Zentrum für Technomathematik

Fachbereich 3 – Mathematik und Informatik

Equivalences between necessary optimality conditions for $\mathcal{H}_2$-norm optimal model reduction

Daniel Wilczek   Angelika Bunse-Gerstner

Dorota Kubalińska   Georg Vossen

Report 07–06

Berichte aus der Technomathematik

Report 07–06   Oktober 2007

EQUIVALENCES BETWEEN NECESSARY OPTIMALITY CONDITIONS FOR $\mathcal{H}_2$-NORM OPTIMAL MODEL REDUCTION

D. WILCZEK, A. BUNSE-GERSTNER, G. VOSSEN AND D. KUBALINSKA

Abstract. In this paper the equivalences between necessary optimality conditions for $\mathcal{H}_2$-norm optimal model reduction for linear time invariant continuous MIMO systems will be proven. Initially three main optimality conditions, namely the Interpolation conditions, Wilson conditions and Hyland-Bernstein conditions, were introduced. While the equivalence proof between Wilson and Hyland-Bernstein conditions is already published and valid for MIMO systems and multiple poles within the system matrix A, the equivalence between Wilson and Interpolation conditions still has to be proven for this most general case. This is done in the main part of this paper.

Key words. Model reduction, $\mathcal{H}_2$-norm, necessary optimality conditions

1. Problem formulation. Consider the following linear time invariant (LTI) descriptor system in frequency space

$$\Sigma := \left(\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right) := \left\{ \begin{array}{rl} sX(s) - X(0) & = AX(s) + BU(s), \\ Y(s) & = CX(s), \end{array} \right\} \tag{1.1}$$

where $X \in \mathbb{C}^n$, $U \in \mathbb{C}^m$ and $Y \in \mathbb{C}^p$ are called the *state variable*, the *input variable* and the *output variable*, respectively. The matrices $A \in \mathbb{C}^{n,n}$, $B \in \mathbb{C}^{n,m}$ and $C \in \mathbb{C}^{p,n}$ are constant matrices w.r.t. the frequency variable $s \in \mathbb{C}$. For simplicity let $X(0) = 0$, which means that the initial state of the system is zero.

Another way to describe a system is the input/output behaviour. The quotient of output devided by input is called *transfer function* $H(s) = \frac{Y(s)}{U(s)}$. With the help of equations (1.1) it could be also written using the system matrices

$$H(s) = C(sI_n - A)^{-1}B$$

with I_n being the n-th order identity matrix.

All systems occuring in this paper are stable, reachable and controllable, i.e. all eigenvalues λ_j of A satisfy $\mathrm{Re}(\lambda_j) < 0$ and the *reachability matrix* defined by

$$\mathcal{R}_n(\Sigma) := K_n(A, B) := [B, AB, \dots, A^{n-1}B] \in \mathbb{C}^{n,nm}$$

and the *observability matrix* defined by

$$\mathcal{O}_n(\Sigma) := K_n^*(A^*, C^*) := [C^*, A^*C^*, \dots, (A^*)^{n-1}C^*]^* \in \mathbb{C}^{pn,n}$$

have full rank. $K_n(A, B)$ is called *Krylov matrix.* Let the matrices $\mathcal{P}$ and $\mathcal{Q}$ be solutions of the so called Lyapunov equations

$$A\mathcal{P} + \mathcal{P}A^* + BB^* = 0 \tag{1.2}$$

$$\mathcal{Q}A + A^*\mathcal{Q} + C^*C = 0\ . \tag{1.3}$$

They are defined as reachability and observability gramian, respectively.

The goal of model reduction via projection is to find an oblique projection $\Pi = VZ^*$ with projection matrices $V, Z \in \mathbb{C}^{n,r}$ and $Z^*V = I_r$ such that $\hat{Y}$ from

$$\hat{X}(s) = Z^*AV\hat{X}(s) + Z^*B\,U(s)\,, \qquad \hat{Y}(s) = CV\,\hat{X}(s)$$

approximates output $Y(s)$ of the original system. The model reduction problem could be solved by projecting the state $\Pi X(s)$. Defining $\hat{X}(s) := Z^* X(s)$ leads to the system above.

Thus the reduced order system $\hat{\Sigma}$ is obtained by the projection matrices Z and V

$$\hat{\Sigma} = \left(\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & 0 \end{array}\right) = \left(\begin{array}{c|c} Z^*AV & Z^*B \\ \hline CV & 0 \end{array}\right) . \tag{1.4}$$

with $\hat{A} \in \mathbb{C}^{r,r}$, $\hat{B} \in \mathbb{C}^{r,m}$ and $\hat{C} \in \mathbb{C}^{p,r}$.

For $\mathcal{H}_2$-norm model reduction the $\mathcal{H}_2$-norm is used as a measure of approximation. For the system Σ it is defined with the help of the transfer function $H(s)$ [1]

$$||H||^2_{\mathcal{H}_2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \operatorname{trace}\left(H(\mathrm{i}w)^* H(\mathrm{i}w)\right) dw, \tag{1.5}$$

where i is the imaginary number with $\mathrm{i}^2 = -1$.

Hence, the aim of $\mathcal{H}_2$-norm optimal model reduction, namely the approximation of the output $\hat{Y}(s)$ of the projected (reduced) system to the output $Y(s)$ of the original system, is

$$\begin{aligned} \min_{\hat{Y}} J(\hat{Y}) &= ||Y(s) - \hat{Y}(s)||^2_{\mathcal{H}_2} \\ &= ||H(s)U(s) - \hat{H}(s)U(s)||^2_{\mathcal{H}_2} = ||H - \hat{H}||^2_{\mathcal{H}_2} U \\ \text{or equivalently} \qquad \min_{\hat{\Sigma}} J(\hat{\Sigma}) &= ||\Sigma - \hat{\Sigma}||^2_{\mathcal{H}_2} . \end{aligned} \tag{1.6}$$

The difference between original and reduced system is the so-called *error system*

$$\Sigma - \hat{\Sigma} = \left(\begin{array}{cc|c} A & 0 & B \\ 0 & \hat{A} & \hat{B} \\ \hline C & -\hat{C} & 0 \end{array}\right) . \tag{1.7}$$

As long as the representation (1.1) of the system Σ is unique except for state basis transformations we could either identify the system with its matrices A, B and C or with its transfer function H. If system Σ fullfilles certain properties it is called a *real system.*

Definition 1.1 (Real system). *A system H is called real if there exist real matrices A, B and C such that $H = C(s\,\mathrm{Id} - A)^{-1}B$ holds where* Id *is the identity mapping.*

2. First-order $\mathcal{H}_2$ optimality conditions. In this section we will briefly review three different necessary optimality conditions namely the Interpolation, the Wilson and the Hyland-Bernstein conditions for $\mathcal{H}_2$-norm optimal model reduction. While the Interpolation conditions describe a kind of Hermite-interpolation of the transfer function in special points called mirror images, the optimality conditions of Wilson and Hyland-Bernstein are connected with Lyapunov equations.

2.1. Interpolation conditions. Before regarding the Interpolation conditions it is helpful to briefly introduce different representations of the transfer function $\hat{H}(s)$ (for more details view [4]). A transfer function of a stable system could also be written as a quotient of two polynomials

$$\hat{H}(s) = \frac{\sum_{k=0}^{n-1} \alpha_k s^k}{\sum_{k=0}^{n} \beta_k s^k}, \quad \beta_n \neq 0, \tag{2.1}$$

with complex coefficients α_k, $k = 0, \ldots, n-1$ and β_k, $k = 0, \ldots, n$. The eigenvalues of the matrix A, i.e. the poles of the system (1.1) correspond to the zeros of the denominator. Expanding $H(s)$ into its Laurent series around each pole λ_j, $j = 1, \ldots, R$ yields

$$\hat{H}(s) = \sum_{k=-\infty}^{\infty} \gamma_k (s - \hat{\lambda}_j)^k \ .$$

Here, γ_k are called the Laurent coefficients of $H(s)$ at λ_j and γ_{-1} is called the residue. The order k_{0_j} of a pole λ_j is defined as the lowest index k_0 such that $\gamma_k = 0$ holds for all $k > -k_{0_j}$. For simplicity we define $r_j := k_{0_j}$. Therewith r_j is the algebraic multiplicity of the j-th eigenvalue.

Now we could introduce the following conditions.

THEOREM 2.1 (Interpolation conditions). *Necessary conditions for $\mathcal{H}_2$-norm optimal model reduction problem (1.6) for reduced systems with R pairwise different poles ($\sum_{j=1}^{R} r_j = r$) are given by [4]*

$$\sum_{q=0}^{r_j-k_j} \frac{(-1)^q}{q!} H^{(q)}(-\hat{\lambda}_j^*) \hat{b}^*_{\mathtt{r}_{j-1}+k_j+q} = \sum_{q=0}^{r_j-k_j} \frac{(-1)^q}{q!} \hat{H}^{(q)}(-\hat{\lambda}_j^*) \hat{b}^*_{\mathtt{r}_{j-1}+k_j+q} \tag{2.2}$$

$$\sum_{q=0}^{k_j-1} \frac{(-1)^q}{q!} \hat{c}^*_{\mathtt{r}_{j-1}+k_j-q} H^{(q)}(-\hat{\lambda}_j^*) = \sum_{q=0}^{k_j-1} \frac{(-1)^q}{q!} \hat{c}^*_{\mathtt{r}_{j-1}+k_j-q} \hat{H}^{(q)}(-\hat{\lambda}_j^*) \tag{2.3}$$

and

$$\begin{aligned} &\sum_{q=1}^{r_j} \tfrac{(-1)^q}{q!} \sum_{p=0}^{r_j-q} \hat{c}^*_{\mathtt{r}_{j-1}+p+1} H^{(q)}(-\hat{\lambda}_j^*) \hat{b}^*_{\mathtt{r}_j+p+q} \\ = &\sum_{q=1}^{r_j} \tfrac{(-1)^q}{q!} \sum_{p=0}^{r_j-q} \hat{c}^*_{\mathtt{r}_{j-1}+p+1} \hat{H}^{(q)}(-\hat{\lambda}_j^*) \hat{b}^*_{\mathtt{r}_j+p+q} \end{aligned} \tag{2.4}$$

where $\mathtt{r}_l := \sum_{i=1}^{l} r_i$, $k_j = 1, \ldots, r_j$, $j = 1, \ldots, R$, $\hat{b}_l :=$ l-th row of $\hat{B}$ and $\hat{c}_l :=$ l-th column of $\hat{C}$.

Thus the sum of certain derivatives of $H(s)$ and $\hat{H}(s)$ in special points pre- or postmultiplied with certain columns of $\hat{C}$ or rows of $\hat{B}$, respectively, must coincide. These special points $-\hat{\lambda}_j^*$, $j = 1, \ldots, r$ are called mirror images and we should remind that they are unknown a-priori.

The above conditions describe the most general case, namely MIMO systems with multiple, complex poles. For SISO systems or generic systems (i.e. all poles are distinct ($r_j = 1 \ \forall \hat{\lambda}_j, \ \ j = 1, \ldots, r$)) the conditions become much more simple.

The Interpolation conditions for SISO systems were first pointed out by Meier and Luenberger [8]. A new proof for the SISO case and single, real poles was given by Gugercin, Antoulas and Beattie [3]. A generalization for MIMO systems with multiple poles is given in [4]. The following remark presents the Interpolation conditions for simple poles.

Remark 2.2 (Interpolation conditions for simple poles). *The necessary Interpolation conditions for $\mathcal{H}_2$-norm optimal model reduction problem (1.6) for reduced generic systems are given by [4]*

$$\left.\begin{aligned} H(-\hat{\lambda}_j^*)\hat{b}_j^* &= \hat{H}(-\hat{\lambda}_j^*)\hat{b}_j^*, \\ \hat{c}_j^* H(-\hat{\lambda}_j^*) &= \hat{c}_j^*\hat{H}(-\hat{\lambda}_j^*), \\ \hat{c}_j^* H'(-\hat{\lambda}_j^*)\hat{b}_j^* &= \hat{c}_j^*\hat{H}'(-\hat{\lambda}_j^*)\hat{b}_j^*, \end{aligned}\right\} \quad j = 1, \ldots, r, \tag{2.5}$$

i.e. one-sided tangential interpolation of the transfer functions in the first moments and two-sided tangential interpolation in the second moment at the mirror images.

Consider that in the SISO case the rows $\hat{b}_j$ and the columns $\hat{c}_j$ $(j = 1, \ldots, r)$ are scalars. Hence they can be coated and the Interpolation conditions simplify to

$$H(-\hat{\lambda}_j^*) = \hat{H}(-\hat{\lambda}_j^*) \quad H'(-\hat{\lambda}_j^*) = \hat{H}'(-\hat{\lambda}_j^*), \qquad j = 1, \ldots, r\,. \tag{2.6}$$

Remark 2.3. *For the first m derivatives of a transfer function in the point $\hat{\lambda}_j^*$ we obtain*

$$\begin{aligned} &H'(-\hat{\lambda}_j^*) = H'(s)\big|_{s=-\hat{\lambda}_j^*} = -C(-A - \hat{\lambda}_j^* I)^{-2}B \\ &H^{(q)}(-\hat{\lambda}_j^*) = (-1)^q\, q!\, C(-A - \hat{\lambda}_j^* I)^{-(q+1)}B \quad \textit{for} \quad q \geq 0\,. \end{aligned} \tag{2.7}$$

2.2. Wilson conditions. The Wilson optimality conditions are framed in terms of Lyapunov equations. Therefore it is necessary to introduce some related coherences. The system matrices of the error system (1.7) are

$$A_e = \begin{pmatrix} A & 0 \\ 0 & \hat{A} \end{pmatrix} \qquad B_e = \begin{pmatrix} B \\ \hat{B} \end{pmatrix} \qquad C_e = \begin{pmatrix} C & -\hat{C} \end{pmatrix}\,.$$

Now consider the two Lyapunov equations of the error system

$$\begin{aligned} A_e\mathcal{P}_e + \mathcal{P}_e A_e^* + B_e B_e^* = 0 \\ \mathcal{Q}_e A_e + A_e^*\mathcal{Q}_e + C_e^* C_e = 0 \end{aligned}$$

where the symmetric matrices $\mathcal{P}_e$ and $\mathcal{Q}_e$ are the reachability and observability gramians of the error system, respectively. Partitioning $\mathcal{P}_e$ and $\mathcal{Q}_e$ leads to

$$\mathcal{P}_e = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{bmatrix} \qquad \mathcal{Q}_e = \begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{21} & \mathcal{Q}_{22} \end{bmatrix},$$

whereas $\mathcal{P}_{11}, \mathcal{Q}_{11} \in \mathbb{C}^{n,n}; \mathcal{P}_{12}, \mathcal{P}_{21}^*, \mathcal{Q}_{12}, \mathcal{Q}_{21}^* \in \mathbb{C}^{n,r}$ and $\mathcal{P}_{22}, \mathcal{Q}_{22} \in \mathbb{C}^{r,r}$. The full-rank submatrices $\mathcal{P}_{11}, \mathcal{Q}_{11}, \mathcal{P}_{22}$ and $\mathcal{Q}_{22}$ solve the Lyapunov equations (1.2), (1.3) and

$$\hat{A}\mathcal{P}_{22} + \mathcal{P}_{22}\hat{A}^* + \hat{B}\hat{B}^* = 0 \tag{2.8}$$

$$\mathcal{Q}_{22}\hat{A} + \hat{A}^*\mathcal{Q}_{22} + \hat{C}^*\hat{C} = 0 \tag{2.9}$$

and hence, they are the gramians of the original and the reduced system, respectively. Due to the fact that gramians are symmetric we obtain $\mathcal{P}_{12} = \mathcal{P}_{21}^*$ and $\mathcal{Q}_{12} = \mathcal{Q}_{21}^*$ and both matrices are the solutions of the following Sylvester-equations

$$A\mathcal{P}_{12} + \mathcal{P}_{12}\hat{A}^* + B\hat{B}^* = 0 \tag{2.10}$$

$$A^*\mathcal{Q}_{12} + \mathcal{Q}_{12}\hat{A} - C^*\hat{C} = 0\ . \tag{2.11}$$

Finding an $\mathcal{H}_2$-norm optimal reduced model for a real system Σ requires to determine the first derivatives of the error functional $J(\hat{A}, \hat{B}, \hat{C})$. The derivatives of J with respect to the elements of $\hat{A}, \hat{B}$ and $\hat{C}$ namely $\hat{a}, \hat{b}$ and $\hat{c}$ give the following necessary conditions [10].

Theorem 2.4 (Wilson Conditions). *The necessary conditions of Wilson for $\mathcal{H}_2$-norm optimal model reduction problem (1.6) are*

$$\mathcal{P}_{12}^*\mathcal{Q}_{12} + \mathcal{P}_{22}\mathcal{Q}_{22} = 0 \tag{2.12}$$

$$\mathcal{Q}_{12}^*B + \mathcal{Q}_{22}\hat{B} = 0 \tag{2.13}$$

$$\hat{C}\mathcal{P}_{22} - C\mathcal{P}_{12} = 0\ . \tag{2.14}$$

Remark 2.5. *Wilson conditions for real systems are proved in [10]. The idea of a generalization for complex systems is given in [2]. Remember that the reduced system is always a projection of the original system with the projection matrix $\Pi = VZ^*$. The projection matrices Z and V could be deduced via a comparison of the conditions (2.13) and (2.14) with the reduced system (1.4)*

$$V := \mathcal{P}_{12}\mathcal{P}_{22}^{-1} \qquad Z := -\mathcal{Q}_{12}\mathcal{Q}_{22}^{-1}\ .$$

Condition (2.12) assures $Z^*V = I$.

2.3. Hyland-Bernstein conditions. Similar to the Wilson conditions we provide the Hyland-Bernstein conditions by means of the gramians and the Lyapunov equations.

Theorem 2.6 (Hyland-Bernstein Conditions). *Suppose the system $\hat{\Sigma}$ solves the $\mathcal{H}_2$-norm optimal model reduction problem (1.6). Then there exist two nonnegative-definite matrices $\mathcal{P}, \mathcal{Q} \in \mathbb{C}^{n,n}$ and a positive-definite matrix $M \in \mathbb{C}^{r,r}$ such that [6]*

$$\mathcal{P}\mathcal{Q} = VMZ^* \tag{2.15}$$

$$\operatorname{rank}\mathcal{P} = \operatorname{rank}\mathcal{Q} = \operatorname{rank}\mathcal{P}\mathcal{Q}\ . \tag{2.16}$$

Furthermore the projection matrix Π of the reduced system $\hat{\Sigma}$ satisfies the following two conditions

$$\Pi\,[A\mathcal{P} + \mathcal{P}A^* + BB^*] = 0$$

$$[A^*\mathcal{Q} + \mathcal{Q}A + CC^*]\,\Pi = 0\ .$$

3. Equivalence between the necessary conditions. All conditions presented in the last section are equivalent to each other. This was already pointed out by Gugercin, Antoulas and Beattie [3] for the SISO systems with single poles. Here we expand those equivalence proofs for the MIMO case and multiple poles.

3.1. Equivalence between Interpolation and Wilson conditions. The equivalence between Interpolation and Wilson conditions could be verified by a proper analysis of the projection $\Pi = VZ^*$. The following lemma reveals how to qualify the projection matrix V.

LEMMA 3.1. *The following statements are equivalent.*

$$\begin{array}{rrl} (i) & V = & \mathcal{P}_{12}\mathcal{P}_{22}^{-1} \\ (ii) & \operatorname{Ran} V = & \operatorname{colspan}\{v_1, v_2, \dots, v_r\} \quad \textit{with} \\ & v_{\mathtt{r}_{j-1}+k_j} := & \sum_{q=0}^{r_j-k_j} (-A-\hat{\lambda}_j^* I)^{-(q+1)} B \hat{b}^*_{\mathtt{r}_{j-1}+k_j+q} \\ & = & K_{r_j-k_j+1}\Big((-A-\hat{\lambda}_j^* I)^{-1}, (-A-\hat{\lambda}_j^* I)^{-1} B\Big) \\ & & \cdot \Big[\hat{b}_{\mathtt{r}_{j-1}+k_j}, \hat{b}_{\mathtt{r}_{j-1}+k_j+1}, \dots, \hat{b}_{\mathtt{r}_{j-1}+r_j}\Big]^* \end{array}$$

where $\mathtt{r}_l := \sum_{i=1}^{l} r_i$, $1 \le j \le R$, $1 \le k_j \le r_j$ *and* $\hat{B}^* = \Big[\hat{b}_1^*, \dots, \hat{b}_r^*\Big]$.

Proof: Without loss of generality it is applicable to assume that $\hat{A} = \operatorname{diag}[J_1, \dots, J_R]$, $\hat{B}$ and $\hat{C}$ build an eigenvector basis. The Jordan matrices J_j of the R pairwise different eigenvalues $\hat{\lambda}_j$, $j = 1, \dots, R$, each of order r_j, is a $r_j \times r_j$-dimensional matrix with $\hat{\lambda}_j$ on its diagonal, ones on the super-diagonal and zeros elsewhere.
Consider Silvester-equation (2.10) with

$$\mathcal{P}_{12} = \begin{bmatrix} p_{1,1} \cdots p_{1,\mathtt{r}_1} & p_{1,\mathtt{r}_1+1} \cdots p_{1,\mathtt{r}_2} & \cdots & p_{1,\mathtt{r}_{R-1}+1} \cdots p_{1,\mathtt{r}_R} \\ \vdots \ \ \ddots \ \ \vdots & \vdots \ \ \ddots \ \ \vdots & \vdots & \vdots \ \ \ddots \ \ \vdots \\ p_{n,1} \cdots p_{n,\mathtt{r}_1} & p_{n,\mathtt{r}_1+1} \cdots p_{n,\mathtt{r}_2} & \cdots & p_{n,\mathtt{r}_{R-1}+1} \cdots p_{n,\mathtt{r}_R} \end{bmatrix} \quad \text{and}$$

$$\hat{A} = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_R \end{bmatrix} .$$

A rearrangement of this matrix equation leads to

$$\begin{array}{lllllllll} Ap_1 & + & p_1\hat{\lambda}_1^* & + & p_2 & + & B\hat{b}_1^* & = 0 \\ Ap_2 & + & p_2\hat{\lambda}_1^* & + & p_3 & + & B\hat{b}_2^* & = 0 \\ & & \vdots & & & & \vdots & \\ Ap_{\mathtt{r}_1-1} & + & p_{\mathtt{r}_1-1}\hat{\lambda}_1^* & + & p_{\mathtt{r}_1} & + & B\hat{b}_{\mathtt{r}_1-1}^* & = 0 \\ Ap_{\mathtt{r}_1} & + & p_{\mathtt{r}_1}\hat{\lambda}_1^* & + & & & B\hat{b}_{\mathtt{r}_1}^* & = 0 \\ & & \vdots & & & & \vdots & \\ Ap_{\mathtt{r}_{R-1}+1} & + & p_{\mathtt{r}_{R-1}+1}\hat{\lambda}_R^* & + & p_{\mathtt{r}_{R-1}+2} & + & B\hat{b}_{\mathtt{r}_{R-1}+1}^* & = 0 \\ & & \vdots & & & & \vdots & \\ Ap_{\mathtt{r}_R} & + & p_{\mathtt{r}_R}\hat{\lambda}_R^* & + & & & B\hat{b}_{\mathtt{r}_R}^* & = 0 \, , \end{array}$$

where p_k ($k = 1, \ldots, r = \mathbb{r}_R$) are the columns of $\mathcal{P}_{12}$. Now dissolve these equation with respect to p_k

$$\begin{aligned}
p_1 &= (-A - \hat{\lambda}_1^* I)^{-1} \quad (B\hat{b}_1^* + p_2) \\
p_2 &= (-A - \hat{\lambda}_1^* I)^{-1} \quad (B\hat{b}_2^* + p_3) \\
&\vdots \\
p_{\mathbb{r}_1 - 1} &= (-A - \hat{\lambda}_1^* I)^{-1} \quad (B\hat{b}_{\mathbb{r}_1-1}^* + p_{\mathbb{r}_1}) \\
p_{\mathbb{r}_1} &= (-A - \hat{\lambda}_1^* I)^{-1} \quad B\hat{b}_{\mathbb{r}_1}^* \\
&\vdots \\
p_{\mathbb{r}_{R-1}+1} &= (-A - \hat{\lambda}_R^* I)^{-1} \quad (B\hat{b}_{\mathbb{r}_{R-1}+1}^* + p_{\mathbb{r}_R}) \\
&\vdots \\
p_{\mathbb{r}_R} &= (-A - \hat{\lambda}_R^* I)^{-1} \quad B\hat{b}_{\mathbb{r}_R}^* \; .
\end{aligned}$$

We could expand the above equations and substitute each result successively in the next equation

$$p_{\mathbb{r}_{j-1}+k_j} = (-A - \hat{\lambda}_j^* I)^{-1} B\hat{b}_{\mathbb{r}_{j-1}+k_j}^* + \cdots + (-A - \hat{\lambda}_j^* I)^{-(r_j - k_j + 1)} B\hat{b}_{\mathbb{r}_{j-1}+r_j}^* \; .$$

Up to this point we transformed only algebraic equations. For the implication $(i) \Longrightarrow (ii)$ the Wilson condition (2.14) comes into operation. Remember that $V := \mathcal{P}_{12}\mathcal{P}_{22}^{-1}$ and $\mathcal{P}_{22}^{-1}$ has full-rank. Hence it holds that the image of $\mathcal{P}_{12}$ is a subset of the image of $\mathcal{P}_{22}^{-1}$ and therefore is equal to the image of V which is an intersection of both

$$\begin{aligned}
\operatorname{Ran} V = \operatorname{Ran} \mathcal{P}_{12} \cap \operatorname{Ran} \mathcal{P}_{22}^{-1} &\underset{\operatorname{Ran} \mathcal{P}_{12} \subseteq \operatorname{Ran} \mathcal{P}_{22}^{-1}}{=} \operatorname{Ran} \mathcal{P}_{12} \\
&= \operatorname{colspan}\{p_1, \ldots, p_r\} \\
&= \operatorname{colspan} \left\{ \sum_{q=0}^{r_1 - 1} (-A - \hat{\lambda}_1^* I)^{-(q+1)} B\hat{b}_{q+1}^*, \ldots, (-A - \hat{\lambda}_R^* I)^{-1} B\hat{b}_r^* \right\} .
\end{aligned}$$

For the proof of the reverse implication $(ii) \Longrightarrow (i)$ we have to show that

$$\operatorname{Ran} V = \operatorname{colspan}\{p_1, \ldots, p_r\} \quad \Longrightarrow \quad V = \mathcal{P}_{12}\mathcal{P}_{22}^{-1} \; .$$

The left side leads to $V = \mathcal{P}_{12} * K$, where $K \in \mathbb{C}^{r,r}$ is a nonsingular matrix. Premultiply equation (2.10) with Z^* yields

$$Z^* A \mathcal{P}_{12} + Z^* \mathcal{P}_{12} \hat{A}^* + Z^* B \hat{B}^* = 0 \; .$$

Because the matrices V and Z describe an oblique projection we get the following results

$$\begin{aligned}
Z^* V = I_r &\quad \Longrightarrow \quad Z^* \mathcal{P}_{12} = K^{-1} \\
Z^* A V = \hat{A} &\quad \Longrightarrow \quad Z^* A \mathcal{P}_{12} = \hat{A} K^{-1}
\end{aligned}$$

Thus we obtain

$$\hat{A} K^{-1} + K^{-1} \hat{A}^* + \hat{B}\hat{B}^* = 0$$

which is indeed the Lyapunov equation (2.8) for the reachability gramian of the reduced system. Consequently K^{-1} equals $\mathcal{P}_{22}$ which completes the proof.

Equivalently the projection matrix Z could be determined with the following lemma.

LEMMA 3.2. *The following statements are equivalent.*

$$\begin{array}{lrl} (i) & Z = & -\mathcal{Q}_{12}\mathcal{Q}_{22}^{-1} \\ (ii) & \operatorname{Ran} Z^* = & \operatorname{rowspan}\{z_1^*, z_2^*, \dots, z_r^*\} \quad \textit{with} \\ & z^*_{\mathtt{r}_{j-1}+k_j} := & \sum_{q=0}^{k_j-1} \hat{c}^*_{\mathtt{r}_{j-1}+k_j-q} C(-A-\hat{\lambda}_j^* I)^{-(q+1)} \\ & = & \left[\hat{c}^*_{\mathtt{r}_{j-1}+k_j}, \hat{c}^*_{\mathtt{r}_{j-1}+k_j-1}, \dots, \hat{c}^*_{\mathtt{r}_{j-1}+1}\right] \\ & & \cdot K^*_{k_j}\left((-A^*-\hat{\lambda}_j I)^{-1}, (-A^*-\hat{\lambda}_j I)^{-1} C^*\right) \end{array}$$

where $\mathtt{r}_l := \sum_{i=1}^{l} r_i$, $1 \le j \le R$, $1 \le k_j \le r_j$ *and* $\hat{C} = [\hat{c}_1, \dots, \hat{c}_r]$.

Proof: Analogues to the preceeding proof using (2.11) instead of (2.10) we get a similar expression of the columns q_k $(k = 1, \dots, r)$ of $\mathcal{Q}_{12}$

$$\begin{cases} q_1 & = (A^* + \hat{\lambda}_1 I)^{-1} \quad C^*\hat{c}_1 \\ q_2 & = (A^* + \hat{\lambda}_1 I)^{-1} \quad (C^*\hat{c}_2 - q_1) \\ & \vdots \\ q_{\mathtt{r}_1} & = (A^* + \hat{\lambda}_1 I)^{-1} \quad (C^*\hat{c}_{\mathtt{r}_1} - q_{\mathtt{r}_1-1}) \\ & \vdots \\ q_{\mathtt{r}_{R-1}+1} & = (A^* + \hat{\lambda}_R I)^{-1} \quad C^*\hat{c}_{\mathtt{r}_{R-1}+1} \\ & \vdots \\ q_{\mathtt{r}_R} & = (A^* + \hat{\lambda}_R I)^{-1} \quad (C^*\hat{c}_{\mathtt{r}_R} - q_{\mathtt{r}_R-1}) \, . \end{cases}$$

These equations could be expanded to

$$\begin{aligned} q_{\mathtt{r}_{j-1}+k_j} =& (A^* + \hat{\lambda}_j I)^{-1} C^* \hat{c}_{\mathtt{r}_{j-1}+k_j} - \dots + (-1)^{k_j-1}(A^* + \hat{\lambda}_j I)^{-k_j} C^* \hat{c}_{\mathtt{r}_{j-1}+1} \\ =& -\sum_{q=0}^{k_j-1} (-A^* - \hat{\lambda}_j I)^{-(q+1)} C^* \hat{c}_{\mathtt{r}_{j-1}+k_j-q} \end{aligned}$$

and we get analogous results for the projection matrix $Z = -\mathcal{Q}_{12}\mathcal{Q}_{22}^{-1}$

$$\begin{aligned} \operatorname{Ran} Z &= \operatorname{colspan}\{-q_1, \dots, -q_r\} \\ &= \operatorname{colspan}\{(-A^* - \hat{\lambda}_1 I)^{-1} C^* \hat{c}_1, \dots, \sum_{q=0}^{r_R-1} (-A^* - \hat{\lambda}_R I)^{-1} C^* \hat{c}_{\mathtt{r}_R-q}\} \\ \operatorname{Ran} Z^* &= \operatorname{rowspan}\{\hat{c}_1^* C(-A - \hat{\lambda}_1^* I)^{-1}, \dots, \sum_{q=0}^{r_R-1} \hat{c}^*_{\mathtt{r}_R-q} C(-A - \hat{\lambda}_R^* I)^{-1}\} \, . \end{aligned}$$

On the other hand the equations above lead to $Z = \mathcal{Q}_{12}L$, where $L \in \mathbb{R}^{r,r}$ is a nonsingular matrix. Now by premultiplying equation (2.11) with $-V^*$ and concerning that V and Z describe an oblique projection we get

$$-\underbrace{V^* A^* \mathcal{Q}_{12}}_{=\hat{A}^* L^{-1}} - \underbrace{V^* \mathcal{Q}_{12}}_{L^{-1}} \hat{A} + \underbrace{V^* C^*}_{\hat{C}^*} \hat{C} = 0 \, .$$

A comparison with the Lyapunov equation (2.9) of the observability gramian of the reduced systems yields $-L^{-1} = \mathcal{Q}_{22}$.

REMARK 3.3. *If the system has only one single input the columnspan of* $\operatorname{Ran} V$ *simplifies to*

$$\begin{aligned}
\operatorname{Ran} V =& \textit{colspan}\{v_1, v_2, \ldots, v_r\} \quad \textit{with} \\
v_{\mathtt{r}_{j-1}+k_j} :=& \sum_{m=0}^{r_j-k_j} (-A - \hat{\lambda}_j^* I)^{-(m+1)} B \qquad \textit{with} \quad B \in \mathbb{C}^{n,1} \\
=& K_{r_j-k_j+1}\Big((-A - \hat{\lambda}_j^* I)^{-1}, (-A - \hat{\lambda}_j^* I)^{-1} B\Big) \cdot \mathbb{1}_{k_j} \\
\textit{with} \quad \mathbb{1}_{k_j} =& (1, 1, \ldots, 1)^* \in \mathbb{R}^{k_j,1} \ ,
\end{aligned}$$

for $1 \le j \le R$ *and* $1 \le k_j \le r_j$.

Else if the system has only one single output the rowspan of $\operatorname{Ran} Z^*$ *simplifies to*

$$\begin{aligned}
\operatorname{Ran} Z^* =& \textit{rowspan}\{z_1^*, z_2^*, \ldots, z_r^*\} \quad \textit{with} \\
z_{\mathtt{r}_{j-1}+k_j}^* :=& \sum_{m=0}^{k_j-1} C(-A - \hat{\lambda}_j^* I)^{-(m+1)} \qquad \textit{and} \quad C \in \mathbb{C}^{1,n} \\
=& \mathbb{1}_{k_j}^* \cdot K_{k_j}^*\Big((-A^* - \hat{\lambda}_j I)^{-1}, (-A^* - \hat{\lambda}_j I)^{-1} C^*\Big)
\end{aligned}$$

for $1 \le j \le R$ *and* $1 \le k_j \le r_j$.

The following lemma connects the previous results with the Interpolation conditions. For simple poles it was proven in [9]. Here we expand the proof to multiple poles.

LEMMA 3.4. *Let* $V \in \mathbb{C}^{n,r}$ *and* $Z \in \mathbb{C}^{n,r}$ *be matrices of full rank* r *such that* $Z^* V = I_r$. *Let* $\sigma_l \in \mathbb{C}$, $l = 1, \ldots, R$, *be given points and let* $\ell_l \in \mathbb{C}^{1\times p}$ *and* $\rho_l \in \mathbb{C}^{m\times 1}$, $l = 1, \ldots, r$, *be given vectors. If*

$$\begin{aligned}
\operatorname{Ran} V =& \operatorname{colspan}\{v_1, v_2, \ldots, v_r\} \quad \textit{with} \\
v_{\mathtt{r}_{j-1}+k_j} :=& \sum_{q=0}^{r_j-k_j} (-A + \sigma_j I)^{-(q+1)} B \rho_{\mathtt{r}_{j-1}+k_j+q} \quad \textit{and} \\
\operatorname{Ran} Z^* =& \operatorname{rowspan}\{z_1^*, z_2^*, \ldots, z_r^*\} \quad \textit{with} \\
z_{\mathtt{r}_{j-1}+k_j}^* :=& \sum_{q=0}^{k_j-1} \ell_{\mathtt{r}_{j-1}+k_j-q} C(-A + \sigma_j I)^{-(q+1)} \ ,
\end{aligned}$$

where $\mathtt{r}_l := \sum_{i=1}^{l} r_i$, $\mathtt{r}_R = r$, *holds for* $1 \le j \le R$ *and* $1 \le k_j \le r_j$ *the following* *tangential Hermite interpolation conditions are satisfied*

$$\sum_{q=0}^{r_j-k_j} \frac{(-1)^q}{q!} H^{(q)}(\sigma_j)\rho_{\mathtt{r}_{j-1}+k_j+q} = \sum_{q=0}^{r_j-k_j} \frac{(-1)^q}{q!} \hat{H}^{(q)}(\sigma_j)\rho_{\mathtt{r}_{j-1}+k_j+q}$$
$$\sum_{q=0}^{k_j-1} \frac{(-1)^q}{q!} \ell_{\mathtt{r}_{j-1}+k_j-q} H^{(q)}(\sigma_j) = \sum_{q=0}^{k_j-1} \frac{(-1)^q}{q!} \ell_{\mathtt{r}_{j-1}+k_j-q} \hat{H}^{(q)}(\sigma_j)$$
$$\sum_{q=1}^{r_j} \frac{(-1)^q}{q!} \sum_{p=0}^{r_j-q} \ell_{\mathtt{r}_{j-1}+p+1} H^{(q)}(\sigma_j)\rho_{\mathtt{r}_{j-1}+p+q}$$
$$= \sum_{q=1}^{r_j} \frac{(-1)^q}{q!} \sum_{p=0}^{r_j-q} \ell_{\mathtt{r}_{j-1}+p+1} \hat{H}^{(q)}(\sigma_j)\rho_{\mathtt{r}_{j-1}+p+q} .$$

Proof: First of all we define two variables

$$M_j := (-A + \sigma_j I)$$
$$Y_j^* := (Z^* M_j V)^{-1} Z^* M_j .$$

Obviously it holds $Y^*V = I_r$. Now consider the right side of the first equation of the Hermite interpolation conditions and keep in mind that the reduced system (1.4) is constructed by an oblique projection

$$\sum_{q=0}^{r_j-k_j} \frac{(-1)^q}{q!} \hat{H}^{(q)}(\sigma_j)\rho_{\mathtt{r}_{j-1}+k_j+q} \underset{(2.7)}{=} \sum_{q=0}^{r_j-k_j} CV\left[Z^*M_jV\right]^{-(q+1)} Z^*B\rho_{\mathtt{r}_{j-1}+k_j+q}$$
$$= \sum_{q=0}^{r_j-k_j} CV\left[Z^*M_jV\right]^{-q} Z^*VY_j^*M_j^{-1}B\rho_{\mathtt{r}_{j-1}+k_j+q} .$$

Since $B\rho_{\mathtt{r}_j} = M_j\, v_{\mathtt{r}_j}$ with regular M_j it yields $B\rho_{\mathtt{r}_j} \in \operatorname{colspan}(V)$. The same holds for

$$B\rho_{\mathtt{r}_{j-1}+k_j} = M_j \left(v_{\mathtt{r}_{j-1}+k_j} - \sum_{q=k_j+1}^{r_j} M_j^{-(q-k_j)} B\rho_{\mathtt{r}_{j-1}+q} \right) \qquad k_j = r_j - 1, \ldots, 1$$

as a linear combination of vectors in colspan(V). Additionally $\mathbb{v} = VY_j^*\mathbb{v}$ for $\mathbb{v} \in$ colspan(V) [9]. Thus the above equation simplifies to

$$= \sum_{q=0}^{r_j-k_j} CV\left[Z^*M_jV\right]^{-q} Z^*M_j^{-1}B\rho_{\mathtt{r}_{j-1}+k_j+q}$$
$$= \sum_{q=0}^{r_j-k_j} CV\underbrace{\left[Z^*M_jV\right]^{-1} Z^*M_j}_{=Y_j^*} M_j^{-(q+1)}B\rho_{\mathtt{r}_{j-1}+k_j+q}$$
$$= \sum_{q=0}^{r_j-k_j} CM_j^{-(q+1)}B\rho_{\mathtt{r}_{j-1}+k_j+q} \underset{(2.7)}{=} \sum_{q=0}^{r_j-k_j} \frac{(-1)^q}{q!} H^{(q)}(\sigma_j)\rho_{\mathtt{r}_{j-1}+k_j+q}$$

The performance of the second interpolation condition could be proved similarly. Define $X_j := M_jV(Z^*M_jV)^{-1}$. X_j is a right inverse of Z^*. Now consider a vector $\mathbb{z}$

belonging to the linear span of Z. Hence there exist a vector $\tilde{\mathbb{z}} \in \mathbb{C}$ such that $\mathbb{z}^* = \tilde{\mathbb{z}} Z^*$. Postmultiplying this equation with XZ^* implies the needful result $\mathbb{z}^* X Z^* = \mathbb{z}^*$. Now we could show the identity of both sides of the second interpolation condition

$$\sum_{q=0}^{k_j-1} \frac{(-1)^q}{q!} \ell_{\mathtt{r}_{j-1}+k_j-q} \hat{H}^{(q)}(\sigma_j) \underset{(2.7)}{=} \sum_{q=0}^{k_j-1} \ell_{\mathtt{r}_{j-1}+k_j-q} CV \left[Z^* M_j V\right]^{-(q+1)} Z^* B$$

$$= \sum_{q=0}^{k_j-1} \underbrace{\ell_{\mathtt{r}_{j-1}+k_j-q} C M_j^{-t}}_{\in \operatorname{colspan}(Z^*)} X_j Z^* V \left[Z^* M_j V\right]^{-(q-t+1)} Z^* B \qquad \text{with } t = 1, \dots, q$$

$$= \sum_{q=0}^{k_j-1} \ell_{\mathtt{r}_{j-1}+k_j-q} C M_j^{-(q+1)} B = \sum_{q=0}^{k_j-1} \frac{(-1)^q}{q!} \ell_{\mathtt{r}_{j-1}+k_j-q} H^{(q)}(\sigma_j)\ .$$

The preceding two discussions lead directly to the proof of the third interpolation condition.

$$\sum_{q=1}^{r_j} \frac{(-1)^q}{q!} \sum_{p=0}^{r_j-q} \ell_{\mathtt{r}_{j-1}+p+1} \hat{H}^{(q)}(\sigma_j) \rho_{\mathtt{r}_{j-1}+p+q}$$

$$\underset{(2.7)}{=} \sum_{q=1}^{r_j} \sum_{p=0}^{r_j-q} \ell_{\mathtt{r}_{j-1}+p+1} CV \left[Z^* M_j V\right]^{-(q+1)} Z^* B \rho_{\mathtt{r}_{j-1}+p+q}$$

Interchange the two sums

$$= \sum_{p=1}^{r_j} \ell_{\mathtt{r}_{j-1}+p} \sum_{q=0}^{r_j-p} CV \left[Z^* M_j V\right]^{-(q+2)} Z^* B \rho_{\mathtt{r}_{j-1}+p+q}\ .$$

Analogous to the proof of the right sided tangential interpolation it follows

$$= \sum_{p=1}^{r_j} \ell_{\mathtt{r}_{j-1}+p} CV \left[Z^* M_j V\right]^{-1} \sum_{q=0}^{r_j-p} \underbrace{\left[Z^* M_j V\right]^{-1} Z^* M_j}_{=Y_j^*} M_j^{-(q+1)} B \rho_{\mathtt{r}_{j-1}+p+q}\ .$$

The backmost sum conforms the definition of $v_{\mathtt{r}_{j-1}+p}$

$$= \sum_{p=1}^{r_j} \underbrace{\ell_{\mathtt{r}_{j-1}+p} C(M_j^{-1}}_{\in \operatorname{colspan}(Z^*)} \underbrace{M_j) V \left[Z^* M_j V\right]^{-1}}_{=X_j} (Z^* V) Y_j^* v_{\mathtt{r}_{j-1}+p}\ .$$

Since $\mathbb{z}^* X Z^* = \mathbb{z}^*$ and $\mathbb{v} = V Y_j^* \mathbb{v}$ for $\mathbb{v} \in \operatorname{colspan}(V)$ and $\mathbb{z} \in \operatorname{colspan}(Z)$ it follows

$$= \sum_{p=1}^{r_j} \ell_{\mathtt{r}_{j-1}+p} C M_j^{-1} \sum_{q=1}^{r_j-p} M_j^{-(q+1)} B \rho_{\mathtt{r}_{j-1}+p+q}\ .$$

Interchanging the two sums back finally leads to the required term

$$= \sum_{q=1}^{r_j} \sum_{p=1}^{r_j-q} \ell_{\mathtt{r}_{j-1}+p+1} C M_j^{-(q+1)} B \rho_{\mathtt{r}_{j-1}+p+q}$$
$$\underset{(2.7)}{=} \sum_{q=1}^{r_j} \frac{(-1)^q}{q!} \sum_{p=0}^{r_j-q} \ell_{\mathtt{r}_{j-1}+p+1} H^{(q)}(\sigma_j) \rho_{\mathtt{r}_j+p+q} \, . \qquad \square$$

REMARK 3.5. *Hence, setting* $\sigma_j = -\hat{\lambda}_j^*$, $\rho_j = \hat{b}_j^*$ *and* $\ell_j = \hat{c}_j^*$ *for* $j = 1, \dots, r$ *in Lemma 3.4 shows the implication of the Wilson conditions in Theorem 2.4 to the Interpolation conditions presented in Theorem 2.1.*

The reverse direction of the preceeding lemma completes the equivalence proof between the Wilson and the Interpolation conditions.

LEMMA 3.6. *Let* $\hat{\Sigma}$ *be a reduced system which satisfies the interpolation conditions (2.2) - (2.4).* $\hat{\Sigma}$ *can always be derived from* Σ *by a projection* $\Pi = VZ^*$ *with* $\operatorname{Ran} V =$ *colspan*$\{\tilde{v}_1, \dots, \tilde{v}_r\}$ *and* $\operatorname{Ran} Z =$ *colspan*$\{\tilde{z}_1, \dots, \tilde{z}_r\}$ *and*

$$\tilde{v}_{\mathtt{r}_{j-1}+k_j} := \sum_{q=0}^{r_j-k_j} (-A - \hat{\lambda}_j^* I)^{-(q+1)} B \hat{b}^*_{\mathtt{r}_{j-1}+k_j+q}$$
$$\tilde{z}^*_{\mathtt{r}_{j-1}+k_j} := \sum_{q=0}^{k_j-1} \hat{c}^*_{\mathtt{r}_{j-1}+k_j-q} C (-A - \hat{\lambda}_j^* I)^{-(q+1)}$$
$$\text{for } j = 1, \dots, R \quad \text{and} \quad 1 \le k_j \le r_j \, .$$

Proof. The system $\hat{\Sigma}$ is completely described by its matrix valued transfer function $\hat{H}(s) = \hat{C}(sI_n - \hat{A})^{-1}\hat{B}$ and hence, under assumption that $\hat{A}$ is represented in its Jordan normal form, comprises $R+rm+rp$ specific elements, namely the entries of its system matrices. Thus the Interpolation conditions (2.2) - (2.4) supply $R+rm+rp$ constraints which can be met by the same number of restrictions imposed on the columnspaces of the projection matrices V and Z. See also [7]. □

At least together with the previous conclusions it is possible to imply the equivalence between Interpolation and Wilson conditions for multiple poles. Even though it is unknown a priori wether the reduced system has multiple poles or not we showed that the equivalences hold.

PROPOSITION 3.7. *The necessary Interpolation conditions (2.2) - (2.4) for multiple poles pointed out in Theorem 2.1 are equivalent to the Wilson conditions (2.12) - (2.14) presented in Theorem 2.4.*

3.2. Equivalence between Hyland-Bernstein and Wilson conditions. The idea of the proof of the following theorem can be found in [3].

THEOREM 3.8. *Let* $\mathcal{P}$, $\mathcal{Q}$ *and* M *be positive-definit and consequently symmetric matrices.* $\mathcal{P}_{22}$, $\mathcal{Q}_{22}$, $\mathcal{P}_{12}$ *and* $\mathcal{Q}_{12}$ *are solutions of the equations (2.8) - (2.11), respectively. Then the necessary conditions of Wilson (Theorem 2.4) and Hyland-Bernstein (Theorem 2.6) are equivalent.*

REFERENCES

[1] A.C. Antoulas, *Approximation of large-scale dynamical systems*, Advances in Design and Control **6**, SIAM, Philadelphia, 2005.

[2] A. Bunse-Gerstner, G. Vossen, D. Kubalinska and D. Wilczek, *h_2-norm optimal model reduction for large-scale discrete dynamical MIMO systems* submitted to Journal of Computational and Applied Mathematics dedicated to Bill Gragg, 2007.

[3] S. Gugercin, C. Beattie and A.C. Antoulas, *Rational Krylov Methods for Optimal $\mathcal{H}_2$ Model Reduction*, ICAM Technical Report, Virginia Tech, 2006 and submitted to SIAM Journal on Matrix Analysis and Applications.

[4] G. Vossen, A. Bunse-Gerstner, D. Kubalinska and D. Wilczek, *Necessary optimality conditions for $\mathcal{H}_2$-optimal model reduction*, ZeTeM Technical Report, University of Bremen, to appear.

[5] E.J. Grimme, *Krylov projection methods for model reduction*, PhD Thesis, ECE Department, University of Illinois, Urbana-Champaign, 1997.

[6] D.C. Hyland and D.S. Bernstein, *The optimal projection equations for model reduction and the relationship among the methods of Wilson, Skelton and Moore*, IEEE Trans. Automatic Control. **30** (1985) 1201-1211.

[7] A.J. Mayo and A.C. Antoulas, *A framework for the solution of the generalized realization problem*, Linear Algebra Appl. (2007), doi:10.1016/j.laa.2007.03.008

[8] L. Meier and D.G. Luenberger, *Approximation of Linear Constant Systems*, IEEE Trans. Automatic Control. **12** (1967) 585-588.

[9] A. Vandendorpe, *Model reduction of linear systems, an interpolation point of view*, PhD Thesis, Universite Catholique De Louvain, December 2004.

[10] D.A. Wilson, *Optimum Solution of Model Reduction Problem*, Proc. Inst. Elec. Eng. **117**:6 (1970) 1161-1165.