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

On a Finiteness Result for the Number of Critical Points of the H2 Approximation Criterion

Bernard Hanzon [*,1] **Ralf Peeters** [**] **Ivo Bleylevens** [***]

[*] *Edgeworth Centre for Financial Mathematics, School of Mathematical Sciences, University College, Cork, Ireland, b.hanzon@ucc.ie*
[**] *Department of Knowledge Engineering, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands, ralf.peeters@maastrichtuniversity.nl*
[***] *Department of Knowledge Engineering, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands, ivo.bleylevens@maastrichtuniversity.nl*

Abstract: The long-standing open problem about whether the number of critical points in the H_2 SISO real model order reduction problem is finite is answered in the positive in the case the transfer function of the to-be-reduced model has distinct poles (ie. only has poles of algebraic multiplicity one). This has important implications for various search methods for finding critical points or local minima of the criterion function for this model reduction problem. In fact more is shown namely that in a particular parametrization the number of complex solutions of the first order conditions of the H_2 real model order reduction problem is finite and lies in a bounded set of which the bound can be determined from information about the to-be-reduced model. This implies that the H_2 model order reduction problem can be solved in principle by constructive algebra methods (such as Groebner basis methods) in case the to-be-reduced model has distinct poles. It simplifies the methods for co-order three reduction as described in a companion paper.

Keywords: H_2-model order reduction, Gröbner basis, Stetter-Möller matrix method, algebraic optimization, polynomial eigenvalue problem, generalized eigenvalue problem

1. INTRODUCTION

When studying model order reduction and system approximation from a geometric point of view, it is convenient to start from an enveloping inner product space which contains both the system to be approximated and the subset of systems in which an approximation is to be found. In the H_2-model order reduction problem, this enveloping space is taken to be the Hardy space $\mathcal{H}_2(\Pi^+)$ of complex functions which are square integrable on the imaginary axis $i\mathbb{R}$ and analytic on the closure of the right half plane $\Pi^+ = \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$. (For a definition of the Hardy spaces see e.g. Walter Rudin [1987]). The associated H_2-inner product of two functions $g(s)$ and $h(s)$ is given by

$$\langle g, h\rangle = \frac{1}{2\pi i}\int_{i\mathbb{R}} g(s)h^*(-s)ds = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(i\omega)h(i\omega)^* d\omega, \tag{1}$$

where the asterisk denotes complex conjugation. An important subset of $\mathcal{H}_2(\Pi^+)$ is formed by the class of strictly proper rational functions which have all their poles in the open left half plane $\Pi^- = \{s \in \mathbb{C} \mid \mathrm{Re}(s) < 0\}$. Such functions correspond to transfer functions of continuous-time stable and (strictly) causal LTI systems of finite order.

[1] Supported by Science Foundation Ireland under grant numbers 07/MI/008 and RFP2007-MATF802.

In this paper we shall restrict to *real* such transfer functions, both for the given system $h(s)$ and for their approximations $g(s)$. Following the approach and notation of Hanzon, Maciejowski, Chou [1998], we start from a given such function

$$h(s) = \frac{e(s)}{d(s)} = \frac{e_{N-1}s^{N-1} + \ldots + e_1 s + e_0}{s^N + d_{N-1}s^{N-1} + \ldots + d_1 s + d_0} \tag{2}$$

of McMillan degree $N \geq 1$, in which the real polynomials $d(s)$ and $e(s)$ are co-prime (i.e., they have no zeros in common and $e(s)$ does not vanish identically). The polynomial $d(s)$ is assumed to be Hurwitz (i.e., its zeros are all in Π^-). For a given approximation order $n < N$, the submanifold of candidate approximations S_n consists of all the functions

$$g(s) = \frac{b(s)}{a(s)} = \frac{b_{n-1}s^{n-1} + \ldots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0} \tag{3}$$

of McMillan degree n, in which the real polynomials $a(s)$ and $b(s)$ are co-prime, and $a(s)$ is monic and Hurwitz.

The H_2-model order reduction problem consists of finding an approximation $g \in S_n$ to the given function h of degree N, which minimizes the H_2-criterion function

$$V_h(g) = \|h - g\|^2_{H_2} \tag{4}$$

where $\|\cdot\|_{H_2}$ denotes the H_2-norm, as induced by the H_2-inner product on $\mathcal{H}_2(\Pi^+)$.

978-3-902823-06-9/12/$20.00 © 2012 IFAC
10.3182/20120711-3-BE-2027.00156

In this paper we further develop and investigate an algebraic approach to compute a global minimizer to the H_2-model order reduction problem, first introduced in Hanzon, Maciejowski, Chou [1998]. (See also Hanzon, Maciejowski, Chou [2007] and Bleylevens [2010]). Conceptually, one way to proceed is to compute all the stationary points of the criterion, to compare their associated criterion values, and to select a best one. From a geometric point of view, the stationary points of the H_2-criterion on S_n are those points g for which the difference $h - g$ is orthogonal to the tangent space of S_n at g. As the tangent space to S_n at $g(s) = \frac{b(s)}{a(s)}$ is known to have the orthogonal complement $\{a(-s)^2 R(s) \mid R(s) \in \mathcal{H}_2(\Pi^+)\}$, see Meier and Luenberger [1967], it follows that g is a stationary point of V_h on S_n if and only if there exists an H_2-function R for which it holds that:

$$\frac{e(s)}{d(s)} - \frac{b(s)}{a(s)} = a(-s)^2 R(s). \tag{5}$$

Obviously, in such a case R is a real *rational* function, which does not vanish (since $n < N$) and which has its poles in Π^-. Since $a(s)$ is required to be Hurwitz, $a(-s)$ has all of its zeros in Π^+ so that $a(-s)^2$ does not cancel any poles of $R(s)$. The poles of $R(s)$ are therefore among the poles of the left hand side expression, i.e., among the zeros of $a(s)$ and $d(s)$. Upon multiplication by $a(s)d(s)$, it is obtained that

$$e(s)a(s) - b(s)d(s) = a(-s)^2 q(s), \tag{6}$$

where $q(s) = a(s)d(s)R(s)$ is a non-vanishing real *polynomial.* The left hand side expression is polynomial of degree $\leq N + n - 1$. Hence, the degree of q is $\leq k - 1$, where $k = N - n$ is referred to as the *co-order* of the H_2-model order reduction problem. We shall denote $L = k-1 = N-n-1$ and employ the notation $q(s) = q_L s^L + \ldots + q_1 s + q_0$. Note that to solve this equation, one could proceed to equate the coefficients of the polynomial $a(-s)^2 q(s) - e(s)a(s) + b(s)d(s)$ to zero, which then gives rise to a system of $N + n$ polynomial equations in the $N + n$ unknown coefficients of $a(s)$, $b(s)$ and $q(s)$. However, even while the total degrees of these equations do not exceed 3, such a system of equations may still be hard to solve algebraically, as the equations are not in Gröbner basis form.

Still following the approach of Hanzon, Maciejowski, Chou [1998], we therefore proceed alternatively by eliminating the polynomial $b(s)$ (which occurs linearly in the equation) and by reparameterizing the polynomial $a(s)$ in terms of its values at the poles $\delta_1, \ldots, \delta_N$ of the given function $h(s)$. As a first step, this leads to a system of N equations in the N unknown coefficients of $a(s)$ and $q(s)$:

$$e(\delta_i)a(\delta_i) = a(-\delta_i)^2 q(\delta_i), \qquad (i = 1, \ldots, N). \tag{7}$$

Note that this requires us to make the technical assumption that the N poles of $h(s)$ are all distinct. Note also that $b(s)$ can be recovered uniquely from $a(s)$ and $q(s)$ as:

$$b(s) = \frac{e(s)a(s) - a(-s)^2 q(s)}{d(s)}. \tag{8}$$

Proceeding with the case of distinct poles $\delta_1, \ldots, \delta_N$, we transfer one degree of freedom from $q(s)$ to the monic polynomial $a(s)$, by choosing a nonzero normalizing quantity q_*, and we introduce the polynomials $\tilde{a}(s)$ and $\rho(s)$ as:

$$\rho(s) = \rho_L s^L + \ldots + \rho_1 s + \rho_0 = \frac{q(s)}{q_*}, \tag{9}$$

$$\tilde{a}(s) = \tilde{a}_{N-1} s^{N-1} + \ldots + \tilde{a}_1 s + \tilde{a}_0 = q_* a(-s). \tag{10}$$

Note that the polynomial $\tilde{a}(s)$ is regarded here as a polynomial of degree $\leq N - 1$, while its degree is in fact $n = N - L - 1$. The choice of the nonzero quantity q_* is made to eliminate one degree of freedom from $q(s)$: we choose $q_* = q_j$ for a selected index $j \in \{0, 1, \ldots, L\}$. The fact that $q(s)$ does not vanish identically guarantees the existence of at least one such nonzero coefficient of $q(s)$ for each stationary point of the H_2-criterion. We let j iterate over all possible values to cover all situations.

The system of equations now takes the form:

$$e(\delta_i)\tilde{a}(-\delta_i) = \tilde{a}(\delta_i)^2 \rho(\delta_i), \qquad (i = 1, \ldots, N). \tag{11}$$

From Lagrange interpolation theory it is well known that the polynomial $\tilde{a}(s)$ is uniquely characterized as a polynomial of degree $\leq N - 1$ by its values at the N distinct points $\delta_1, \ldots, \delta_N$. Denoting these values by

$$x_i = \tilde{a}(\delta_i), \qquad (i = 1, \ldots, N), \tag{12}$$

it is possible to express the value of $\tilde{a}(s)$ at any given point $s = s_0$ as a linear combination of the values $x_1, \ldots, x_N$. Let the Vandermonde matrix $V(\delta_1, \ldots, \delta_N)$ be defined as

$$V(\delta_1, \delta_2, \ldots, \delta_N) = \begin{pmatrix} 1 & \delta_1 & \ldots & \delta_1^{N-1} \\ 1 & \delta_2 & \ldots & \delta_2^{N-1} \\ \vdots & \vdots & & \vdots \\ 1 & \delta_N & \ldots & \delta_N^{N-1} \end{pmatrix}, \tag{13}$$

then the quantities $x_1, \ldots, x_N$ are related to the coefficients $\tilde{a}_0, \ldots, \tilde{a}_{N-1}$ by:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = V(\delta_1, \ldots, \delta_N) \begin{pmatrix} \tilde{a}_0 \\ \vdots \\ \tilde{a}_{N-1} \end{pmatrix}. \tag{14}$$

The inverse of the Vandermonde matrix is well known to exist for distinct $\delta_1, \ldots, \delta_N$, see also Section 2. Therefore it holds that

$$\tilde{a}(s_0) = \begin{pmatrix} 1 & s_0 & \ldots & s_0^{N-1} \end{pmatrix} V(\delta_1, \ldots, \delta_N)^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}. \tag{15}$$

It follows that

$$\begin{pmatrix} \tilde{a}(-\delta_1) \\ \tilde{a}(-\delta_2) \\ \vdots \\ \tilde{a}(-\delta_N) \end{pmatrix} = V(-\delta_1, \ldots, -\delta_N) V(\delta_1, \ldots, \delta_N)^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}. \tag{16}$$

In this way we arrive at a system of parameterized quadratic equations of the form

$$\begin{pmatrix} \frac{\rho(\delta_1)}{e(\delta_1)} x_1^2 \\ \frac{\rho(\delta_2)}{e(\delta_2)} x_2^2 \\ \vdots \\ \frac{\rho(\delta_N)}{e(\delta_N)} x_N^2 \end{pmatrix} = M(\delta_1, \delta_2, \ldots, \delta_N) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \tag{17}$$

in which the matrix $M(\delta_1, \ldots, \delta_N)$ is defined by:

$$M(\delta_1, \ldots, \delta_N) = V(-\delta_1, \ldots, -\delta_N) V(\delta_1, \ldots, \delta_N)^{-1}. \tag{18}$$

Note that the quantities $e(\delta_i)$, $(i = 1, \ldots, N)$, are all nonzero because $e(s)$ and $d(s)$ are co-prime. The structured matrices $V(\delta_1, \ldots, \delta_N)$ and $M(\delta_1, \ldots, \delta_N)$ have a number of useful properties which are addressed in the next section.

For the case of co-order $k = 1$, the number of free parameters in $\rho(s)$ equals $L = k - 1 = 0$. This case was studied in Hanzon, Maciejowski, Chou [1998], where it was noted that the quadratic system of equations is in Gröbner basis form for any total degree monomial ordering, and has a special structure which implies that it has a finite number of solutions. It was subsequently cast into the form of a large eigenvalue problem by the Stetter-Möller matrix method. Here we intend to proceed for co-orders $k \geq 2$ by solving for the quantities $x_1, x_2, \ldots, x_N$ in a similar fashion. However, the quadratic system of equations now involves $L \geq 1$ free parameters in $\rho(s)$, and in addition it is subject to L constraints which express that the degree of $\tilde{a}(s)$ is in fact $n = N - L - 1$ rather than $N - 1$. Note that according to our observations above, the coefficients $\tilde{a}_{N-1}, \ldots, \tilde{a}_{N-L}$ can all be expressed *linearly* in terms of $x_1, \ldots, x_N$, with coefficients that constitute the last L rows of the matrix $V(\delta_1, \ldots, \delta_N)^{-1}$.

To conclude this section, a summary of the system of equations is provided for further reference, which describes the H_2-model order reduction problem of co-order $L + 1$. The quadratic system of equations is given by $f_k = 0, (k = 1, \ldots, N)$ with

$$f_k(x_1, \ldots, x_N) = r_k x_k^2 - (m_{k1} x_1 + \ldots + m_{kN} x_N). \tag{19}$$

Here, the quantities r_k do not vanish identically, and depend linearly on the parameters $\rho_0, \rho_1, \ldots, \rho_L$, according to

$$r_k = \frac{\rho_L \delta_k^L + \ldots + \rho_1 \delta_1 + \rho_0}{e(\delta_k)}. \tag{20}$$

The parameters are normalized such that $\rho_j = 1$ is fixed for some selected index j (we iterate over $j = 0, 1, \ldots, L$ to cover all possible cases). This parameterized quadratic system is subject to L homogeneous linear constraints defined by $g_j = 0, (j = 1, \ldots, L)$ with

$$g_j(x_1, \ldots, x_N) = g_{j1} x_1 + \ldots + g_{jN} x_N. \tag{21}$$

The matrices M and G are fully determined by the distinct roots $\delta_1, \ldots, \delta_N$ of the real Hurwitz polynomial $d(s)$. The quantities $e(\delta_k)$ are all nonzero.

2. PROPERTIES OF THE VANDERMONDE MATRIX, ITS INVERSE, AND THE MATRIX M

We consider the Vandermonde matrix $V(\delta_1, \ldots, \delta_N)$. It is well-known that

$$\det V(\delta_1, \ldots, \delta_N) = \Pi_{1 \leq \ell < k \leq N} (\delta_k - \delta_\ell), \tag{22}$$

which is nonzero if and only if the numbers δ_k are all distinct, whence $V(\delta_1, \ldots, \delta_N)$ is invertible. The Vandermonde matrix is well studied in the literature and plays an important role in interpolation theory.

The inverse $V(\delta_1, \ldots, \delta_N)^{-1}$ is given by

$$V(\delta_1, \delta_2, \ldots, \delta_N)^{-1} = \begin{pmatrix} \ell_0^{(1)} & \ell_0^{(2)} & \cdots & \ell_0^{(N)} \\ \ell_1^{(1)} & \ell_1^{(2)} & \cdots & \ell_1^{(N)} \\ \vdots & \vdots & & \vdots \\ \ell_{N-1}^{(1)} & \ell_{N-1}^{(2)} & \cdots & \ell_{N-1}^{(N)} \end{pmatrix} \tag{23}$$

in which the symbols $\ell_k^{(j)}$ denote the coefficients of the Lagrange interpolation polynomial $L^{(j)}(s) = \ell_{N-1}^{(j)} s^{N-1} + \ldots + \ell_1^{(j)} s + \ell_0^{(j)}$, defined by

$$L^{(j)}(s) = \frac{(s - \delta_1) \cdots (s - \delta_{j-1})(s - \delta_{j+1}) \cdots (s - \delta_N)}{(\delta_j - \delta_1) \cdots (\delta_j - \delta_{j-1})(\delta_j - \delta_{j+1}) \cdots (\delta_j - \delta_N)} \tag{24}$$

which has the characterizing properties to be of degree $\leq N - 1$, and such that $L^{(j)}(\delta_j) = 1$, while $L^{(j)}(\delta_k) = 0$ for all $k \neq j$. An explicit expression for $\ell_{i-1}^{(j)}$ is given in the following result.

Proposition 2.1. Let $\delta_1, \ldots, \delta_N$ be N distinct complex numbers, for which the associated Vandermonde matrix $V(\delta_1, \ldots, \delta_N)$ is defined by Eqn. (13). Let the entry in row i and column j of $V(\delta_1, \ldots, \delta_N)^{-1}$ be denoted by $\ell_{i-1}^{(j)}$, as in Eqn. (23). Denote by $A_j^{(N-i)}$ the set of multi-indices $\alpha = (\alpha_1, \ldots, \alpha_N) \in \{0, 1\}^N$ with $\alpha_j = 0$ and $\Sigma_{k=1}^N \alpha_k = N - i$. Then $\ell_{i-1}^{(j)}$ is equal to

$$\frac{(-1)^{N-i} \sum_{(\alpha_1, \ldots, \alpha_N) \in A_j^{(N-i)}} \delta_1^{\alpha_1} \cdots \delta_{j-1}^{\alpha_{j-1}} \delta_{j+1}^{\alpha_{j+1}} \cdots \delta_N^{\alpha_N}}{(\delta_j - \delta_1) \cdots (\delta_j - \delta_{j-1})(\delta_j - \delta_{j+1}) \cdots (\delta_j - \delta_N)}. \tag{25}$$

If $\delta_1, \ldots, \delta_N$ are the zeros of a real Hurwitz polynomial, then the coefficients $\ell_{i-1}^{(j)}$ are all nonzero, with the numerator of (25) having a strictly positive real part for all i and j.

The $L \times N$ matrix G featuring in the linear constraint expressions, is defined to consist of the last L rows of $V(\delta_1, \ldots, \delta_N)^{-1}$, in reverse order:

$$G = \left(\begin{array}{ccc|ccc} 0 & & 0 & 0 & & 1 \\ \vdots & & \vdots & & \cdot^{\cdot^{\cdot}} & \\ 0 & \ldots & 0 & 1 & & 0 \end{array}\right) V(\delta_1, \ldots, \delta_N)^{-1}. \tag{26}$$

For the square submatrices of G of maximal size $L \times L$ we have the following result.

Proposition 2.2. Every $L \times L$ submatrix of G is nonsingular.

Proof. First, note that a permutation of the columns of $V(\delta_1, \ldots, \delta_N)^{-1}$ corresponds to the same permutation on the numbers $\delta_1, \ldots, \delta_N$. Without loss of generality, we may therefore restrict to the last L columns of G. Also, note that we may disregard the reordering of the rows, as they do not affect (non)singularity.

Next, partition $V(\delta_1, \ldots, \delta_N) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$ and $V(\delta_1, \ldots, \delta_N)^{-1} = \left(\begin{array}{c|c} W & X \\ \hline Y & Z \end{array}\right)$. Note that $G = \begin{pmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{pmatrix} (\, Y \,|\, Z \,)$ and the selected block of interest is $\begin{pmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{pmatrix} Z$. Note also that A is the $(N-L) \times (N-L)$ Vandermonde matrix for $\delta_1, \ldots, \delta_{N-L}$, hence invertible. Now, $V(\delta_1, \ldots, \delta_N)$ can be factored as:

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) = \left(\begin{array}{c|c} A & 0 \\ \hline C & I_L \end{array}\right) \left(\begin{array}{c|c} I_{N-L} & A^{-1}B \\ \hline 0 & D - CA^{-1}B \end{array}\right)$$

Inversion leads to the following expression for $V(\delta_1, \ldots, \delta_N)^{-1}$:

$$\left(\begin{array}{c|c} W & X \\ \hline Y & Z \end{array}\right) =$$

$$\left(\begin{array}{c|c} I_{N-L} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ \hline 0 & (D - CA^{-1}B)^{-1} \end{array}\right) \left(\begin{array}{c|c} A^{-1} & 0 \\ \hline -CA^{-1} & I_L \end{array}\right)$$

from which we have that $Z = (D - CA^{-1}B)^{-1}$. Invertibility of $V(\delta_1, \ldots, \delta_N)$ and A implies invertibility of $D - CA^{-1}B$, so that Z is a well-defined invertible matrix. □

We proceed to discuss the $N \times N$ matrix $M = M(\delta_1, \ldots, \delta_N)$ which was defined by (18). Note that $V(-\delta_1, \ldots, -\delta_N) = V(\delta_1, \ldots, \delta_N) \times \operatorname{diag}\{1, -1, \ldots, (-1)^{N-1}\}$, which shows that M is in fact diagonalizable, with its eigenvalues equal to 1 (with multiplicity $\lceil \frac{N}{2} \rceil$) and -1 (with multiplicity $\lfloor \frac{N}{2} \rfloor$). Its eigenvectors show up in the corresponding columns of $V(\delta_1, \ldots, \delta_N)$ and involve the even powers (for the eigenvalue 1) and the odd powers (for the eigenvalue -1) of the parameters δ_k. The determinant of M equals

$$\det(M) = (-1)^{\lfloor \frac{N}{2} \rfloor} = (-1)^{N(N-1)/2}. \tag{27}$$

Proposition 2.3. Let $\delta_1, \ldots, \delta_N$ be distinct complex numbers, and let M be defined by Eqn. (18). Then the entry m_{ij} in row i and column j of M, is given by:

$$m_{ij} = \prod_{\ell \neq j} \frac{(-\delta_i - \delta_\ell)}{(\delta_j - \delta_\ell)}. \tag{28}$$

If $\delta_1, \ldots, \delta_N$ are the zeros of a real Hurwitz polynomial, then the coefficients m_{ij} are all nonzero.

Proof. From the earlier considerations on the relationship between the Lagrange polynomials and the Vandermonde matrix, it follows that $m_{ij} = L^{(j)}(-\delta_i)$, which produces the expression for m_{ij} given in the proposition.
If $d(s) = (s - \delta_1) \cdots (s - \delta_N)$ is Hurwitz, we have for all i and j that $-\delta_i - \delta_j$ has a strictly positive real part, hence is nonzero. □

Next, let K be an arbitrary subset of indices from the set $\{1, 2, \ldots, N\}$ and denote its complement by K'. The number of elements in K is denoted by $|K|$. It obviously holds that $|K'| = N - |K|$. We construct the submatrix $M(K)$ of $M(\delta_1, \ldots, \delta_N)$ to consist of those entries m_{ij} for which both i and j are in K:

$$M(K) = (m_{ij})_{i,j \in K}. \tag{29}$$

Formally, we have that for the empty set $K = \phi$ no entries are selected, and we shall define $\det(M(\phi)) = 1$. When K has a single element, say $K = \{k\}$, we have that $M(\{k\}) = m_{kk}$ which we have shown to satisfy the expression $m_{kk} = \prod_{\ell \neq k} \frac{(-\delta_k - \delta_\ell)}{(\delta_k - \delta_\ell)}$ and which is nonzero when $d(s)$ is Hurwitz. For $K = \{1, 2, \ldots, N\}$ we have that $M(\{1, 2, \ldots, N\}) = M$ for which $\det(M) = (-1)^{\frac{1}{2}N(N-1)}$ which is also nonzero. These results are special cases of the following more general statement.

Proposition 2.4. For K an arbitrary subset of $\{1, 2, \ldots, N\}$ of cardinality $|K|$ it holds that

$$\det(M(K)) = (-1)^{\frac{1}{2}|K|(|K|-1)} \prod_{k \in K, \ell \in K'} \frac{(-\delta_k - \delta_\ell)}{(\delta_k - \delta_\ell)}. \tag{30}$$

If $d(s)$ is Hurwitz, then $\det(M(K)) \neq 0$.

Proof. Denoting the cardinality of K by n, we restrict without loss of generality to the case where $K = \{1, 2, \ldots, n\}$ and $K' = \{n+1, \ldots, N\}$, as the other cases can be obtained from this one through a suitable permutation of the indices.

Note from Eqn. (28) that m_{ij} with $i, j \in K$ can be expressed as:

$$m_{ij} = \tilde{m}_{ij} \frac{(-\delta_i - \delta_{n+1}) \cdots (-\delta_i - \delta_N)}{(\delta_j - \delta_{n+1}) \cdots (\delta_j - \delta_N)}$$

where $\tilde{m}_{ij}$ is defined by

$$\tilde{m}_{ij} = \frac{(-\delta_i - \delta_1) \cdots (-\delta_i - \delta_{j-1})(-\delta_i - \delta_{j+1}) \cdots (-\delta_i - \delta_n)}{(\delta_j - \delta_1) \cdots (\delta_j - \delta_{j-1})(\delta_j - \delta_{j+1}) \cdots (\delta_j - \delta_n)}.$$

Comparing with Eqn. (28), the quantity $\tilde{m}_{ij}$ represents the entry in position (i, j) of $M(\delta_1, \ldots, \delta_n)$ of size $n \times n$. Therefore:

$$M(K) = \Lambda(-\delta_1, \ldots, -\delta_n) M(\delta_1, \ldots, \delta_n) \Lambda(\delta_1, \ldots, \delta_n)^{-1}$$

where

$$\Lambda(\delta_1, \ldots, \delta_n) := \begin{pmatrix} \prod_{\ell \in K'}(\delta_1 - \delta_\ell) & & \\ & \ddots & \\ & & \prod_{\ell \in K'}(\delta_n - \delta_\ell) \end{pmatrix}.$$

Since $\det(M(\delta_1, \ldots, \delta_n)) = (-1)^{\frac{1}{2}n(n-1)}$ it therefore holds that

$$\det(M(K)) = (-1)^{\frac{1}{2}n(n-1)} \prod_{k \in K, \ell \in K'} \frac{(-\delta_k - \delta_\ell)}{(\delta_k - \delta_\ell)},$$

which proves the first part of the theorem. Since none of the factors in this product are zero when $d(s)$ is Hurwitz, nonsingularity of all the submatrices $M(K)$ of $M(\delta_1, \delta_2, \ldots, \delta_N)$ follows. □

Note that the definition for $\det(M(K))$ for $K = \phi$ and the result for $K = \{1, 2, \ldots, N\}$ are consistent with the expression (30) when the usual convention is applied to regard an empty product to be equal to 1.

3. A UNIFORM BOUND ON THE SOLUTIONS OF THE LINEARLY CONSTRAINED PARAMETERIZED QUADRATIC SYSTEM OF EQUATIONS

When considering the parameterized quadratic system of equations subject to the linear constraints (19)-(21), we have the following main result on the boundedness of its solutions $(x_1, \ldots, x_N)$.

Theorem 3.1. Consider the quadratic system of equations Eqn. (19) subject to the linear constraints Eqn. (21) and parameterized by L parameters according to Eqn. (20), with $\rho_j = 1$ for a selected index j, and with the matrices M and G defined as in Eqns. (18) and (26), with $e(\delta_i) \neq 0$ for all $i = 1, \ldots, N$, and with $\delta_1, \ldots, \delta_N$ all distinct, denoting the roots of a real Hurwitz polynomial $d(s)$.
There exists a real number $\mu > 0$, which depends only on the given transfer function $h(s) = e(s)/d(s)$, such that for every solution $(x_1, \ldots, x_N)$ of the linearly constrained quadratic system and for any associated choice of the parameters $\rho_0, \rho_1, \ldots, \rho_L$, (with one selected parameter $\rho_j = 1$ fixed), it holds that:

$$\|x\|_\infty \leq \mu \tag{31}$$

where $\|x\|_\infty = \max_{k=1,\ldots,N}\{|x_k|\}$.

The proof of this theorem is given below. It employs two lemmas.

Lemma 3.1. Consider the parameterized quantities $r_1, .., r_N$ defined in Eqn. (20). There exists a real number $\eta > 0$, which depends only on the given transfer function $h(s) = e(s)/d(s)$, such that for any choice of the parameters $\rho_0, \rho_1, \ldots, \rho_L$, (with one selected parameter $\rho_j = 1$ fixed), there are at least $N - L$ different indices k for which $|r_k| \geq \eta$.

Proof. Note that for all $k = 1, \ldots, N$ we have that $r_k e(\delta_k) = \rho_L \delta_k^L + \ldots + \rho_1 \delta_1 + \rho_0$. If we select any $L+1$ distinct indices $k_1, \ldots, k_{L+1}$ from $\{1, 2, \ldots, N\}$, we have that

$$\begin{pmatrix} r_{k_1} e(\delta_{k_1}) \\ r_{k_2} e(\delta_{k_2}) \\ \vdots \\ r_{k_{L+1}} e(\delta_{k_{L+1}}) \end{pmatrix} = \begin{pmatrix} 1 & \delta_{k_1} & \ldots & \delta_{k_1}^L \\ 1 & \delta_{k_2} & \ldots & \delta_{k_2}^L \\ \vdots & \vdots & & \vdots \\ 1 & \delta_{k_{L+1}} & \ldots & \delta_{k_{L+1}}^L \end{pmatrix} \begin{pmatrix} \rho_0 \\ \rho_1 \\ \vdots \\ \rho_L \end{pmatrix}. \tag{32}$$

Note that the matrix of this linear system is nonsingular, as it is the Vandermonde matrix for $\delta_{k_1}, \ldots, \delta_{k_{L+1}}$. Hence, its singular values are all strictly positive. Denote the smallest singular value of the Vandermonde matrices for all the possible selections of the $L+1$ indices (there is only a finite number of possibilities) by $\sigma_{\min} > 0$. In addition, let $e_{\max} = \max_{k=1,\ldots,N}\{|e(\delta_k)|\}$. Then:

$$\left\| \begin{pmatrix} r_{k_1} \\ \vdots \\ r_{k_{L+1}} \end{pmatrix} \right\|_2 \geq \frac{1}{e_{\max}} \left\| \begin{pmatrix} r_{k_1} e(\delta_{k_1}) \\ \vdots \\ r_{k_{L+1}} e(\delta_{k_{L+1}}) \end{pmatrix} \right\|_2 \geq$$

$$\frac{\sigma_{\min}}{e_{\max}} \left\| \begin{pmatrix} \rho_0 \\ \vdots \\ \rho_L \end{pmatrix} \right\|_2 \geq \frac{\sigma_{\min}}{e_{\max}}. \tag{33}$$

The last inequality follows from the fact that one selected parameter ρ_j is kept fixed at the value 1. This implies that at least one of the quantities $r_{k_1}, \ldots, r_{k_{L+1}}$ has an absolute value $\geq \frac{\sigma_{\min}}{\sqrt{L+1} e_{\max}}$. Setting η equal to:

$$\eta = \frac{\sigma_{\min}}{\sqrt{N} e_{\max}}, \tag{34}$$

we have that $\eta > 0$ is fully determined by the transfer function $h(s) = e(s)/d(s)$, and it holds *a fortiori* that $max_{j=1,\ldots,L+1}\{|r_{k_j}|\} \geq \eta$ for any choice of $L \in \{0, 1, \ldots, N-1\}$. Finally, note that the largest subset of indices k for which $|r_k| < \eta$, cannot have more than L elements, as any subset of $L+1$ indices contains an index k for which $|r_k| \geq \eta$. This implies that for at least $N - L$ different indices k we have that $|r_k| \geq \eta$. □

Lemma 3.2. Consider the (unconstrained) quadratic equations $f_k = 0$, $(k = 1, \ldots, _N)$, defined in Eqn. (19). There exists a real number $\epsilon > 0$, which depends only on the given transfer function $h(s) = e(s)/d(s)$, such that for any choice of the parameters $\rho_0, \rho_1, \ldots, \rho_L$, (with one selected parameter $\rho_j = 1$ fixed), and for any corresponding solution $(x_1, \ldots, x_N)$ of the quadratic system, there are at least $N - L$ different indices k for which

$$|x_k| \leq \epsilon \|x\|_\infty^{1/2}, \tag{35}$$

where $\|x\|_\infty = \max_{i=1,\ldots,N}\{|x_i|\}$.

Proof. The claim clearly holds for the trivial solution for any choice of $\epsilon > 0$. Now consider a nontrivial solution $(x_1, \ldots, x_N)$ of the unconstrained quadratic system of equations, for arbitrary complex values of the parameters. Then for all $k = 1, \ldots, N$ we have:

$$|r_k| \frac{|x_k|^2}{\|x\|_\infty^2} \leq$$

$$\frac{1}{\|x\|_\infty} \left| m_{k1} \frac{x_1}{\|x\|_\infty} + \ldots + m_{kN} \frac{x_N}{\|x\|_\infty} \right| \leq \frac{N\|M\|}{\|x\|_\infty}$$

where $\|M\| = \max_{i,j=1,\ldots,N}\{|m_{ij}|\}$, and where it is noted that $\frac{|x_k|}{\|x\|_\infty} \leq 1$ for all k. In view of Lemma 3.1, it therefore holds that there exists a real number $\eta > 0$, which depends only on the given transfer function $h(s) = e(s)/d(s)$, such that for at least $N - L$ different indices k we have:

$$\frac{|x_k|^2}{\|x\|_\infty^2} \leq \frac{N\|M\|}{|r_k|\|x\|_\infty} \leq \frac{N\|M\|}{\eta\|x\|_\infty}.$$

Define $\epsilon > 0$ as:

$$\epsilon = \sqrt{\frac{N\|M\|}{\eta}}, \tag{36}$$

then it holds for those same indices k that $\frac{|x_k|}{\|x\|_\infty} \leq \frac{\epsilon}{\|x\|_\infty^{1/2}}$, whence:

$$|x_k| \leq \epsilon \|x\|_\infty^{1/2}.$$

Note that also ϵ depends only on the given transfer function $h(s) = e(s)/d(s)$, because both M and η do. □

We are now in a position to provide a proof of Thm. 3.1.

Proof of Thm. 3.1. The claim clearly holds for the trivial solution for any choice of $\mu > 0$. Now, consider a nontrivial solution $(x_1, \ldots, x_N)$ of the linearly unconstrained quadratic system of equations. According to Lemma 3.2 there exists a real constant $\epsilon > 0$, which depends only on the given transfer function $h(s) = e(s)/d(s)$, for which, for any choice of parameters $\rho_0, \rho_1, \ldots, \rho_L$ (with one selected parameter $\rho_j = 1$ fixed), there are at least $N - L$ different indices k for which

$$|x_k| \leq \epsilon \|x\|_\infty^{1/2}. \tag{37}$$

Denote $N - L$ of such indices by $k_1, \ldots, k_{N-L}$, and the remaining L indices by $j_1, \ldots, j_L$.

If $\|x\|_\infty > \epsilon^2$, then any index j for which $|x_j| = \|x\|_\infty$ is among the indices $j_1, \ldots, j_L$ and not among the other indices $k_1, \ldots, k_{N-L}$. In this situation, consider the linear constraints $g_{j1}x_1 + \ldots + g_{jN}x_N = 0$, for all $j = 1, \ldots, L$. Let them be rewritten as:

$$\begin{pmatrix} g_{1j_1} & \cdots & g_{1j_L} \\ \vdots & & \vdots \\ g_{Lj_1} & \cdots & g_{Lj_L} \end{pmatrix} \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_L} \end{pmatrix} +$$

$$\begin{pmatrix} g_{1k_1} & \cdots & g_{1k_{N-L}} \\ \vdots & & \vdots \\ g_{Lk_1} & \cdots & g_{Lk_{N-L}} \end{pmatrix} \begin{pmatrix} x_{k_1} \\ \vdots \\ x_{k_{N-L}} \end{pmatrix} = 0$$

For the vector resulting from the first matrix-vector product on the left hand side, we have from the fact that one of the indices j_i is such that $|x_{j_i}|$ attains the value $\|x\|_\infty$, that a lower bound for its 2-norm is provided by

$$\left\| \begin{pmatrix} g_{1j_1} & \cdots & g_{1j_L} \\ \vdots & & \vdots \\ g_{Lj_1} & \cdots & g_{Lj_L} \end{pmatrix} \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_L} \end{pmatrix} \right\|_2 \geq \tau_{\min} \|x\|_\infty, \qquad (38)$$

where $\tau_{\min}$ denotes the smallest singular value of all the $L \times L$ submatrices of G. Note that according to Prop. 2.2, all these $L \times L$ submatrices of G are nonsingular, so that $\tau_{\min} > 0$, and like G it depends only on the given transfer function $h(s) = e(s)/d(s)$.

For the vector resulting from the second matrix-vector product, we have from the fact that $|x_{k_1}|, \ldots, |x_{k_{N-L}}|$ are all $\leq \epsilon \|x\|_\infty^{1/2}$, that an upper bound for its 2-norm is given by

$$\left\| \begin{pmatrix} g_{1k_1} & \cdots & g_{1k_{N-L}} \\ \vdots & & \vdots \\ g_{Lk_1} & \cdots & g_{Lk_{N-L}} \end{pmatrix} \begin{pmatrix} x_{k_1} \\ \vdots \\ x_{k_{N-L}} \end{pmatrix} \right\|_2 \leq \tau_{\max} \sqrt{N} \epsilon \|x\|_\infty^{1/2}, \qquad (39)$$

where $\tau_{\max}$ denotes the largest singular value of all the $L \times (N-L)$ submatrices of G. Note that $\tau_{\max} > 0$, because the largest singular value of a matrix is zero only for the zero matrix, but G is nonzero. Like G it depends only on the given transfer function $h(s) = e(s)/d(s)$.

From the triangle inequality we have for any two vectors u and v that $\|u + v\|_2 \geq \|u\|_2 - \|v\|_2$. Thus, the two bounds imply that:

$$\|Gx\|_2 \geq \tau_{\min} \|x\|_\infty - \tau_{\max} \sqrt{N} \epsilon \|x\|_\infty^{1/2}. \qquad (40)$$

If the solution $(x_1, \ldots, x_N)$ is such that it satisfies the linear constraints, it follows that the left hand side expression equals zero, which then implies:

$$\tau_{\max} \sqrt{N} \epsilon \|x\|_\infty^{1/2} \geq \tau_{\min} \|x\|_\infty. \qquad (41)$$

It follows that

$$\|x\|_\infty^{1/2} \leq \frac{\tau_{\max} \sqrt{N} \epsilon}{\tau_{\min}}. \qquad (42)$$

If instead $\|x\|_\infty \leq \epsilon^2$, we have that $\|x\|_\infty$ is clearly also bounded. Let

$$\mu := \max\{\epsilon^2, \frac{N \tau_{\max}^2 \epsilon^2}{\tau_{\min}^2}\}, \qquad (43)$$

which depends only on the given transfer function $h(s) = e(s)/d(s)$. Then for any solution $(x_1, \ldots, x_N)$ to the linearly constrained quadratic system of equations, with some associated choice of the parameters, one has that $\|x\|_\infty \leq \mu$. □

This theorem has remarkable consequences.

Theorem 3.2. Let S denote the set of solutions (x, ρ) of the equations (19), (21). Then the subset of points $(x, \rho) \in S$ with $x \neq 0$ is finite.

A short outline of the proof is as follows: For each $i \in \{1, 2, \ldots, N\}$ the values that x_i can take in the solution set is bounded. However, due to the closure theorem from elimination theory in algebraic geometry (cf Cox, Little, O'Shea [1992]), this implies that the number of possible values of x_i is finite (the solution set has to be a zero-dimensional variety). But this implies that the number of solutions vectors x has to be finite. It can be shown that for each non-zero solution vector x there is a unique corresponding value of ρ. It follows that the number of real critical points on the manifold of systems of order n is finite.

4. CONCLUSIONS

Our results have various consequences. Firstly it follows that the number of critical points of the H_2 distance function of a given real SISO system of order N to the manifold of real SISO systems of given order $n < N$ is finite in case the to-be-approximated system has distinct poles (i.e. all poles have multiplicity one). This has consequences for search methods for the H_2 model reduction problem based on finding critical points of the criterion function. Secondly it follows that in the particular parametrization used, once the "zero solutions" (which do not correspond to critical points) are excluded, which can be done by standard constructive algebra methods, a polynomial system of equations is obtained that has a zero dimensional solution variety over the complex numbers. This implies that standard constructive algebra methods can be used in principle to solve the equations and hence the model reduction problem (compare e.g, Hanzon and Maciejowski [1996]). Thirdly the result is used to simplify the generalized eigenvalue method for co-order three H_2 model order reduction as described in Peeters, Bleylevens and Hanzon [2012].

REFERENCES

L. Meier, D. Luenberger, *Approximation of Linear Constant Systems,* IEEE Trans. Aut. Control, Oct. 1967, pp.585–588.

B. Hanzon, J.M. Maciejowski, *Constructive Algebra Methods for the L2-Problem for Stable Linear Systems,* Automatica, Vol. 32, No. 12, pp. 1645-1657, 1996.

B. Hanzon, J.M. Maciejowski, C.T. Chou, *Model reduction in H2 using matrix solutions of polynomial equations* Univ of Cambridge, Dept Eng Technical Report, CUED/F-INFENG/TR.314, March 1998

Bernard Hanzon, Jan M. Maciejowski, Chun Tung Chou, *Optimal H2 order-one reduction by solving eigenproblems for polynomial equations*, arXiv: 0706.1862v1, arxiv.org/abs/0706.1862, 13 June 2007.

W. Rudin *Real and Complex Analysis.* McGraw-Hill, 1987

D.A. Cox, J.B. Little, D. O'Shea, *Ideals, Varieties and Algorithms* Springer-Verlag, 1992.

R. Peeters, I. Bleylevens and B. Hanzon, *On algebraic and linear algebraic aspects of co-order three H2-model order reduction,* Proc. SYSID, July 11-13, 2012, Brussels.

I. Bleylevens, *Algebraic Polynomial System Solving and Applications,* PhD thesis http://arno.unimaas.nl/show.cgi?fid=20542