j+1}, z_{j+1}^*)$ and $\Sigma_{0,j+1}$ is an upper triangular matrix with positive diagonal elements.

Initialization: $T_1 = I_2$ $(H_1 = Z_{\text{real},1},\ \Sigma_{0,1} = W_{\text{real},1})$. ■

The matrix T_{new} in Algorithm 17 consists of a judiciously chosen sequence of zeroing operations through Givens reflections, *cf.* [Golub and Van Loan, 1989, Sect. 5.1].

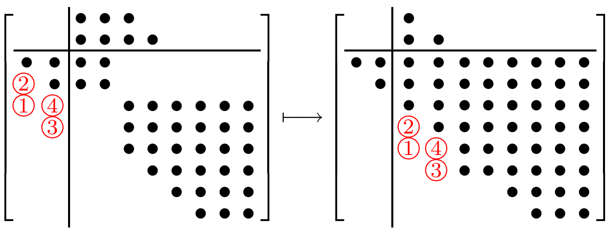
Proposition 18. Consider a matrix $M \in \mathbb{R}^{k\times k}$. Element $M_{i,j}$, where $i \geq j+2$, can be zeroed by applying the similarity transformation GMG, where the symmetric Givens reflector G is defined as:

$$G = \begin{bmatrix} I_{i-2} & & & \\ & -\alpha & \beta & \\ & \beta & \alpha & \\ & & & I_{k-i} \end{bmatrix}, \begin{cases} \alpha = -M_{i-1,j}/\sqrt{M_{i-1,j}^2 + M_{i,j}^2}, \\ \beta = M_{i,j}/\sqrt{M_{i-1,j}^2 + M_{i,j}^2}. \end{cases} \quad (34)$$

■

Observe from the second term of (33) that there are 4 elements that break the desired block-Hessenberg matrix, *viz.* element (2, 1) of $W_{\text{real},j+1}$ and the nonzero elements of $\Sigma_{0,j}$. Indeed, these elements can be zeroed by applying Givens reflections. As a byproduct, however, nonzero elements appear below the second subdiagonal in the next 2 columns of the matrix, which in turn should be zeroed using Givens reflections. This strategy, known as *chasing down the diagonal*, see Rutishauser [1963], is illustrated below.

Example 19. Consider a given 2 × 2 block-Hessenberg matrix H_3, which is extended with a new node-weight block. The initial matrix in step 4 of Algorithm 17, *cf.* second term of (33), has the form as shown on the left:


(35)

By means of 4 Givens reflections, the indicated matrix elements are zeroed under similarity. This yields the following matrix on the right of (35). Indeed, 4 *new* undesired elements are introduced on the 3rd and 4th subdiagonal in column 3–4, which in turn need to be *chased down* by means of new Givens reflections. ■

Note that in each chasing step, see (35), 4 rows and columns of the matrix at hand are modified only. This observation is used to carry out Algorithm 17 more efficiently.

5.3 Updating in Schur parametric form

The crucial step for efficient solving of the inverse eigenvalue problem in Prop. 11 is to perform extension of a block-Hessenberg matrix with a new node-weight block, as described in Algorithm 17, in *Schur parametric form.* Recall that every unitary block-Hessenberg matrix can be characterized by means of its block-Schur parameters, see Prop. 12. It is key to exploit this property explicitly.

Goal 20. Let $H_j = G_1 G_2 \dots G_j$ be a given unitary block-Hessenberg in Schur parametric form, see Prop. 12. Let $H_{j+1} = \bar{G}_1 \bar{G}_2 \dots \bar{G}_j \bar{G}_{j+1}$ denote the new Schur parametric form of the block-Hessenberg matrix that is obtained after extension of H_j with a new node-weight block pair $(Z_{\text{real},j+1}, W_{\text{real},j+1})$, *cf.* Algorithm 17. Determine:

$$(G_1, G_2, \dots, G_j) \longmapsto (\bar{G}_1, \bar{G}_2, \dots, \bar{G}_j, \bar{G}_{j+1}) \quad (36)$$

without manipulating the entire Hessenberg matrix. ■

As mentioned, in each chasing step, see (35), 4 rows and columns of the matrix at hand are modified only. Hence, one *single* block-Givens-reflector is updated in each step.

Proposition 21. [Van Barel and Bultheel, 1995, Sect. 10] Denote with $\hat{H}_{j+1}^{(i)}$, $i = 1, \dots, j+1$ the intermediate matrices that, through application of j chasing steps, lead to the extended block-Hessenberg matrix H_{j+1}. For example, chasing step 1 yields for the lower-right block of (35): $\hat{H}_{j+1}^{(1)} \longmapsto \hat{H}_{j+1}^{(2)}$. The matrix $\hat{H}_{j+1}^{(i)}$ can be decomposed as:

$$\hat{H}_{j+1}^{(i)} = \bar{G}_1 \bar{G}_2 \dots \bar{G}_{i-1} E_{i-1} G_i G_{i+1} \dots G_j, \quad (37)$$

where E_{i-1} has a nontrivial 6 × 6 subblock that contains the undesired matrix elements that are to be chased in a subsequent step.

Initialization: $E_0 = \operatorname{diag}(\operatorname{diag}(Z_{\text{real},j+1}, I_4), I_{2(j-2)})$. ■

The gist of efficiently performing a chasing step $\hat{H}_{j+1}^{(i)} \longmapsto \hat{H}_{j+1}^{(i+1)}$ is to determine the mapping $E_{i-1} G_i \longmapsto \bar{G}_i E_i$, which has a nontrivial 8 × 8 part that directly follows from the 4 Givens reflections that need to be performed.

Remark 22. Having determined $\bar{G}_i E_i$, it is straightforward to isolate the term E_i for the next chasing step. Namely, block $([2j-1, 2j], [2j-1, 2j])$ of $\bar{G}_i E_i$ is equal to the block Schur parameter $\bar{\Gamma}_i$. Using Prop. 9, $\bar{\Sigma}_i$, $\hat{\bar{\Gamma}}_i$, and $\hat{\bar{\Sigma}}_i$ can be determined, which characterize $\bar{G}_i$, *cf.* Prop. 12. Since this block-Givens-reflector is unitary, $E_i = \bar{G}_i^T (\bar{G}_i E_i)$. ■

To summarize, instead of deriving block-Schur parameters from a Hessenberg matrix obtained by manipulating the entire node-weight *matrix* in (31), the *sequence of* 2 × 2 *block-Schur parameters* is updated directly each time a new node-weight block is introduced. This improves the efficiency of the algorithm from $\mathcal{O}(m^2\ell)$ to $\mathcal{O}(m\ell)$. Given the block-Schur parameters, an OBP basis can be determined from Prop. 7, which in turn enables reliable solution of the LS problem in Algorithm 2, *cf.* Prop. 5.

6. EXPERIMENTAL EXAMPLE

In this section, the algorithm described in this paper is applied to parametric identification of an industrial motion system. In particular, a lightweight mechanical positioning device is considered. Typically, lightweight systems show structural deformations upon application of large acceleration forces. Since these deformations are generally not aligned with the motion degrees of freedom of the system, the behavior of the system is intrinsically multivariable. A systematic way of designing high-performance controllers for multivariable systems is model-based control design. Clearly, the succes of this control strategy crucially depends on the capability to construct highly accurate models of the system.

Identifying accurate models of a lightweight system is challenging. The reason for this is threefold. Associated with structural deformations of the system are lightly damped resonance phenomena. i) In order to detect these resonances, a FRF P_o of the system has to be identified on a large number of frequencies ii) A large model order is required to be able to include resonances in the model. iii) Multivariable models are essential. For clarity of the exposition, however, in this paper a *single* motion degree of freedom of the system is considered only.

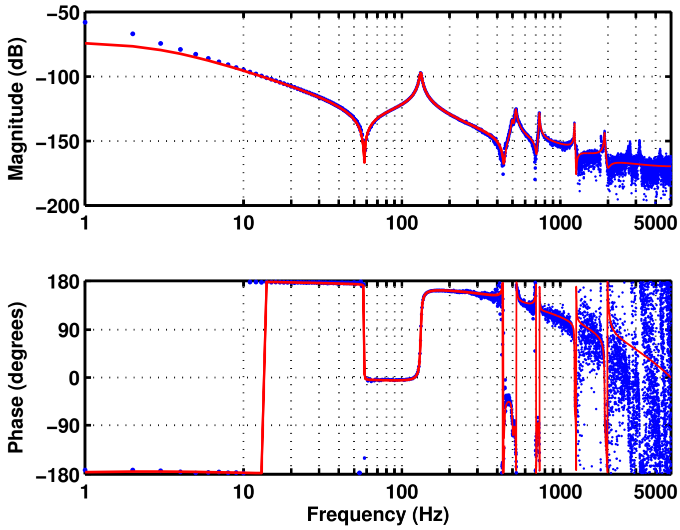


Fig. 1. Measured $m = 5000$ point FRF P_o (dotted) and fitted $\ell = 15^{\text{th}}$ order nominal model $\hat{P}$ (solid).

Figure 1 shows a parametric model $\hat{P}$ of order 15 that has been fitted to a 5000 point FRF of the system using *control-relevant* identification, see Oomen and Bosgra [2008]. The heart of the algorithm proposed therein is a weighted LS optimization problem, in which the weights are inhereted from a subsequent model-based control design step. Using the approach described in this paper, a highly accurate parametric model is obtained indeed, that accurately described the rigid-body behavior, 6 resonance phenomena and phase delay of the system, see Fig. 1. Moreover, this model has been obtained in less than 5 minutes using a first implementation on a desktop pc.

Finally, the above-mentioned results are put in perspective. It has also been tried to solve the weighted least squares problem using in a conventional polynomial power basis. However, resulting condition numbers of 10^{30} were not exceptional, hence, no reliable fit could be obtained. This underlines the need to use an OBP basis. However, when constructing this basis without using the underlying structure of the problem, a computation time of several hours resulted when selecting a subset of the considered FRF points. Even worse, memory problems prevented the actual solution of the full problem considered here.

7. CONCLUSIONS

This paper provides a thorough description of the algorithmic steps that are needed to perform parametric system identification using a data-dependent orthonormal polynomial basis. Herewith, it extends the existing literature on the subject, in that it elaborates on specific steps that are crucial for actual implementation of the proposed polynomial basis in system identification procedures. In particular, this paper provides a detailed derivation of the construction of *block-polynomials in real coefficients*, which is an essential step towards modeling of physical systems using real-rational transfer functions. Using the methodology described in this paper, a highly accurate parametric model of a lightweight motion system is obtained successfully. The capability of building such highly accurate models in a computationally efficient way forms a cornerstone for successful model-based control design.

REFERENCES

A. Bultheel and M. Van Barel. Vector orthogonal polynomials and least squares approximation. *SIAM Journ. Matrix Analysis and Applications*, 16(3):863–885, 1995.

A. Bultheel, M. Van Barel, Y. Rolain, and R. Pintelon. Numerically robust transfer function modeling from noisy frequency domain data. *IEEE Trans. Automatic Control*, 50(11):1835–1839, 2005.

P. Delsarte, Y. V. Genin, and Y. G. Kamp. Orthogonal polynomial matrices on the unit circle. *IEEE Trans. Circuits and Systems*, 25(3):149–160, 1978.

H. Faßbender. Inverse unitary eigenproblems and related orthogonal functions. *Numerische Mathematik, Springer-Verlag*, 77:323–345, 1997.

G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, USA, 1989.

M. Morf, A. Vieira, and T. Kailath. Covariance characterization by partial autocorrelation matrices. *The Annals of Statistics*, 6(3):643–648, 1978.

T. Oomen and O. Bosgra. Robust-control-relevant coprime factor identification: A numerically reliable frequency domain approach. In *Proc. American Control Conference, Seattle, WA, USA*, pages 625–631, 2008.

A. V. Oppenheim and R. W. Schafer. *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, New Jersey, NJ, USA, 1975.

R. Pintelon, P. Guillaume, Y. Rolain, J. Schoukens, and H. Van hamme. Parametric identification of transfer functions in the frequency domain—A survey. *IEEE Trans. Automatic Control*, 39(11):2245–2260, 1994.

L. Reichel, G. S. Ammar, and W. B. Gragg. Discrete least squares approximation by trigonometric polynomials. *Mathematics of Computation*, 57(195):273–289, 1991.

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. *Proc. 33rd Conference on Decision and Control, Lake Buena Vista, FL, USA*, pages 3365–3373, 1994.

H. Rutishauser. On jacobi rotation patterns. *Experimental Arithmetic, High Speed Computing and Mathematics, Proc. Symp. in Applied Math., American Mathematical Society, Providence, R.I., USA*, 15:219–239, 1963.

C. K. Sanathanan and J. Koerner. Transfer function synthesis as a ratio of two complex polynomials. *IEEE Trans. Automatic Control*, 8(1):56–58, 1963.

G. Szegö. *Orthogonal Polynomials*. American Mathematical Society, Colloquium Publications, Vol. 23, 1939.

M. Van Barel and A. Bultheel. Discrete linearized least-squares rational approximation on the unit circle. *Journ. Computational Applied Mathematics*, 50:545–563, 1994.

M. Van Barel and A. Bultheel. Orthonormal polynomial vectors and least squares approximation for a discrete inner product. *Electronic Trans. Numerical Analysis, Kent State University*, 3:1–23, 1995.

A. Wills and B. Ninness. On gradient-based search for multivariable system estimates. *IEEE Trans. Automatic Control*, 53(1):298–306, 2008.

D. C. Youla and N. N. Kazanjian. Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle. *IEEE Trans. Circuits and Systems*, 25(2):57–69, 1978.