$\mathcal{H}_2$-OPTIMAL APPROXIMATION OF MIMO LINEAR DYNAMICAL SYSTEMS*

PAUL VAN DOOREN[†‡], KYLE A. GALLIVAN[¶‖], AND P.-A. ABSIL[†§]

Submitted on 30 JUL 2008

Abstract. We consider the problem of approximating a multiple-input multiple-output (MIMO) $p\times m$ rational transfer function $H(s)$ of high degree by another $p\times m$ rational transfer function $\widehat{H}(s)$ of much smaller degree, so that the $\mathcal{H}_2$ norm of the approximation error is minimized. We characterize the stationary points of the $\mathcal{H}_2$ norm of the approximation error by tangential interpolation conditions and also extend these results to the discrete-time case. We analyze whether it is reasonable to assume that lower-order models can always be approximated arbitrarily closely by imposing only first-order interpolation conditions. Finally, we analyze the $\mathcal{H}_2$ norm of the approximation error for a simple case in order to illustrate the complexity of the minimization problem.

Key words. Multivariable systems, model reduction, optimal $\mathcal{H}_2$ approximation, tangential interpolation.

AMS subject classifications. 41A05, 65D05, 93B40

1. Introduction. In this paper, we consider the problem of approximating a real $p \times m$ rational transfer function $H(s)$ of McMillan degree N by a real $p \times m$ rational transfer function $\widehat{H}(s)$ of lower McMillan degree n using the $\mathcal{H}_2$-norm as the approximation criterion. We refer, e.g., to [Che99, Ant05] for the relevant background on linear system theory and model reduction.

Since a transfer function has an unbounded $\mathcal{H}_2$-norm if it is not strictly proper, we will constrain both $H(s)$ and $\widehat{H}(s)$ to be strictly proper (i.e., they are zero at $s = \infty$). Such transfer functions have minimal (i.e., controllable and observable) state-space realizations $(A, B, C) \in \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times m} \times \mathbb{R}^{p\times N}$ and $(\widehat{A}, \widehat{B}, \widehat{C}) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ satisfying

$$\left\{ \begin{array}{l} \dot{x} = Ax + Bu, \\ y = Cx, \end{array} \right. \qquad H(s) := C(sI_N - A)^{-1}B, \tag{1.1}$$

and

$$\left\{ \begin{array}{l} \dot{\widehat{x}} = \widehat{A}\widehat{x} + \widehat{B}u, \\ \widehat{y} = \widehat{C}\widehat{x}, \end{array} \right. \qquad \widehat{H}(s) := \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}, \tag{1.2}$$

where $u \in \mathbb{R}^m$, $y, \widehat{y} \in \mathbb{R}^p$, $x \in \mathbb{R}^N$, $\widehat{x} \in \mathbb{R}^n$.

We also look at the equivalent formulation in the discrete-time case where the dynamical systems become

$$\left\{ \begin{array}{l} x_{k+1} = Ax_k + Bu_k \\ y_k = Cx_k \end{array} \right. \qquad H(z) := C(zI_N - A)^{-1}B, \tag{1.3}$$

*This paper presents research supported by the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office and by the National Science Foundation under contract OCI-03-24944. The scientific responsibility rests with its authors.

[†]Université catholique de Louvain (UCL), Department of Mathematical Engineering, Bâtiment Euler, Avenue Georges Lemaître 4, B-1348 Louvain-la-Neuve, Belgium.

[‡]http://www.inma.ucl.ac.be/~vdooren/

[¶]Department of Mathematics, Florida State University, Tallahassee FL 32306, USA.

[‖]http://www.scs.fsu.edu/~gallivan/

[§]http://www.inma.ucl.ac.be/~absil/

and

$$\begin{cases} \widehat{x}_{k+1} = \widehat{A}\widehat{x}_k + \widehat{B}u \\ \widehat{y}_k = \widehat{C}\widehat{x}_k \end{cases} \qquad \widehat{H}(z) := \widehat{C}(zI_n - \widehat{A})^{-1}\widehat{B}. \tag{1.4}$$

Expressions for the gradients of the squared $\mathcal{H}_2$-norm error function

$$\mathcal{J}_{(A,B,C)} : (\widehat{A}, \widehat{B}, \widehat{C}) \mapsto \|C(sI_N - A)^{-1}B - \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}\|^2_{\mathcal{H}_2}$$

have been known since the work of Wilson [Wil70] (the expressions are recalled in Theorem 3.2). One can object, however, that the *full parameterization*

$$(\widehat{A}, \widehat{B}, \widehat{C}) \mapsto \widehat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B} \tag{1.5}$$

is not one to one, since the triple

$$(\widehat{A}_T, \widehat{B}_T, \widehat{C}_T) := (T^{-1}\widehat{A}T, T^{-1}\widehat{B}, \widehat{C}T)$$

for any matrix $T \in GL(n, \mathbb{R})$ defines the same transfer function :

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

or

$$\widehat{H}(z) = \widehat{C}(zI_n - \widehat{A})^{-1}\widehat{B} = \widehat{C}_T(zI_n - \widehat{A}_T)^{-1}\widehat{B}_T.$$

If one could eliminate the n^2 degrees of freedom of the invertible transformation T, one could hope to fully parameterize the target system $\widehat{H}(s)$ or $\widehat{H}(z)$ with only $n(m+p)$ independent parameters, and to turn Wilson's conditions into $n(m+p)$ nonredundant scalar conditions. Concerning the parameterization task, Byrnes and Falb [BF79, Th. 4.7] show that the set $\mathrm{Rat}^n_{p,m}$ of $p \times m$ strictly proper rational transfer functions of degree n can be parameterized with only $n(m+p)$ real parameters in a locally smooth manner; but it is also shown there that there exists no *globally* smooth parameterization of $\mathrm{Rat}^n_{p,m}$ if $\min(p, m) > 1$. The task of extracting $n(m+p)$ nonredundant conditions out of Wilson's conditions of stationarity is more delicate, as we shall see.

It has been shown in [VGA08] and stated in [GAB07] that, when they have only first-order poles, the stationary points $\widehat{H}(s)$ of the $\mathcal{H}_2$-norm error function (i.e., the points where the gradient of $J_{(A,B,C)}$ vanishes) can be characterized in diagonal canonical form

$$\widehat{H}(s) = \sum_{i=1}^{n} \frac{\widehat{c}_i \widehat{b}_i^H}{s - \widehat{\lambda}_i}, \tag{1.6}$$

via tangential interpolation conditions which can be formulated as

$$\begin{aligned} [H^T(s) - \widehat{H}^T(s)]\widehat{c}_i &= O(s + \widehat{\lambda}_i), \\ \widehat{b}_i^H[H^T(s) - \widehat{H}^T(s)] &= O(s + \widehat{\lambda}_i), \\ \widehat{b}_i^H[H^T(s) - \widehat{H}^T(s)]\widehat{c}_i &= O(s + \widehat{\lambda}_i)^2. \end{aligned}$$

Notice that the interpolation points are the negative of the poles of $\widehat{H}(s)$. These results are, in fact, a consequence of the relation between the equations of the gradients

of the $\mathcal{H}_2$-norm error (as derived originally by Wilson in [Wil70]) and tangential interpolation based on Sylvester equations (as derived in [BGR90],[GVV04],[GVV05]). Similar conditions can be found in [BKVW07] for the discrete-time case. Observe that the diagonal canonical form (1.6) uses the minimal number, $n(m+p)$, of parameters once the $\widehat{b}_i$'s or $\widehat{c}_i$'s are normalized to remove the scaling invariance. The tangential interpolation conditions also impose the correct number of nonredundant scalar conditions (see Section 4.1). However, in view of the result of Byrnes and Falb, the diagonal canonical form (1.6)—as well as any other canonical form—cannot yield a globally smooth one-to-one parameterization of $\mathrm{Rat}^n_{p,m}$ when $\min(p,m) > 1$. Singularities appear when $\widehat{H}(s)$ has higher-order poles. This is also true for discrete-time systems.

In this paper, we characterize the stationary points $\widehat{H}(s)$ or $\widehat{H}(z)$ of the $\mathcal{H}_2$-norm error function in Jordan canonical form, i.e., *without* the assumption that they have only first-order poles. The stationarity conditions elegantly generalize to higher-order tangential interpolation conditions of degree $k_i - 1$ (in the sense of [GVV05]), where k_i is the size of the ith Jordan block. The interpolation points remain the negative of the poles $\widehat{\lambda}_i$ of $\widehat{H}(s)$, and the interpolation directions are polynomial vectors of degree $k_i - 1$, built from the Jordan-form equivalents of $\widehat{b}_i$ and $\widehat{c}_i$; see Theorem 4.8. We also show that these tangential interpolation conditions contain $n(m+p)$ nonredundant scalar conditions. The result in Theorem 4.8 has several precursors: Aigrain and Williams [AW49] for the SISO case with simple real poles, Meier and Luenberger [ML67] for the general SISO case (see also the alternative derivation in [GAB07]), and [VGA08] for the MIMO case with simple poles (see also the remark in [GAB07]).

Since the set of systems with higher-order poles is nowhere dense in $\mathrm{Rat}^n_{p,m}$, the generalization of the stationarity conditions to higher-order poles seems to be chiefly of theoretical interest. Nevertheless, we argue that the case of higher-order poles cannot be simply brushed aside. First, we show on an example that $\mathcal{H}_2$-optimal reduced-order models with higher-order poles do occur. Second, we point out that the Jordan canonical form changes in a nonsmooth manner at the higher-order poles and that the tangential interpolation conditions for $\mathcal{H}_2$-norm stationary points become ill conditioned around the systems $\widehat{H}(s)$ with higher-order poles. Therefore, insisting on the Jordan canonical form parameterization of the $\mathcal{H}_2$-optimal reduced-order model may seriously affect the sensitivity of any numerical algorithm using such a parameterization. When the influence of a nearby higher-order pole becomes problematic, a possible remedy is to exploit the full parameterization (1.5).

It should be kept in mind that the above discussion only concerns stationarity conditions for the $\mathcal{H}_2$-norm error function. The stationary points may be local minima, saddle points, or local maxima of the $\mathcal{H}_2$-norm error function. When a descent iteration is employed, convergence to saddle points and local maxima is not expected to occur. However, the method can still be trapped in local, nonglobal minima. Such spurious local minima exist in the $\mathcal{H}_2$-optimal model reduction problem, as we show on a simple example. Computing an $\mathcal{H}_2$-optimal reduced-order model is thus a tough (obviously nonconvex) optimization task. Nevertheless, the computed local minima tend to yield approximations that are considered satisfactory in practice, hence the interest for interpolation-based fixed-point type algorithms as revived recently in, e.g., [BG07, GAB07, Gug02].

The paper is organized as follows. After presenting in Section 2 the necessary background material on the $\mathcal{H}_2$ approximation problem, in Section 3 we recall Wil-

son's formulas for the gradient of the $\mathcal{H}_2$-norm error function. In Section 4, Wilson's first-order optimality conditions are expressed in a tangential interpolation form obtained by representing the reduced-order model in Jordan canonical form—thus covering the case of higher-order poles in the reduced-order model. The link to tangential interpolation by means of projection matrices that solve Sylvester equations is discussed in Section 5. The importance of dealing with the case of higher-order poles is illustrated in Section 6. Section 7 shows on a simple example that the $\mathcal{H}_2$-optimal model reduction problem is a difficult optimization problem, with spurious local minimizers in which local optimization algorithms may get trapped. An overview of algorithms for solving the $\mathcal{H}_2$-optimal approximation problem is given in Section 8. The discrete-time case is covered in Section 9, and conclusions are drawn in Section 10.

2. The $\mathcal{H}_2$ approximation problem. Much of the material in this section is standard and can be found in, e.g., [Ant05]. Let $E(s)$ be an arbitrary strictly proper transfer function, with realization triple (A_e, B_e, C_e). If $E(s)$ is unstable, its $\mathcal{H}_2$-norm is defined to be ∞. Otherwise, its squared $\mathcal{H}_2$-norm is defined as the trace of a matrix integral :

$$\|E(s)\|_{\mathcal{H}_2}^2 := \mathrm{tr} \int_{-\infty}^{\infty} E(j\omega)^H E(j\omega) \frac{d\omega}{2\pi} = \mathrm{tr} \int_{-\infty}^{\infty} E(j\omega) E(j\omega)^H \frac{d\omega}{2\pi}. \tag{2.1}$$

By Parseval's identity, this can also be expressed using the state space realization as

$$\begin{aligned} \|E(s)\|_{\mathcal{H}_2}^2 &= \mathrm{tr} \int_0^{\infty} [C_e \exp^{A_e t} B_e][C_e \exp^{A_e t} B_e]^T dt \\ &= \mathrm{tr} \int_0^{\infty} [C_e \exp^{A_e t} B_e]^T [C_e \exp^{A_e t} B_e] dt. \end{aligned}$$

This can also be related to an expression involving the gramians P_e and Q_e defined as

$$P_e := \int_0^{\infty} [\exp^{A_e t} B_e][\exp^{A_e t} B_e]^T dt, \quad Q_e := \int_0^{\infty} [\exp^{A_e t} B_e]^T [C_e \exp^{A_e t}] dt,$$

which are also known to be the solutions of the Lyapunov equations

$$A_e P_e + P_e A_e^T + B_e B_e^T = 0, \quad Q_e A_e + A_e^T Q_e + C_e^T C_e = 0. \tag{2.2}$$

Using these, it easily follows that the squared $\mathcal{H}_2$-norm of $E(s)$ can be expressed as

$$\|E(s)\|_{\mathcal{H}_2}^2 = \mathrm{tr}\ B_e^T Q_e B_e = \mathrm{tr}\ C_e P_e C_e^T. \tag{2.3}$$

REMARK 2.1. *It is easy to show that if A_e has a single real eigenvalue λ that tends to zero, i.e., A_e tends to lose its stability :*

$$A_e x = \lambda x, \quad y^T A_e = \lambda y^T, \quad \lambda \to 0$$

then P_e and Q_e tend to a rank one matrix of infinite norm, since

$$P_e \to x x^T \beta/(2\lambda), \quad Q_e \to y y^T \gamma/(2\lambda), \quad \text{where} \quad \beta = y^T B_e B_e^T y, \quad \gamma = x^T C_e^T C_e^T x.$$

It then follows that $\mathcal{J} \to \beta\gamma/(2\lambda)$ also becomes infinite. Similar behavior is also found for complex conjugate pairs of eigenvalues tending to the imaginary axis. It thus follows that the squared $\mathcal{H}_2$-norm of $E(s)$ tends to infinity as soon as A_e looses

its stability. This explains why this norm is typically defined to be infinite when $E(s)$ is unstable.

We now apply this to the error function

$$E(s) := H(s) - \widehat{H}(s) = C(sI_N - A)^{-1}B - \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}.$$

A realization of $E(s)$ in partitioned form is given by

$$(A_e, B_e, C_e) := \left(\begin{bmatrix} A & \\ & \widehat{A} \end{bmatrix}, \begin{bmatrix} B \\ \widehat{B} \end{bmatrix}, \begin{bmatrix} C & -\widehat{C} \end{bmatrix} \right), \tag{2.4}$$

and the Lyapunov equations (2.2) become

$$P_e := \begin{bmatrix} P & X \\ X^T & \widehat{P} \end{bmatrix}, \quad \begin{bmatrix} A & \\ & \widehat{A} \end{bmatrix} \begin{bmatrix} P & X \\ X^T & \widehat{P} \end{bmatrix} + \begin{bmatrix} P & X \\ X^T & \widehat{P} \end{bmatrix} \begin{bmatrix} A^T & \\ & \widehat{A}^T \end{bmatrix} + \begin{bmatrix} B \\ \widehat{B} \end{bmatrix} \begin{bmatrix} B^T & \widehat{B}^T \end{bmatrix} = 0, \tag{2.5}$$

and

$$Q_e := \begin{bmatrix} Q & Y \\ Y^T & \widehat{Q} \end{bmatrix}, \quad \begin{bmatrix} A^T & \\ & \widehat{A}^T \end{bmatrix} \begin{bmatrix} Q & Y \\ Y^T & \widehat{Q} \end{bmatrix} + \begin{bmatrix} Q & Y \\ Y^T & \widehat{Q} \end{bmatrix} \begin{bmatrix} A & \\ & \widehat{A} \end{bmatrix} + \begin{bmatrix} C^T \\ -\widehat{C}^T \end{bmatrix} \begin{bmatrix} C & -\widehat{C} \end{bmatrix} = 0. \tag{2.6}$$

In order to minimize the $\mathcal{H}_2$-distance $\|H(s) - \widehat{H}(s)\|^2_{\mathcal{H}_2}$ of the low-order system $\widehat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}$ to a given the full-order model $H(s) = C(sI_N - A)^{-1}B$, we must minimize the function $\mathcal{J}_{(A,B,C)}$ defined by

$$\mathcal{J}_{(A,B,C)}(\widehat{A}, \widehat{B}, \widehat{C}) = \|C(sI_N - A)^{-1}B - \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}\|^2_{\mathcal{H}_2}. \tag{2.7a}$$

We will frequently omit the subscript in $\mathcal{J}_{(A,B,C)}(\widehat{A}, \widehat{B}, \widehat{C})$ when the full-order model is clear from the context. In view of (2.3), $\mathcal{J}(\widehat{A}, \widehat{B}, \widehat{C})$ admits the formulation

$$\mathcal{J}(\widehat{A}, \widehat{B}, \widehat{C}) = \operatorname{tr}\left(\begin{bmatrix} B^T & \widehat{B}^T \end{bmatrix} \begin{bmatrix} Q & Y \\ Y^T & \widehat{Q} \end{bmatrix} \begin{bmatrix} B \\ \widehat{B} \end{bmatrix} \right) = \operatorname{tr}\left(B^T Q B + 2B^T Y \widehat{B} + \widehat{B}^T \widehat{Q} \widehat{B} \right), \tag{2.7b}$$

where Q, Y and $\widehat{Q}$ depend on A, $\widehat{A}$, C and $\widehat{C}$ through the Lyapunov equation (2.6), or equivalently

$$\mathcal{J}(\widehat{A}, \widehat{B}, \widehat{C}) = \operatorname{tr}\left(\begin{bmatrix} C & -\widehat{C} \end{bmatrix} \begin{bmatrix} P & X \\ X^T & \widehat{P} \end{bmatrix} \begin{bmatrix} C^T \\ -\widehat{C}^T \end{bmatrix} \right) = \operatorname{tr}\left(CPC^T - 2CX\widehat{C}^T + \widehat{C}\widehat{P}\widehat{C}^T \right), \tag{2.7c}$$

where P, X and $\widehat{P}$ depend on A, $\widehat{A}$, B and $\widehat{B}$ through the Lyapunov equation (2.5). Note that the terms B^TQB and CPC^T in the above expressions are constant, and hence can be discarded in the optimization.

REMARK 2.2. *The Sylvester equations* (2.5) *and* (2.6) *are nonsingular if and only if the union of the spectra of A and $\widehat{A}$ does not contain any pair of opposite points (see [Gan59, Ch. VI]). In particular, they are nonsingular if the transfer functions $H(s) = C(sI_N - A)^{-1}B$ and $\widehat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}$ are stable. In fact, the function*

$$(A, B, C, \widehat{A}, \widehat{B}, \widehat{C}) \mapsto \mathcal{J}_{(A,B,C)}(\widehat{A}, \widehat{B}, \widehat{C})$$

is smooth around every point where $H(s)$ and $\widehat{H}(s)$ are stable. In particular, when $H(s)$ is stable, the function

$$(\widehat{A}, \widehat{B}, \widehat{C}) \mapsto \mathcal{J}_{(A,B,C)}(\widehat{A}, \widehat{B}, \widehat{C})$$

is smooth around every point where $\widehat{H}(s)$ is stable.

3. Gradients of the squared $\mathcal{H}_2$-norm error function. The expansions above can be used to obtain formulas for the gradients of the squared $\mathcal{H}_2$-norm error function $\mathcal{J}$ versus $\widehat{A}$, $\widehat{B}$, and $\widehat{C}$. We define the gradients as follows.

DEFINITION 3.1. *The gradients of a real-valued function $f(\widehat{A},\widehat{B},\widehat{C})$ of a real matrix variables $\widehat{A}\in\mathbb{R}^{n\times n}$, $\widehat{B}\in\mathbb{R}^{n\times m}$, $\widehat{C}\in\mathbb{R}^{p\times n}$, are the real matrices $\nabla_{\widehat{A}}f(\widehat{A},\widehat{B},\widehat{C})\in\mathbb{R}^{n\times n}$, $\nabla_{\widehat{B}}f(\widehat{A},\widehat{B},\widehat{C})\in\mathbb{R}^{n\times m}$, $\nabla_{\widehat{C}}f(\widehat{A},\widehat{B},\widehat{C})\in\mathbb{R}^{p\times n}$, defined by*

$$[\nabla_{\widehat{A}}f(\widehat{A},\widehat{B},\widehat{C})]_{i,j}=\frac{\partial}{\partial\widehat{A}_{i,j}}f(\widehat{A},\widehat{B},\widehat{C}),\quad i=1,\ldots,n,\quad j=1,\ldots,n,$$
$$[\nabla_{\widehat{B}}f(\widehat{A},\widehat{B},\widehat{C})]_{i,j}=\frac{\partial}{\partial\widehat{B}_{i,j}}f(\widehat{A},\widehat{B},\widehat{C}),\quad i=1,\ldots,n,\quad j=1,\ldots,m,$$
$$[\nabla_{\widehat{C}}f(\widehat{A},\widehat{B},\widehat{C})]_{i,j}=\frac{\partial}{\partial\widehat{C}_{i,j}}f(\widehat{A},\widehat{B},\widehat{C}),\quad i=1,\ldots,p,\quad j=1,\ldots,n.$$

We will write $\nabla_{\widehat{A}}f$ as a compact notation for $\nabla_{\widehat{A}}f(\widehat{A},\widehat{B},\widehat{C})$ when the argument is clear from the context.

Starting from the characterizations (2.5,2.7c) and (2.6,2.7b) of the $\mathcal{H}_2$ norm, one can derive succinct forms of the gradients. This theorem is originally due to Wilson [Wil70], but we state here the version derived in [VGA08], where a proof based on inner products and traces is given.

THEOREM 3.2. *The gradients $\nabla_{\widehat{A}}\mathcal{J}$, $\nabla_{\widehat{B}}\mathcal{J}$ and $\nabla_{\widehat{C}}\mathcal{J}$ of the squared $\mathcal{H}_2$-norm error $\mathcal{J}$ (2.7), where both (A,B,C) and $(\widehat{A},\widehat{B},\widehat{C})$ are minimal (i.e., controllable and observable), are given by*

$$\nabla_{\widehat{A}}\mathcal{J}=2(\widehat{Q}\widehat{P}+Y^TX),\quad \nabla_{\widehat{B}}\mathcal{J}=2(\widehat{Q}\widehat{B}+Y^TB),\quad \nabla_{\widehat{C}}\mathcal{J}=2(\widehat{C}\widehat{P}-CX),\quad (3.1)$$

where

$$A^TY+Y\widehat{A}-C^T\widehat{C}=0,\quad \widehat{A}^T\widehat{Q}+\widehat{Q}\widehat{A}+\widehat{C}^T\widehat{C}=0,\quad (3.2)$$
$$X^TA^T+\widehat{A}X^T+\widehat{B}B^T=0,\quad \widehat{P}\widehat{A}^T+\widehat{A}\widehat{P}+\widehat{B}\widehat{B}^T=0.\quad (3.3)$$

The gradient forms of Theorem 3.2 allowed us to derive in [VGA08] a theorem that also provides an important link to tangential interpolation by projection.

4. Stationarity conditions in Jordan form. In this section, we revisit Wilson's conditions (Theorem 3.2) with $\widehat{H}(s)$ in Jordan canonical form. We first consider the continuous-time case and discuss the discrete-time case in Section 9.

We will assume that both transfer functions $H(s)$ and $\widehat{H}(s)$ have real minimal (controllable and observable) realizations (A,B,C) and $(\widehat{A},\widehat{B},\widehat{C})$.

4.1. First-order poles. We first assume that all the poles of $\widehat{H}(s)$ are distinct (but possibly complex), which implies that the Jordan canonical form reduces to a diagonal form.

Since the number of parameters in the full parameterization (1.5) is not minimal, the gradient conditions of Theorem 3.2 must be redundant. This is made explicit in the theorem below, proved in [VGA08]. For this we will need s_i, t_i^H, the (complex) left and right eigenvectors of the (real) matrix $\widehat{A}$ corresponding to the (complex) eigenvalue $\widehat{\lambda}_i$. We then have :

$$\widehat{A}s_i=\widehat{\lambda}_is_i,\quad \widehat{C}s_i=:\widehat{c}_i,\quad t_i^H\widehat{A}=\widehat{\lambda}_it_i^H,\quad t_i^H\widehat{B}=:\widehat{b}_i^H,\quad i=1,\ldots,n,\quad (4.1)$$

and $\widehat{H}(s)$ has the partial fraction expansion

$$\widehat{H}(s) = \sum_{i=1}^{n} \frac{\widehat{c}_i \widehat{b}_i^H}{s - \widehat{\lambda}_i}, \tag{4.2}$$

where $\widehat{b}_i \in \mathbb{C}^m$ and $\widehat{c}_i \in \mathbb{C}^p$ and where $\{(\widehat{\lambda}_i, \widehat{b}_i, \widehat{c}_i) : i = 1, \ldots, n\}$ is a self-conjugate set. The form (4.2) corresponds to the diagonal canonical form of $\widehat{H}(s)$, a particular case of the Jordan canonical form when all the Jordan blocks have dimension one. It involves the minimal number $n(m+p)$ of parameters once normalization conditions are imposed on either the $\widehat{b}_i$'s or the $\widehat{c}_i$'s.

THEOREM 4.1. *Let $H(s) = C(sI_N - A)^{-1}B$ and $\widehat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}$ be real minimal realizations, and let $\widehat{\lambda}_i$, $\widehat{b}_i$, $\widehat{c}_i$, s_i, and t_i, $i = 1, \ldots, n$, be as in* (4.1)*. Assume that $-\widehat{\lambda}_i$ is not a pole of $H(s)$, $i = 1, \ldots, n$. Then*

$$\frac{1}{2}(\nabla_{\widehat{B}}\mathcal{J})^T s_i = [H^T(-\widehat{\lambda}_i) - \widehat{H}^T(-\widehat{\lambda}_i)]\widehat{c}_i \tag{4.3}$$

$$\frac{1}{2} t_i^H (\nabla_{\widehat{C}}\mathcal{J})^T = \widehat{b}_i^H [H^T(-\widehat{\lambda}_i) - \widehat{H}^T(-\widehat{\lambda}_i)] \tag{4.4}$$

$$\frac{1}{2} t_i^H (\nabla_{\widehat{A}}\mathcal{J})^T s_i = \widehat{b}_i^H \frac{d}{ds} [H^T(s) - \widehat{H}^T(s)]\Big|_{s=-\widehat{\lambda}_i} \widehat{c}_i \tag{4.5}$$

$$\frac{1}{2} t_i^H (\nabla_{\widehat{A}}\mathcal{J})^T s_j = \frac{1}{2(\widehat{\lambda}_i - \widehat{\lambda}_j)} [\widehat{b}_i^H (\nabla_{\widehat{B}}\mathcal{J})^T s_j - t_i^H (\nabla_{\widehat{C}}\mathcal{J})^T \widehat{c}_j],\ i \neq j, \tag{4.6}$$

where $\mathcal{J}$ is the squared $\mathcal{H}_2$-norm error defined in (2.7)*.*

Let $S := \begin{bmatrix} s_1 & \ldots & s_n \end{bmatrix}$, then the above theorem shows that the off-diagonal elements of $S^{-1}(\nabla_{\widehat{A}}\mathcal{J})^T S$ actually depend on $(\nabla_{\widehat{B}}\mathcal{J})^T$ and $(\nabla_{\widehat{C}}\mathcal{J})^T$. Therefore one need only impose conditions on $\mathrm{diag}(S^{-1}(\nabla_{\widehat{A}}\mathcal{J})^T S)$ and on $(\nabla_{\widehat{B}}\mathcal{J})^T$ and $(\nabla_{\widehat{C}}\mathcal{J})^T$ to characterize stationary points of $\mathcal{J}$. The following corollary easily follows. It is derived independently in [BKVW07] for the discrete-time case, and also suggested in [GAB07].

COROLLARY 4.2. *With the notation and assumptions of Theorem 4.1, if $(\nabla_{\widehat{B}}\mathcal{J})^T = 0$, $(\nabla_{\widehat{C}}\mathcal{J})^T = 0$ and $\mathrm{diag}(S^{-1}(\nabla_{\widehat{A}}\mathcal{J})^T S) = 0$, then $\nabla_{\widehat{A}}\mathcal{J} = 0$ and the following tangential interpolation conditions are satisfied for all $\widehat{\lambda}_i$, $i = 1, \ldots, n$:*

$$[H^T(-\widehat{\lambda}_i) - \widehat{H}^T(-\widehat{\lambda}_i)]\widehat{c}_i = 0, \tag{4.7}$$

$$\widehat{b}_i^H [H^T(-\widehat{\lambda}_i) - \widehat{H}^T(-\widehat{\lambda}_i)] = 0, \tag{4.8}$$

$$\widehat{b}_i^H \frac{d}{ds} [H^T(s) - \widehat{H}^T(s)]\Big|_{s=-\widehat{\lambda}_i} \widehat{c}_i = 0. \tag{4.9}$$

These tangential interpolation conditions contain $n(m+p)$ nonredundant conditions. To see this, fix i and consider first the case where $\widehat{\lambda}_i$ is real. The first two equations (4.7) and (4.8) impose that the determinant of $[H^T(-\widehat{\lambda}_i) - \widehat{H}^T(-\widehat{\lambda}_i)]$ vanishes, which accounts for one real scalar condition. Next, (4.7) and (4.8) require that $\widehat{c}_i$ and $\widehat{b}_i$ belong to the kernel of $[H^T(-\widehat{\lambda}_i) - \widehat{H}^T(-\widehat{\lambda}_i)]$, which imposes $p-1$ and $m-1$ real scalar conditions. Finally, the last equation (4.9) imposes one real scalar condition, for a total of $m+p$ conditions corresponding to the fixed i. In the complex case, we have a pair of complex-conjugate poles $\widehat{\lambda}_i$ and $\widehat{\lambda}_{i+1}$. The constraint $\det[H^T(-\widehat{\lambda}_i) - \widehat{H}^T(-\widehat{\lambda}_i)] = 0$ imposes two real scalar conditions, the first two equations impose further $2(p-1)$ and $2(m-1)$ real scalar conditions, and the last equation

imposes two real scalar conditions, for a total of $2(m+p)$ real scalar conditions. The equations for $\widehat{\lambda}_{i+1}$ impose the same conditions since equations (4.7)–(4.9) are then just the complex conjugate ones as for $\widehat{\lambda}_i$. The total for $\widehat{\lambda}_i$ and $\widehat{\lambda}_{i+1}$ is thus $2(m+p)$ real scalar conditions. To conclude, observe that i ranges from 1 to n, which yields a total of $n(m+p)$ real scalar conditions. This matches the number, $n(m+p)$, of independent parameters.

The above conditions can also be expressed in terms of the Taylor expansion of $H(s) - \widehat{H}(s)$:

$$[H^T(s) - \widehat{H}^T(s)]\widehat{c}_i = O(s+\widehat{\lambda}_i), \quad \widehat{b}_i^H[H^T(s) - \widehat{H}^T(s)] = O(s+\widehat{\lambda}_i),$$

$$\widehat{b}_i^H[H^T(s) - \widehat{H}^T(s)]\widehat{c}_i = O(s+\widehat{\lambda}_i)^2.$$

That formulation is in fact easier to extend to higher-order poles. Observe also that we retrieve the conditions of Meier and Luenberger [ML67] for the single-input single-output (SISO) case since then $\widehat{b}_i^H$ and $\widehat{c}_i$ are just nonzero scalars that can be divided out. The above conditions then become the $2n$ conditions

$$H(-\widehat{\lambda}_i) = \widehat{H}(-\widehat{\lambda}_i), \quad \frac{d}{ds}\, H(s)\big|_{s=-\widehat{\lambda}_i} = \frac{d}{ds}\, \widehat{H}(s)\Big|_{s=-\widehat{\lambda}_i}, \quad i = 1, \ldots, n.$$

When the transfer function $\widehat{H}(s)$ has repeated first-order poles, the results are essentially the same except that there are bases S_i and T_i^H of right and left invariant subspaces corresponding to a single eigenvalue $\widehat{\lambda}_i$. We then have

$$\widehat{A}S_i = \widehat{\lambda}_i S_i, \quad \widehat{C}S_i = \widehat{C}_i, \quad T_i^H\widehat{A} = \widehat{\lambda}_i T_i^H, \quad T_i^H\widehat{B} = \widehat{B}_i^H, \quad T_i^H S_i = I_k.$$

Theorems 4.1 and 4.2 still hold but with the vectors $\widehat{c}_i$ and $\widehat{b}_i^H$ replaced by the matrices $\widehat{C}_i$ and $\widehat{B}_i^H$. It may seem that this implies that we then impose more than $n(m+p)$ conditions, but in fact one can choose the individual vectors of S_i and T_i^H such that the off diagonal elements of $T_i^H(\nabla_{\widehat{A}}\mathcal{J})^T S_i$ are zero. Only its diagonal elements need then to be constrained to be zero to force the stationarity conditions.

4.2. Higher-order poles. Let us now allow $\widehat{H}(s)$ to have multiple and higher-order poles. The main result is given in Theorem 4.8, where we show that the stationary points of the $\mathcal{H}_2$-norm error function are characterized by tangential interpolation conditions whose degree depends on the size of the Jordan blocks of $\widehat{H}(s)$. The result generalizes Corollary 4.2.

Let $\widehat{H}(s)$ then have the following minimal (controllable and observable) representation

$$\widehat{H}(s) = \sum_{i=1}^{\ell} \widehat{H}_i(s), \quad \widehat{H}_i(s) := \widehat{C}_i(sI - \widehat{A}_i)^{-1}\widehat{B}_i^H, \quad \widehat{A}_i := \begin{bmatrix} \widehat{\lambda}_i & -1 & & \\ & \widehat{\lambda}_i & \ddots & \\ & & \ddots & -1 \\ & & & \widehat{\lambda}_i \end{bmatrix}, \tag{4.10}$$

where $\widehat{A}_i \in \mathbb{C}^{k_i \times k_i}$, $\widehat{B}_i^H \in \mathbb{C}^{k_i \times m}$, $\widehat{C}_i \in \mathbb{C}^{p \times k_i}$ and where $\{(\widehat{A}_i, \widehat{B}_i^H, \widehat{C}_i) : i = 1, \ldots, \ell\}$ is a self-conjugate set. Notice that this is essentially the partial fraction expansion of $\widehat{H}(s)$ and that there may be more than one Jordan block $\widehat{A}_i$ associated with the

same complex eigenvalue $\widehat{\lambda}_i$. The minimality of the representation implies linear independence of the leading columns in each block $\widehat{B}_i$ and of the trailing rows in each block $\widehat{C}_i$ that correspond to the same eigenvalue $\widehat{\lambda}_i$, since these blocks appear as subblocks of a minimal realization of $\widehat{H}(s)$.

We will need S_i, T_i^H, the (complex) left and right eigenspaces of the (real) matrix $\widehat{A}$ corresponding to the (complex) eigenvalue $\widehat{\lambda}_i$. Because of the expansion (4.2), we then have :

$$\widehat{A}S_i = S_i\widehat{A}_i, \quad \widehat{C}S_i = \widehat{C}_i, \quad T_i^H\widehat{A} = \widehat{A}_iT_i^H, \quad T_i^H\widehat{B} = \widehat{B}_i^H, \quad T_i^HS_i = I_k. \tag{4.11}$$

Note also that the matrices S_i and T_i^H are not unique. When there is only one Jordan block associated with an eigenvalue $\widehat{\lambda}_i$, its degree of freedom is just a block scaling S_iD_i and $D_i^{-1}T_i^H$ with $D_i \in \mathbb{C}^{k_i\times k_i}$ invertible. When there is more than one Jordan block associated with $\widehat{\lambda}_i$, the degrees of freedom are more involved, but we associate below right and left bases S_i, T_i with each individual Jordan block A_i.

We will also need the following lemmas in preparation for the main theorem.

LEMMA 4.3. *If $-\lambda$ is not an eigenvalue of A, the solution of the matrix equation*

$$A^TY + YF - C^TL = 0 \quad \text{with} \quad F := \begin{bmatrix} \lambda & -1 & & \\ & \lambda & \ddots & \\ & & \ddots & -1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{k\times k},$$

with $L := \begin{bmatrix} \ell_0 & \ell_1 & \dots & \ell_{k-1} \end{bmatrix}$, is given by

$$Y = \begin{bmatrix} (A^T+\lambda I)^{-1}C^T & (A^T+\lambda I)^{-2}C^T & \dots & (A^T+\lambda I)^{-k}C^T \end{bmatrix} \begin{bmatrix} \ell_0 & \ell_1 & \dots & \ell_{k-1} \\ & \ell_0 & \ddots & \vdots \\ & & \ddots & \ell_1 \\ & & & \ell_0 \end{bmatrix}.$$

Moreover, let

$$\phi_\lambda(s) := \begin{bmatrix} 1 & (s+\lambda) & \dots & (s+\lambda)^{k-1} \end{bmatrix}^T, \quad y(s) := Y\phi_\lambda(s),$$

then

$$y(s) = (A^T - sI)^{-1}C^TL\phi_\lambda(s) + O(s+\lambda)^k$$

which means that the ith column y_i of Y is also the coefficient of $(s+\lambda)^{i-1}$ in the Taylor expansion of $(A^T - sI)^{-1}C^TL\phi_\lambda(s)$.

Proof. The first part easily follows from $(A^T+\lambda I)y_1 = C^T\ell_0$ and $(A^T+\lambda I)y_i = C^T\ell_{i-1} + y_{i-1}$, $i > 1$. The second part follows from the identity

$$(A^T - sI)^{-1}C^T = \sum_{i=1}^{\infty}(s+\lambda)^{i-1}(A^T+\lambda I)^{-i}C^T$$

and from the convolution of this formal series with the polynomial vector $L\phi_\lambda(s)$. ∎

We also give the dual version of this lemma.

LEMMA 4.4. *If $-\lambda$ is not an eigenvalue of A, the solution of the matrix equation*

$$X^H A^T + F X^H - R^H B^T = 0$$

with $F \in \mathbb{C}^{k\times k}$ as above and $R := \begin{bmatrix} r_{k-1} & r_{k-2} & \dots & r_0 \end{bmatrix}$, is given by

$$X^H = \begin{bmatrix} r_0^H & r_1^H & \dots & r_{k-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} B^T(A^T+\lambda I)^{-k} \\ \vdots \\ B^T(A^T+\lambda I)^{-2} \\ B^T(A^T+\lambda I)^{-1} \end{bmatrix}.$$

Moreover, let

$$\psi_\lambda(s) := \begin{bmatrix} (s+\lambda)^{k-1} & \dots & (s+\lambda) & 1 \end{bmatrix}, \quad x^H(s) := \psi_\lambda(s) X^H,$$

then

$$x^H(s) = \psi_\lambda(s) R^H B^T (A^T - sI)^{-1} + O(s+\lambda)^k$$

which means that the ith row x_i^H of X^H is also the coefficient of $(s+\lambda)^{i-1}$ in the Taylor expansion of $\psi_\lambda(s) R^H B^T (A^T - sI)^{-1}$.

Proof. The proof is just the dual of the previous lemma. □

We first obtain an expression for $\nabla_{\widehat{B}}\mathcal{J}$ and $\nabla_{\widehat{C}}\mathcal{J}$ that exploits the Jordan canonical form. The result generalizes formulas (4.3) and (4.4) to higher-order poles.

THEOREM 4.5. *Let $H(s) = C(sI_N - A)^{-1}B$ and $\widehat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}$ be real minimal realizations, and let $\widehat{A}_i$, $\widehat{B}_i$, $\widehat{C}_i$, S_i, and T_i, $i = 1, \dots, \ell$, describe the Jordan canonical form of $\widehat{H}(s)$ as in* (4.10) *and* (4.11)*. Assume that $-\widehat{\lambda}_i$ is not a pole of $H(s)$, $i = 1, \dots, \ell$. Define*

$$\psi_{\widehat{\lambda}_i}(s) := \begin{bmatrix} (s+\widehat{\lambda}_i)^{k_i-1} & \dots & (s+\widehat{\lambda}_i) & 1 \end{bmatrix}, \quad \phi_{\widehat{\lambda}_i}(s) := \begin{bmatrix} 1 & (s+\widehat{\lambda}_i) & \dots & (s+\widehat{\lambda}_i)^{k_i-1} \end{bmatrix}^T.$$

Then we have

$$\frac{1}{2}(\nabla_{\widehat{B}}\mathcal{J})^T S_i \phi_{\widehat{\lambda}_i}(s) = [H^T(s) - \widehat{H}^T(s)]\widehat{C}_i \phi_{\widehat{\lambda}_i}(s) + O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.12}$$

$$\frac{1}{2}\psi_{\widehat{\lambda}_i}(s) T_i^H (\nabla_{\widehat{C}}\mathcal{J})^T = \psi_{\widehat{\lambda}_i}(s)\widehat{B}_i^H [H^T(s) - \widehat{H}^T(s)] + O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.13}$$

where $\mathcal{J}$ is the squared $\mathcal{H}_2$-norm error defined in (2.7)*.*

Proof. Define $Y_i := Y S_i$, $\widehat{Q}_i := -\widehat{Q} S_i$, $X_i := -X T_i$ and $\widehat{P}_i := -\widehat{P} T_i$. Then we have

$$A^T Y_i + Y_i \widehat{A}_i = C^T \widehat{C}_i, \quad \widehat{A}^T \widehat{Q}_i + \widehat{Q}_i \widehat{A}_i = \widehat{C}^T \widehat{C}_i,$$

$$X_i^H A^T + \widehat{A}_i X_i^H = \widehat{B}_i^H B^T, \quad \widehat{P}_i^H \widehat{A}^T + \widehat{A}_i \widehat{P}_i^H = \widehat{B}_i^H \widehat{B}^T.$$

If $-\widehat{\lambda}_i$ is not an eigenvalue of A or $\widehat{A}$, both $(A^T - sI)^{-1}$ and $(\widehat{A}^T - sI)^{-1}$ have Taylor expansions in $(s+\widehat{\lambda}_i)$. It then follows from Lemmas 4.3 and 4.4 that

$$Y_i \phi_{\widehat{\lambda}_i}(s) = (A^T - sI)^{-1} C^T \widehat{C}_i \phi_{\widehat{\lambda}_i}(s) + O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.14}$$

$$\widehat{Q}_i \phi_{\widehat{\lambda}_i}(s) = (\widehat{A}^T - sI)^{-1} \widehat{C}^T \widehat{C}_i \phi_{\widehat{\lambda}_i}(s) + O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.15}$$

$$\psi_{\widehat{\lambda}_i}(s)X_i^H = \psi_{\widehat{\lambda}_i}(s)\widehat{B}_i^H B^T(A^T - sI)^{-1} + O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.16}$$

$$\psi_{\widehat{\lambda}_i}(s)\widehat{P}_i^H = \psi_{\widehat{\lambda}_i}(s)\widehat{B}_i^H \widehat{B}^T(\widehat{A}^T - sI)^{-1} + O(s+\widehat{\lambda}_i)^{k_i}. \tag{4.17}$$

This then yields

$$\frac{1}{2}(\nabla_{\widehat{B}}\mathcal{J})^T S_i \phi_{\widehat{\lambda}_i}(s) = (\widehat{B}^T\widehat{Q}+B^TY)S_i\phi_{\widehat{\lambda}_i}(s) = [H^T(s)-\widehat{H}^T(s)]\widehat{C}_i\phi_{\widehat{\lambda}_i}(s)+O(s+\widehat{\lambda}_i)^{k_i},$$

$$\frac{1}{2}\psi_{\widehat{\lambda}_i}(s)T_i^H(\nabla_{\widehat{C}}\mathcal{J})^T = \psi_{\widehat{\lambda}_i}(s)T_i^H(\widehat{P}\widehat{C}^T-X^TC^T) = \psi_{\widehat{\lambda}_i}(s)\widehat{B}_i^H[H^T(s)-\widehat{H}^T(s)]+O(s+\widehat{\lambda}_i)^{k_i}.$$

□

REMARK 4.6. *The condition that $-\widehat{\lambda}_i$ is not a pole of $H(s)$ is satisfied when choosing stable interpolation points $\widehat{\lambda}_i$, which is typically the case in the algorithms we discuss below.*

The following generalization of the tangential interpolation conditions (4.7) and (4.8) immediately follows from the previous theorem.

COROLLARY 4.7. *With the notation and assumptions of Theorem 4.5, if $\nabla_{\widehat{B}}\mathcal{J} = 0$ and $\nabla_{\widehat{C}}\mathcal{J} = 0$, then the following tangential interpolation conditions are satisfied for all $\widehat{\lambda}_i, i = 1,\ldots,n$:*

$$[H^T(s) - \widehat{H}^T(s)]\widehat{c}_i(s) = O(s+\widehat{\lambda}_i)^{k_i}, \quad \widehat{b}_i(s)^H[H^T(s) - \widehat{H}^T(s)] = O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.18}$$

where $\widehat{b}_i^H(s) := \psi_{\widehat{\lambda}_i}(s)\widehat{B}_i^H$ and $\widehat{c}_i(s) := \widehat{C}_i\phi_{\widehat{\lambda}_i}(s)$.

We now turn to the gradient of $\mathcal{J}$ versus $\widehat{A}$. We do not have expressions for $T_i^H(\nabla_{\widehat{A}}\mathcal{J})^T S_j$ that are clean extensions of (4.5) and (4.6), however, we do generalize the two-sided tangential interpolation condition (4.9) that follows from $\nabla_{\widehat{A}}\mathcal{J} = 0$. This yields the following main theorem, which states the complete generalization of Corollary 4.2 to higher-order poles, i.e., the characterization of stationary points by means of tangential interpolation conditions.

THEOREM 4.8. *With the notation and assumptions of Theorem 4.5, if $\nabla_{\widehat{B}}\mathcal{J} = 0$, $\nabla_{\widehat{C}}\mathcal{J} = 0$ and $\nabla_{\widehat{A}}\mathcal{J} = 0$, then the following tangential interpolation conditions are satisfied for $i = 1,\ldots,\ell$:*

$$[H^T(s) - \widehat{H}^T(s)]\widehat{c}_i(s) = O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.19}$$

$$\widehat{b}_i(s)^H[H^T(s) - \widehat{H}^T(s)] = O(s+\widehat{\lambda}_i)^{k_i}, \tag{4.20}$$

$$\widehat{b}_i(s)^H[H^T(s) - \widehat{H}^T(s)]\widehat{c}_i(s) = O(s+\widehat{\lambda}_i)^{2k_i}, \tag{4.21}$$

where $\widehat{b}_i^H(s) := \psi_{\widehat{\lambda}_i}(s)\widehat{B}_i^H$ and $\widehat{c}_i(s) := \widehat{C}_i\phi_{\widehat{\lambda}_i}(s)$.

Proof. Conditions (4.19) and (4.20) were obtained in Corollary 4.7. It remains to show that (4.21) holds.

We can interpret conditions (4.19)–(4.21) in terms of Taylor expansions of the error function $E(s) := H(s) - \widehat{H}(s)$. Let

$$E(s) := \sum_{j=0}^{\infty} E_i(s+\widehat{\lambda}_i)^j, \quad \widehat{c}_i(s) := \sum_{j=0}^{k_i} l_j(s+\widehat{\lambda}_i)^j, \quad \widehat{b}_i^H(s) := \sum_{j=0}^{k_i} r_j^H(s+\widehat{\lambda}_i)^j,$$

be the Taylor expansions around $s = -\widehat{\lambda}_i$ of the rational function $E(s)$ and of the polynomials $\widehat{c}_i(s)$ and $\widehat{b}_i(s)^H$. Then conditions (4.19)–(4.21) are respectively equivalent to

$$\begin{bmatrix} E_0^H & E_1^H & \dots & E_{k_i-1}^H \\ & E_0^H & \ddots & \vdots \\ & & \ddots & E_1^H \\ & & & E_0^H \end{bmatrix} \begin{bmatrix} l_0 & l_1 & \dots & l_{k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix} = 0, \tag{4.22}$$

$$\begin{bmatrix} r_0^H & r_1^H & \dots & r_{k_i-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} E_0^H & E_1^H & \dots & E_{k_i-1}^H \\ & E_0^H & \ddots & \vdots \\ & & \ddots & E_1^H \\ & & & E_0^H \end{bmatrix} = 0, \tag{4.23}$$

and

$$\begin{bmatrix} r_0^H & r_1^H & \dots & r_{2k_i-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} E_0^H & E_1^H & \dots & E_{2k_i-1}^H \\ & E_0^H & \ddots & \vdots \\ & & \ddots & E_1^H \\ & & & E_0^H \end{bmatrix} \begin{bmatrix} l_0 & l_1 & \dots & l_{2k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix} = 0. \tag{4.24}$$

The condition that the first k_i or $2k_i$ terms of the Taylor expansion vanish is indeed equivalent to the fact that the above partial convolutions are zero. We know that (4.22) and (4.23) hold, since (4.19) and (4.20) hold; it remains to show (4.24) to conclude the proof.

We will need the identity

$$\begin{bmatrix} E_{k_i}^H & \dots & E_{2k_i-1}^H \\ \vdots & \ddots & \vdots \\ E_1^H & \dots & E_{k_i}^H \end{bmatrix} =$$
$$\begin{bmatrix} B^T(A^T + \widehat{\lambda}_i I)^{-k_i} \\ \vdots \\ B^T(A^T + \widehat{\lambda}_i I)^{-1} \end{bmatrix} \begin{bmatrix} (A^T + \widehat{\lambda}_i I)^{-1} C^T & \dots & (A^T + \widehat{\lambda}_i I)^{-k_i} C^T \end{bmatrix}$$
$$- \begin{bmatrix} \widehat{B}^T(\widehat{A}^T + \widehat{\lambda}_i I)^{-k_i} \\ \vdots \\ \widehat{B}^T(\widehat{A}^T + \widehat{\lambda}_i I)^{-1} \end{bmatrix} \begin{bmatrix} (\widehat{A}^T + \widehat{\lambda}_i I)^{-1} \widehat{C}^T & \dots & (\widehat{A}^T + \widehat{\lambda}_i I)^{-k_i} \widehat{C}^T \end{bmatrix}, \tag{4.25}$$

which holds since

$$E_{f+g-1}^H = B^T(A^T + \lambda_i I)^{-f}(A^T + \lambda_i I)^{-g} C^T - \widehat{B}^T(\widehat{A}^T + \lambda_i I)^{-f}(\widehat{A}^T + \lambda_i I)^{-g} \widehat{C}^T.$$

Define

$$Y_i := Y S_i,\ \widehat{Q}_i := -\widehat{Q} S_i,\ X_i^H = -T_i^H X^T,\ \widehat{P}_i^H := -T_i^H \widehat{P}. \tag{4.26}$$

Using Wilson's formulas (Theorem 3.2) for the first equality, Lemmas 4.3 and 4.4 for the second one, and the identity (4.25) for the third, we have

$$
\begin{aligned}
& T_i^H(\nabla_{\widehat{A}}\mathcal{J})^T S_i && (4.27)\\
=& \widehat{P}_i^H \widehat{Q}_i - X_i^H Y_i \\
=& \begin{bmatrix} r_0^H & r_1^H & \dots & r_{k_i-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} B^T(A^T+\widehat{\lambda}_i I)^{-k_i} \\ \vdots \\ B^T(A^T+\widehat{\lambda}_i I)^{-1} \end{bmatrix} && (4.28)\\
& \begin{bmatrix} (A^T+\widehat{\lambda}_i I)^{-1}C^T & \dots & (A^T+\widehat{\lambda}_i I)^{-k_i}C^T \end{bmatrix} \begin{bmatrix} l_0 & l_1 & \dots & l_{k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix} \\
& - \begin{bmatrix} r_0^H & r_1^H & \dots & r_{k_i-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} \widehat{B}^T(\widehat{A}^T+\widehat{\lambda}_i I)^{-k_i} \\ \vdots \\ \widehat{B}^T(\widehat{A}^T+\widehat{\lambda}_i I)^{-1} \end{bmatrix} \\
& \begin{bmatrix} (\widehat{A}^T+\widehat{\lambda}_i I)^{-1}\widehat{C}^T & \dots & (\widehat{A}^T+\widehat{\lambda}_i I)^{-k_i}\widehat{C}^T \end{bmatrix} \begin{bmatrix} l_0 & l_1 & \dots & l_{k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix} \\
=& - \begin{bmatrix} r_0^H & r_1^H & \dots & r_{k_i-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} E_{k_i}^H & \dots & E_{2k_i-1}^H \\ \vdots & \ddots & \vdots \\ E_1^H & \dots & E_{k_i}^H \end{bmatrix} \begin{bmatrix} l_0 & l_1 & \dots & l_{k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix}. && (4.29)
\end{aligned}
$$

We are now ready to show (4.24). Since (4.22) and (4.23) hold, the left-hand side of (4.24) satisfies

$$
\begin{aligned}
& \begin{bmatrix} r_0^H & r_1^H & \dots & r_{2k_i-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} E_0^H & E_1^H & \dots & E_{2k_i-1}^H \\ & E_0^H & \ddots & \vdots \\ & & \ddots & E_1^H \\ & & & E_0^H \end{bmatrix} \begin{bmatrix} l_0 & l_1 & \dots & l_{2k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix} \\
=& \begin{bmatrix} & & & \\ 0 & \begin{bmatrix} r_0^H & r_1^H & \dots & r_{k_i-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} & \begin{bmatrix} E_{k_i}^H & \dots & E_{2k_i-1}^H \\ \vdots & \ddots & \vdots \\ E_1^H & \dots & E_{k_i}^H \end{bmatrix} & \begin{bmatrix} l_0 & l_1 & \dots & l_{k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix} \\ 0 & & & \end{bmatrix} \\
=& \begin{bmatrix} 0 & -T_i^H(\nabla_{\widehat{A}}\mathcal{J})^T S_i \\ 0 & 0 \end{bmatrix}, && (4.30)
\end{aligned}
$$

where the first equality follows from a careful blockwise inspection, and the second equality uses (4.29). Since $\nabla_{\widehat{A}}\mathcal{J} = 0$, it follows that (4.24) holds, and thus (4.21) holds. □

4.3. Number of parameters and conditions. In this subsection, we show that the tangential interpolation conditions obtained in Theorem 4.8—i.e., (4.19)–(4.21)—impose the correct number, $n(m+p)$, of nonredundant scalar conditions.

To this end, fix i and consider the Jordan block of size k_i associated to λ_i. The tangential interpolation conditions are equivalent to (4.22)–(4.24). Both (4.22) and (4.23) agree on imposing that

$$\begin{bmatrix} E_0^H & E_1^H & \dots & E_{k_i-1}^H \\ & E_0^H & \ddots & \vdots \\ & & \ddots & E_1^H \\ & & & E_0^H \end{bmatrix}$$

has a kernel of dimension k_i. Indeed, the fact that the realization is observable imposes that $\ell_0 \neq 0$, and thus the k_i columns of

$$\begin{bmatrix} l_0 & l_1 & \dots & l_{k_i-1} \\ & l_0 & \ddots & \vdots \\ & & \ddots & l_1 \\ & & & l_0 \end{bmatrix}$$

are linearly independent. This counts for k_i conditions. Next, in (4.22), the equations in columns 1 to $k_i - 1$ are redundant with the equations in column k_i. There are thus $k_i p$ conditions, but the left-hand matrix is known to have a kernel of dimension k_i; this reduces the number of nontrivial conditions to $k_i p - k_i$. The same reasoning on (4.23) leads to $k_i m - k_i$ conditions. Finally, once (4.22) and (4.23) hold, the two-sided condition (4.21), equivalent to (4.24), imposes k_i additional conditions. This is because the left-hand side of (4.24) reduces to (4.30), a Toeplitz matrix with only k_i nonzero diagonals. In total for i, we have $k_i(m+p)$ nonredundant conditions. The overall total is thus $\sum_{i=1}^{\ell} k_i(m+p) = n(m+p)$, which is the dimension of $\mathrm{Rat}_{p,m}^n$.

5. Relation with tangential interpolation by projection. The gradient forms of Theorem 3.2 yields the following theorem (proved in [VGA08]) that provides an important link to tangential interpolation by projection.

THEOREM 5.1. *At every stationary point of $\mathcal{J}$ (2.7) where $\widehat{P}$ and $\widehat{Q}$ are invertible, we have the following identities*

$$\widehat{A} = W^T A V, \quad \widehat{B} = W^T B, \quad \widehat{C} = CV, \quad W^T V = I_n \tag{5.1}$$

where $W := -Y\widehat{Q}^{-1}$, $V := X\widehat{P}^{-1}$ and X, Y, $\widehat{P}$ and $\widehat{Q}$ satisfy the Sylvester equations (3.2,3.3).

If we rewrite the above theorem as a projection problem, then we are constructing a projector $\Pi := VW^T$ (implying $W^T V = I_n$) where V and W are given by the following (transposed) Sylvester equations

$$(\widehat{Q}W^T)A + \widehat{A}^T(\widehat{Q}W^T) + \widehat{C}^T C = 0, \quad A(V\widehat{P}) + (V\widehat{P})\widehat{A}^T + B\widehat{B}^T = 0. \tag{5.2}$$

Note that $\widehat{P}$ and $\widehat{Q}$ can be interpreted as normalizations to ensure that $W^TV = I_n$.

Rewriting the Sylvester equations (5.2) as

$$W^TA + (\widehat{Q}^{-1}\widehat{A}\widehat{Q})W^T + (\widehat{C}\widehat{Q}^{-1})C = 0, \tag{5.3a}$$

$$AV + V(\widehat{P}\widehat{A}^T\widehat{P}^{-1}) + B(\widehat{B}^T\widehat{P}^{-1}) = 0, \tag{5.3b}$$

shows the relation with the tangential interpolation described in [GVV05]. There it is shown that when solving two Sylvester equations for the unknowns $W, V \in \mathbb{R}^{N\times n}$

$$W^TA - \Sigma_\mu^TW^T + L^TC = 0, \tag{5.4}$$

$$AV - V\Sigma_\sigma + BR = 0, \tag{5.5}$$

and constructing the reduced-order model (of degree n) as follows

$$(\widehat{A}, \widehat{B}, \widehat{C}) := ((W^TV)^{-1}W^TAV, (W^TV)^{-1}W^TB, CV), \tag{5.6}$$

amounts to a tangential interpolation problem (provided the matrix W^TV is invertible). The "interpolation conditions" (Σ_σ, R) and (Σ_μ, L) (where $\Sigma_\mu, \Sigma_\sigma \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times n}$ and $L \in \mathbb{R}^{p\times n}$) are known to uniquely determine the projected system $(\widehat{A}, \widehat{B}, \widehat{C})$ [GVV05]. Moreover, they reproduce exactly the conditions derived in the previous section since they can be expressed in another coordinate system by applying invertible transformations of the type $\left(Q^{-1}\Sigma_\sigma Q, RQ\right)$ and $\left(P^{-1}\Sigma_\mu P, LP\right)$ to the interpolation conditions. This yields transformed matrices VP and WQ but does not affect the transfer function of the reduced-order model $(\widehat{A}, \widehat{B}, \widehat{C})$ (see [GVV05] for more details). The novelty of the derivation in this paper is the case of higher-order poles: the tangential interpolation conditions in Theorem 4.8 contain fewer redundant equations than those that would follow from [GVV05].

6. First-order versus higher-order poles. In this section we show that $\mathcal{H}_2$-optimal reduced-order models with repeated poles can indeed occur and that in their neighborhood one can expect the tangential interpolation approach to have serious numerical difficulties. We start with a lemma that will allow us to demonstrate this.

LEMMA 6.1. *A stable n-th degree transfer function $\hat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}$ is a stationary point of the error function $\|\widehat{H}(s) - H(s)\|_{\mathcal{H}_2}$ if and only if $H(s)$ can be realized as follows*

$$A = \begin{bmatrix} \widehat{A} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} \widehat{B} \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} \widehat{C} & C_2 \end{bmatrix}, \tag{6.1}$$

where moreover

$$\widehat{A}\widehat{P} + \widehat{P}\widehat{A}^T + \widehat{B}\widehat{B}^T = 0, \quad A_{21}\widehat{P} + B_2\widehat{B}^T = 0, \tag{6.2}$$

$$\widehat{Q}\widehat{A} + \widehat{A}^T\widehat{Q} + \widehat{C}^T\widehat{C} = 0, \quad \widehat{Q}A_{12} + \widehat{C}^TC_2 = 0. \tag{6.3}$$

Proof. The proof follows from the stationarity conditions in Theorem 3.2. The "if" part is direct: the stationarity conditions hold with $X = \left[\begin{smallmatrix} \widehat{P} \\ 0 \end{smallmatrix}\right]$ and $Y = -\left[\begin{smallmatrix} \widehat{Q} \\ 0 \end{smallmatrix}\right]$. For the "only if" part, the assumption that $\widehat{H}(s)$ is stable and of degree n, guarantees

that the matrices $\widehat{P}$ and $\widehat{Q}$ exist and are invertible. Using $Y^T X = -\widehat{P}\widehat{Q}$ one can then always choose a coordinate system for the realization of $H(s)$ in which

$$X = \begin{bmatrix} \widehat{P} \\ 0 \end{bmatrix}, \quad Y = -\begin{bmatrix} \widehat{Q} \\ 0 \end{bmatrix}$$

and hence

$$W = X\widehat{P}^{-1} = \begin{bmatrix} I_n \\ 0 \end{bmatrix}, \quad V = -Y\widehat{Q}^{-1} = \begin{bmatrix} I_n \\ 0 \end{bmatrix}.$$

Therefore we have $A_{11} = \widehat{A},\ B_1 = \widehat{B},\ C_1 = \widehat{C}$. □

This special coordinate system can be used to construct a transfer function $H(s)$ for which a *given* $\widehat{H}(s)$ is the best $\mathcal{H}_2$ norm approximation of $H(s)$.

THEOREM 6.2. *Let $\hat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}$ be a given stable n-th degree transfer function, then there always exists a stable N-th degree transfer function $H(s) = C(sI_N - A)^{-1}B$ with $N > n$, for which $\widehat{H}(s)$ is a stationary point of the $\mathcal{H}_2$ error function.*

Proof. It suffices to construct $\widehat{P}$ and $\widehat{Q}$ satisfying the Lyapunov equations in (6.2) and (6.3), and then choose $A_{21} = -B_2\widehat{B}^T\widehat{P}^{-1}$ and $A_{12} = -\widehat{Q}^{-1}\widehat{C}^T C_2$ to satisfy the conditions of Lemma 6.1. Notice that this always has a solution since $\widehat{P}$ and $\widehat{Q}$ are invertible because $\widehat{H}(s)$ is stable and minimal. In order to guarantee that $H(s)$ is also stable, one needs to choose the remaining degrees of freedom, i.e. A_{22}, B_2 and C_2 to satisfy this condition. This can be achieved in several ways, but the simplest one is to choose A_{22} stable, and the matrices B_2 and C_2 sufficiently small. The matrices $A_{21} = -B_2\widehat{B}^T\widehat{P}^{-1}$ and $A_{12} = -\widehat{Q}^{-1}\widehat{C}^T C_2$ will then also be small, and A will then be essentially block diagonal and hence stable. □

The above theorem does not show that the constructed stationary point is also a local minimum, but the following example shows that this is not too difficult to construct. Choose $\widehat{H}(s) = 1/(s-a)^2$ with $a = -1$ and a realization

$$\widehat{A} = \left[\begin{smallmatrix} a & 1 \\ 0 & a \end{smallmatrix}\right], \widehat{B} = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right], \widehat{C} = \left[\begin{smallmatrix} 1 & 0 \end{smallmatrix}\right]$$

then the transfer function $H(s) = (0.25s^2 - 0.5s + 9.25)/(s^3 + 7s^2 + 19s + 9)$ with realization

$$A = \left[\begin{smallmatrix} a & 1 & d \\ 0 & a & e \\ e & d & f \end{smallmatrix}\right], B = \left[\begin{smallmatrix} 0 \\ 1 \\ g \end{smallmatrix}\right], C = \left[\begin{smallmatrix} 1 & 0 & g \end{smallmatrix}\right]$$

with $f = -5$, $g = .5$, $d = 4ag$, $e = 4a^2 g$, is stable and satisfies the stationarity conditions of Lemma 6.1. Moreover, 1000 random perturbations of the stationary point $\widehat{H}(s)$ show that this is clearly a local minimum of the error function $\|H - \widehat{H}\|_{\mathcal{H}_2}$.

This example shows that if we aim for an $\mathcal{H}_2$-optimal reduced-order model $\widehat{H}(s)$ with multiple poles, the model reduction technique that restricts itself to first-order poles will not be able to produce that solution. However, what happens if we perturb $H(s)$ or $\widehat{H}(s)$? What can we say about the mapping from one to the other? This is addressed in the following theorem, which shows that if $\widehat{H}(s)$ is a stationary point of the $\mathcal{H}_2$-distance to $H(s)$, then every sufficiently nearby transfer function $\widehat{H}_\Delta(s)$ is a stationary point of a nearby system $H_\Delta(s)$.

THEOREM 6.3. *Let $\hat{H}(s) = \widehat{C}(sI_n - \widehat{A})^{-1}\widehat{B}$ and $H(s) = C(sI_N - A)^{-1}B$ be stable and minimal transfer functions such that $\widehat{H}(s)$ is a stationary point (resp.,*

nondegenerate local minimum) of the error function $\|H(s) - \widehat{H}(s)\|_{\mathcal{H}_2}$. *Then, for every neighborhood* $\mathcal{U}$ *of* $H(s)$ *in* $\mathrm{Rat}_{p,m}^n$, *there exists a neighborhood* $\hat{\mathcal{U}}$ *of* $\widehat{H}(s)$ *in* $\mathrm{Rat}_{p,m}^N$ *such that, for all* $\widehat{H}_\Delta(s) \in \hat{\mathcal{U}}$, *there exists* $H_\Delta(s) \in \mathcal{U}$ *for which* $\widehat{H}_\Delta(s)$ *is a stationary point (resp., nondegenerate local minimum) of the* $\mathcal{H}_2$*-distance to* $H_\Delta(s)$.

Proof. The proof consists of constructing a continuous mapping ψ from a neighborhood $\mathcal{V}$ of $\widehat{H}(s)$ in $\mathrm{Rat}_{p,m}^n$ into $\mathrm{Rat}_{p,m}^N$ such that $\widehat{H}_\Delta(s)$ is a stationary point of the $\mathcal{H}_2$-distance to $\psi(\widehat{H}_\Delta(s))$ for all $\widehat{H}_\Delta(s)$ in $\mathcal{V}$. We use Lemma 6.1 to do this. Let $(\widehat{A}_\Delta, \widehat{B}_\Delta, \widehat{C}_\Delta)$ be a nearby realization of the nearby system $\widehat{H}_\Delta(s)$. The solution $\widehat{P}_\Delta$ and $\widehat{Q}_\Delta$ of the perturbed Lyapunov equations in (6.2) and (6.3), will be close to $\widehat{P}$ and $\widehat{Q}$ by continuity of the solution of a non-singular system of equations. For the same reason we can construct nearby solutions $A_{21\Delta} = -B_2\widehat{B}_\Delta^T\widehat{P}_\Delta^{-1}$ and $A_{12\Delta} = -\widehat{Q}_\Delta^{-1}\widehat{C}_\Delta^T C_2$ to finally yield a realization

$$A_\Delta = \begin{bmatrix} \widehat{A}_\Delta & A_{12\Delta} \\ A_{21\Delta} & A_{22} \end{bmatrix}, \quad B_\Delta = \begin{bmatrix} \widehat{B}_\Delta \\ B_2 \end{bmatrix}, \quad C_\Delta = \begin{bmatrix} \widehat{C}_\Delta & C_2 \end{bmatrix},$$

for a transfer function $H_\Delta(s) =: \psi(\widehat{H}_\Delta(s))$ which is close to $H(s)$ and satisfies the conditions of Lemma 6.1. Since, in view of its expression (2.1), the $\mathcal{H}_2$-norm error function is locally smooth in terms of the coefficients of system parameters of $H(s)$ and $\widehat{H}(s)$, every stationary point that is a nondegenerate local minimum remains a local minimum for sufficiently small perturbations. The proof therefore applies to such points. □

This theorem implies that the set of full-order models $H(s)$ that have $\mathcal{H}_2$-stationary reduced-order models with only simple poles, is open and dense in $\mathrm{Rat}_{p,m}^N$. This follows from the following reasoning. From the continuity of the mapping from $H(s)$ to $\widehat{H}(s)$ and from the fact that the set of systems with only simple poles is open, it follows that, around a system $H(s)$ with reduced-order models with only simple poles, there is an neighborhood of systems with reduced-order models with only simple poles. If $H(s)$ has a reduced-order model $\widehat{H}(s)$ with multiple poles, then, because the "reduction" map is an open map and the set of systems with only simple poles has an empty interior, it follows that any neighborhood of $H(s)$ contains a full-order model with a reduced-order model with only simple poles. One could conclude from this that one need only consider first-order interpolation techniques, but, when one approaches a system for which the target function $\widehat{H}(s)$ has multiple poles, the interpolation conditions change in a non-smooth manner in its neighborhood. The first-order conditions will become linearly dependent and they will no longer define the reduced-order model uniquely. This is obvious in the SISO case. In the MIMO case, observe that the tangential interpolation conditions (4.7) involve the interpolation direction $\widehat{c}_i = \widehat{C}s_i$, where s_i is the eigenvector of $\widehat{A}$ related to $\widehat{\lambda}_i$; if $\widehat{\lambda}_i$ and $\widehat{\lambda}_{i+1}$ coalesce to form a nontrivial Jordan block, then the eigenvectors s_i and s_{i+1} merge (see [Wil65]) and hence the tangential interpolation directions merge, too. This implies that the systems of equations that one solves become ill-conditioned in the neighborhood of a point where the solution has higher-order poles. The same ill-condioned behavior can be expected for any target system $\widehat{H}(s)$ which has no higher-order poles but is near a system with higher-order poles.

7. First- and complex second-order approximation. In this section we consider how the error function changes with the interpolation conditions. In order to analyze this, we look at first- and second-order approximations only, i.e., approxima-

tion by systems with one real pole or two complex conjugate poles. If we are looking for a (real) first-order approximation

$$\widehat{H}(s) = cb^T/(s-\lambda)$$

then according to the formulas of Section 4, it should satisfy the following properties at every stationary point of $\mathcal{J}$:

$$H^T(-\lambda)c = -b\frac{c^Tc}{2\lambda}, \quad b^TH^T(-\lambda) = -c^T\frac{b^Tb}{2\lambda}, \quad b^T\frac{d}{ds}H^T(s)c|_{s=-\lambda} = -\frac{b^Tbc^Tc}{4\lambda^2}.$$

If we are looking for a second-order approximation with complex conjugate poles

$$\widehat{H}(s) = cb^H/(s-\lambda) + \overline{c}\overline{b}^H/(s-\overline{\lambda})$$

then it should satisfy the following properties at every stationary point of $\mathcal{J}$:

$$H^T(-\lambda)c = -b\frac{c^Hc}{2\lambda}, \quad b^HH^T(-\lambda) = -c^H\frac{b^Hb}{2\lambda}, \quad b^H\frac{d}{ds}H^T(s)c|_{s=-\lambda} = -\frac{b^Hbc^Hc}{4\lambda^2}.$$

In both cases, the first two equations express that for every interpolation point $-\lambda$ (real or complex) one should choose left and right singular vectors of $H^T(-\lambda)$ as tangential interpolation directions b and c for constructing the first- and second-order section. The third equation (combined with the two previous ones) expresses that the interpolation point is a stationary point of the error function versus λ.

If we keep the interpolation point as a parameter, we can plot the error function versus $-\lambda$, but where b and c are chosen optimal for that interpolation point. In other words, the optimal approximation $\widehat{H}(s)$ is then completely defined by the interpolation point $-\lambda$. We can therefore have a look at the function we need to optimize by plotting the error function $\|H(s)-\widehat{H}(s)\|^2_{\mathcal{H}_2}$ as a function of λ. It follows from the optimality conditions on b or c that $\|H(s)-\widehat{H}(s)\|^2_{\mathcal{H}_2} = \|H(s)\|^2_{\mathcal{H}_2} - \|\widehat{H}(s)\|^2_{\mathcal{H}_2}$. Indeed, let $\nabla_{\widehat{B}}\mathcal{J} = Y^TB + \widehat{Q}\widehat{B} = 0$ then

$$\begin{aligned}\|H(s)-\widehat{H}(s)\|^2_{\mathcal{H}_2} &= \mathrm{tr}\left(B^TQB + B^TY\widehat{B} + \widehat{B}^TY^TB + \widehat{B}^T\widehat{Q}\widehat{B}\right)\\ &= \mathrm{tr}\left(B^TQB\right) - \mathrm{tr}\left(\widehat{B}^T\widehat{Q}\widehat{B}\right)\\ &= \|H(s)\|^2_{\mathcal{H}_2} - \|\widehat{H}(s)\|^2_{\mathcal{H}_2}.\end{aligned}$$

The development for $\nabla_{\widehat{C}}\mathcal{J} = CX - \widehat{C}\widehat{P} = 0$ is essentially the same. In the real case we then have

$$\|\widehat{H}(s)\|^2_{\mathcal{H}_2} = b^T\widehat{H}^T(-\lambda)c = \frac{b^Tbc^Tc}{-2\lambda}$$

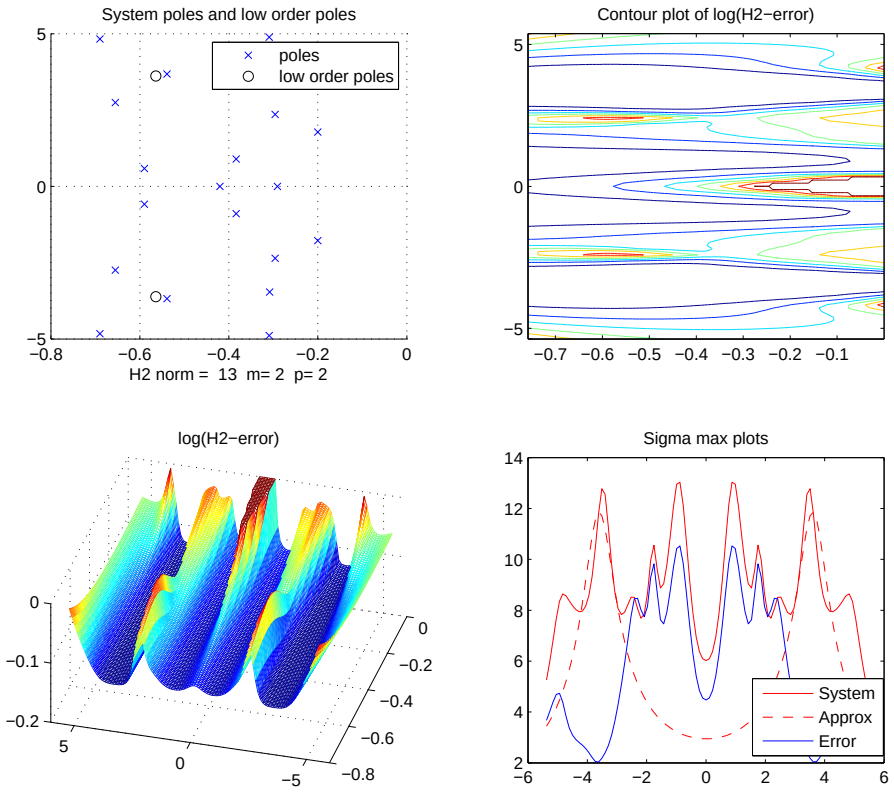
which implies $\|\widehat{H}(s)\|^2_{\mathcal{H}_2} = \frac{\sigma^2(H(-\lambda))}{-2\lambda}$ because of the above formulas. This indicates that we need to choose the vectors b and c corresponding to the largest singular value of $H(\lambda)$. In the complex case we have

$$\|\widehat{H}(s)\|^2_{\mathcal{H}_2} = 2\Re\left(b^H\widehat{H}^T(-\lambda)c + b^T\widehat{H}^T(-\overline{\lambda})\overline{c}\right)$$

and the same conclusion follows after some manipulation.

In Figure 7.1 we show this function for a MIMO example with $m = p = 2$ and $N = 20$, for which the optimum is reached at a pair of complex conjugate interpolation points. Subplot 1 shows the poles of $H(s)$ (blue crosses) and the poles of the $\mathcal{H}_2$-optimal reduced-order model $\widehat{H}(s)$ (black circles). Subplots 2 and 3 show the log of the $\mathcal{H}_2$ norm of the error as a function of the interpolation point $-\lambda$ (both in contour and in 3D view). Subplot 4 shows the frequency response norms $\sigma_{\max}(G(j\omega))$, where $G(s)$ is the system $H(s)$, the optimal second-order approximation $\widehat{H}(s)$ and the error $H(s) - \widehat{H}(s)$. This system was generated randomly, but the function is not so simple to optimize. It is clearly not convex and there are several basins of attraction to local minima that are not optimal. One often recommends to start with the poles closest to the $j\omega$ axis as interpolation points (or the largest peaks in the frequency response), but for this example that would converge to local minima, as one can see from the $\mathcal{H}_2$ error plot.

FIG. 7.1. *Second-order approximation of MIMO case*



8. Algorithms for solving the interpolation problem. One can view (3.2,3.3) and (5.1) as two coupled systems of equations

$$(X, Y, \widehat{P}, \widehat{Q}) = F(\widehat{A}, \widehat{B}, \widehat{C}) \quad \text{and} \quad (\widehat{A}, \widehat{B}, \widehat{C}) = G(X, Y, \widehat{P}, \widehat{Q})$$

for which we have a fixed point $(\widehat{A},\widehat{B},\widehat{C}) = G(F(\widehat{A},\widehat{B},\widehat{C}))$ at every stationary point of $\mathcal{J}(\widehat{A},\widehat{B},\widehat{C})$. This automatically suggests an iterative procedure

$$(X,Y,\widehat{P},\widehat{Q})_{i+1} = F(\widehat{A},\widehat{B},\widehat{C})_{i+1}, \quad (\widehat{A},\widehat{B},\widehat{C})_{i+1} = G(X,Y,\widehat{P},\widehat{Q})_i,$$

which is expected to converge to a nearby fixed point. This is essentially the idea behind existing algorithms using Sylvester equations in their iterations (see [Ant05]). Specifically, this is the idea behind the IRKA algorithm of [GAB07], except that one has to adapt the formulas to make sure that the matrices V and W satisfy $W^TV = I_n$. Another approach would be to use the gradients (or the interpolation conditions of Theorem 4.1) to develop descent methods or even Newton-like methods, as was done for the SISO case in [GAB07]. Quasi-Newton methods where the optimal variables are the interpolation points were developed in [BG07]. Such local optimization methods allow for local superlinear convergence to local minimizers of the error function, but cannot guarantee global convergence to the global minimizer. The analysis of Section 6 also shows that using the diagonal canonical form for such algorithms may lack the required robustness properties.

9. The discrete-time case. Now consider the equivalent formulation in the discrete-time case. We then have the dynamical systems

$$\left\{ \begin{array}{l} x_{k+1} = Ax_k + Bu_k \\ y_k = Cx_k \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} \widehat{x}_{k+1} = \widehat{A}\widehat{x}_k + \widehat{B}u \\ \widehat{y}_k = \widehat{C}\widehat{x}_k \end{array} \right.$$

with transfer functions

$$H(z) = C(zI - A)^{-1}B, \quad \text{and} \quad \widehat{H}(z) = \widehat{C}(zI - \widehat{A})^{-1}\widehat{B}.$$

The squared $\mathcal{H}_2$-norm of the error function $E(z) := H(z) - \widehat{H}(z)$ is then defined as

$$\mathcal{J} := \| E(z) \|_{\mathcal{H}_2}^2 := \mathrm{tr} \int_{-\infty}^{\infty} E(e^{j\omega})E(e^{j\omega})^H \frac{d\omega}{2\pi} = \mathrm{tr} \sum_{k=0}^{\infty} (C_eA_e^kB_e)(C_eA_e^kB_e)^T \quad (9.1)$$

where (A_e, B_e, C_e) defined in (2.4) is again a realization of the error transfer function $E(z)$. The $\mathcal{H}_2$-norm can now be rewritten in terms of the solutions of the Stein equations

$$A_eP_eA_e^T + B_eB_e^T = P_e, \quad A_e^TQ_eA_e + C_e^TC_e = Q_e \quad (9.2)$$

as

$$\mathcal{J} = \mathrm{tr}\left(C_eP_eC_e^T\right) = \mathrm{tr}\left(B_e^TQ_eB_e\right).$$

Partition again the solutions

$$P_e := \left[\begin{array}{cc} P & X \\ X^T & \widehat{P} \end{array} \right], \quad Q_e := \left[\begin{array}{cc} Q & Y \\ Y^T & \widehat{Q} \end{array} \right],$$

to obtain the Stein equations in the form

$$\left[\begin{array}{cc} A & \\ & \widehat{A} \end{array} \right] \left[\begin{array}{cc} P & X \\ X^T & \widehat{P} \end{array} \right] \left[\begin{array}{cc} A^T & \\ & \widehat{A}^T \end{array} \right] + \left[\begin{array}{c} B \\ \widehat{B} \end{array} \right] \left[B^T \;\; \widehat{B}^T \right] = \left[\begin{array}{cc} P & X \\ X^T & \widehat{P} \end{array} \right],$$

$$\left[\begin{array}{cc} A^T & \\ & \widehat{A}^T \end{array} \right] \left[\begin{array}{cc} Q & Y \\ Y^T & \widehat{Q} \end{array} \right] \left[\begin{array}{cc} A & \\ & \widehat{A} \end{array} \right] + \left[\begin{array}{c} C^T \\ -\widehat{C}^T \end{array} \right] \left[C \;\; -\widehat{C} \right] = \left[\begin{array}{cc} Q & Y \\ Y^T & \widehat{Q} \end{array} \right].$$

THEOREM 9.1. *The gradients $\nabla_{\widehat{A}}\mathcal{J}$, $\nabla_{\widehat{B}}\mathcal{J}$ and $\nabla_{\widehat{C}}\mathcal{J}$ of $\mathcal{J} := \|E(s)\|_{\mathcal{H}_2}^2$ are given by*

$$\nabla_{\widehat{A}}\mathcal{J} = 2(\widehat{Q}\widehat{A}\widehat{P} + Y^T A X), \quad \nabla_{\widehat{B}}\mathcal{J} = 2(\widehat{Q}\widehat{B} + Y^T B), \quad \nabla_{\widehat{C}}\mathcal{J} = 2(\widehat{C}\widehat{P} - CX), \tag{9.3}$$

where

$$A^T Y \widehat{A} - C^T \widehat{C} = Y, \quad \widehat{A}^T \widehat{Q} \widehat{A} + \widehat{C}^T \widehat{C} = \widehat{Q}, \tag{9.4}$$

$$\widehat{A} X^T A^T + \widehat{B} B^T = X^T, \quad \widehat{A}\widehat{P}\widehat{A}^T + \widehat{B}\widehat{B}^T = \widehat{P}. \tag{9.5}$$

Setting the gradient of $\mathcal{J}$ to zero yields the stationarity conditions derived in [BKVW07]. These are the discrete-time counterpart of Wilson's conditions (see [Wil70] or Theorem 3.2).

Again, at a stationary point (where all gradients are zero) we have that the projection matrices

$$W := -Y\widehat{Q}^{-1}, \quad V := X\widehat{P}^{-1}$$

satisfy $\widehat{A} = W^T A V$, $\widehat{B} = W^T B$, $\widehat{C} = CV$, $W^T V = I$ and the Sylvester equations

$$\begin{cases} \widehat{A}^T(\widehat{Q}W^T)A + \widehat{C}^T C = (\widehat{Q}W^T) \\ A(V\widehat{P})\widehat{A}^T + B\widehat{B}^T = (V\widehat{P}) \end{cases}$$

indicating that we are solving a tangential interpolation problem in the inverses of the eigenvalues of $\widehat{A}$, and this both left and right.

Let us now look at the tangential interpolation conditions for the discrete-time case. We treat immediately the higher-order case and specialize afterward to the case of order 1 interpolation conditions. Lemmas 4.3 and 4.4 have the following analogues.

LEMMA 9.2. *If λ^{-1} is not an eigenvalue of A, the solution of the matrix equation*

$$A^T Y F - Y = C^T L \quad \text{with} \quad F := \begin{bmatrix} \lambda & -1 & & \\ & \lambda & \ddots & \\ & & \ddots & -1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{k\times k},$$

with $L := \begin{bmatrix} \ell_0 & \ell_1 & \dots & \ell_{k-1} \end{bmatrix}$, is given by

$$Y = \begin{bmatrix} (\lambda A^T - I)^{-1} C^T & \dots & A^{T^{k-1}} (\lambda A^T - I)^{-k} C^T \end{bmatrix} \begin{bmatrix} \ell_0 & \ell_1 & \dots & \ell_{k-1} \\ & \ell_0 & \ddots & \vdots \\ & & \ddots & \ell_1 \\ & & & \ell_0 \end{bmatrix}.$$

Moreover, let

$$\phi_\lambda(z) := \begin{bmatrix} 1 & (\lambda - z) & \dots & (\lambda - z)^{k-1} \end{bmatrix}^T, \quad y(z) := Y\phi_\lambda(z)$$

then

$$y(z) = (zA^T - I)^{-1}C^T L\phi_\lambda(z) + O(\lambda - z)^k$$

which means that the ith column y_i of Y is also the coefficient of $(\lambda - z)^{i-1}$ in the Taylor expansion of $(zA^T - I)^{-1}C^T L\phi_\lambda(z)$.

Proof. The first part easily follows from $(\lambda A^T - I)y_1 = C^T\ell_0$ and $(\lambda A^T - I)y_i = C^T\ell_{i-1} + A^T y_{i-1},\ i > 1$. The second part follows from the identity

$$(zA^T - I)^{-1}C^T = \sum_{i=0}^{\infty}(\lambda - z)^i {A^T}^i(\lambda A^T - I)^{-i-1}C^T$$

and from the convolution of this formal series with the polynomial vector $L\phi_\lambda(z)$. □

We give the dual version of this lemma without proof.

LEMMA 9.3. *If λ^{-1} is not an eigenvalue of A, the solution of the matrix equation*

$$FX^H A^T - X^H = R^H B^T$$

with $F \in \mathbb{C}^{k\times k}$ as above and $R := \begin{bmatrix} r_{k-1} & r_{k-2} & \dots & r_0 \end{bmatrix}$, is given by

$$X^H = \begin{bmatrix} r_0^H & r_1^H & \dots & r_{k-1}^H \\ & r_0^H & \ddots & \vdots \\ & & \ddots & r_1^H \\ & & & r_0^H \end{bmatrix} \begin{bmatrix} B^T {A^T}^{k-1}(\lambda A^T - I)^{-k} \\ \vdots \\ B^T A^T(\lambda A^T - I)^{-2} \\ B^T(\lambda A^T - I)^{-1} \end{bmatrix}.$$

Moreover, let

$$\psi_\lambda(z) := \begin{bmatrix} (\lambda - z)^{k-1} & \dots & (\lambda - z) & 1 \end{bmatrix}, \quad x^H(z) := \psi_\lambda(z)X^H$$

then

$$x^H(z) = \psi_\lambda(z)R^H B^T(zA^T - I)^{-1} + O(\lambda - z)^k$$

which means that the ith row x_i^H of X^H is also the coefficient of $(\lambda - z)^{i-1}$ in the Taylor expansion of $\psi_\lambda(z)R^H B^T(zA^T - I)^{-1}$.

This now leads to the following theorems with interpolation conditions in terms of the transfer function $H_*(z) := z^{-1}H^T(z^{-1})$:

$$H_*(z) := B^T(I - zA^T)^{-1}C^T = -\sum_{i=0}^{\infty}(\lambda - z)^i B^T {A^T}^i(\lambda A^T - I)^{-i-1}C^T.$$

Since the proof is essentially the same as the one for the continuous-time case, it is omitted here.

THEOREM 9.4. *Let $\widehat{H}(z) = \sum_{i=1}^{\ell}\widehat{H}_i(z)$, $\widehat{H}_i(z) := \widehat{C}_i(zI - \widehat{A}_i)^{-1}\widehat{B}_i^H$ where $\{(\widehat{A}_i, \widehat{B}_i^H, \widehat{C}_i) : i = 1, \dots, \ell\}$ is a self-conjugate set and $\widehat{A}_i$ is just one Jordan block of size k_i associated with eigenvalue $\widehat{\lambda}_i$, and where $\widehat{\lambda}_i^{-1}$ is not a pole of $H(z)$ or $\widehat{H}(z)$. Then with*

$$\widehat{b}_i(z)^H := \begin{bmatrix} (\widehat{\lambda}_i - z)^{k_i-1} & \dots & (\widehat{\lambda}_i - z) & 1 \end{bmatrix} \widehat{B}_i^H,$$

$$\widehat{c}_i(z) := \widehat{C}_i \begin{bmatrix} 1 & (\widehat{\lambda}_i - z) & \dots & (\widehat{\lambda}_i - z)^{k_i - 1} \end{bmatrix}^T,$$

we have

$$[H_*^T(z) - \widehat{H}_*^T(z)]\widehat{c}_i(z) = O(\widehat{\lambda}_i - z)^{k_i}, \tag{9.6}$$

$$\widehat{b}_i(z)^H[H_*^T(z) - \widehat{H}_*^T(z)] = O(\widehat{\lambda}_i - z)^{k_i}, \tag{9.7}$$

$$\widehat{b}_i(z)^H[H_*^T(z) - \widehat{H}_*^T(z)]\widehat{c}_i(z) = O(\widehat{\lambda}_i - z)^{2k_i}, \tag{9.8}$$

where S_i, T_i are as defined in (4.11).

In the case of first-order poles, the conditions reduce to the following result, found in [BKVW07] in an equivalent form.

COROLLARY 9.5. *For the case of first-order poles (i.e. $k_i = 1$), the above conditions become :*

$$[H_*^T(z) - \widehat{H}_*^T(z)]\widehat{c}_i = O(\widehat{\lambda}_i - z), \quad \widehat{b}_i^H[H_*^T(z) - \widehat{H}_*^T(z)] = O(\widehat{\lambda}_i - z),$$

$$\widehat{b}_i^H[H_*^T(z) - \widehat{H}_*^T(z)]\widehat{c}_i = O(\widehat{\lambda}_i - z)^2.$$

If, moreover, $m = p = 1$, we retrieve the $2n$ conditions described in the SISO result of [ML67] :

$$H_*(\widehat{\lambda}_i) = \widehat{H}_*(\widehat{\lambda}_i), \quad \frac{d}{dz}\, H_*(z)\big|_{z=\widehat{\lambda}_i} = \frac{d}{dz}\, \widehat{H}_*(z)\Big|_{z=\widehat{\lambda}_i}, \quad i = 1, \dots, n.$$

10. Conclusion. In this paper, we have characterized the stationary points of the $\mathcal{H}_2$-norm approximation error $\|H(s) - \widehat{H}(s)\|_{\mathcal{H}_2}^2$ in the MIMO case, with the reduced-order system $\widehat{H}(s)$ in Jordan canonical form. The stationarity conditions take the form of tangential interpolation conditions—whose degree depend on the size of the Jordan blocks—written in terms of the Jordan parameters of $\widehat{H}(s)$. The conditions are thus implicit, which calls for iterative algorithms. However, we have shown that the Jordan-based approach becomes ill-conditioned in the neighborhood of target transfer functions $\widehat{H}(s)$ with higher-order poles. It is therefore more robust to use the interpolation conditions in the Sylvester equation form (Theorem 5.1) since the $\mathcal{H}_2$ norm is smooth in the parameters $(\widehat{A}, \widehat{B}, \widehat{C})$ of these equations. We have also shown that the underlying optimization problem can have several local minima by just analyzing the approximation problem by systems of McMillan degree one (with a real pole) and two (with complex conjugate poles). The case of discrete-time systems has also been considered.

REFERENCES

[Ant05] Athanasios C. Antoulas. *Approximation of large-scale dynamical systems*, volume 6 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. With a foreword by Jan C. Willems.

[AW49] P. R. Aigrain and E. M. Williams. Synthesis of n-reactance networks for desired transient response. *J. Appl. Phys.*, 20:597–600, 1949.

[BF79] Christopher I. Byrnes and Peter L. Falb. Applications of algebraic geometry in system theory. *Amer. J. Math.*, 101(2):337–363, 1979.

[BG07] Christopher A. Beattie and Serkan Gugercin. Krylov-based minimization for optimal H_2 model reduction. In *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007.

[BGR90] Joseph A. Ball, Israel Gohberg, and Leiba Rodman. *Interpolation of rational matrix functions*, volume 45 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1990.

[BKVW07] A. Bunse-Gerstner, D. Kubalińska, G. Vossen, and D. Wilczek. h_2-norm optimal model reduction for large-scale discrete dynamical MIMO systems. Technical Report 07-04, Universität Bremen, Zentrum für Technomathematik, August 2007. http://www.math.uni-bremen.de/zetem/reports/reports-liste.html.

[Che99] Chi-Tsong Chen. *Linear System Theory and Design*. Oxford University Press, New York, NY, 1999.

[GAB07] S. Gugercin, A. C. Antoulas, and C. Beattie. H_2 model reduction for large-scale linear dynamical systems. accepted for publication in SIAM J. Matrix Anal. Appl., 2007.

[Gan59] F. R. Gantmacher. *The Theory of Matrices I, II*. Chelsea, New-York, 1959.

[Gug02] Serkan Gugercin. *Projection methods for model reduction of large-scale dynamical systems*. PhD thesis, ECE Dept., Rice University, December 2002.

[GVV05] K. Gallivan, A. Vandendorpe, and P. Van Dooren. Model reduction of MIMO systems via tangential interpolation. *SIAM J. Matrix Anal. Appl.*, 26(2):328–349, 2004/05.

[GVV04] K. Gallivan, A. Vandendorpe, and P. Van Dooren. Sylvester equations and projection-based model reduction. *J. Comput. Appl. Math.*, 162(1):213–229, 2004.

[ML67] L. Meier and D. G. Luenberger. Approximation of linear constant systems. *IEEE Trans. Automatic Control*, 12:585–588, 1967.

[VGA08] P. Van Dooren, K. A. Gallivan, and P.-A. Absil. H_2-optimal model reduction of MIMO systems. *Appl. Math. Lett.*, 2008. to appear.

[Wil65] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.

[Wil70] D. A. Wilson. Optimum solution of model-reduction problem. *Proc. Inst. Elec. Eng.*, 117:1161–1165, 1970.