Globally Optimal H_2-Norm Model Reduction: A Numerical Linear Algebra Approach ★

Oscar Mauricio Agudelo * **Christof Vermeersch** *
Bart De Moor, ***Fellow, IEEE & SIAM*** *

* *Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS), Dept. of Electrical Engineering (ESAT), KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium (e-mail: {mauricio.agudelo,christof.vermeersch,bart.demoor}@esat.kuleuven.be)*

Abstract: We show that the H_2-norm model reduction problem for single-input/single-output (SISO) linear time-invariant (LTI) systems is essentially an eigenvalue problem (EP), from which the globally optimal solution(s) can be retrieved. The first-order optimality conditions of this model reduction problem constitute a system of multivariate polynomial equations that can be converted to an affine (or inhomogeneous) multiparameter eigenvalue problem (AMEP). We solve this AMEP by using the so-called augmented block Macaulay matrix, which is introduced in this paper and has a special (block) multi-shift-invariant null space. The set of all stationary points of the optimization problem, i.e., the $(2r)$-tuples (r is the order of the reduced model) of affine eigenvalues and eigenvectors of the AMEP, follows from a standard EP related to the structure of that null space. At least one of these $(2r)$-tuples corresponds to the globally optimal solution of the H_2-norm model reduction problem. We present a simple numerical example to illustrate our approach.

Keywords: model reduction, multivariate polynomials, augmented block Macaulay matrix, multiparameter eigenvalue problems, numerical algorithms.

1. INTRODUCTION

Model reduction aims to approximate a large high-order model by a model of lower order (less states). Large models may be too complicated for simulation or for control system design; hence, model reduction in these scenarios is of crucial importance (see Antoulas (2005) for some motivating examples). In general, the model reduction problem for single-input/single-output (SISO) linear time-invariant (LTI) systems can be cast in the following way:

For a given nth-order LTI continuous-time stable system with transfer function

$$G(s) = C\,(sI_n - A)^{-1}\,B,$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n}$, and $C^T \in \mathbb{R}^{n}$ are the system matrices, we look for an rth-order stable reduced model

$$G_r(s) = C_r\,(sI_r - A_r)^{-1}\,B_r,$$

with $r < n$, $A_r \in \mathbb{R}^{r\times r}$, $B_r \in \mathbb{R}^{r}$, and $C_r^T \in \mathbb{R}^{r}$, such that $G_r(s)$ is a "good approximation" of $G(s)$. In the particular setting of H_2-norm model reduction, we seek to minimize the squared H_2-norm of $G_e(s) = G(s) - G_r(s)$, i.e.,

$$G_r(s) = \arg\min \left\| G(s) - \hat{G}_r(s) \right\|_{H_2}^2, \tag{1}$$

where the H_2-norm of $G_e(s)$ is defined as

$$\begin{aligned} \|G_e(s)\|_{H_2} &= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} |G_e(j\omega)|^2 d\omega} \\ &= \sqrt{\int_0^{\infty} g_e(t)^2 dt}. \end{aligned} \tag{2}$$

Here, $g_e(t)$ is the impulse response of $G_e(s)$. Note that the H_2-norm is only defined (bounded) for stable and strictly proper transfer functions.

The optimization problem (1) is nonconvex and obtaining the global minimizer is known to be a very challenging task. In general, the available methods that address H_2-norm model reduction can be divided into two main groups: Lyapunov-based methods (e.g., Spanos et al. (1992); Žigić et al. (1993); Yan and Lam (1999)) and interpolation-based methods (e.g., Meier and Luenberger (1967); Gugercin et al. (2006, 2008); Antoulas et al.

★ This work was supported in part by the KU Leuven: Research Fund (projects C16/15/059, C3/19/053, C24/18/022, C3/20/117), Industrial Research Fund (fellowships 13-0260, IOF/16/004), and several Leuven Research and Development bilateral industrial projects, in part by Flemish Government agencies: FWO (EOS project G0F6718N (SeLMA), SBO project S005319, infrastructure project I013218N, TBM project T001919N, and PhD grants (SB/1SA1319N, SB/1S93918, SB/1S1319N)), EWI (Flanders AI Research Program), and VLAIO (City of Things (COT.2018.018), Baekeland PhD mandate (HBC.20192204), and innovation mandate (HBC.2019.2209)), and in part by the European Commission (EU Research Council under the European Union's Horizon 2020 research and innovation programme (ERC Adv. Grant under grant 885682). Other funding: Foundation "Kom op tegen Kanker", CM (Christelijke Mutualiteit). The work of Christof Vermeersch was supported by the FWO Strategic Basic Research fellowship under grant SB/1SA1319N. *(Corresponding author: Oscar Mauricio Agudelo.)*

(2010); Anić et al. (2013)). Unlike Lyapunov-based methods, which rapidly become infeasible when the dimensions increase, interpolation approaches (in which $G_r(s)$ interpolates $G(s)$ at some points in the frequency domain) have proved to be numerically very effective (Antoulas et al., 2010). Although the literature typically makes a distinction between both approaches, the two frameworks are actually equivalent, as shown by Gugercin et al. (2008). Interpolation-based H_2-norm optimality conditions were originally derived in Meier and Luenberger (1967) for SISO systems and extended later to the multiple-input/multiple-output (MIMO) case by both Gugercin et al. (2008) and Van Dooren et al. (2008). Based on the conditions and results from rational interpolation (Beattie and Gugercin, 2017), several iterative numerical algorithms have been proposed (e.g., Gugercin et al. (2006); Bunse-Gerstner et al. (2010); Anić et al. (2013)). However, none of these algorithms are guaranteed to converge to the globally optimal solution, despite the use of several heuristic rules during their initialization. For the particular cases of first-order and second-order SISO approximants, the global optimum can be found by solving a polynomial system in one and two variables, respectively (see Ahmad et al. (2010, 2011)). However, to the best of the authors' knowledge, to this day, there is not a single methodology that is guaranteed to provide the globally optimal solution of the H_2-norm model reduction problem for an approximant of arbitrary order.

In this paper, we follow the subsequent procedure to derive the globally optimal solution of (1) for any n and r (with $r < n$):

- Given that the H_2-norm can be computed algebraically from the solution of a Lyapunov equation, we exploit this fact to rewrite the objective function of (1) in terms of the unknown parameters of $G_r(s)$.
- By deriving the first-order optimality conditions of this redefined objective function, we generate a system of multivariate polynomial equations, whose common roots comprise all the stationary points of the optimization problem.
- We convert these multivariate polynomial equations to an affine (or inhomogeneous) multiparameter eigenvalue problem (AMEP).
- By using a special matrix construction that we refer to as the augmented block Macaulay matrix, we transform the AMEP into a standard eigenvalue problem (EP), from which we determine the $(2r)$-tuples of affine eigenvalues and eigenvectors of the AMEP.
- Finally, we pick the real-valued $(2r)$-tuple that leads to the stable reduced model with the smallest H_2-error, which corresponds to the globally optimal solution of the H_2-norm model reduction problem.

The main claim of this paper is that the H_2-norm model reduction problem for SISO LTI systems is an AMEP. Furthermore, in this work, we provide a new solution method for this kind of problems based on the augmented block Macaulay matrix, which can be seen as a generalization of the block Macaulay matrix introduced by De Moor (2019) and Vermeersch and De Moor (2019) to solve homogeneous multiparameter eigenvalue problems.

The remainder of this paper is organized as follows: In Section 2, we derive the first-order optimality conditions of an appropriately redefined objective function in order to generate a system of multivariate polynomial equations. Section 3 explains how this system can be transformed into an AMEP, and in Section 4, we show how to solve it using the augmented block Macaulay matrix. Section 5 presents a numerical example, and finally in Section 6, we provide some concluding remarks and future research directions.

2. MULTIVARIATE POLYNOMIAL EQUATIONS

In this section, we show that finding the optimal and suboptimal solutions of the H_2-norm model reduction problem (1) is equivalent to finding the common roots of a system of multivariate polynomial equations. These multivariate polynomial equations correspond to the first-order optimality conditions of a conveniently rewritten objective function.

2.1 Redefinition of the objective function

The H_2-norm of the error transfer function $G_e(s)$ can be computed algebraically via its state space realization, instead of evaluating the integral in (2). Hence, as shown by Antoulas (2005) and Van Dooren et al. (2008), we can conveniently express the objective function of the optimization problem (1) as

$$J = \|G_e(s)\|_{H_2}^2 = C_e W C_e^T, \tag{3}$$

where $W = W^T$ is the controllability Gramian of $G_e(s)$ satisfying the Lyapunov equation

$$A_e W + W A_e^T + B_e B_e^T = 0, \tag{4}$$

with

$$A_e = \begin{bmatrix} A & 0 \\ 0 & A_r \end{bmatrix}, B_e = \begin{bmatrix} B \\ B_r \end{bmatrix}, \text{ and } C_e = [C \ -C_r]$$

the system matrices of $G_e(s) = C_e(sI_{n+r} - A_e)^{-1}B_e$. In the remainder of this subsection, we rewrite the objective function (3) only in terms of the unknown parameters ($a_i \in \mathbb{R}$ and $b_i \in \mathbb{R}$, $\forall i = 1, \ldots, r$) of the transfer function of the reduced-order model (keep in mind that W is also unknown)

$$G_r(s) = \frac{b_1 s^{r-1} + b_2 s^{r-2} + \cdots + b_{r-1}s + b_r}{s^r + a_1 s^{r-1} + \cdots + a_{r-1}s + a_r} .$$

As state space representation of $G_r(s)$, we use the control canonical form:

$$A_r = \begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{r-1} & -a_r \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, B_r = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$
$$\text{and } C_r = [b_1 \ b_2 \ b_3 \ \cdots \ b_r] .$$

By partitioning W, we can rewrite the objective function (3) as

$$\begin{aligned} J &= C_e W C_e^T \\ &= [\, C \ -C_r \,] \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} C^T \\ -C_r^T \end{bmatrix} \\ &= C_r W_{22} C_r^T - 2C_r W_{21} C^T + C W_{11} C^T \end{aligned} \tag{5}$$

and the Lyapunov equation (4) as

$$\begin{bmatrix} A & 0 \\ 0 & A_r \end{bmatrix}\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}+\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}\begin{bmatrix} A^T & 0 \\ 0 & A_r^T \end{bmatrix}+\begin{bmatrix} B \\ B_r \end{bmatrix}\begin{bmatrix} B^T & B_r^T \end{bmatrix}=$$

$$\begin{bmatrix} AW_{11}+W_{11}A^T+BB^T & AW_{12}+W_{12}A_r^T+BB_r^T \\ A_rW_{21}+W_{21}A^T+B_rB^T & A_rW_{22}+W_{22}A_r^T+B_rB_r^T \end{bmatrix}$$

$$=0. \tag{6}$$

Here, W_{11} and W_{22} are the controllability Gramians of $G(s)$ and $G_r(s)$, respectively, and $W_{12} = W_{21}^T$, since $W = W^T$. Observe that the term $CW_{11}C^T$ in (5) can be dropped because it does not depend on the parameters of $G_r(s)$. Thus, we can use

$$\tilde{J} = C_rW_{22}C_r^T - 2C_rW_{21}C^T \tag{7}$$

as a new objective function.

In what follows, we eliminate W_{22} and W_{21} from $\tilde{J}$ by using (6). It is not difficult to see that $C_rW_{22}C_r^T$ and $C_rW_{21}C^T$ can be written as $\operatorname{vec}(C_r^TC_r)^T\operatorname{vec}(W_{22})$ and $\operatorname{vec}(C_r^TC)^T\operatorname{vec}(W_{21})$, respectively [1]. If we introduce the vectors $g_r = \operatorname{vec}(C_r^TC_r)^T \in \mathbb{R}^{r^2}$ and $g_m = \operatorname{vec}(C_r^TC)^T \in \mathbb{R}^{nr}$, then we can compactly write $\tilde{J}$ as

$$\tilde{J} = g_r\operatorname{vec}(W_{22}) - 2g_m\operatorname{vec}(W_{21}). \tag{8}$$

Notice that the Lyapunov equation $A_rW_{22} + W_{22}A^T + B_rB_r^T = 0$ from (6) can be expressed as (Horn and Johnson, 1994)

$$\begin{aligned}(A_r \otimes I_r + I_r \otimes A_r)\operatorname{vec}(W_{22}) &= -\operatorname{vec}(B_rB_r^T)\\ \underbrace{(A_r \oplus A_r)}_{T_r}\operatorname{vec}(W_{22}) &= -\underbrace{\operatorname{vec}(B_rB_r^T)}_{f_r},\end{aligned} \tag{9}$$

and the equation $A_rW_{21} + W_{21}A^T + B_rB^T = 0$ as

$$\begin{aligned}(A \otimes I_r + I_n \otimes A_r)\operatorname{vec}(W_{21}) &= -\operatorname{vec}(B_rB^T)\\ \underbrace{(A \oplus A_r)}_{T_m}\operatorname{vec}(W_{21}) &= -\underbrace{\operatorname{vec}(B_rB^T)}_{f_m},\end{aligned} \tag{10}$$

with $T_r \in \mathbb{R}^{r^2\times r^2}$, $f_r \in \mathbb{R}^{r^2}$, $T_m \in \mathbb{R}^{nr\times nr}$, and $f_m \in \mathbb{R}^{nr}$. The operators $\otimes$ and $\oplus$ denote the Kronecker product and the Kronecker sum, respectively. Finally, from (9) and (10), we have that $\operatorname{vec}(W_{22}) = -T_r^{-1}f_r$ and $\operatorname{vec}(W_{21}) = -T_m^{-1}f_m$, and, by substituting them into (8), we get $\tilde{J}$ only in terms of the parameters a_i and b_i of $G_r(s)$:

$$\tilde{J} = -g_rT_r^{-1}f_r + 2g_mT_m^{-1}f_m. \tag{11}$$

This objective function has to be minimized over the unknown parameters a_i and $b_i, \forall i = 1,\ldots,r$.

From Theorem 4.4.5 in Horn and Johnson (1994), we know that the eigenvalues of the Kronecker sum of two matrices $X \in \mathbb{R}^{n_x\times n_x}$ and $Y \in \mathbb{R}^{n_y\times n_y}$ correspond to all possible pairwise sums of the eigenvalues of X and Y, that is, if $\sigma(X) = \{\lambda_1,\ldots,\lambda_{n_x}\}$ and $\sigma(Y) = \{\mu_1,\ldots,\mu_{n_y}\}$, then $\sigma(X \oplus Y) = \{\lambda_i + \mu_j : i = 1,\ldots,n_x, j = 1,\ldots,n_y\}$. A sufficient condition for T_r and T_m to be invertible can be drawn from this result: If all the eigenvalues of A and A_r have a negative real part (which is the case for the optimal and suboptimal solutions of (1)), then all the eigenvalues of $T_r = A_r \oplus A_r$ and $T_m = A \oplus A_r$ also have a negative real part, implying in this way the non-singularity of the matrices T_r and T_m.

[1] The vec-operation $\operatorname{vec}(\cdot)$ stacks the columns of a matrix $M = \begin{bmatrix} m_1 & m_2 \end{bmatrix}$ into a vector $m = \operatorname{vec}(M) = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$.

2.2 First-order optimality conditions

Keeping in mind that g_r, f_r, g_m, and f_m are only a function of b_i, and T_r and T_m are only a function of a_i, the first-order optimality conditions of $\tilde{J}$, $\forall i = 1,\ldots,r$ are given by

$$\begin{aligned}\frac{\partial \tilde{J}}{\partial a_i} &= -g_r\frac{\partial T_r^{-1}}{\partial a_i}f_r + 2g_m\frac{\partial T_m^{-1}}{\partial a_i}f_m = 0\\ \frac{\partial \tilde{J}}{\partial b_i} &= -\frac{\partial g_r}{\partial b_i}T_r^{-1}f_r + 2\frac{\partial g_m}{\partial b_i}T_m^{-1}f_m = 0.\end{aligned}$$

Since $\frac{\partial T_r^{-1}}{\partial a_i} = -T_r^{-1}\frac{\partial T_r}{\partial a_i}T_r^{-1}$ and $\frac{\partial T_m^{-1}}{\partial a_i} = -T_m^{-1}\frac{\partial T_m}{\partial a_i}T_m^{-1}$, the previous equations become

$$\begin{aligned}\frac{\partial \tilde{J}}{\partial a_i} &= g_rT_r^{-1}T_r^{a_i}T_r^{-1}f_r - 2g_mT_m^{-1}T_m^{a_i}T_m^{-1}f_m = 0\\ \frac{\partial \tilde{J}}{\partial b_i} &= -g_r^{b_i}T_r^{-1}f_r + 2g_m^{b_i}T_m^{-1}f_m = 0,\end{aligned} \tag{12}$$

with $T_r^{a_i} = \frac{\partial T_r}{\partial a_i}$, $T_m^{a_i} = \frac{\partial T_m}{\partial a_i}$, $g_r^{b_i} = \frac{\partial g_r}{\partial b_i}$, and $g_m^{b_i} = \frac{\partial g_m}{\partial b_i}$.

As $T_r^{-1} = \operatorname{adj}(T_r)/\det(T_r)$ and $T_m^{-1} = \operatorname{adj}(T_m)/\det(T_m)$, where $\operatorname{adj}(T_r)$ and $\operatorname{adj}(T_m)$ are the adjugate matrices of T_r and T_m, respectively, and given that $\det(T_r) \neq 0$ and $\det(T_m) \neq 0$, the partial derivatives in (12) define a system of $2r$ multivariate polynomial equations in $2r$ unknowns $(a_i, b_i, \forall i = 1,\ldots,r)$, after "multiplying out" $\det(T_r)$ and $\det(T_m)$. Their common roots comprise all the global and local minima as well as all the maxima and saddle points of $\tilde{J}$ and J. In the next section, we reformulate (12) to obtain a new set of multivariate polynomial equations from which we can formulate an affine multiparameter eigenvalue problem in a straightforward way.

3. AFFINE MULTIPARAMETER EIGENVALUE PROBLEM

Now, we introduce two auxiliary vectors, $h = T_r^{-1}f_r \in \mathbb{R}^{r^2}$ and $p = T_m^{-1}f_m \in \mathbb{R}^{nr}$, to partially linearize (12). The vectors $h^{a_i} = -T_r^{-1}T_r^{a_i}h \in \mathbb{R}^{r^2}$ and $p^{a_i} = -T_m^{-1}T_m^{a_i}p \in \mathbb{R}^{nr}$ are the partial derivatives of h and p with respect to the unknown parameters a_i $(\forall i = 1,\ldots,r)$. With these definitions, we can rewrite (12) as

$$\begin{aligned}\frac{\partial \tilde{J}}{\partial a_i} &= -g_rh^{a_i} + 2g_mp^{a_i} = 0\\ \frac{\partial \tilde{J}}{\partial b_i} &= -g_r^{b_i}h + 2g_m^{b_i}p = 0.\end{aligned} \tag{13}$$

The first-order optimality conditions given in (13), together with the definitions of the vectors h, p, h^{a_i}, and p^{a_i}, conform a new system of multivariate polynomial equations from which the optimal solution(s) can be retrieved:

$$\begin{aligned}-g_rh^{a_i} + 2g_mp^{a_i} &= 0, \quad \forall i = 1,\ldots,r,\\ -g_r^{b_i}h + 2g_m^{b_i}p &= 0, \quad \forall i = 1,\ldots,r,\\ T_rh^{a_i} + T_r^{a_i}h &= 0, \quad \forall i = 1,\ldots,r,\\ T_mp^{a_i} + T_m^{a_i}p &= 0, \quad \forall i = 1,\ldots,r,\\ T_rh - f_r &= 0,\\ T_mp - f_m &= 0.\end{aligned} \tag{14}$$

This system consists of $r^3 + r^2(n+1) + r(n+2)$ cubic polynomial equations in the same number of unknowns, which are h, p, h^{a_i}, p^{a_i}, a_i, and b_i $(\forall i = 1,\ldots,r)$.

Given that h, p, h^{a_i}, and p^{a_i} only appear linearly, we can compactly rewrite (14) as follows [2]:

$$\underbrace{\begin{bmatrix} -(I_r \otimes g_r) & 2\,(I_r \otimes g_m) & 0 & 0 \\ 0 & 0 & \{-g_r^{b_i}\}_i & \{2g_m^{b_i}\}_i \\ I_r \otimes T_r & 0 & \{T_r^{a_i}\}_i & 0 \\ 0 & I_r \otimes T_m & 0 & \{T_m^{a_i}\}_i \\ 0 & 0 & T_r & 0 \\ 0 & 0 & 0 & T_m \end{bmatrix}}_{\mathcal{A}(a_i,b_i)} \underbrace{\begin{bmatrix} \{h^{a_i}\}_i \\ \{p^{a_i}\}_i \\ h \\ p \end{bmatrix}}_{z} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -f_r \\ -f_m \end{bmatrix}}_{q} = 0. \tag{15}$$

The rectangular matrix $\mathcal{A}$ has $r^3 + r^2(n+1) + r(n+2)$ rows and $r^3 + r^2(n+1) + rn$ columns ($2r$ more rows than columns) and it is a function of the unknown parameters a_i and b_i, which appear quadratically in g_r and linearly in $g_r^{b_i}$, g_m, T_r, and T_m. The vector q is a constant column vector of appropriate length. The equation $\mathcal{A}(a_i, b_i)z + q = 0$ is basically an affine (or inhomogeneous) quadratic multiparameter eigenvalue problem (AMEP), where the parameters a_i and b_i constitute the $(2r)$-tuples $(a_1, \ldots, a_r, b_1, \ldots, b_r)$ of affine eigenvalues and the vectors h, h^{a_i}, p, and p^{a_i} generate the affine eigenvectors z. This statement becomes more clear when we rewrite $\mathcal{A}(a_i, b_i)z + q = 0$ as

$$\left(\mathcal{A}_1 + \sum_{\omega \neq 1} \mathcal{A}_\omega \omega\right) z + q = 0,$$

or as

$$-\mathcal{A}_1 z = \sum_{\omega \neq 1} \mathcal{A}_\omega \omega z + q, \tag{16}$$

where the matrix $\mathcal{A}_\omega$ (e.g., $\mathcal{A}_{a_1}$ or $\mathcal{A}_{b_1^2}$) contains the coefficients of the monomial $\omega = a_1^{k_1} \cdots a_r^{k_r} b_1^{l_1} \cdots b_r^{l_r}$ with non-negative integer exponents k_i and l_i in the matrix $\mathcal{A}$. The structure of (16) is that of a homogeneous multiparameter eigenvalue problem (MEP) (e.g., Volkmer (1988); De Moor (2019); Vermeersch and De Moor (2019)) except for the constant vector q. This modified structure is comparable to the one of an affine (or inhomogeneous) 1-parameter eigenvalue problem $Ax = \mu x + b$, $\|x\|_2 = 1$, as defined by Mattheij and Söderlind (1987), where A is a square matrix, b is a constant vector of appropriated length, and μ and x are the affine (or inhomogeneous) eigenvalues and eigenvectors of A with respect to b, respectively.

In order to solve (16), we introduce in the next section the so-called augmented block Macaulay matrix, which iteratively linearizes the AMEP. This matrix can be seen as a generalization of the block Macaulay matrix presented by De Moor (2019) and Vermeersch and De Moor (2019).

[2] The curly brackets $\{M_i\}_i$ indicate a vertical stack of matrices M_i over the index i, e.g., for $i = 1, 2$, $\{M_i\}_i = \begin{bmatrix} M_1^T & M_2^T \end{bmatrix}^T$.

4. SOLUTIONS OF THE AFFINE MULTIPARAMETER EIGENVALUE PROBLEM

4.1 Augmented block Macaulay matrix

For the sake of simplicity, and without loss of generality, let us consider the case when $r = 1$, that is, when we look for an H_2-norm optimal first-order approximant of $G(s)$. In this case, $G_r(s)$ only has two parameters (a_1 and b_1) and the system in (14) consists of $2n+4$ multivariate polynomial equations in $2n+4$ unknowns, namely, h, p, h^{a_1}, p^{a_1}, a_1, and b_1. We can rewrite (15) for $r = 1$ as

$$\begin{bmatrix} -g_r & 2g_m & 0 & 0 \\ 0 & 0 & -g_r^{b_1} & 2g_m^{b1} \\ T_r & 0 & T_r^{a_1} & 0 \\ 0 & T_m & 0 & T_m^{a_1} \\ 0 & 0 & T_r & 0 \\ 0 & 0 & 0 & T_m \end{bmatrix} \begin{bmatrix} h^{a_1} \\ p^{a_1} \\ h \\ p \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -f_r \\ -f_m \end{bmatrix} = 0, \tag{17}$$

or, in terms of the model parameters, as

$$\begin{bmatrix} -b_1^2 & 2b_1 C & 0 & 0 \\ 0 & 0 & -2b_1 & 2C \\ -2a_1 & 0 & -2 & 0 \\ 0 & -a_1 I_n + A & 0 & -I_n \\ 0 & 0 & -2a_1 & 0 \\ 0 & 0 & 0 & -a_1 I_n + A \end{bmatrix} \begin{bmatrix} h^{a_1} \\ p^{a_1} \\ h \\ p \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ -B \end{bmatrix} = 0,$$

where a_1 and b_1 constitute the 2-tuples of affine eigenvalues, while h, p, h^{a_1}, and p^{a_1} generate the affine eigenvectors z. We can now recast (17) as an AMEP:

$$\left(\mathcal{A}_1 + a_1 \mathcal{A}_{a_1} + b_1 \mathcal{A}_{b_1} + a_1^2 \mathcal{A}_{a_1^2} + a_1 b_1 \mathcal{A}_{a_1 b_1} + b_1^2 \mathcal{A}_{b_1^2}\right) z + q = 0,$$

or written as a matrix-vector product:

$$\left[\, q \,\vdots\, \mathcal{A}_1 \;\; \mathcal{A}_{a_1} \;\; \mathcal{A}_{b_1} \;\; \mathcal{A}_{a_1^2} \;\; \mathcal{A}_{a_1 b_1} \;\; \mathcal{A}_{b_1^2} \,\right] \begin{bmatrix} 1 \\ \hdashline z \\ \hline a_1 z \\ b_1 z \\ \hline a_1^2 z \\ a_1 b_1 z \\ b_1^2 z \end{bmatrix} = 0, \tag{18}$$

where the coefficient matrices $\mathcal{A}_\omega \in \mathbb{R}^{(2n+4)\times(2n+2)}$ and the vector q can be obtained from (17) in a straightforward fashion. We can "enlarge" this AMEP by multiplying it with monomials in a_1 and b_1 of increasing total degree, a process that we call forward shift recursions (FSRs). Hence, in a first iteration we multiply (18) with shifts of first total degree (i.e., a_1 and b_1), in a second iteration with shifts of second total degree (i.e., a_1^2, $a_1 b_1$, and b_1^2), etc., so that we get

$$\underbrace{\left[\, M_A \,\vdots\, M_B \,\right]}_{M} \underbrace{\begin{bmatrix} k_A \\ \hdashline k_B \end{bmatrix}}_{k} = 0, \tag{19}$$

with matrices

$$M_A = \begin{bmatrix} q & 0 & 0 & 0 & \cdots \\ 0 & q & 0 & 0 & \cdots \\ 0 & 0 & q & 0 & \cdots \\ 0 & 0 & 0 & q & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \text{ and}$$

$$M_B = \begin{bmatrix} \mathcal{A}_1 & \mathcal{A}_{a_1} & \mathcal{A}_{b_1} & \mathcal{A}_{a_1^2} & \mathcal{A}_{a_1 b_1} & \mathcal{A}_{b_1^2} & 0 & \cdots \\ 0 & \mathcal{A}_1 & 0 & \mathcal{A}_{a_1} & \mathcal{A}_{b_1} & 0 & \mathcal{A}_{a_1^2} & \cdots \\ 0 & 0 & \mathcal{A}_1 & 0 & \mathcal{A}_{a_1} & \mathcal{A}_{b1} & 0 & \cdots \\ 0 & 0 & 0 & \mathcal{A}_1 & 0 & 0 & \mathcal{A}_{a_1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

and vectors

$$k_A = \begin{bmatrix} 1 \\ \hline a_1 \\ b_1 \\ \hline a_1^2 \\ a_1 b_1 \\ b_1^2 \\ \hline \vdots \end{bmatrix} \text{ and } k_B = \begin{bmatrix} z \\ \hline a_1 z \\ b_1 z \\ \hline a_1^2 z \\ a_1 b_1 z \\ b_1^2 z \\ \hline \vdots \end{bmatrix}. \quad (20)$$

The matrix M_A is clearly a block diagonal matrix, which accounts for the shifts of the affine part of the original AMEP, while the matrix M_B is the quasi-Toeplitz block Macaulay matrix introduced by De Moor (2019) and Vermeersch and De Moor (2019) to solve homogeneous multiparameter eigenvalue problems. This block Macaulay matrix is an extension of the classical Macaulay matrix used to determine the common roots of a system of multivariate polynomial equations (Dreesen, 2013; Dreesen et al., 2018). The matrix M, which is basically a block Macaulay matrix to which we have appended to its left a block diagonal matrix, will be referred to as the augmented block Macaulay matrix.

The initial augmented block Macaulay matrix in (18) has degree $d_0 = 3$, because of the cubic polynomials, and after every iteration its degree increases by one. We keep iterating (adding equations) until the nullity, which is the dimension of the null space of M, stabilizes (this happens when both the finite solutions and the solutions at infinity form a zero-dimensional solution set). In the next two subsections, we show how to compute the finite solutions of the AMEP by exploiting the (block) multi-shift-invariance property of the null space of M. Initially, in Subsection 4.2, we consider an AMEP with only finite solutions, and afterwards, in Subsection 4.3, we deal with the general case, in which the AMEP has both finite solutions and solutions at infinity.

Note that the above-mentioned rationale can be straightforwardly generalized to higher-order approximants of $G(s)$, i.e., to $r > 1$.

4.2 Finite solutions in multi-shift-invariant subspaces

From didactical point of view, we assume first that all solutions are finite and simple (i.e., they have algebraic multiplicity equal to one). Then, the multivariate Vandermonde vectors k evaluated at each finite solution (i.e., $k^{(i)}, \forall i = 1, \ldots, m_a$, with m_a the number of finite solutions) form a basis K that spans the (right) null space of M:

$$K = \begin{bmatrix} K_A \\ \hdashline K_B \end{bmatrix} = \begin{bmatrix} k_A^{(1)} & k_A^{(2)} & \cdots & k_A^{(m_a)} \\ \hdashline k_B^{(1)} & k_B^{(2)} & \cdots & k_B^{(m_a)} \end{bmatrix}.$$

The structure of the matrix K_A and K_B are identical to the (block) Vandermonde basis for the null space of the classical Macaulay matrix (Dreesen et al., 2018) and the block Macaulay matrix (De Moor, 2019), respectively. Hence, K_A and K_B span vector spaces that are multi-shift-invariant and block multi-shift-invariant, respectively. In the remainder of this subsection, we exploit the multi-shift-invariant structure of the space spanned by K_A, to recover the finite $(2r)$-tuples of affine (or inhomogeneous) eigenvalues. Notice that is also possible to work with the space spanned by K_B, as shown by De Moor (2019).

The multi-shift-invariant structure of the space spanned by K_A can be understood in the following way: If we select one row of a multivariate Vandermonde basis K_A and multiply (or shift) it by one of the affine eigenvalues, i.e., the unknown parameters a_i and b_i, we find another row of that basis K_A. Notice that multi-shift-invariance is a property of the space spanned by K_A and not of the specific basis (Dreesen et al., 2018).

Let us apply this rationale to the case where we look for an H_2-norm optimal first-order approximant. Take for instance a multivariate Vandermonde vector k_A of degree $d = 2$, i.e.,

$$k_A = \begin{bmatrix} 1 \\ \hline a_1 \\ b_1 \\ \hline a_1^2 \\ a_1 b_1 \\ b_1^2 \end{bmatrix},$$

and multiply the first three elements by a_1. The multiplied elements correspond to the second, fourth, and fifth element of the same vector:

$$\begin{bmatrix} 1 \\ a_1 \\ b_1 \end{bmatrix} \xrightarrow{a_1} \begin{bmatrix} a_1 \\ a_1^2 \\ a_1 b_1 \end{bmatrix}.$$

Alternatively, we can write this multiplication using row selection matrices S_1 and S_2, as $S_1 k_A a_1 = S_2 k_A$, with

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

and

$$S_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

By multiplying each column of the basis K_A with a_1, we get

$$(S_1 K_A)\, D_{a_1} = (S_2 K_A), \quad (21)$$

where D_{a_1} is a diagonal matrix containing the evaluations of the shift function, i.e., a_1, at the different solutions. Note that in general, we can also use a_i or b_i, $\forall i = 1, \ldots, r$ as alternative shift functions. We recognize in (21) a generalized eigenvalue problem, with as its matrix of eigenvectors the identity matrix. In order to ensure that this eigenvalue problem is not degenerate (i.e., it does not have infinite eigenvalues), the matrix $S_1 K_A$ needs to be of full column rank, which requires the selection matrix S_1 to include m_a linearly independent rows. Consequently, we have to increase the degree of the augmented block Macaulay matrix at least until its nullity equals the number of finite roots m_a.

In practice, we do not know the Vandermonde basis K in advance, since it is constructed from the unknown solutions. Given that the choice of the basis of the null space of M is not unique, a numerical basis Z obtained

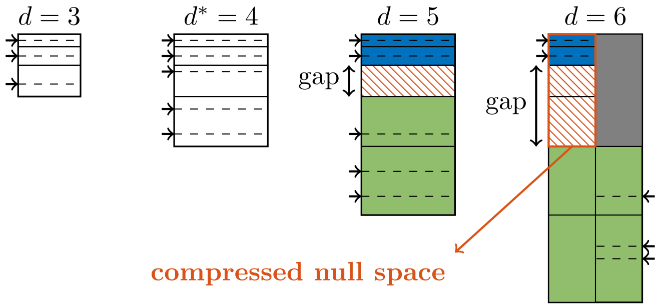

Fig. 1. The upper part of the numerical basis of the null space Z_A of the augmented block Macaulay matrix M grows as its degree d increases. At a certain degree d^*, the nullity stabilizes at total number of solutions m_b, if the solution set is zero-dimensional. From that degree on, some linearly independent rows (corresponding to the finite solutions) stabilize, while the other linearly independent rows (corresponding to the solutions at infinity) move to higher degree blocks. Eventually, a gap (of at least one degree block) of only linearly dependent rows emerges, which separates these linearly independent rows. The influence of the solutions at infinity can be removed via a column compression. The finite solution approach can then be applied straightforwardly on this compressed null space. (We adapted this figure from Vermeersch and De Moor (2019).)

for example via the singular value decomposition does not have the Vandermonde structure as in (20). However, this numerical basis Z is related to the multivariate Vandermonde basis K via $K = ZT$, with $T \in \mathbb{R}^{m_a \times m_a}$ a non-singular matrix. This relation also holds for the upper part of K and Z, i.e., $K_A = Z_A T$, which reduces (21) to

$$(S_1 Z_A)\, T D_g = (S_2 Z_A)\, T, \tag{22}$$

and transforms the AMEP into a standard eigenvalue problem,

$$T D_g T^{-1} = (S_1 Z_A)^{\dagger} (S_2 Z_A)\,. \tag{23}$$

Once we have solved (23) and obtained the matrix of eigenvectors T, we can use $K_A = Z_A T$ to determine K_A. From this matrix K_A, we can obtain the different $(2r)$-tuples of finite affine (or inhomogeneous) eigenvalues of the AMEP, which correspond to the set of all stationary points of the optimization problem (1).

4.3 Solutions at infinity in multi-shift-invariant subspaces

In the previous subsection, it was assumed that the AMEP only has finite affine (or inhomogeneous) eigenvalues. However, due to algebraic relationships among the coefficients of the matrices, solutions at infinity often emerge (Dreesen et al., 2018). In that situation, the total number of solutions $m_b = m_a + m_\infty$ corresponds to both the finite solutions and solutions at infinity. The solutions at infinity also generate vectors in the basis K of the null space of the augmented block Macaulay matrix M. Therefore, when the augmented block Macaulay matrix M reaches a sufficient degree d^*, the nullity of M stabilizes at m_b (instead of m_a). Let us define a degree block as the collection of all the rows that correspond to monomials of the same total degree. When the degree of the augmented block Macaulay matrix increases beyond d^*, some linearly independent rows (corresponding to the finite solutions) stabilize at a certain position in the null space, while the others (corresponding to solutions at infinity) move to higher positions, i.e., to monomials of higher total degree. Consequently, a gap (of at least one degree block) without any linearly independent rows, i.e., solutions, emerges between the finite solutions and the solutions at infinity, as visualized in Fig. 1.

The upper part of the null space actually consists of three zones after stabilization ($d > d^*$). These zones are determined by checking the rank of the basis Z_A row-wise from the top to the bottom:

(1) Finite zone: The first zone of the null space contains at least one linearly independent row per degree block, up to the number of finite roots m_a. This zone accommodates all the finite solutions.
(2) Gap zone: At a certain point, the rank does not increase anymore and all the rows are linearly dependent on some rows of the first zone. There is a so-called gap (of at least one degree block) of linearly dependent rows. Hence, no solutions live in this zone.
(3) Infinite zone: Eventually, in the third zone, the rank increases again, by at least one per degree block, until it reaches the total number of solutions m_b. The linearly independent rows in this zone correspond to the solutions at infinity.

Because of this behavior, we can remove the influence of the solutions at infinity via a column compression:

Theorem 1. (Column compression (Dreesen et al., 2018)). The upper part of the numerical basis of the null space $Z_A = \begin{bmatrix} Z_1^T & Z_2^T \end{bmatrix}^T$ is a $l \times m_b$ matrix, which can be partitioned into an $s \times m_b$ matrix Z_1 (which contains the finite and gap zones) and an $(l-s) \times m_b$ matrix Z_2 (which contains the infinite zone), with $\operatorname{rank}(Z_1) = m_a < m_b$. Furthermore, let the singular value decomposition of $Z_1 = U\Sigma Q^{\mathrm{T}}$. Then, $V = Z_A Q$ is called the column compression of Z_A and can be partioned as

$$V = \begin{bmatrix} V_{11} & 0 \\ V_{21} & V_{22} \end{bmatrix}, \tag{24}$$

where V_{11} is the $s \times m_a$ matrix that corresponds to the compressed null space.

This compressed null space allows us to straightforwardly use the above-described finite solution approach and to find only the finite $(2r)$-tuples of affine eigenvalues of the AMEP.

A positive-dimensional solution set at infinity Sometimes it may happen that, although the finite solution set is zero-dimensional, the solution set at infinity is positive-dimensional. In contrary to problems with a zero-dimensional solution set, where the nullity stabilizes at total number of solutions m_b, the nullity of a positive-dimensional solution set does not stabilize. For example, if the set of infinite solutions is one-dimensional, then the nullity increases, but the nullity change stabilizes. Even in this case, we can still use the algorithm described in this section to correctly retrieve the finite solutions of the AMEP (see Dreesen (2013) for an example of this

Table 1. The stabilization diagram for the numerical example, showing the properties of the augmented block Macaulay matrix M as a function of its degree d.

degree	size	rank	nullity	nullity change
3	10×49	10	39	-
4	30×83	30	53	14
5	60×126	60	66	13
6	100×178	100	78	12
7	150×239	150	89	11
8	210×309	210	99	10
9	280×388	280	108	9
10	360×476	360	116	8
11	450×573	450	123	7
12	550×679	549	130	7
13	660×794	657	137	7
14	780×918	774	144	7

Table 2. The two stable reduced-order models $G_r(s)$ with real coefficients and nonzero numerators found via the augmented block Macaulay matrix approach and their associated H_2-error.

$\mathbf{G_r(s)}$	$\|\mathbf{G(s) - G_r(s)}\|_{\mathbf{H_2}}$
$G_1(s) = \frac{1.2799}{s+9.6796}$	0.2784
$G_2(s) = \frac{-0.0437}{s+0.2671}$	0.3982

positive-dimensional situation when rooting systems of multivariate polynomial equations).

5. NUMERICAL EXAMPLE

In this section, we present a small numerical proof-of-concept to illustrate the novel model reduction approach from this paper. We consider the transfer function

$$G(s) = \frac{s^2 + 9s - 10}{s^3 + 12s^2 + 49s + 78} ,$$

for which we want to compute the H_2-norm globally optimal first-order approximant $G_r(s) = \frac{b_1}{s+a_1}$.

For this example, the system in (14) consists of 10 multivariate polynomial equations in 10 unknowns, of which 8 appear only linearly. This translates into an affine quadratic 2-parameter eigenvalue problem with coefficient matrices $\mathcal{A}_\omega \in \mathbb{R}^{10\times 8}$ and a constant vector $q \in \mathbb{R}^{10}$, where the unknown parameters a_1 and b_1 constitute the 2-tuples of affine eigenvalues.

We observe that an augmented block Macaulay matrix $M \in \mathbb{R}^{660\times 794}$ of degree $d = 13$ suffices to find the gap in its null space. In this particular example, the nullity does not stabilize, but the nullity change does, which indicates that the solutions at infinity form a one-dimensional variety (see Table 1). Since in the first (finite) zone of Z_A we detect 8 linearly independent rows, the system of multivariate polynomial equations, and therefore the AMEP, has $m_a = 8$ finite solutions. Starting from these linearly independent rows, we construct the standard EP in (23), from which we retrieve the 8 finite solutions, i.e., the different 2-tuples (a_1, b_1), of the AMEP: $(9.6796, 1.2799)$, $(-16.6189, 1.9263)$, $(0.26711, -0.043711)$, $(-4.1639-0.90269i, 24.93+6.5394i)$, $(-4.1639+0.90269i, 24.93-6.5394i)$, $(1, -6.4513\times 10^{-15})$, $(-9.9999, -2.4217\times 10^{-4})$, and $(-6.2697 \times 10^{-12}, 1.5735 \times 10^{-12})$. Only two of these solutions lead to stable transfer functions with real coefficients and nonzero numerators, and they are shown in Table 2 together with their associated H_2-error. Fig. 2 visualizes the contour plot of the H_2-error for this numerical example. Clearly, the globally optimal first order approximant of $G(s)$ is [3]

$$G_1(s) = \frac{1.2799}{s + 9.6796} .$$

In order to corroborate the previous results, we used the iterative rational Krylov algorithm (IRKA) (Gugercin et al., 2008), available in the sssMOR (Sparse State-Space and Model Order Reduction) MATLAB toolbox (Castagnotto et al., 2017). We observe that, depending on the initialization, the algorithm can converge to one of the two solutions in Table 2 or to a solution that does not lead to a stable reduced model (e.g., $a_1 = -16.6189$, $b_1 = 1.9265$).

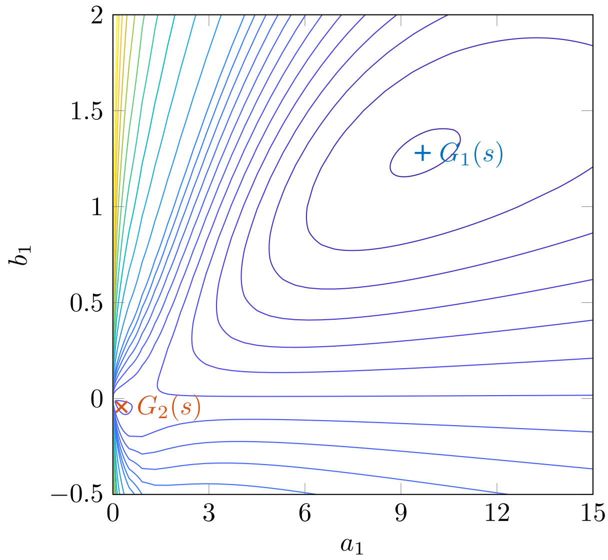

Fig. 2. The contour plot of the H_2-error $\|G(s) - G_r(s)\|_{H_2}$ for the numerical example. Here, $G_2(s)$ is a local minimizer (×) and $G_1(s)$ corresponds to the globally optimal solution (+).

6. CONCLUSIONS AND FUTURE RESEARCH

In this paper, we showed that the globally optimal H_2-norm model reduction problem for SISO LTI systems is essentially an eigenvalue problem. We proposed a novel numerical linear algebra algorithm to retrieve the globally optimal solution(s). This algorithm can be briefly summarized as follows: First, we translate the H_2-norm model reduction problem into a system of multivariate polynomial equations via the first-order optimality conditions of a conveniently redefined objective function. Then, we exploit the fact that in these equations several variables appear only linearly to formulate an affine (or inhomogeneous) quadratic multiparameter eigenvalue problem (AMEP). By using the augmented block Macaulay matrix introduced in this work, we take advantage of the (block) multi-shift-invariance property of its null space to transform the

[3] In Table 2, we only show the first four decimals of the H_2-error.

AMEP into a standard eigenvalue problem (EP), of which the solutions correspond to the set of all stationary points of the optimization problem. Finally, from the $(2r)$-tuples of affine eigenvalues, we select the real-valued tuple that leads to the stable transfer function with the smallest H_2-error. We provided a proof-of-concept with a numerical example, in which we computed the globally optimal first-order approximant of a third-order transfer function.

Notice that, as the order of the model $G(s)$ and its approximant $G_r(s)$ increase, the number of stationary points grows rapidly. Hence, solving the AMEP and evaluating all the solutions become very quickly impractical. Consequently, one of our current research efforts is to modify the algorithm in such a way that it only computes the optimal $(2r)$-tuple of affine eigenvalues. Exploiting the structure and sparsity of the augmented block Macaulay matrix to tackle large model reduction problems is also part of our future work, as well as a rigorous study of the properties of these AMEPs.

Although we did not address large practical problems in this paper, the mathematical claim that the H_2-norm model reduction problem for SISO LTI systems is an AMEP and the proposed solution approach for this type of problems are both important contributions to the field of systems and control.

REFERENCES

Ahmad, M.I., Frangos, M., and Jaimoukha, I.M. (2011). Second order $\mathcal{H}_2$ optimal approximation of linear dynamical systems. *IFAC Proceedings Volumes*, 44(1), 12767–12771.

Ahmad, M.I., Jaimoukha, I., and Frangos, M. (2010). $\mathcal{H}_2$ optimal model reduction of linear dynamical systems. In *Proc. of the 49th IEEE Conference on Decision and Control (CDC)*, 5368–5371. Atlanta, GA, USA.

Anić, B., Beattie, C., Gugercin, S., and Antoulas, A.C. (2013). Interpolatory weighted-$\mathcal{H}_2$ model reduction. *Automatica*, 49(5), 1275–1280.

Antoulas, A.C. (2005). *Approximation of Large-Scale Dynamical Systems.* Advances in Design and Control. SIAM, Philadelphia, PA, USA.

Antoulas, A.C., Beattie, C.A., and Gugercin, S. (2010). Interpolatory model reduction of large-scale dynamical systems. In J. Mohammadpour and K.M. Grigoriadis (eds.), *Efficient Modeling and Control of Large-Scale Systems*, 3–58. Springer, Boston, MA, USA.

Beattie, C. and Gugercin, S. (2017). Model reduction by rational interpolation. In P. Benner, A. Cohen, M. Ohlberger, and K. Willcox (eds.), *Model Reduction and Approximation*, 297–334. SIAM, Philadelphia, PA, USA.

Bunse-Gerstner, A., Kubalińska, D., Vossen, G., and Wilczek, D. (2010). h_2-norm optimal model reduction for large scale discrete dynamical MIMO systems. *Journal of Computational and Applied Mathematics*, 233(5), 1202–1216.

Castagnotto, A., Cruz Varona, M., Jeschek, L., and Lohmann, B. (2017). sss & sssMOR: Analysis and reduction of large-scale dynamic systems in MATLAB. *Automatisierungstechnik*, 65(2), 134–150.

De Moor, B. (2019). Least squares realization of LTI models is an eigenvalue problem. In *Proc. of the 18th European Control Conference (ECC)*, 2270–2275. Naples, Italy.

Dreesen, P. (2013). *Back to the Roots: Polynomial System Solving Using Linear Algebra.* Ph.D. thesis, KU Leuven, Leuven, Belgium.

Dreesen, P., Batselier, K., and De Moor, B. (2018). Multidimensional realisation theory and polynomial system solving. *International Journal of Control*, 91(12), 2692–2704.

Gugercin, S., Antoulas, A., and Beattie, C. (2008). $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30(2), 609–638.

Gugercin, S., Antoulas, A.C., and Beattie, C.A. (2006). A rational Krylov iteration for optimal $\mathcal{H}_2$ model reduction. In *Proc. of the 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, 1665–1667. Kyoto, Japan.

Horn, R.A. and Johnson, C.R. (1994). *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, UK.

Mattheij, R.M.M. and Söderlind, G. (1987). On inhomogeneous eigenvalue problems. *Linear Algebra and its Applications*, 88–89, 507–531.

Meier, L. and Luenberger, D. (1967). Approximation of linear constant systems. *IEEE Transactions on Automatic Control*, 12(5), 585–588.

Spanos, J.T., Milman, M.H., and Mingori, D.L. (1992). A new algorithm for L_2 optimal model reduction. *Automatica*, 28(5), 897–909.

Van Dooren, P., Gallivan, K.A., and Absil, P.A. (2008). $\mathcal{H}_2$-optimal model reduction of MIMO systems. *Applied Mathematics Letters*, 21(12), 1267–1273.

Vermeersch, C. and De Moor, B. (2019). Globally optimal least-squares ARMA model identification is an eigenvalue problem. *IEEE Control Systems Letters*, 3(4), 1062–1067.

Volkmer, H. (1988). *Multiparameter Eigenvalue Problems and Expansion Theorems*, volume 1356 of *Lecture Notes in Mathematics.* Springer, Berlin, Germany.

Žigić, D., Watson, L.T., and Beattie, C. (1993). Contragredient transformations applied to the optimal projection equations. *Linear Algebra and its Applications*, 188–189, 665–676.

Yan, W.Y. and Lam, J. (1999). An approximate approach to h_2 optimal model reduction. *IEEE Transactions on Automatic Control*, 44(7), 1341–1358.