International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

A Tutorial on Hankel-Norm Approximations

Glover, K.

IIASA Working Paper

WP-89-079

October 1989

WORKING PAPER

A TUTORIAL ON HANKEL-NORM APPROXIMATIONS

K. Glover

October 1989
WP-89-079

IIASA
International Institute
for Applied Systems Analysis

A TUTORIAL ON HANKEL-NORM APPROXIMATIONS

K. Glover

October 1989
WP-89-079

Cambridge University Engineering Department,
Control and Management Systems Division, England.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

This is a contribution to the activity on the topic *From Data to Model* initiated at the Systems and Decision Sciences Program of IIASA by Professor J. C. Willems.

A. Kurzhanski
Program Leader
System and Decision Sciences Program.

A TUTORIAL ON HANKEL-NORM APPROXIMATION

KEITH GLOVER

Abstract

A self-contained derivation is presented of the characterization of all optimal Hankel-norm approximations to a given matrix-valued transfer function. The approach involves a state-space characterization of all-pass systems as in the author's previous work, but has been greatly simplified. A section of preliminary results is included giving general results on linear fractional transformations, Hankel operators and all-pass systems. These results then can be applied to give the characterization of all optimal Hankel-norm approximations of a given stable transfer function. Frequency response bounds for these approximations are then derived from finite rank perturbation results.

Keywords

Hankel norm, Hardy spaces, $\mathcal{H}_\infty$, model order reduction, rational approximation.

1 INTRODUCTION

An important question when modelling dynamic systems is whether a model can be simplified without undue loss of accuracy. A measure of the complexity of a linear state-space model,

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (1.1)$$
$$y(t) = Cx(t) + Du(t) \quad (1.2)$$

is the dimension, n, of its state vector $x(t)$. In (1.1)-(1.2), $u(t) \in \mathbf{C}^m$, $x(t) \in \mathbf{C}^n$, $y(t) \in \mathbf{C}^p$ for all t, and A, B, C are complex matrices of compatible dimensions. Low order models will give more efficient simulations and, for example, control system design calculations.

Approximating (1.1)-(1.2) by a reduced order system,

$$\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}u(t) \quad (1.3)$$
$$\hat{y}(t) = \hat{C}\hat{x}(t) + \hat{D}u(t) \quad (1.4)$$

where $\hat{x}(t) \in \mathbf{C}^k$, $k < n$, is termed a *model reduction problem*. Substantial progress has been made on problems of this type in recent years by the use of truncated balanced realizations as introduced by Moore (1981) and optimal Hankel-norm approximations as given in Adamjan, Arov and Krein (1971). The first method truncates states from a particular realization but has not been shown to be optimal in any sense; whereas the second method minimizes a specific norm of the error between (1.1)-(1.2) and (1.3)-(1.4). Both methods have been shown to give excellent results in many application areas. If we define the corresponding transfer functions as

$$G(s) = D + C(sI - A)^{-1}B$$
$$\hat{G}(s) = \hat{D} + \hat{C}(sI - \hat{A})^{-1}\hat{B}$$

then we might consider minimizing a variety of norms on the error system, $G(s) - \hat{G}(s)$. The induced norm corresponding to $\mathcal{L}_2$-norms on the signals is the H_∞-norm of $(G(s) - \hat{G}(s))$, denoted $\|G - \hat{G}\|_\infty$. One of the reasons for the success of the above two methods is that both have been shown to be close to optimal with respect to the H_∞-norm [see Enns (1984) and Glover (1984)].

Glover (1984) gave a characterizations of all optimal Hankel-norm approximations of a given $G(s)$ together with an upper bound on $\|G - \hat{G}\|_\infty$. The approach taken involved some lengthy calculations and it is the primary purpose of the present paper to re-derive many of these results in a self-contained but more efficient manner, hence giving greater insight into the technique and its derivation.

Background to the problem can be found in Glover (1984) and reference to more recent works, especially that of Ball and Ran (1986), can be found in Francis (1987) together with its application to H_∞-control problems. The approach to be described here was also partly presented in Glover(1987) and Glover, Curtain, and Partington (1988).

In section 2 a number of background results will be stated and for completeness most will be derived. Section 3 then considers a sub-optimal Hankel-norm approximation problem, whereas section 4 considers the optimal case. Section 5 then derives the H_∞-norm upper bounds.

The following notation will be used. For $A \in \mathbf{C}^{n\times m}$, A' denotes its complex conjugate transpose and $A^\dagger$ denotes its pseudo inverse. $\mathbf{C}_+$ and $\mathbf{C}_-$ denote the open right and left half planes respectively. $\mathcal{RH}^{p\times m}_{\infty,+}$ denotes the space of proper rational $p\times m$-matrix-valued functions of $s \in \mathbf{C}$, analytic in $\mathbf{C}_+$ (i.e. poles in $\mathbf{C}_-$). Similarly $\mathcal{RH}^{p\times m}_{\infty,-}$ functions are analytic in $\mathbf{C}_-$, and $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$ implies that $Q = G + F$ for $G \in \mathcal{RH}^{p\times m}_{\infty,+}$, $F \in \mathcal{RH}^{p\times m}_{\infty,-}$ with G of McMillan degree $\leq k$. For $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$

$$\|Q\|_\infty := \sup_\omega \bar{\sigma}(Q(j\omega))$$

where $\bar{\sigma}$ denotes the maximum singular value. State-space realizations are denoted

$$G(s) = D + C(sI - A)^{-1}B = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

and

$$G(s)^\sim = G(-\bar{s})' = \left[\begin{array}{c|c} -A' & C' \\ \hline -B' & D' \end{array}\right].$$

[Note that in contrast to Francis (1987), $\mathcal{RH}_\infty$ will include rational functions with complex coefficients.]

2 PRELIMINARIES

A number of results on the manipulation of matrices and systems will be required. Since some results are not easily accessible, proofs or outlines of proofs will also be included for tutorial reasons when appropriate.

2.1 Unitary dilation of matrices

A standard approach to representing a contraction is to imbed it in a unitary operator (a dilation of the contraction). Constant matrices will be considered first.

Lemma 2.1 *Let $D_{11} \in \mathbf{C}^{p\times m}$ satisfy $D_{11}'D_{11} \le I$ with the nullity of $I - D_{11}'D_{11} = r$. Then there exist $D_{12} \in \mathbf{C}^{p\times(p-r)}$, $D_{21} \in \mathbf{C}^{(m-r)\times m}$, $D_{22} \in \mathbf{C}^{(m-r)\times(p-r)}$ such that*

$$\begin{aligned} D_{12}D_{12}' &= I - D_{11}D_{11}', \\ D_{21}'D_{21} &= I - D_{11}'D_{11}, \\ D_{22} &= -D_{21}D_{11}'(D_{12}^L)' = -(D_{21}^R)'D_{11}'D_{12}, \quad \text{where } D_{12}^L D_{12} = I = D_{21}D_{21}^R, \end{aligned}$$

and $D := \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$ *is unitary.*

Proof. With $D_{22} := -D_{21}D_{11}'(D_{12}^L)'$ note that

$$\begin{aligned} D_{21}'D_{22} &= -D_{21}'D_{21}D_{11}'(D_{12}^L)' \\ &= (D_{11}'D_{11} - I)D_{11}'(D_{12}^L)' \\ &= -D_{11}'D_{12}D_{12}'(D_{12}^L)' = -D_{11}'D_{12} \\ D_{22}'D_{22} &= D_{12}^L D_{11}D_{21}'D_{21}D_{11}'(D_{12}^L)' \\ &= D_{12}^L D_{11}(I - D_{11}'D_{11})D_{11}'(D_{12}^L)' \\ &= D_{12}^L D_{12}D_{12}'D_{11}D_{11}'(D_{12}^L)' \\ &= I - D_{12}'D_{12} \end{aligned}$$

which, together with the definition of D_{21}, verifies that $D'D = I$. $DD' = I$ then gives the other expression for D_{22}. □

Lemma 2.2 *Let $B \in \mathbf{C}^{n\times m}$, $C \in \mathbf{C}^{p\times n}$ have rank r and satisfy $C'C = BB'$. Then there exists a unitary $D \in \mathbf{C}^{(m+p-r)\times(m+p-r)}$ where $D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & 0 \end{bmatrix}$, $D_{11} = -C'^{\dagger}B = -CB'^{\dagger}$ such that $\begin{bmatrix} C' & 0 \end{bmatrix} D + \begin{bmatrix} B & 0 \end{bmatrix} = 0$.*

Proof. Let $C = U_1\Sigma_1 V_1'$ with $U_1'U_1 = I$, $V_1'V_1 = I$, $\det\Sigma_1 \neq 0$ be the SVD of C. Define $W_1 = B'V_1\Sigma_1^{-1}$. Then

$$\begin{aligned}
W_1'W_1 &= \Sigma_1^{-1}V_1'BB'V_1\Sigma_1^{-1} = I, \\
D_{11} &= -C'^{\dagger}B = -U_1\Sigma_1^{-1}V_1'B = -U_1W_1', \\
D_{11}B' &= -U_1\Sigma_1^{-1}V_1'C'C = -U_1\Sigma_1^{-1}V_1'V_1\Sigma_1^2V_1' = -C, \\
C'D_{11}D_{11}'C &= V_1\Sigma_1^2V_1' = C'C \\
&\Rightarrow (C'D_{11} + B)(D_{11}'C + B') = 0 \Rightarrow C'D_{11} + B = 0, \\
B &= V_1\Sigma_1W_1'; \;\; B'^{\dagger} = V_1\Sigma_1^{-1}W_1', \\
D_{11} &= -CB'^{\dagger}
\end{aligned}$$

The construction is completed by choosing D_{12}, D_{21} such that $\begin{bmatrix} U_1, & D_{12} \end{bmatrix}$ and $\begin{bmatrix} W_1, & D_{21}' \end{bmatrix}$ are unitary. □

2.2 Linear fractional maps

Consider the feedback system of Fig. 1.

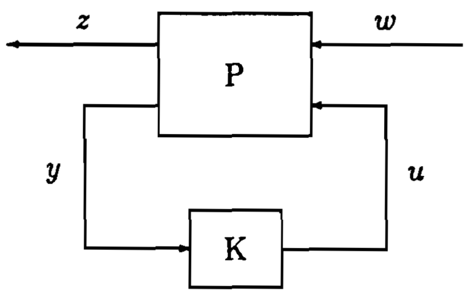

Figure 1: The linear fractional map

We will refer to the transfer function from w to z as the *linear fractional map* of K with coefficient matrix P, denoted

$$\mathcal{F}_l(P,K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}$$

where

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, \; P_{11} : p_1 \times m_1, \; P_{22} : p_2 \times m_2$$

and dimensions are compatible. Similarly feedback around the upper loop is denoted $\mathcal{F}_u(P,J) = P_{22} + P_{21}J(I - P_{11}J)^{-1}P_{12}$. Redheffer (1960) proves many results on such transformations, some of which are now given. Notice that for the feedback loop to be well-posed, the condition $\det(I - P_{22}(\infty)K(\infty)) \neq 0$ is required.

Theorem 2.3 *Let* $\det(I - P_{22}K)(\infty) \neq 0$. *Then*

(a) If $\|P\|_\infty \leq 1$, $\|K\|_\infty \leq 1$, *then*

$$\|\mathcal{F}_l(P,K)\|_\infty \leq 1.$$

(b) If $P^\sim P = I$ *and* $K^\sim K = I$ *then*

$$[\mathcal{F}_l(P,K)]^{\sim}\mathcal{F}_l(P,K) = I.$$

(c) If $PP^{\sim} = I$ and $KK^{\sim} = I$ then

$$\mathcal{F}_l(P,K)[\mathcal{F}_l(P,K)]^{\sim} = I.$$

(d) If P_{21} has generically full row rank with $P^{\sim}P = I$ and $\|K\|_\infty > 1$ then

$$\|\mathcal{F}_l(P,K)\|_\infty > 1.$$

(e) If P_{12} has generically full column rank with $PP^{\sim} = I$ and $\|K\|_\infty > 1$ then

$$\|\mathcal{F}_l(P,K)\|_\infty > 1.$$

(f) If rank $P_{21}(j\omega) = p_2 \ \forall\ \omega \in \mathbf{R} \cup \infty$ with $P^{\sim}P = I$ then $\|\mathcal{F}_l(P,K)\|_\infty < 1$ if and only if $\|K\|_\infty < 1$.

(g) If rank $P_{12}(j\omega) = m_2 \ \forall\ \omega \in \mathbf{R} \cup \infty$ with $PP^{\sim} = I$ then $\|\mathcal{F}_l(P,K)\|_\infty < 1$ if and only if $\|K\|_\infty < 1$.

Proof. Consider the system of Fig. 1, which will be well-posed by assumption, and consider the signals w, u, z, y at frequency ω.

(a) $|w|^2 + |u|^2 \geq |z|^2 + |y|^2$ since $\|P\|_\infty \leq 1$. Further, $\|K\|_\infty \leq 1 \Rightarrow |u|^2 \leq |y|^2$ and hence $|w|^2 + |u|^2 \geq |z|^2 + |u|^2$ for all ω and the result follows.

(b) $u = Ky \Rightarrow |u|^2 = y'K'Ky = |y|^2$, also $|z|^2 + |y|^2 = |w|^2 + |u|^2 = |w|^2 + |y|^2 \Rightarrow |z|^2 = |w|^2 \ \forall\ \omega, y$ and the result follows.

(c) is the dual of *(b)*

(d) Consider a frequency ω such that $\bar{\sigma}(K(j\omega)) > 1$ and $P_{21}(j\omega)$ has full row rank. Then there exists $\hat{y}$ such that $\hat{u} = K(j\omega)\hat{y}$, $|\hat{u}| > |\hat{y}|$. Now let $w = P_{21}(j\omega)^{\dagger}(\hat{y} - P_{22}(j\omega)\hat{u})$; then $y = P_{21}w + P_{22}u = P_{21}P_{21}^{\dagger}(\hat{y} - P_{22}K\hat{y}) + P_{22}Ky \Rightarrow y = \hat{y}, u = \hat{u}$, since $P_{21}P_{21}^{\dagger} = I$. Hence $|z|^2 + |y|^2 = |w|^2 + |u|^2 > |w|^2 + |y|^2 \Rightarrow |z| > |w|$ for this w and ω and the result follows.

(e) is the dual of *(d)*

(f) and (g) Follow in the same way as (d).

□

It is therefore seen that if P is an all-pass system then the feedback system will have norm strictly less than unity if and only if the feedback term satisfies $\|K\|_\infty < 1$. The location of the closed-loop poles can also be deduced as follows:

Lemma 2.4 *Let $P \in \mathcal{RH}_{\infty,-(k)}$ and $K \in \mathcal{RH}_{\infty,-(\ell)}$ satisfy $\|P_{22}K\|_\infty < 1$. Then*

$$\mathcal{F}_l(P,K) \in \mathcal{RH}_{\infty,-(k+\ell)}.$$

Proof. A proof is given in Glover *et al.* (1988) and just observes that the open-loop poles move continuously to the closed-loop poles as the feedback gain is increased, but cannot cross the imaginary axis due to the condition on $P_{22}K$. □

Furthermore the location of any cancellations in the feedback can be examined as follows [a similar result is in Limebeer and Hung (1987)].

Lemma 2.5 *Let P have the state-space realization*

$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right],$$

where rank $D_{12} = m_2$, rank $D_{21} = p_2$, $B_2 = B_{20}D_{12}$, $C_2 = D_{21}C_{20}$, and let K have a minimal realization. Then

(a) All unobservable modes of the natural realization of $\mathcal{F}_l(P,K)$ are contained in $\lambda_i(A - B_{20}C_1)$.

(b) All uncontrollable modes of the natural realization of $\mathcal{F}_l(P,K)$ are contained in $\lambda_i(A - B_1C_{20})$.

The natural realization of $\mathcal{F}_l(P,K)$ refers to the feedback connection of the realizations for P and K.

Proof. Let K have the minimal realization $K = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right]$. Then the state-space equations for the closed loop are:

$$\begin{aligned} \mathcal{F}_l(P,K) &= \left[\begin{array}{cc|c} A + B_2\hat{D}L_1C_2 & B_2L_2\hat{C} & B_1 + B_2\hat{D}L_1D_{21} \\ \hat{B}L_1C_2 & \hat{A} + \hat{B}L_1D_{22}\hat{C} & \hat{B}L_1D_{21} \\ \hline C_1 + D_{12}L_2\hat{D}C_2 & D_{12}L_2\hat{C} & D_{11} + D_{12}\hat{D}L_1D_{21} \end{array}\right] \\ &=: \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & D_c \end{array}\right], \end{aligned}$$

where $L_1 := (I - D_{22}\hat{D})^{-1}$, $L_2 := (I - \hat{D}D_{22})^{-1}$.

Suppose $\mathcal{F}_l(P,K)$ has unobservable state $(x', y')'$ and mode λ; then the $P-B-H$ test [Kailath (1980)] gives

$$\begin{aligned} & \begin{bmatrix} A_c - \lambda I \\ C_c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \\ \Rightarrow\ & \begin{bmatrix} A - \lambda I & B_2 \\ C_1 & D_{12} \end{bmatrix} \begin{bmatrix} x \\ \hat{D}L_1C_2x + L_2\hat{C}y \end{bmatrix} = 0 \\ \Rightarrow\ & \begin{bmatrix} A - \lambda I & B_{20} \\ C_1 & I \end{bmatrix} \begin{bmatrix} x \\ D_{12}\hat{D}L_1C_2x + D_{12}L_2\hat{C}y \end{bmatrix} = 0 \\ \Rightarrow\ & (A - B_{20}C_1 - \lambda I)x = 0 \end{aligned}$$

If $x = 0$ then $\hat{C}y = 0$ and $\hat{A}y = 0$ which contradicts $(\hat{A}, \hat{C})$ being completely observable. Hence $x \neq 0$ and $\lambda \in \lambda_i(A - B_{20}C_1)$ and part *(a)* is proven. Part *(b)* is a dual result. □

The following corollary is now an immediate consequence of Lemmas 2.4 and 2.5.

Corollary 2.6 *Let $P \in \mathcal{RH}_{\infty,-(k)}$, $P \notin \mathcal{RH}_{\infty,-(k-1)}$, have a state-space realization as in Lemma 2.5 with*

$$\begin{aligned} \operatorname{Re}\lambda_i(A - B_{20}C_1) &\geq 0 \\ \operatorname{Re}\lambda_i(A - C_{20}B_1) &\geq 0 \end{aligned}$$

and let $K \in \mathcal{RH}_{\infty,-(\ell)}$, $K \notin \mathcal{RH}_{\infty,-(\ell-1)}$, and $\|KP_{22}\|_\infty < 1$. Then

$$\begin{aligned} \mathcal{F}_l(P,K) &\in \mathcal{RH}_{\infty,-(k+\ell)} \\ \mathcal{F}_l(P,K) &\notin \mathcal{RH}_{\infty,-(k+\ell-1)}. \end{aligned}$$

The following lemma concerns the inversion of a linear fractional map.

Lemma 2.7 *Let P and K be rational transfer function matrices, and let $G = \mathcal{F}_l(P,K)$. Then*

(a) If P and K are proper with $\det(I - P_{22}K)(\infty) \neq 0$ then G is proper.

(b) If P_{12} and P_{21} have generically full column and row rank respectively, then $\mathcal{F}_l(P,K) = \mathcal{F}_l(P,K_2)$ implies that $K_1 = K_2$.

(c) If P and G are proper, $\det P(\infty) \neq 0$, $\det\left(P + \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix}\right)(\infty) \neq 0$ and P_{12} and P_{21} are square and invertible for almost all s, then K is proper and

$$K = \mathcal{F}_u(P^{-1}, G)$$

Proof.

(a) is immediate from the definition of $\mathcal{F}_l(P,K)$.

(b) follows from the identity

$$\mathcal{F}_l(P,K_1) - \mathcal{F}_l(P,K_2) = P_{12}(I - K_2P_{22})^{-1}(K_1 - K_2)(I - P_{22}K_1)^{-1}P_{21}$$

(c) Let $Q = P^{-1}$, which will be proper since $\det P(\infty) \neq 0$, and define

$$\begin{aligned} K &= \mathcal{F}_u(Q,G) = Q_{22} + Q_{21}G(I - Q_{11}G)^{-1}Q_{12} \\ &= [Q_{22}Q_{12}^{-1}(I - Q_{11}G) + Q_{21}G](I - Q_{11}G)^{-1}Q_{12} \\ &= P_{12}^{-1}(G - P_{11})(I - Q_{11}G)^{-1}Q_{12} \end{aligned}$$

This expression is well-posed and proper since at $s = \infty$

$$\begin{aligned} \det(I - Q_{11}G) &= \det\left(I - \begin{bmatrix} I & 0 \end{bmatrix} P^{-1} \begin{bmatrix} I & 0 \end{bmatrix}' G\right) \\ &= \det\left[P^{-1}\left(P - \begin{bmatrix} I & 0 \end{bmatrix}' G \begin{bmatrix} I & 0 \end{bmatrix}\right)\right] \\ &\neq 0. \end{aligned}$$

We also need to ensure that $\mathcal{F}_l(P,K)$ is well-posed:

$$\begin{aligned} I - P_{22}K &= (I - P_{22}Q_{22}) - P_{22}Q_{21}G(I - Q_{11}G)^{-1}Q_{12} \\ &= P_{21}Q_{12} + P_{21}Q_{11}G(I - Q_{11}G)^{-1}Q_{12} \\ &= P_{21}(I - Q_{11}G)^{-1}Q_{12} \end{aligned}$$

and $\det(I - P_{22}K) \neq 0$ since P_{21}^{-1} exists and $Q_{12}^{-1} = P_{12} - P_{11}P_{21}^{-1}P_{22}$. Hence the LFT are both well-posed and we immediately obtain that $\mathcal{F}_l(P,K) = G$ as required on substituting for K and $(I - P_{22}K)$ as above. □

Remark 2.1 The proof of part *(c)* was primarily to show that the feedback systems were well-posed. A simple interpretation of the result is given by considering the signals in the feedback systems, assuming they are well-posed, as follows:

$$\begin{aligned} \begin{bmatrix} z \\ y \end{bmatrix} &= P\begin{bmatrix} w \\ u \end{bmatrix}, \quad u = Ky \\ z &= \mathcal{F}_l(P,K)w = Gw \end{aligned}$$

hence

$$\begin{aligned} \begin{bmatrix} w \\ u \end{bmatrix} &= P^{-1}\begin{bmatrix} z \\ y \end{bmatrix}, \quad z = Gw \\ u &= \mathcal{F}_u(P^{-1},G)w \\ \Rightarrow K &= \mathcal{F}_u(P^{-1},G) \end{aligned}$$

2.3 Hankel Operators

It is now well-known that Hankel operators play an important role in model reduction and H_∞ design [see Francis (1987), Glover (1984), and the references therein]. General results on Hankel operators, particularly for the infinite rank case, can be found in the books by Power (1982) and Partington (1988).

Let the Hankel operator, Γ, corresponding to the stable system $G(s) = C(sI - A)^{-1}B$ be defined as

$$\Gamma_G : L^2(0,\infty) \to L^2(0,\infty): \quad u \to \int_0^\infty h(t+\tau)u(\tau)\,d\tau$$

where $h(t) = Ce^{At}B$.

The rank of Γ_G is the McMillan degree of G and it will have a singular value decomposition

$$\Gamma_G(u) = \sum_{i=1}^{n} \sigma_i \langle u, v_i \rangle w_i$$

where the σ_i are the ordered (Hankel) singular values, also denoted $\sigma_i(G)$, and (v_i, w_i) the corresponding Schmidt pairs. Let the controllability and observability Gramians X, Y be given by the unique solutions to the Lyapunov equations,

$$\begin{aligned} AX + XA' + BB' &= 0 \\ A'Y + YA + C'C &= 0. \end{aligned}$$

Further, let $XYx_i = \sigma_i^2 x_i$, $x_i'Yx_i = 1$. Then it is easily verified [see Glover (1984)] that

$$\begin{aligned} v_i(t) &= B'\exp(A't)Yx_i\sigma_i^{-1} \\ w_i(t) &= C\exp(At)x_i. \end{aligned}$$

Now let us consider the Hankel-norm approximation, that is, approximating Γ_G by $\Gamma_{\hat{G}}$ of rank $k < n$. The main result of Adamjan, Arov and Krein (1971) is that $\inf \|\Gamma_G - \Gamma_{\hat{G}}\| = \sigma_{k+1}(G)$, and the derivation in Glover (1984) to derive all solutions to this problem is quite involved. Sections 3 and 4 will give a much more economical derivation based on the results in this section. The present derivation will, however, still be based on the central all-pass construction in Glover (1984).

First, a general result on approximating operators (not necessarily Hankel operators) is given.

Lemma 2.8 *Let $\Gamma : X_1 \rightarrow X_2$ be an operator on the Hilbert spaces X_1 and X_2 with Schmidt vectors (v_i, w_i):*

$$\begin{aligned} \Gamma v_i &= \sigma_i w_i, \quad i = 1, 2, \ldots \\ \Gamma^* w_i &= \sigma_i v_i, \quad i = 1, 2, \ldots \end{aligned}$$

where

$$\sigma_i \geq \sigma_{i+1}, |w_i| = |v_i| = 1 \quad \textit{for all } i$$

If $\hat{\Gamma} : X_1 \rightarrow X_2$ is of rank k then $\|\Gamma - \hat{\Gamma}\| \geq \sigma_{k+1}$. Further if $\sigma_k > \sigma_{k+1}$ and $\|\Gamma - \hat{\Gamma}\| = \sigma_{k+1}$ then $(\Gamma - \hat{\Gamma})v_{k+1} = \sigma_{k+1} w_{k+1}$.

Proof. The proof is taken from Partington (1988, Theorem 6.14). Let P be the projection from X_2 onto span$(w_1, w_2, ..w_{k+1})$; then

$$\|P(\Gamma - \hat{\Gamma})\| \leq \|\Gamma - \hat{\Gamma}\|.$$

Consider the following restriction of $P\hat{\Gamma}$:

$$P\hat{\Gamma} : \text{lin span } (v_1, .., v_{k+1}) \rightarrow \text{lin span } (w_1, ..w_{k+1})$$

which has rank $\leq k$ and hence there exists $x \in \ker(P\hat{\Gamma})$, $\|x\| = 1$ say $x = \sum_{i=1}^{k+1} a_i v_i$ with $\sum_{i=1}^{k+1} a_i^2 = 1$.

$$\begin{aligned} P\Gamma(x) &= \sum_{i=1}^{k+1} a_i \sigma_i w_i \\ \|\Gamma - \hat{\Gamma}\|^2 &\geq \|P\Gamma(x) - P\hat{\Gamma}(x)\|^2 = \|P\Gamma(x)\|^2 = \sum_{i=1}^{k+1} \sigma_i^2 a_i^2 \geq \sigma_{k+1}^2 \end{aligned}$$

Further if $\|\Gamma - \hat{\Gamma}\| = \sigma_{k+1}$ and $\sigma_k > \sigma_{k+1}$, then

$$a_1 = a_2 = \ldots = a_k = 0, \qquad |a_{k+1}| = 1 \;\Rightarrow\; x = a_{k+1} v_{k+1}$$

Also since $\|\Gamma x - \hat{\Gamma} x\| \leq \sigma_{k+1}$ and $\Gamma x = \Gamma v_{k+1} = \sigma_{k+1} w_{k+1}$, $\langle w_{k+1}, \hat{\Gamma} v_{k+1} \rangle = 0$. Then $\hat{\Gamma} v_{k+1} = 0$ and the result follows. □

Specialising this result to Hankel operators and interpreting it in the frequency domain gives the following result [see Francis (1987, page 71) for the $k = 0$ case].

Lemma 2.9 *Let the Hankel operator Γ_G have Schmidt pairs as above with $\sigma_k > \sigma_{k+1}$. Let $Q \in \mathcal{RH}_{\infty,-(k)}$ be such that*

$$\|G + Q\|_\infty = \sigma_{k+1} \;;$$

then

$$\begin{aligned} (G + Q)V(-s) &= \sigma_{k+1} W(s) \\ W^\sim (G + Q) &= \sigma_{k+1} V^T(s) \end{aligned}$$

where

$$\begin{aligned} V(s) &= \text{Laplace transform of } v_{k+1}(t) \in H_2(\text{rhp}) \\ W(s) &= \text{Laplace transform of } w_{k+1}(t) \in H_2(\text{rhp}). \end{aligned}$$

Note that for $XYx_i = \sigma_i^2 x_i$,

$$\begin{aligned} V(-s) &= B'(-sI - A')^{-1} Y x_{k+1} \sigma_{k+1}^{-1} \\ W(s) &= C(sI - A)^{-1} x_{k+1} \end{aligned}$$

Proof. Let $Q = -\hat{G} + F$ with $\hat{G}$ rational of McMillan degree k and $F \in \mathcal{RH}_{\infty,-}$. $\|G - \hat{G} + F\|_\infty = \sigma_{k+1}$ implies that $\|\Gamma_G - \Gamma_{\hat{G}}\| \le \sigma_{k+1}$ and hence by Lemma 2.8,

$$(\Gamma_G - \Gamma_{\hat{G}})\, v_{k+1} = \sigma_{k+1} w_{k+1},$$

and recalling that the Hankel operator, Γ_G, is equivalent to a Toeplitz operator with symbol, G, followed by a projection [Francis (1987)] we have in the frequency domain that

$$(G(s) - \hat{G}(s))V(-s) = \sigma_{k+1} W(s) + U(-s)$$

where

$$U(s), V(s), W(s) \in H_2, \quad \|V\|_2 = \|W\|_2 = 1.$$

Hence

$$(G + Q)V(-s) = \sigma_{k+1} W(s) + U(-s) + F(s)V(-s)$$

and $\|G + Q\|_\infty = \sigma_{k+1}$ implies

$$\begin{aligned}
\|(G+Q)V(-s)\|_2 &\le \sigma_{k+1} \\
U(-s) + F(s)V(-s) &\in H_2^{\perp} \text{ implies that} \\
\|(G+Q)V(-s)\|_2^2 &= \sigma_{k+1}^2 \|W\|_2^2 + \|U(-s) + F(s)V(-s)\|_2^2 \\
&= \sigma_{k+1}^2 + \|U(-s) + F(s)V(-s)\|_2^2 \\
&\le \sigma_{k+1}^2
\end{aligned}$$

Therefore, $U(-s) + F(s)V(-s) = 0$ and the result follows. Similarly for the dual result. □

Note that in the case when G is scalar that Lemma 2.9 implies that

$$G + Q = \sigma_{k+1} W(s)/V(-s)$$

and the difficulty is to demonstrate that $Q \in H^{\infty}_{-(k)}$ [Adamjan *et al.* (1971)].

2.4 All-pass systems

The approach taken in Glover (1984) to optimal Hankel norm approximation is to construct an augmented all-pass error system, and then to connect a contraction around the augmented system to generate all solutions. A characterization of all-pass systems is given in Glover (1984, Theorem 5.1) and is now re-stated.

Lemma 2.10

(a) Let $G(s) = D + C(sI - A)^{-1}B$ be a minimal realization. Then $GG^{\sim} = G^{\sim}G = I$ if and only if $\exists\ X = X'$, $Y = Y'$ such that

(i) $XY = I$

(ii) $DD' = I$

(iii) $AX + XA' + BB' = 0$

(iv) $DB' + CX = 0$

(v) $D'D = I$

(vi) $A'Y + YA + C'C = 0$

(vii) $D'C + B'Y = 0$

(b) Conditions (ii) - (iv) above imply $GG^{\sim} = I$

(c) Conditions (v) - (vii) above imply $G^{\sim}G = I$

Note that stability is not assumed and parts *(b)* and *(c)* do not need minimality.

An all-pass dilation of transfer functions can be obtained as follows and entirely analogously to Lemma 2.1.

Lemma 2.11 *Let* $\|G_{11}\|_\infty \leq 1$ *then defining*

$$G_{12}: \quad G_{12}G_{12}^{\sim} = I - G_{11}G_{11}^{\sim}$$
$$G_{21}: \quad G_{21}^{\sim}G_{21} = I - G_{11}^{\sim}G_{11}$$

where G_{12} *and* $G_{21}^{\sim}$ *are of generically full column rank. Then*

$$G_{22} = -G_{21}G_{11}^{\sim}(G_{12}^{L})^{\sim} = -(G_{21}^{R})^{\sim}G_{11}^{\sim}G_{12}$$

makes

$$G := \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \text{all-pass}$$

Proof. The proof is identical to Lemma 2.1 except that we take a generic point on $s = j\omega$. This then gives $G^{\sim}G = I$ for almost all $s = j\omega$ and hence for all s. □

2.5 Alternative Linear Fractional Transformations

An alternative approach to many of the results stated in this section is via coprime factorizations over $\mathcal{RH}_{\infty,-}$ (see Vidyasagar(1985)), although it is usual to consider factorizations over $\mathcal{RH}_{\infty,+}$ in control problems. A *right coprime factorization* of G over $\mathcal{RH}_{\infty,-}$ is given by $G = NM^{-1}$ where $N, M \in \mathcal{RH}_{\infty,-}$ and there exist $X, Y \in \mathcal{RH}_{\infty,-}$ such that the following right Bezout identity or right Diophantine identity is satisfied:

$$XN + YM = I$$

If $G \in \mathcal{RH}_{\infty,-(k)}$, $G \notin \mathcal{RH}_{\infty,-(k-1)}$ with $G = NM^{-1}$ as above, then $\det M(s)$ will have precisely k zeros (including multiplicities) in $\mathbf{C}_-$, or equivalently, since M has no poles in $\mathbf{C}_-$ the principle of the argument gives that the winding number of $\det M(s)$ about the origin, as s traverses the Nyquist D contour, is equal to k, (see Vidyasagar(1985) for more details). Hence the McMillan degree of the stable part of G can be determined.

When P_{21} is invertible for almost all s then the following alternative form of the linear fractional transformation can be used:

$$\begin{aligned} \mathcal{F}_l(P,K) &= T_\Theta(K) \\ &:= (\theta_{11}K + \theta_{12})(\theta_{21}K + \theta_{22})^{-1} \\ &= (\theta_{11}U + \theta_{12}V)(\theta_{21}U + \theta_{22}V)^{-1} \end{aligned}$$

where $K = UV^{-1}$ is a right coprime factorization over $\mathcal{RH}_{\infty,-}$. It is straightforward to verify that P and Θ are related as follows:

$$\begin{aligned}\Theta &= \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix} \\ &= \begin{bmatrix} P_{12} - P_{11}P_{21}^{-1}P_{22} & P_{11}P_{21}^{-1} \\ -P_{21}^{-1}P_{22} & P_{21}^{-1} \end{bmatrix} \\ &= \left[\begin{array}{c|cc} A - B_1D_{21}^{-1}C_2 & B_2 - B_1D_{21}^{-1}D_{22} & B_1D_{21}^{-1} \\ \hline C_1 - D_{11}D_{21}^{-1}C_2 & D_{12} - D_{11}D_{21}^{-1}D_{22} & D_{11}D_{21}^{-1} \\ -D_{21}^{-1}C_2 & -D_{21}^{-1}D_{22} & D_{21}^{-1} \end{array}\right].\end{aligned}$$

This representation is used extensively in the literature and the monographs of Dym(1989) and Helton(1987) contain a wealth of results in this area.

Now let us consider Corollary 2.6 in this framework. The assumptions that $B_2 = B_{20}D_{12}$ and $C_2 = D_{21}C_{20}$ imply that,

$$G = \mathcal{F}_l(P, K) = T_\Theta(D_{12}KD_{21})$$

where,

$$\Theta = \left[\begin{array}{c|cc} A - B_1C_{20} & B_{20} - B_1D_{220} & B_1 \\ \hline C_1 - D_{11}C_{20} & I - D_{11}D_{220} & D_{11} \\ -C_{20} & -D_{220} & I \end{array}\right]$$

and

$$D_{220} := D_{21}^{\dagger}D_{22}D_{12}^{\dagger} \Rightarrow D_{22} = D_{21}D_{220}D_{12}$$

and it is easy to verify that,

$$\Theta^{-1} = \left[\begin{array}{c|cc} A - B_{20}C_1 & B_{20} & B_1 - B_{20}D_{11} \\ \hline -C_1 & I & -D_{11} \\ C_{20} + D_{220}C_1 & D_{220} & I - D_{220}D_{11} \end{array}\right].$$

The assumptions that Re $\lambda_i(A - B_1C_{20}) > 0$ and Re $\lambda_i(A - B_{20}C_1) > 0$ imply that Θ, $\Theta^{-1} \in \mathcal{RH}_{\infty,-}$, which are the fundamental assumptions being made. Now let $D_{12}KD_{21}$ have right coprime factorization UV^{-1} with $XU+YV = I$ and $U, V, X, Y \in \mathcal{RH}_{\infty,-}$, then

$$G = (\theta_{11}U + \theta_{12}V)(\theta_{21}U + \theta_{22}V)^{-1}$$

is a right coprime factorization of G since

$$\begin{bmatrix} X & Y \end{bmatrix}\Theta^{-1}\begin{bmatrix} \theta_{11}U + \theta_{12}V \\ \theta_{21}U + \theta_{22}V \end{bmatrix} = I.$$

The above winding number result, together with the identities,

$$\begin{aligned}\det(\theta_{21}U + \theta_{22}V) &= \det(\theta_{22})\det(V)\det(I + \theta_{22}^{-1}\theta_{21}UV^{-1}) \\ &= \det(\theta_{22})\det(V)\det(I - P_{22}K) \\ \det(\theta_{22}) &= \frac{\det(sI - A)}{\det(sI - A + B_1C_{20})}\end{aligned}$$

and $\|P_{22}K\|_\infty < 1$ gives that the number of poles in $\mathbf{C}_-$ for G is precisely $k + \ell$.

Results analagous to Theorem 2.3 involve so so-called J-inner functions ($J := \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$). Θ is J-inner if $\Theta^{\sim}J\Theta = J$ for $s = j\omega$, with $\Theta \in \mathcal{RH}_{\infty,-}$, (again $\mathcal{RH}_{\infty,+}$ is more commonly used). For such Θ then T_Θ maps the unit ball onto the unit ball. This representation is more natural in, for example, the work of Ball and Ran (1986) on Hankel-norm approximation and Green *et al.* (1988) on a generalization to the Nehari problem.

3 SUB-OPTIMAL HANKEL-NORM APPROXIMATIONS

In this section the problem of approximating a Hankel operator Γ_G of rank n by $\Gamma_{\hat{G}}$ of rank $k < n$ will be considered. Indeed, all solutions $\Gamma_{\hat{G}}$ to the problem

$$\|G - \hat{G}\|_H = \|\Gamma_G - \Gamma_{\hat{G}}\| < \sigma \tag{3.5}$$

for some σ will be solved. It will be shown that this problem is equivalent to finding all $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$ such that

$$\|G + Q\|_\infty < \sigma. \tag{3.6}$$

This equivalence is a consequence of a theorem due to Nehari (1957) for which we will give an independent derivation.

3.1 All-pass dilations

Firstly note that for $G \in \mathcal{RH}^{p\times m}_{\infty,+}$ of degree n and $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$, $Q = -F - \hat{G}$ for $F \in \mathcal{RH}^{p\times m}_{\infty,-}$, $\hat{G} \in \mathcal{RH}^{p\times m}_{\infty,+}$ we have

$$\|G + Q\|_\infty \geq \|G - \hat{G}\|_H = \|\Gamma_G - \Gamma_{\hat{G}}\| \geq \sigma_{k+1}(G). \tag{3.7}$$

The first inequality is standard since the Hankel operator is a restriction of the convolution operator [see for example Francis (1987) or Glover (1984, Lemma 6.2)]. The final inequality follows from Lemma 2.8 since $\text{rank}(\Gamma_{\hat{G}}) \leq k$. Hence in order for (3.5) to have a solution, $\sigma > \sigma_{k+1}$ is required. Further it will be assumed that $\sigma < \sigma_k$ and without loss of generality that $\sigma = 1$ (a scaling of G can achieve this). That is,

$$\sigma_k > 1 > \sigma_{k+1}.$$

We will now construct $J \in H^{(p+m)\times(m+p)}_{\infty,-(k)}$ such that $G_a + J$ is all-pass, where

$$G_a = \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix} = \left[\begin{array}{c|cc} A & B & 0 \\ \hline C & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} A & B_a \\ \hline C_a & 0 \end{array}\right] \tag{3.8}$$

$$J = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & D_e \end{array}\right] = \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1 & \hat{B}_2 \\ \hline \hat{C}_1 & D_{11} & D_{12} \\ \hat{C}_2 & D_{21} & D_{22} \end{array}\right] \tag{3.9}$$

$$E := G_a + J = \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & D_e \end{array}\right] = \left[\begin{array}{cc|c} A & 0 & B_a \\ 0 & \hat{A} & \hat{B} \\ \hline C_a & \hat{C} & D_e \end{array}\right] \tag{3.10}$$

Now from Lemma 2.10, E will be all-pass if there exists $X_e = X_e'$ such that

$$A_eX_e + X_eA_e' + B_eB_e' = 0 \tag{3.11}$$

$$D_eD_e' = I \tag{3.12}$$

$$D_eB_e' + C_eX_e = 0 \tag{3.13}$$

Now let X and Y be the controllability and observability Gramians of G satisfying

$$AX + XA' + BB' = 0 \tag{3.14}$$

$$A'Y + YA + C'C = 0 \tag{3.15}$$

so that $\sigma_i^2(G) = \lambda_i(XY)$. The (1,1) block of (3.11), bearing in mind the form of A_e in (3.10), gives that $\begin{bmatrix} I & O \end{bmatrix} X_e \begin{bmatrix} I & O \end{bmatrix}' = X$. Further $X_e^{-1}(3.11)X_e^{-1}$ and (3.13) give

$$X_e^{-1}A_e + A_e'X_e^{-1} + C_e'C_e = 0 \tag{3.16}$$

and hence $\begin{bmatrix} I & O \end{bmatrix} X_e^{-1} \begin{bmatrix} I & O \end{bmatrix}' = Y$. Let us now postulate a form for X_e, given by

$$X_e = \begin{bmatrix} X & I \\ I & YZ^{-1} \end{bmatrix}; \quad X_e^{-1} = \begin{bmatrix} Y & -Z' \\ -Z & ZX \end{bmatrix} \tag{3.17}$$

where $Z := XY - I$.

Although this form for X_e is apparently taken 'out of the air', its form is fixed once the dimension of $\hat{A}$ is chosen to be that of A and the (1,2) block of X_e is assumed to be nonsingular (which is then transformed to the identity by a similarity transformation on the realization of J). Lemma 8.2 in Glover (1984) in fact generates all possible X_e but the present approach does not require this. All that is required is the particular candidate solution in (3.17). Now let us solve for $\hat{A}$, $\hat{B}$, and $\hat{C}$ given some unitary D_e.

$\hat{C}$ is obtained from the (1,1) block of (3.13); $\hat{B}$ from the (1,1) block of (3.13) $\times X_e^{-1}$, $\hat{A}$ from the (2,1) block of (3.11).

$$\begin{aligned} \hat{C} &= -C_a X - D_e B_a' && (3.18) \\ \hat{B} &= Z'^{-1}(YB_a + C_a' D_e) && (3.19) \\ \hat{A} &= -A' - \hat{B}B_a' && (3.20) \\ &= -Z'^{-1}A'Z' + Z'^{-1}C_a'\hat{C} && (3.21) \end{aligned}$$

(3.21) is obtained from the (1,2) block of (3.16) and will be valid once (3.11) and (3.13) are verified.

(3.18) and (3.19) give that

$$(3.13) \times \begin{bmatrix} I & Y \\ O & -Z \end{bmatrix} = 0 \;\Rightarrow\; (3.13) = 0.$$

(3.14) and (3.19) give (3.11) $\times \begin{bmatrix} I \\ O \end{bmatrix} = 0$, and (3.15) gives that

$$\begin{bmatrix} I & O \end{bmatrix} X_e^{-1}(3.11)X_e^{-1} \begin{bmatrix} I & O \end{bmatrix}' = 0$$

and hence

$$\begin{bmatrix} I & O \\ Y & -Z' \end{bmatrix} (3.11) \begin{bmatrix} I & Y \\ O & -Z \end{bmatrix} = 0$$

which implies that (3.11) is satisfied. Hence the required all-pass equations are satisfied, and given X_e there are precisely the correct number of equations to generate $\hat{A}$, $\hat{B}$, and $\hat{C}$. Furthermore $\hat{A}$ will have $\leq k$ eigenvalues in the open left half plane since YZ'^{-1} has k positive eigenvalues,

$$\hat{A}YZ^{-1} + Z'^{-1}Y\hat{A}' + \hat{B}\hat{B}' = 0 \tag{3.22}$$

and by Theorem 3.3(2) in Glover (1984). A final property of J that will be required in Theorem 3.2 to characterize all solutions is that, for D_{12} and D_{21} invertible,

$$\hat{A} - \hat{B}_1 D_{21}^{-1}\hat{C}_2 = -A' - \hat{B}_1 B' - \hat{B}_1 D_{21}^{-1}(-D_{21}B') = -A' \tag{3.23}$$

from (3.20) and (3.18). Similarly (3.21) and (3.19) give

$$\hat{A} - \hat{B}_2 D_{12}^{-1}\hat{C}_1 = -Z'^{-1}A'Z' + Z'^{-1}C'\hat{C}_1 + Z'^{-1}C'D_{12}D_{12}^{-1}\hat{C}_1 = -Z'^{-1}A'Z' \tag{3.24}$$

The following theorem can now be stated:

Theorem 3.1 *Given $G \in \mathcal{RH}^{p\times m}_{\infty,+}$ defined by (3.8) then:*

(a) There exists $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$ such that $\|G+Q\|_\infty < 1$ iff $\sigma_{k+1}(G) = \lambda^{1/2}_{k+1}(XY) < 1$, where X and Y are given by (3.14) and (3.15).

(b) If $\sigma_k(G) > 1 > \sigma_{k+1}(G)$ then J defined by (3.9), (3.14)–(3.20) satisfies $J \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$.

Proof. If $\|G+Q\|_\infty < 1$ then (3.7) implies that $\sigma_{k+1}(G) < 1$. Conversely, if $\sigma_{k+1}(G) < 1 < \sigma_k(G)$ then the construction of J has been shown to yield $J \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$, with $(G_a + J)$ all-pass. Furthermore $J_{12}(j\omega)$ is full rank for all ω (including ∞) since J_{12}^{-1} has 'A-matrix' $(\hat{A} - \hat{B}_2 D_{12}^{-1}\hat{C}_1) = -A'$ by (3.23) and hence J_{12} has no zeros on the imaginary axis since A is stable. Hence $\|G - J_{11}\|_\infty < 1$. If $\sigma_i > 1 > \sigma_{i+1} = \sigma_k = \sigma_{k+1}$ for some $i < k$ then the same construction can be used with k replaced by i, again giving J_{11} as a suitable Q. □

3.2 Characterization of all solutions

Once the all-pass dilation of Theorem 3.1 has been constructed, the results of section 2 can be applied to show that all solutions are characterized as follows.

Theorem 3.2 *Given $G \in \mathcal{RH}^{p\times m}_{\infty,+}$ defined by (3.8) with $\sigma_k(G) > 1 > \sigma_{k+1}(G)$, then all $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$ such that*

$$\|G+Q\|_\infty < 1 \tag{3.25}$$

are given by

$$Q = \mathcal{F}_l(J, \Phi), \quad \Phi \in \mathcal{RH}^{p\times m}_{\infty,-}, \|\Phi\|_\infty < 1. \tag{3.26}$$

where J is defined in (3.9), (3.14)–(3.20) with D_{12} and D_{21} invertible.

Proof. Let $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$ be such that (3.25) holds. Then (3.26) has a solution for some rational proper Φ by Lemma 2.7 on noting that

$$\det D_e' \det\left(D_e + \begin{bmatrix} G(\infty) & 0 \\ 0 & 0 \end{bmatrix}\right) = \det \begin{bmatrix} I + D_{11}'G(\infty) & 0 \\ D_{12}'G(\infty) & I \end{bmatrix} \neq 0$$

since $\bar{\sigma}(D_{11}'G(\infty)) < 1$. Furthermore, (3.25) and (3.26) imply that

$$G + Q = \mathcal{F}_l(J + G_a, \Phi)$$

with $\|G+Q\|_\infty < 1$ and $J + G_a$ all-pass. Hence Theorem 2.3 implies that $\|\Phi\|_\infty < 1$. Finally Corollary 2.6 can be applied to $Q = \mathcal{F}_l(J,\Phi)$ to give that $\Phi \in \mathcal{RH}_{\infty,-(0)}$ since $Q \in \mathcal{RH}_{\infty,-(k)}$, $J \in \mathcal{RH}_{\infty,-(k)}$ and $J \notin H_{\infty,-(k-1)}$ (since $\|G + J_{11}\|_\infty < 1 < \sigma_{k-1}$, and the realization of J satisfies (3.23) and (3.24)). □

4 OPTIMAL HANKEL-NORM APPROXIMATIONS

In the limit as $\sigma_{k+1}(G) \to 1$ the characterization of all solutions in Theorem 3.2 becomes degererate because the term $Z = (XY - I)$ becomes singular. It is possible to rewrite the equations for J in descriptor form as in Safonov *et al.* (1987), and this will show that the optimal solutions are no longer strictly proper. The characterization of all-pass systems can also be done for descriptor systems and this approach is taken in Glover *et al* (1989) for an $\mathcal{H}_\infty$ control problem. To characterize all optimal solutions we will exploit the constraint given by Lemma 2.8 on all $(G+Q)$ such that $\|G+Q\|_\infty = \sigma_{k+1}$, where $Q \in \mathcal{RH}_{\infty,-(k)}$, and involving the Schmidt vectors of Γ_G. Suppose that σ_{k+1} has multiplicity r and that $\sigma_{k+1} = 1$.

Let the corresponding controllability and observability Gramians be

$$X = \begin{bmatrix} I_r & 0 \\ 0 & X_2 \end{bmatrix}, \quad Y = \begin{bmatrix} I_r & 0 \\ 0 & Y_2 \end{bmatrix} \tag{4.1}$$

after a suitable change of state coordinates, with

$$\{\lambda_i(X_2 Y_2)\} = \left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_{k-1}^2, \sigma_{k+r+1}^2, \ldots, \sigma_n^2\right\}$$

The Laplace transforms of the Schmidt vectors of Γ_e corresponding to σ_{k+1} are then

$$\begin{aligned} W_i(s) &= C(sI-A)^{-1}e_i, \quad i = 1,2,\ldots,r \\ V_i(-s) &= B'(-sI-A')^{-1}e_i \quad i = 1,2,\ldots r \end{aligned}$$

where e_i are the standard basis vectors. Hence from Lemma 2.9 if $\|G+Q_i\| = \sigma_{k+1}$ for $Q_i \in \mathcal{RH}_{\infty,-(k)}$ and $i = 1,2$, then for $W := [W_1, W_2, \ldots, W_r]$, $V := [V_1, \ldots, V_r]$,

$$W^\sim(Q_1 - Q_2) = 0 \tag{4.2}$$

$$(Q_1 - Q_2)V(-s) = 0 \tag{4.3}$$

In order to characterize all optimal solutions, suppose that we can construct $J^o \in \mathcal{RH}_{\infty,-(k)}^{(p+m-\ell)\times(p+m-l)}$, where ℓ is assumed to be the generic ranks of both W and V, with $J_{22}^o(\infty) = 0$, such that $G_a^o + J^o$ is all-pass, where $G_a^o = \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{RH}_{\infty,+}^{(p+m-\ell)\times(p+m-\ell)}$. A set of solutions would then be given by

$$Q = \mathcal{F}_l(J^o, \Phi), \quad \Phi \in \mathcal{RH}_{\infty,-}, \|\Phi\|_\infty \le 1,$$

since $G + Q = \mathcal{F}_l(G_a^o + J^o, \Phi)$ so that $\|G+Q\|_\infty \le 1$ by Theorem 2.3 and $Q \in \mathcal{RH}_{\infty,-(k)}$ by Lemma 2.4. Now suppose that $Q \in \mathcal{RH}_{\infty,-(k)}$ and $\|G+Q\|_\infty \le 1$; then (4.2) and (4.3) together with $\|G_a^o + J^o\|_\infty = 1$ imply that

$$\begin{aligned} &W^\sim(Q - J_{11}^o) = 0 \\ &W^\sim J_{12}^o = 0 \\ &(Q - J_{11}^o)V(-s) = 0 \\ &J_{21}^o V(-s) = 0 \end{aligned}$$

Furthermore J_{12}^o and J_{21}^o have generically full column and row ranks respectively, so that for a generic point s, $Q - J_{11}^o \in \{\text{null space of } W^\sim\} \supset \{\text{range space } J_{12}^o\}$, but these two spaces will both have dimension $p - \ell$ and are hence equal; similarly for J_{21}^o. Hence the equation

$$Q - J_{11}^o = J_{12}^o \Psi J_{21}^o$$

has a rational solution Ψ, which will be proper. $\Phi(I-J_{22}^{o}\Phi)^{-1}=\Psi$ is achieved by setting $\Psi=(I+\Psi J_{22}^{o})^{-1}\Psi$, which is well-posed since $J_{22}^{o}(\infty)=0$ and this satisfies $Q=\mathcal{F}(J^{o},\Phi)$. Theorem 2.3 and Corollary 2.6 can then be applied to prove that $\Phi\in\mathcal{RH}_{\infty,-}$, $\|\Phi\|_{\infty}\leq 1$.

It only remains to construct J^{o} and verify its properties and this is a minor variation of the all-pass construction of section 3 and gives the following results.

Let the realization of G be partitioned conformally with X and Y as

$$G=\left[\begin{array}{c|c}A&B\\\hline C&0\end{array}\right]=\left[\begin{array}{cc|c}A_{11}&A_{12}&B_1\\A_{21}&A_{22}&B_2\\\hline C_1&C_2&0\end{array}\right]$$

The Lyapunov equations for X and Y then give

$$-A_{11}-A_{11}'=B_1B_1'=C_1'C_1$$

and hence by Lemma 2.2 there exists a unitary $D_e^{o}=\begin{bmatrix}D_{11}&D_{12}\\D_{21}&0\end{bmatrix}\in\mathbb{C}^{(p+m-\ell)\times(p+m-\ell)}$ where $\ell=\operatorname{rank}C_1=\operatorname{rank}B_1$, such that

$$\begin{bmatrix}C_1'&0\end{bmatrix}D_e^{o}+\begin{bmatrix}B_1&0\end{bmatrix}=0$$

A suitable value for X_e^{o}, the solution to the all-pass equations, is given by

$$X_e^{o}=\begin{bmatrix}I&0&0\\0&X_2&I\\0&I&Y_2Z_2^{-1}\end{bmatrix},\quad (X_e^{o})^{-1}=\begin{bmatrix}I&0&0\\0&Y_2&-Z_2'\\0&-Z_2&Z_2X_2\end{bmatrix}\tag{4.4}$$

It is then a straightforward exercise to verify that the all-pass equations are satisfied by the following realization of J^{o}:

$$J^{o}=\left[\begin{array}{c|cc}-A_{22}'-\hat{B}_1^{o}B_2'&\hat{B}_1^{o}&Z_2'^{-1}C_2'D_{12}\\\hline -C_2X_2-D_{11}B_2'&D_{11}&D_{12}\\-D_{21}B_2'&D_{21}&0\end{array}\right]$$

$$\hat{B}_1^{o}=Z_2'^{-1}(Y_2B_2+C_2'D_{11})$$

This realization of J^{o} clearly satisfies the required stability assumptions for Corollary 2.6. Furthermore, the generic rank of $W\geq\operatorname{rank}\lim_{s\to\infty}sW=\operatorname{rank}C_1=\ell$ and since $W^{\sim}J_{12}^{o}=0$, W has generic rank ℓ. Hence the characterization of all solutions is proven. This result is now stated without the $\sigma_{k+1}=1$ assumption which is removed by a simple scaling.

Theorem 4.1 *Let $G\in\mathcal{RH}_{\infty,+}^{p\times m}$ satisfy $\sigma_k(G)>\sigma_{k+1}(G)$. Then there exists a $\dot{Q}\in\mathcal{RH}_{\infty,-(k)}^{p\times m}$ such that $\|G+Q\|_{\infty}\leq\sigma$ if and only if $\sigma\geq\sigma_{k+1}(G)$. Furthermore all solutions to*

$$\|G+Q\|_{\infty}\leq\sigma=\sigma_{k+1}(G)$$

are given by

$$Q=\mathcal{F}_l(J^{o},\Phi),\qquad \Phi\in\mathcal{RH}_{\infty,-}^{(m-\ell)\times(p-\ell)},\ \|\Phi\|_{\infty}\leq\gamma$$

where J^{o} is constructed as follows. Let $G=\left[\begin{array}{cc|c}A_{11}&A_{12}&B_1\\A_{21}&A_{22}&B_2\\\hline C_1&C_2&0\end{array}\right]$ be a realization of G with controllability and observability Gramians given by $\begin{bmatrix}\sigma I&0\\0&X_2\end{bmatrix}$ and $\begin{bmatrix}\sigma I&0\\0&Y_2\end{bmatrix}$, respectively, and with $Z_2=X_2Y_2-\sigma^2 I$ invertible. Define $D_e=\begin{bmatrix}D_{11}&D_{12}\\D_{21}&0\end{bmatrix}\in\mathbb{C}^{(p+m-\ell)\times(p+m-\ell)}$ according to Lemma 2.2 where $\ell=\operatorname{rank}C_1=\operatorname{rank}B_1$, and

$$\begin{bmatrix} C_1' & 0 \end{bmatrix} D_e + \begin{bmatrix} B_1 & 0 \end{bmatrix} = 0.$$

Then J^o is given by

$$J^o = \left[\begin{array}{c|cc} -A_{22}' - \hat{B}_1^o B_2' & \hat{B}_1^o & Z_2'^{-1} C_2' D_{12} \\ \hline -C_2 X_2 - \sigma D_{11} B_2' & \sigma D_{11} & D_{12} \\ -D_{21} B_2' & D_{21} & 0 \end{array}\right]$$

$$\hat{B}_1^o = Z_2'^{-1}(Y_2 B_2 + \sigma C_2' D_{11})$$

The set of all $\hat{G} \in \mathcal{RH}_{\infty,+}$ of McMillan degree k such that $\|G - \hat{G}\|_H = \sigma_{k+1}(G)$ is given by $\hat{G} = -Q + F$ for $F \in \mathcal{RH}_{\infty,-}$, with Q as above.

5 FREQUENCY RESPONSE BOUNDS

Section 4 was concerned with finding $Q \in \mathcal{RH}^{p\times m}_{\infty,-(k)}$ such that $\|G + Q\|_\infty \leq \sigma_{k+1}$, the optimal achievable norm, and by (3.7) this implies that for $Q = -\hat{G} - F$ with $\hat{G} \in \mathcal{RH}_{\infty,+}$, $F \in \mathcal{RH}_{\infty,-}$ we have

$$\sigma_{k+1} \leq \|G - \hat{G}\|_H \leq \|G - \hat{G} - F\|_\infty = \sigma_{k+1}$$

and hence the characterization of all optimal Hankel-norm approximations is given by the causal part of $-Q$. The question now arises as to whether $\hat{G}$ is a good approximation to G in the H_∞-norm. The results of this section will now re-derive some of those of Glover (1984) but in a more efficient manner. The basic approach is to exploit the optimality of $\hat{G} + F$ and to show that $\|F\|_\infty$ can be bounded.

In order to bound $\|F\|_\infty$ we will first re-state Corollary 9.3 from Glover (1984).

Lemma 5.1 *Let $G(s) \in \mathcal{RH}^{p\times m}_{\infty,+}$ have Hankel singular values $\sigma_1 > \sigma_2 \cdots > \sigma_N$, where each σ_i has multiplicity r_i, and let $G(\infty) = 0$. Then*

(a) $\|G\|_\infty \leq 2(\sigma_1 + \sigma_2 + \ldots + \sigma_N)$

(b) there exists a constant D such that $\|G - D\|_\infty \leq (\sigma_1 + \sigma_2 + \ldots + \sigma_N)$

Proof. The proof of this lemma just involves computing J^o in Theorem 4.1 for $k = n - r_N$. The form of X_e and X_e^{-1} then give that $J^o \in \mathcal{RH}_{\infty,+}$ and that $\sigma_i^2(J^o) = \lambda_i(Y_2 \hat{X}_2)$, $\|G_a^o + J^o\|_\infty = \sigma_N$. Now J^o can be approximated in the same way and this repeated until just a constant remains. □

A lemma on all-pass systems is now stated.

Lemma 5.2 *Let $E = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ satisfy the all-pass equations of Lemma 2.10 and let A have dimension n_1+n_2 with n_1 eigen-values strictly in the left half plane and $n_2 < n_1$ eigen-values strictly in the right half plane. If $E = G + F$ with $G \in \mathcal{RH}^{p\times m}_{\infty,+}$ and $F \in \mathcal{RH}^{p\times m}_{\infty,-}$ then,*

$$\sigma_i(G) = \begin{cases} 1 & i = 1, 2, \ldots, n_1 - n_2 \\ \sigma_{i-n_1+n_2}(F^\sim) & i = n_1 - n_2 + 1, \ldots, n_1 \end{cases}$$

In particular this result holds if $E = G + F$ is all-pass with $G \in \mathcal{RH}^{p\times m}_{\infty,+}$ of degree n_1, and $F \in \mathcal{RH}^{p\times m}_{\infty,-}$ of degree $n_2 < n_1$.

Proof. Firstly let the realization be transformed to,

$$E = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad \mathrm{Re}\,\lambda_i(A_1) < 0, \quad \mathrm{Re}\,\lambda_i(A_2) > 0,$$

in which case $G = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D \end{array}\right]$, $F = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & 0 \end{array}\right]$. The all-pass equations of Lemma 2.10 *(i)-(vii)* are then satisfied by a transformed X and Y, partitioned as,

$$X = \begin{bmatrix} X_1 & X_2 \\ X_2' & X_3 \end{bmatrix}, \quad Y = \begin{bmatrix} Y_1 & Y_2' \\ Y_2 & Y_3 \end{bmatrix}$$

$XY = I$ implies that,

$$\begin{aligned} \det(\lambda I - X_1 Y_1) &= \det(\lambda I - (I - X_2 Y_2)) \\ &= \det((\lambda - 1)I + X_2 Y_2) \\ &= (\lambda - 1)^{n_1 - n_2} \det((\lambda - 1)I + Y_2 X_2) \\ &= (\lambda - 1)^{n_1 - n_2} \det(\lambda I - Y_3 X_3) \end{aligned}$$

The result now follows on observing that $\sigma_i(G) = \lambda_i(X_1Y_1)$ and $\sigma_i^2(F^\sim) = \lambda_i(X_3Y_3)$. The final statement then follows from Lemma 2.10 which gives the existence of suitable X and Y when the realization is minimal. □

Corollary 5.3 *Let G_a^o and J^o be as defined in Theorem 4.1 and write $J^o = \hat{G}_a^o + F_a^o$ with $\hat{G}_a^o \in \mathcal{RH}_{\infty,+}^{p\times m}$ and $F_a^o \in \mathcal{RH}_{\infty,-}^{p\times m}$. Then for $i = 1,2,...,2k+r$,*

$$\sigma_i(G_a^o - \hat{G}_a^o) = \sigma_{k+1}(G),$$

and for $i = 1,2,...,n-k-r$,

$$\sigma_{i+3k+r}(G) \le \sigma_i(F_a^{o\sim}) = \sigma_{i+2k+r}(G_a^o - \hat{G}_a^o) \le \sigma_{i+k+r}(G)$$

Proof. The construction of J^o ensures that the all-pass equations are satisfied and an inertia argument easily establishes that the A-matrix has precisely $n+k$ eigen-values in the open lhp and $n-k-r$ in the open rhp. Hence Lemma 5.2 can be applied to give the equalities. The inequalities are standard results on the singular vaues of finite rank perturbations and follow from the mini-max definition of singular values, see for example Theorem 1.4 in Partington(1988). □

The following result can now be derived and is similar to Theorem 9.7 and Corollary 9.9 in Glover (1984).

Theorem 5.4 *Let $Q = \mathcal{F}_l(J^o, \Phi)$ be given by Theorem 4.1 for Φ a* constant *contraction, and let $Q = -\hat{G} - F$ for $\hat{G} \in \mathcal{RH}_{\infty,+}$, $F \in \mathcal{RH}_{\infty,-}$. Then*

(a) $\sigma_i(G - \hat{G}) \le \begin{cases} \sigma_{k+1}(G), & i = 1,2,\ldots,2k+r \\ \sigma_{i-k}(G) & i = 2k+r+1,\ldots,n+k \end{cases}$

(b) $\sigma_i(G - \hat{G}) \ge \sigma_{i+k}(G) \quad i = 1,2,\ldots,n-k$

(c) $\sigma_i(F^\sim) \le \sigma_{i+k+r}(G), \quad i = 1,2,\ldots,n-k-r$

(d) there exists a D_0 such that

(i) $\delta := \|F - D_0\|_\infty \le \sum_{i=1}^{n-k-r} \sigma_i(F^\sim)$

(ii) $\|G - \hat{G} - D_0\|_\infty \le \sigma_{k+1}(G) + \delta \le \sigma_{k+1}(G) + \sum_{i=1}^{n-k-r} \sigma_{i+k+r}(G)$.

Proof.

(a) $\|G-\hat{G}\|_H = \sigma_{k+1}(G) \geq \sigma_i(G-\hat{G})$ for all i. Further, as in Corollary 5.3 for $i > 2k+r$,

$$\begin{aligned} \sigma_i(G-\hat{G}) &= \inf_{\deg(K_1)\leq i-1} \|G-\hat{G}-K_1\|_H \\ &\leq \inf_{\deg(K_2)\leq i-k-1} \|G-K_2\|_H \\ &= \sigma_{i-k}(G) \end{aligned}$$

(b) Standard finite rank perturbation result as in (a).

(c) By Lemma 2.1 we will dilate Φ to a unitary matrix, $\Phi_a = \begin{bmatrix} \Phi & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$, and observe that by Theorem 2.3

$$\begin{bmatrix} G+Q & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} := \mathcal{F}_l\left(\begin{bmatrix} G+J_{11}^o & 0 & J_{12}^o & 0 \\ 0 & 0 & 0 & I \\ J_{21}^o & 0 & J_{22}^o & 0 \\ 0 & I & 0 & 0 \end{bmatrix}, \ \Phi_a \right)$$

is all-pass and satisfies an all-pass equation with X_e^o as in (4.4), with $Q_a = \begin{bmatrix} Q & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$ having the same state dimension as J^o (i.e. $n-k-r$) . Lemma 5.2 can now be applied to $(G_a+Q_a) = (G_a-\hat{G}_a-F_a)$ together with part (a) applied to $\sigma_i(G_a-\hat{G}_a)$, to give the result.

(d) This follows immediately from Lemma 5.1 and part *(c)*. □

Remark 5.1 If a non-constant Φ is used to generate Q, then weaker frequency response bounds are obtained by first dilating Q to an all-pass Q_a of the same degree as Q [see Glover (1984, Theorem 5.2)]. Then the bounds on $\sigma_i(F^{\sim})$ can be derived but will be weaker.

Remark 5.2 Note that Theorem 5.4(a) and Lemma 5.1 could be used to derive a bound on $\|G-\hat{G}-D_0\|_\infty$. However, this would give an extra term of $(2k+r)\sigma_{k+1}(G)$ and would be much weaker but would not depend on Φ being a constant.

Remark 5.3 Trefethen and Gutknetch(1983) have proposed using this method in the scalar case, which they call the Carathéodory Fejér method, for real rational approximation on $[-\epsilon, \epsilon]$ and for complex uniform rational approximation on the disk of radius ϵ. They obtain asymptotic results as $\epsilon \rightarrow 0$, essentially giving that $\sigma_{k+1} \sim O(\epsilon^{2k+1})$ and that $\|F\|_\infty \sim O(\epsilon^{4k+3})$. These estimates show that in this asymptotic sense the term F becomes insignificant and hence $\hat{G}$ gives an essentially optimal approximant in the $\mathcal{H}_\infty$-norm. Note that this asymptotic norm bound is substantially smaller than could be deduced from Theorem 5.4.

Example

We now give an example to illustrate the results of Theorem 5.4. Let

$$G(s) = \sum_{i=0}^{7} \frac{1}{1 + 10^{-i}s}$$

This is an example of a transfer function with a positive semidefinite symmetric Hankel operator, and hence its eigen-values equal its singular values. Also its poles and zeros are interlaced on the negative real axis and $\|G\|_\infty = 2\sum_{i\geq 1}\sigma_i(G)$. The singular values, $\sigma_i(G)$ are given in Table 1. The optimal Hankel norm approximants, $\hat{G}_k$, of degrees $k = 1, \ldots, 7$ were calculated together with the anti-causal terms F_k. Table 1 also gives the $\sigma_{i-k-1}(F_k^{\sim})$, verifying the inequalities of Corollary 5.3 which can in fact be shown to be always strict for systems with interlaced poles and zeros. The frequency response error is in fact given by

$$\|G - \hat{G} - D_0\|_\infty = \sigma_{k+1}(G) + \sum_{i>0} \sigma_i(F_k^{\sim})$$

with the bound of Theorem 5.4 *(d)(i)* and the first inequality of *(d)(ii)* both equalities. For small values of k the error curves, $(G(j\omega) - \hat{G}_k(j\omega) - D_0)$, are far from being circular, in contrast to Remark 5.3, and that for $k = 2$ is plotted in Figure 1.

This example has not been chosen to illustrate the utility of the method, since this is a very difficult system to approximate with its poles spanning 8 orders of magnitude. It has however been chosen to illustrate the theoretical bounds and the fact that they may be tight. The truncated balanced realization technique will give errors equal to $2\sum_{i>k}\sigma_i(G)$ on examples of this type.

REFERENCES

Adamjan, V.M., D.Z. Arov and M.G. Krein (1971). Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem, *Math USSR Sbornik*, vol. 15, pp. 31–73.

Ball, J.A. and A.C.M. Ran (1986), Hankel norm approximation of a rational matrix function in terms of its realizations, in *Modelling, Identification and Robust Control* (C.I. Byrnes and A. Lindquist, eds.). North-Holland.

Dym, H. (1989), *J-Contractive Matrix Functions, Reproducing Kernel Hilbert Spaces and Interpolation*, under preparation.

Enns, D.F. (1984), Model reduction with balanced realisations: an error-bound and frequency-weighted generalization, *Proc. IEEE Conf. on Decision and Control*, Las Vegas NV, pp. 127–132.

Francis, B.A. (1987), *A Course in H_∞ Control Theory*. Springer-Verlag Lecture Notes in Control and Information Sciences, vol. 88.

Glover, K. (1984). All optimal Hankel-norm approximationsof linear multivariable systems and their $\mathcal{L}_\infty$-error bounds, *Int. J. Control*, vol. 39, pp. 1115-1193.

Glover, K. (1987), Model reduction: a tutorial on Hankel-norm methods and lower bounds on L_2 errors, *Proc. Xth Trienniel IFAC World Congress*, Pergamon Press, Munich, vol. X, pp. 288–293.

Glover, K., Curtain R.F. and J.R. Partington (1988), Realisation and approximation of linear infinite dimensional systems with error bounds, *SIAM J. Control and Optim.*, vol. 26, no. 4, pp. 863–898.

Glover, K., D.J.N. Limebeer, J.C. Doyle, E.M. Kasenally, and M.G. Safonov (1989). A characterization of all solutions to the four block general distance problem, under revision.

Green, M., K. Glover, D.J.N. Limebeer and J.C. Doyle (1988), "A J-spectral Factorization Approach to H_∞ Control", submitted.

Helton, J.W. (1987), *Operator Theory, Analytic Functions, Matrices, and Electrical Engineering*, American Mathematical Society CMBS Number 68.

Kailath, T. (1980), *Linear Systems.* Prentice-Hall.

Limebeer, D.J.N. and Y.S. Hung (1987). An analysis of pole-zero cancellations in H_∞-optimal control problems of the first kind, *SIAM J. Control Opt.*, vol.25, pp. 1457-1493.

Moore B.C. (1981), Principal component analysis in linear systems: controllability, observability and model reduction, *IEEE Trans. Auto. Cont.*, vol. AC–26, pp. 17–32.

Nehari, Z. (1957), On bounded bilinear forms, *Annals of Math.*, vol. 65, no. 1, pp. 155–162.

Partington, J.R. (1988), *An Introduction to Hankel Operators.* Cambridge University Press, London Mathematical Society Student Texts, vol. 13.

Power, S.C. (1982), *Hankel Operators on Hilbert Space.* Pitman.

Redheffer, R.M. (1960), On a certain linear fractional transformation, *J. Math. Phys.*, vol. 39, pp. 269–286.

Safonov, M.G., R.Y. Chiang and D.J.N. Limebeer (1987), Hankel model reduction without balancing: a descriptor approach, *Proc. 26th IEEE Conf. Dec. and Cont.*, Los Angeles.

Trefethen, L.N. and M. Gutknetch (1983), The Carathéodory-Fejér method for real rational approximation, *SIAM J. Numer. Anal.*, Vol. 20 No. 2, pp. 420-436.

Table 1: Hankel singular values for the example

i	$\sigma_i(G)$	$\sigma_{i-2}(F_1^{\sim})$	$\sigma_{i-3}(F_2^{\sim})$	$\sigma_{i-4}(F_3^{\sim})$	$\sigma_{i-5}(F_4^{\sim})$	$\sigma_{i-6}(F_5^{\sim})$	$\sigma_{i-7}(F_6^{\sim})$
1	1.2473						
2	0.9714						
3	0.6770	0.4428					
4	0.4428	0.4152	0.1821				
5	0.2812	0.1783	0.1580	0.0940			
6	0.1783	0.1505	0.1460	0.0551	0.0497		
7	0.1170	0.0850	0.0057	0.0071	0.0356	0.0017	
8	0.0850	0.0444	0.0049	0.0070	0.0297	0.0015	0.0118
$\sigma_{k+1}(G)$		0.9714	0.6770	0.4428	0.2812	0.1783	0.1170
$\|G - \hat{G}_k - D_0\|_\infty$		2.2875	1.1738	0.6058	0.3962	0.1815	0.1288
$\sum_{i>k} \sigma_i(G)$		2.7527	1.7813	1.1043	0.6615	0.3803	0.2020

Figure 2: Error curve for $(G(j\omega) - \hat{G}_2(j\omega) - D_0)$