Journal of Mathematical Systems, Estimation, and Control
Vol. 7, No. 2, 1997, pp. 129–155

Globally Convergent Homotopy Algorithms for the Combined H^2/H^∞ Model Reduction Problem*

Yuzhen Ge† Layne T. Watson†

Emmanuel G. Collins, Jr.‡ Dennis S. Bernstein

Abstract

The problem of finding a reduced order model, optimal in the H^2 sense, to a given system model is a fundamental one in control system analysis and design. The addition of a H^∞ constraint to the H^2 optimal model reduction problem results in a more practical yet computationally more difficult problem. Without the global convergence of probability-one homotopy methods the combined H^2/H^∞ model reduction problem is difficult to solve. Several approaches based on homotopy methods have been proposed. The issues are the number of degrees of freedom, the well posedness of the finite dimensional optimization problem, and the numerical robustness of the resulting homotopy algorithm. Homotopy algorithms based on several formulations — input normal form; Ly, Bryson, and Cannon's 2×2 block parametrization; a new nonminimal parametrization — are developed and compared here.

Key words: reduced order model problem, H^∞ control, H^2 control, probability-one homotopy algorithm

AMS Subject Classifications: 93B20, 93B25, 93B40, 65H10

1 Introduction

In a feedback control setting, order reduction techniques may be used either to simplify the plant for control design or to simplify the controller

*Received May 17, 1993. Received in final form July 1, 1994.

†The work of these authors was supported in part by AFOSR Grant F49620-92-J-0236 and by Department of Energy grant DE-FG05-883425068.

‡The work of Collins was sponsored in part by NASA contract NAS8-38575 and Sandia National Laboratories contract 54-7609.

for ease of implementation. In either case, the resulting reduced-order systems must be constructed with their closed loop role in mind. Although numerous order reduction techniques have been proposed, it is clear from small-gain type arguments that the order reduction procedure should be to approximate the system frequency response to the greatest extent possible.

Several order reduction techniques have been proposed for approximating the frequency response of a given system. For example, frequency weighting has been studied in [5] in conjunction with balancing [12]. Moreover, Hankel norm reduction has been shown to have fundamental ramifications for frequency domain approximation [1], [2], [7]. An overview and discussion of these ideas is given in [3].

In the present paper we follow the approach of [8], which is based upon a state space H^∞ formulation. In particular, by using a Riccati equation to enforce an H^∞ constraint on the norm of the reduction error in conjunction with an H^2 upper bound or entropy cost [13], it was shown in [8] that H^∞ constrained reduced order systems can be characterized by necessary conditions for optimality of the H^2 upper bound. The resulting algebraic conditions, which are a generalization of the "pure" H^2 optimality conditions given in [9], consist of nonstandard coupled Riccati and Lyapunov type matrix equations.

The purpose of the present paper is to make significant progress in developing novel, stable, globally convergent numerical algorithms for solving the optimality conditions for H^2/H^∞ order reduction given in [8]. The approach we take is based on the construction of probability-one homotopy maps, similar to those developed for the H^2 order reduction problem in [6].

2 Statement of the Problem

Given the controllable and observable, time invariant, continuous time system

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,Du(t), \\ y(t) &= C\,x(t), \end{aligned} \tag{2.1}$$

where $t \in [0, \infty)$, $A \in \mathbf{R}^{n\times n}$ is asymptotically stable, $B \in \mathbf{R}^{n\times m}$, $C \in \mathbf{R}^{l\times n}$, $D \in \mathbf{R}^{m\times p}$ ($m \le p$) and the input $Du(t)$ is white noise with symmetric and positive definite intensity $V \equiv DD^T$, find a n_m-th order model ($n_m < n$)

$$\begin{aligned} \dot{x}_m(t) &= A_m\,x_m(t) + B_m\,Du(t), \\ y_m(t) &= C_m\,x_m(t), \end{aligned} \tag{2.2}$$

where $A_m \in \mathbf{R}^{n_m\times n_m}$, $B_m \in \mathbf{R}^{n_m\times m}$, $C_m \in \mathbf{R}^{l\times n_m}$, which satisfies the following criteria:

(i) A_m is asymptotically stable;

(ii) the transfer function of the reduced order model lies within γ of the transfer function of the full order model in the H^∞ norm, i.e.,

$$\|H(s) - H_m(s)\|_\infty \leq \gamma \tag{2.3}$$

where

$$H(s) \equiv EC(sI_n - A)^{-1}BD, \qquad H_m(s) \equiv EC_m(sI_m - A_m)^{-1}B_mD,$$

$\gamma > 0$ is a given constant, $E \in \mathbf{R}^{q\times l}$ ($q \geq l$) is a given constant matrix; and (iii) the H^2 model reduction criterion

$$J(A_m, B_m, C_m) \equiv \lim_{t\to\infty} \mathcal{E}\left[(y - y_m)^T R(y - y_m)\right] \tag{2.4}$$

is minimized, where $\mathcal{E}$ is the expected value and $R = E^TE$ is a symmetric and positive definite weighting matrix.

3 The Auxiliary Minimization Problem

Define

$$\tilde{n} \equiv n + n_m, \qquad \tilde{E} \equiv E\tilde{C}, \qquad \tilde{D} \equiv \tilde{B}D,$$

$$\tilde{A} \equiv \begin{pmatrix} A & 0 \\ 0 & A_m \end{pmatrix}, \qquad \tilde{B} \equiv \begin{pmatrix} B \\ B_m \end{pmatrix}, \qquad \tilde{C} \equiv (C \quad -C_m), \tag{3.1}$$

$$\begin{aligned} \tilde{R} &\equiv \tilde{E}^T\tilde{E} = \tilde{C}^TR\tilde{C} = \begin{pmatrix} C^TRC & -C^TRC_m \\ -C_m^TRC & C_m^TRC_m \end{pmatrix}, \\ \tilde{V} &\equiv \tilde{D}\tilde{D}^T = \tilde{B}V\tilde{B}^T = \begin{pmatrix} BVB^T & BVB_m^T \\ B_mVB^T & B_mVB_m^T \end{pmatrix}. \end{aligned} \tag{3.2}$$

The full order system (2.1) and the reduced order system (2.2) can be written as a single augmented system

$$\begin{aligned} \dot{\tilde{x}}(t) &= \tilde{A}\,\tilde{x}(t) + \tilde{D}\,u(t), \\ \tilde{y}(t) &= \tilde{C}\,\tilde{x}(t). \end{aligned} \tag{3.3}$$

Using this notation the cost $J(A_m, B_m, C_m)$ can be written as

$$\begin{aligned} J(A_m, B_m, C_m) &= \lim_{t\to\infty} \mathcal{E}\left[(y - y_m)^T R\,(y - y_m)\right] = \lim_{t\to\infty} \mathcal{E}(\tilde{y}^T R\,\tilde{y}) \\ &= \lim_{t\to\infty} \mathcal{E}(\tilde{x}^T\,\tilde{C}^T\,R\,\tilde{C}\,\tilde{x}) = \lim_{t\to\infty} \mathcal{E}(\tilde{x}^T\,\tilde{R}\,\tilde{x}) \\ &= \mathrm{tr}\,\left[\lim_{t\to\infty} \mathcal{E}(\tilde{x}^T\,\tilde{R}\,\tilde{x})\right] = \lim_{t\to\infty} \mathcal{E}\left[\mathrm{tr}\,(\tilde{x}^T\,\tilde{R}\,\tilde{x})\right] \\ &= \lim_{t\to\infty} \mathcal{E}\left[\mathrm{tr}\,(\tilde{x}\,\tilde{x}^T\,\tilde{R})\right] = \mathrm{tr}\,\left[\lim_{t\to\infty} \mathcal{E}(\tilde{x}\,\tilde{x}^T)\tilde{R}\right] = \mathrm{tr}\,(\tilde{Q}\,\tilde{R}), \end{aligned} \tag{3.4}$$

where the variance matrix $\tilde{Q} = \lim_{t\to\infty} \mathcal{E}(\tilde{x}\,\tilde{x}^T) = \int_0^\infty e^{\tilde{A}t}\tilde{B}V\tilde{B}^T e^{\tilde{A}^T t}\,dt$ (see [9]) satisfies

$$\tilde{A}\,\tilde{Q} + \tilde{Q}\,\tilde{A}^T + \tilde{V} = 0. \tag{3.5}$$

Lemma 1 [8] *Let (A_m, B_m, C_m) be given and assume there exists $\mathcal{Q} \in \mathbf{R}^{\tilde{n}\times\tilde{n}}$ satisfying*

$$\mathcal{Q} \quad \text{is symmetric and nonnegative definite} \tag{3.6}$$

and

$$\tilde{A}\,\mathcal{Q} + \mathcal{Q}\tilde{A}^T + \gamma^{-2}\mathcal{Q}\tilde{R}\mathcal{Q} + \tilde{V} = 0. \tag{3.7}$$

Then

$$(\tilde{A}, \tilde{D}) \quad \text{is stabilizable} \tag{3.8}$$

if and only if

$$A_m \quad \text{is asymptotically stable.}$$

Furthermore, if (3.8) holds, then

$$\|H(s) - H_m(s)\|_\infty \le \gamma, \tag{3.9}$$

$$\tilde{Q} \le \mathcal{Q} \qquad (\mathcal{Q} - \tilde{Q} \text{ is nonnegative definite}),$$

and

$$\text{tr}\ \tilde{Q}\tilde{R} \equiv J(A_m, B_m, C_m) \le \mathcal{J}(A_m, B_m, C_m) \equiv\ \text{tr}\ \mathcal{Q}\tilde{R}.$$

Hence the H^∞ constraint is automatically enforced when a nonnegative definite solution to (3.7) is known to exist. Furthermore, the solution $\mathcal{Q}$ provides an upper bound for the actual state covariance $\tilde{Q}$ along with a bound on the H^2 model reduction.

The satisfaction of (3.6)–(3.8) leads to (i) A_m stable; (ii) a bound on the H^∞ distance between the full order and reduced order systems; and (iii) an upper bound for the H^2 model-reduction criterion. The auxiliary minimization problem is to determine (A_m, B_m, C_m) that minimizes $\mathcal{J}(A_m, B_m, C_m)$ and thus provides a bound for the actual H^2 criterion $J(A_m, B_m, C_m)$.

In order that (3.6)–(3.7) have a solution, and degenerate solutions are ruled out, (A_m, B_m, C_m) is restricted to the set

$$\begin{aligned} \mathcal{S} \equiv \big\{(A_m, B_m, C_m) :\ & \tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R} \text{ is asymptotically stable,} \\ & \mathcal{Q} \text{ solving (3.7) is symmetric positive definite,} \\ & \text{and } (A_m, B_m, C_m) \text{ is controllable and observable } \big\}. \end{aligned}$$

$\mathcal{S}$ is open since asymptotic stability, positive definiteness, controllability, and observability are all preserved under small perturbations. Requiring that $\mathcal{S}$ be open is theoretically convenient, and a solution on its boundary would only occur in rare degenerate cases [6], [8].

4 A Homotopy Approach

A Homotopy Approach Based on the Input Normal Form

Treating all the components of A_m, B_m, C_m as unknowns is redundant, since only $n_m(m+l)$ parameters suffice to describe the reduced order model. There are numerous ways to parametrize the model (A_m, B_m, C_m) — input normal form is one way that uses the minimal number of parameters. As for the H^2 model order reduction problem, a particular minimal parametrization assumes some *structure*, which *the optimal reduced order model may not possess.* A comparison of various minimal parametrizations, and examples of the optimal reduced order models failing to have prescribed structures (e.g., input normal form), are given in [6] for the H^2 problem. Those counterexamples *a priori* extend to the H^2/H^∞ problem considered here.

Theorem 1 [10] *Suppose $\bar{A}_m$ is asymptotically stable. Then for every minimal $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$, i.e., $(\bar{A}_m, \bar{B}_m)$ is controllable and $(\bar{A}_m, \bar{C}_m)$ is observable, there exist a similarity transformation U and a positive definite matrix $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_{n_m})$ such that $A_m = U^{-1}\bar{A}_m U$, $B_m = U^{-1}\bar{B}_m$, and $C_m = \bar{C}_m U$ satisfy*

$$
\begin{aligned}
0 &= A_m + A_m^T + B_m V B_m^T, \\
0 &= A_m^T \Omega + \Omega A_m + C_m^T R C_m.
\end{aligned}
\tag{4.1}
$$

In addition,

$$
\begin{aligned}
(A_m)_{ii} &= -\frac{1}{2}\left(B_m V B_m^T\right)_{ii}, \\
\omega_i &= \frac{\left(C_m^T R C_m\right)_{ii}}{\left(B_m V B_m^T\right)_{ii}}, \\
(A_m)_{ij} &= \frac{\left(C_m^T R C_m\right)_{ij} - \omega_j \left(B_m V B_m^T\right)_{ij}}{\omega_j - \omega_i}, \qquad \text{if } \omega_i \neq \omega_j.
\end{aligned}
\tag{4.2}
$$

Definition 1 The triple (A_m, B_m, C_m) satisfying (4.1) or (4.2) is said to be in *input normal form.*

To optimize $\mathcal{J}(A_m, B_m, C_m)$ over the open set $\mathcal{S}$ under the constraints that symmetric positive definite $\mathcal{Q}$ satisfies (3.7), and (A_m, B_m, C_m) is in input normal form, the following Lagrangian is formed:

$$\begin{aligned}
&\mathcal{L}(A_m, B_m, C_m, \Omega,\ \mathcal{Q}, \mathcal{P}, M_c, M_o) \\
&\quad \equiv \operatorname{tr}\left[\mathcal{Q}\tilde{R} + (\tilde{A}\mathcal{Q} + \mathcal{Q}\tilde{A}^T + \gamma^{-2}\mathcal{Q}\tilde{R}\mathcal{Q} + \tilde{V})\mathcal{P}\right. \\
&\quad \left. + \left(A_m + A_m^T + B_m V B_m^T\right) M_c + \left(A_m^T \Omega + \Omega A_m + C_m^T R C_m\right) M_o\right],
\end{aligned}$$

where the symmetric matrices M_c, M_o, and $\mathcal{P} \in \mathbf{R}^{\tilde{n}\times\tilde{n}}$ are Lagrange multipliers. $\Omega =$ diag $(\omega_1, \ldots, \omega_{n_m})$ is related to the input normal form constraint. Since the constraint matrix equations are symmetric, so are the multipliers written in matrix form, and the trace is a concise way to write the constraint portion of the Lagrangian function. Setting $\partial\mathcal{L}/\partial\mathcal{Q} = 0$ yields

$$0 = (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T \mathcal{P} + \mathcal{P}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}) + \tilde{R}. \tag{4.3}$$

Partition $\mathcal{Q}$, $\mathcal{P} \in \mathbf{R}^{\tilde{n}\times\tilde{n}}$ into

$$\mathcal{Q} = \begin{pmatrix} \mathcal{Q}_1 & \mathcal{Q}_{12} \\ \mathcal{Q}_{12}^T & \mathcal{Q}_2 \end{pmatrix}, \qquad \mathcal{P} = \begin{pmatrix} \mathcal{P}_1 & \mathcal{P}_{12} \\ \mathcal{P}_{12}^T & \mathcal{P}_2 \end{pmatrix} \tag{4.4}$$

where $\mathcal{Q}_1$, $\mathcal{P}_1 \in \mathbf{R}^{n\times n}$ and $\mathcal{Q}_2$, $\mathcal{P}_2 \in \mathbf{R}^{n_m\times n_m}$. Define

$$\mathcal{P}\mathcal{Q} \equiv Z = \begin{pmatrix} Z_1 & Z_{12} \\ Z_{21} & Z_2 \end{pmatrix} \tag{4.5}$$

where

$$\begin{aligned}
Z_1 &\equiv \mathcal{P}_1\mathcal{Q}_1 + \mathcal{P}_{12}\mathcal{Q}_{12}^T, & Z_{12} &\equiv \mathcal{P}_1\mathcal{Q}_{12} + \mathcal{P}_{12}\mathcal{Q}_2, \\
Z_{21} &\equiv \mathcal{P}_{12}^T\mathcal{Q}_1 + \mathcal{P}_2\mathcal{Q}_{12}^T, & Z_2 &\equiv \mathcal{P}_{12}^T\mathcal{Q}_{12} + \mathcal{P}_2\mathcal{Q}_2.
\end{aligned}$$

$\partial L/\partial\Omega = 0$ and $\partial L/\partial A_m = 0$ yield

$$0 = 2M_c + 2\Omega M_o + 2\left(\mathcal{P}_{12}^T\mathcal{Q}_{12} + \mathcal{P}_2\mathcal{Q}_2\right), \qquad 0 = \left(A_m M_o\right)_{ii}, \quad 1 \le i \le n_m.$$

A straightforward calculation shows

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial B_m} &= 2(\mathcal{P}_{12}^T B V + \mathcal{P}_2 B_m V) + 2M_c B_m V, \\
\frac{\partial\mathcal{L}}{\partial C_m} &= 2(RC_m\mathcal{Q}_2 - RC\mathcal{Q}_{12}) + 2RC_m M_o \\
&\quad + \gamma^{-2}\Big[-RC(Z_1^T\mathcal{Q}_{12} + Z_{21}^T\mathcal{Q}_2 + \mathcal{Q}_1 Z_{12} + \mathcal{Q}_{12} Z_2) \\
&\qquad + RC_m(\mathcal{Q}_{12}^T Z_{12} + Z_{12}^T\mathcal{Q}_{12} + \mathcal{Q}_2 Z_2 + Z_2^T\mathcal{Q}_2)\Big].
\end{aligned} \tag{4.6}$$

Theorem 2 [4] *The matrices M_c and M_o in (4.6) satisfy*

$$
\begin{aligned}
M_c &= -\left(\frac{1}{2}S + \Omega M_o\right),\\
\left(M_o\right)_{ii} &= -\frac{1}{\left(A_m\right)_{ii}} \sum_{\substack{j=1\\ j\neq i}}^{n_m} \left(A_m\right)_{ij} \left(M_o\right)_{ji},\\
\left(M_o\right)_{ij} &= \frac{(S)_{ij} - (S)_{ji}}{2(\omega_j - \omega_i)}, \quad \text{if } \omega_j \neq \omega_i,
\end{aligned}
\tag{4.7}
$$

where

$$
S = 2\left(\mathcal{P}_{12}^T \mathcal{Q}_{12} + \mathcal{P}_2 \mathcal{Q}_2\right). \tag{4.8}
$$

At this point the unknowns are B_m and C_m ($n_m(m+l)$ parameters, due to the input normal form structure in (4.2)), and the equations to be solved are the partials in (4.6) set to zero. Choose a problem for which a solution is known, defined by the matrices A_0, B_0, C_0, R_0, V_0, γ_0; exactly how this initial problem is chosen is described in the next section. A homotopy approach based on the input normal form is now described. Let A_f, B_f, C_f , R_f, V_f, and γ_f denote A, B, C, R, V, and γ in the above and define

$$
\begin{aligned}
A(\lambda) &= A_0 + \lambda(A_f - A_0), & R(\lambda) &= R_0 + \lambda(R_f - R_0),\\
B(\lambda) &= B_0 + \lambda(B_f - B_0), & V(\lambda) &= V_0 + \lambda(V_f - V_0),\\
C(\lambda) &= C_0 + \lambda(C_f - C_0), & \gamma(\lambda) &= \gamma_0 + \lambda(\gamma_f - \gamma_0).
\end{aligned}
\tag{4.9}
$$

For brevity, $A(\lambda)$, $B(\lambda)$, $C(\lambda)$, $R(\lambda)$, $V(\lambda)$, and $\gamma(\lambda)$ will be denoted by A, B, C, R, V, and γ respectively in the following. Let

$$
\begin{aligned}
H_{B_m}(\theta,\lambda) &= \frac{\partial L}{\partial B_m} = 2\left(\mathcal{P}_{12}^T B + \mathcal{P}_2 B_m\right)V + 2M_c B_m V,\\
H_{C_m}(\theta,\lambda) &= \frac{\partial L}{\partial C_m} = 2R\left(C_m \mathcal{Q}_2 - C\mathcal{Q}_{12}\right) + 2RC_m M_o\\
&\quad + \gamma^{-2}\big[-RC(Z_1^T \mathcal{Q}_{12} + Z_{21}^T \mathcal{Q}_2 + \mathcal{Q}_1 Z_{12} + \mathcal{Q}_{12} Z_2)\\
&\quad + RC_m(\mathcal{Q}_{12}^T Z_{12} + Z_{12}^T \mathcal{Q}_{12} + \mathcal{Q}_2 Z_2 + Z_2^T \mathcal{Q}_2)\big],
\end{aligned}
$$

where

$$
\theta \equiv \begin{pmatrix} \mathrm{Vec}\,(B_m) \\ \mathrm{Vec}\,(C_m) \end{pmatrix}
$$

denotes the independent variables B_m and C_m, M_o and M_c satisfy (4.7), and $\mathcal{Q}$ and $\mathcal{P}$ satisfy respectively (3.7) and (4.3) with partitioned forms

(4.4). Vec(P) for a matrix $P \in \mathbf{R}^{p\times q}$ is the concatenation of its columns:

$$\mathrm{Vec}(P) \equiv \begin{pmatrix} P_{\cdot 1} \\ P_{\cdot 2} \\ \vdots \\ P_{\cdot q} \end{pmatrix} \in \mathbf{R}^{pq}.$$

The homotopy map is defined as

$$\rho(\theta,\lambda) = \begin{pmatrix} \mathrm{Vec}\,\left[H_{B_m}(\theta,\lambda)\right] \\ \mathrm{Vec}\,\left[H_{C_m}(\theta,\lambda)\right] \end{pmatrix}, \tag{4.10}$$

and its Jacobian matrix is

$$D\rho(\theta,\lambda) = \big(D_\theta\rho(\theta,\lambda), D_\lambda\rho(\theta,\lambda)\big). \tag{4.11}$$

Define

$$\begin{aligned}
\hat{H}_{B_m}(\mathcal{P}^{(j)}, M_c^{(j)}) &= 2(\mathcal{P}_{12}^{T\,(j)}B + \mathcal{P}_2^{(j)}B_m)V + 2M_c^{(j)}B_mV, \\
\hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}, M_o^{(j)}) &= 2R(C_m\mathcal{Q}_2^{(j)} - C\mathcal{Q}_{12}^{(j)}) + 2RC_mM_o^{(j)} \\
&\quad - \gamma^{-2}RC(Z_1^{T\,(j)}\mathcal{Q}_{12} + Z_{21}^{T\,(j)}\mathcal{Q}_2 + Z_1^T\mathcal{Q}_{12}^{(j)} + Z_{21}^T\mathcal{Q}_2^{(j)} \\
&\quad + \mathcal{Q}_1^{(j)}Z_{12} + \mathcal{Q}_1Z_{12}^{(j)} + \mathcal{Q}_{12}^{(j)}Z_2 + \mathcal{Q}_{12}Z_2^{(j)}) \\
&\quad + \gamma^{-2}RC_m(Z_{12}^{T(j)}\mathcal{Q}_{12} + Z_{12}^T\mathcal{Q}_{12}^{(j)} + \mathcal{Q}_{12}^{T(j)}Z_{12} + \mathcal{Q}_{12}^TZ_{12}^{(j)} \\
&\quad + \mathcal{Q}_2^{(j)}Z_2 + Z_2^{T\,(j)}\mathcal{Q}_2 + \mathcal{Q}_2Z_2^{(j)} + Z_2^T\mathcal{Q}_2^{(j)}),
\end{aligned}$$

where the superscript (j) means $\partial/\partial\theta_j$: $Y^{(j)} \equiv \partial Y/\partial\theta_j$. Using the above definitions, we have for $\theta_j = (B_m)_{kl}$,

$$\begin{aligned}
\frac{\partial H_{B_m}}{\partial(B_m)_{kl}} &= \hat{H}_{B_m}(\mathcal{P}^{(j)}, M_c^{(j)}) + 2(\mathcal{P}_2 + M_c)E^{(k,l)}V, \\
\frac{\partial H_{C_m}}{\partial(B_m)_{kl}} &= \hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}, M_o^{(j)}),
\end{aligned} \tag{4.12}$$

and for $\theta_j = (C_m)_{kl}$,

$$\begin{aligned}
\frac{\partial H_{B_m}}{\partial(C_m)_{kl}} &= \hat{H}_{B_m}(\mathcal{P}^{(j)}, M_c^{(j)}), \\
\frac{\partial H_{C_m}}{\partial(C_m)_{kl}} &= \hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}, M_o^{(j)}) + 2RE^{(k,l)}(\mathcal{Q}_2 + M_o) \\
&\quad + \gamma^{-2}RE^{(k,l)}(Z_{12}^T\mathcal{Q}_{12} + \mathcal{Q}_{12}^TZ_{12} + \mathcal{Q}_2^TZ_2 + Z_2^T\mathcal{Q}_2),
\end{aligned} \tag{4.13}$$

where $E^{(k,l)}$ is a matrix of the appropriate dimension whose only nonzero element is $e_{kl} = 1$. $\mathcal{P}^{(j)}$ and $\mathcal{Q}^{(j)}$ can be obtained by solving the Lyapunov equations

$$\begin{aligned}
0 &= (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})\mathcal{Q}^{(j)} + \mathcal{Q}^{(j)}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T + \tilde{V}^{(j)} \\
&\quad + \tilde{A}^{(j)}\mathcal{Q} + \mathcal{Q}\tilde{A}^{T(j)} + \gamma^{-2}\mathcal{Q}\tilde{R}^{(j)}\mathcal{Q}, \\
0 &= (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T\mathcal{P}^{(j)} + \mathcal{P}^{(j)}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}) + \tilde{R}^{(j)} \\
&\quad + (\tilde{A}^{(j)} + \gamma^{-2}\mathcal{Q}^{(j)}\tilde{R} + \gamma^{-2}\mathcal{Q}\tilde{R}^{(j)})^T\mathcal{P} \\
&\quad + \mathcal{P}(\tilde{A}^{(j)} + \gamma^{-2}\mathcal{Q}^{(j)}\tilde{R} + \gamma^{-2}\mathcal{Q}\tilde{R}^{(j)}).
\end{aligned} \tag{4.14}$$

Similarly for λ, using a dot to denote $\partial/\partial\lambda$,

$$\begin{aligned}
\frac{\partial H_{B_m}}{\partial\lambda} &= \hat{H}_{B_m}(\dot{\mathcal{P}}, \dot{M}_c) + 2\mathcal{P}_{12}^T(\dot{B}V + B\dot{V}) + 2(\mathcal{P}_2 + M_c)B_m\dot{V}, \\
\frac{\partial H_{C_m}}{\partial\lambda} &= \hat{H}_{C_m}(\dot{\mathcal{Q}}, \dot{Z}, \dot{M}_o) + 2\dot{R}C_m(\mathcal{Q}_2 + M_o) - 2(\dot{R}C + R\dot{C})\mathcal{Q}_{12} \\
&\quad + \gamma^{-2}\dot{R}h_\lambda - 2\gamma^{-3}\dot{\gamma}Rh_\lambda \\
&\quad - \gamma^{-2}R\dot{C}(Z_1^T\mathcal{Q}_{12} + Z_{21}^T\mathcal{Q}_2 + \mathcal{Q}_1Z_{12} + \mathcal{Q}_{12}Z_2),
\end{aligned} \tag{4.15}$$

where

$$\begin{aligned}
h_\lambda &= -C(Z_1^T\mathcal{Q}_{12} + Z_{21}^T\mathcal{Q}_2 + \mathcal{Q}_1Z_{12} + \mathcal{Q}_{12}Z_2) \\
&\quad + C_m(\mathcal{Q}_{12}^TZ_{12} + Z_{12}^T\mathcal{Q}_{12} + \mathcal{Q}_2Z_2 + Z_2^T\mathcal{Q}_2)
\end{aligned},$$

and $\dot{\mathcal{P}}$ and $\dot{\mathcal{Q}}$ are obtained by solving the Lyapunov equations

$$\begin{aligned}
0 &= (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})\dot{\mathcal{Q}} + \dot{\mathcal{Q}}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T + \dot{\tilde{V}} \\
&\quad + \dot{\tilde{A}}\mathcal{Q} + \mathcal{Q}\dot{\tilde{A}}^T + \gamma^{-2}\mathcal{Q}\dot{\tilde{R}}\mathcal{Q} - 2\gamma^{-3}\dot{\gamma}\mathcal{Q}\tilde{R}\mathcal{Q}, \\
0 &= (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T\dot{\mathcal{P}} + \dot{\mathcal{P}}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}) + \dot{\tilde{R}} \\
&\quad + (\dot{\tilde{A}} + \gamma^{-2}\dot{\mathcal{Q}}\tilde{R} + \gamma^{-2}\mathcal{Q}\dot{\tilde{R}} - 2\gamma^{-3}\dot{\gamma}\mathcal{Q}\tilde{R})^T\mathcal{P} \\
&\quad + \mathcal{P}(\dot{\tilde{A}} + \gamma^{-2}\dot{\mathcal{Q}}\tilde{R} + \gamma^{-2}\mathcal{Q}\dot{\tilde{R}} - 2\gamma^{-3}\dot{\gamma}\mathcal{Q}\tilde{R}).
\end{aligned} \tag{4.16}$$

5 Numerical Algorithm for Normal Form

The initial point $(\theta, \lambda) = (\theta_0, 0) = ((B_m)_0, (C_m)_0, 0)$ is ideally chosen so that the triple $((A_m)_0, (B_m)_0, (C_m)_0)$ is in input normal form and satisfies $\rho(\theta_0, 0) = 0$.

Theorem 3 [12] *Suppose $\bar{A}$ is asymptotically stable. Then for every minimal $(\bar{A}, \bar{B}, \bar{C})$, i.e., $(\bar{A}, \bar{B})$ is controllable and $(\bar{A}, \bar{C})$ is observable, there* *exist a similarity transformation T and a positive definite matrix $\Lambda = \operatorname{diag}(d_1, d_2, \cdots, d_n)$ with $d_i \geq d_{i+1}$ such that $A = T^{-1}\bar{A}T$, $B = T^{-1}\bar{B}$, and $C = \bar{C}T$ satisfy*

$$0 = A\Lambda + \Lambda A^T + BVB^T,$$
$$0 = A^T\Lambda + \Lambda A + C^TRC.$$

Definition 2 The triple (A, B, C) in the above theorem is *balanced.*

According to Moore [12], under certain conditions, the leading principal $n_m \times n_m$ block of A, the leading principal $n_m \times m$ block of B, and the leading principal $l \times n_m$ block of C in balanced form are good approximations to the reduced order model. This suggests that the initial point $(\theta_0, 0)$ be chosen as follows:

1) Transform the given triple (A_f, B_f, C_f) to balanced form (A_b, B_b, C_b).
2) Partition (A_b, B_b, C_b) as

$$A_b = \begin{array}{cc} & \overbrace{\phantom{A_{11}}}^{n_m} \\ n_m\{ & \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \end{array}, \quad B_b = n_m\{ \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad C_b = \begin{array}{c} \overbrace{}^{n_m} \\ \begin{pmatrix} C_1 & C_2 \end{pmatrix} \end{array}.$$

3) (A_0, B_0, C_0) is chosen as

$$A_0 = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}, \qquad B_0 = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}, \qquad C_0 = \begin{pmatrix} C_1 & 0 \end{pmatrix}.$$

4) The initial point for the reduced order model is chosen as

$$\bar{\theta}_0 = \begin{pmatrix} \operatorname{Vec}\,(\bar{B}_m)_0 \\ \operatorname{Vec}\,(\bar{C}_m)_0 \end{pmatrix} = \begin{pmatrix} \operatorname{Vec}\,B_1 \\ \operatorname{Vec}\,C_1 \end{pmatrix},$$

   and $(\bar{A}_m)_0 = A_{11}$ by construction.
5) Transform the initial point $\big((\bar{A}_m)_0, (\bar{B}_m)_0, (\bar{C}_m)_0\big)$ to input normal form so that the initial reduced order model is

$$\big((A_m)_0, (B_m)_0, (C_m)_0\big) = \big(T^{-1}(\bar{A}_m)_0 T, \quad T^{-1}(\bar{B}_m)_0, \quad (\bar{C}_m)_0 T\big).$$

   The initial point for the homotopy map is then $(\theta_0, 0)$, where

$$\theta_0 = \begin{pmatrix} \operatorname{Vec}\,(B_m)_0 \\ \operatorname{Vec}\,(C_m)_0 \end{pmatrix}.$$

(In general, the truncation to obtain the approximate reduced order model should be based on the component costs instead of on the sizes of the balanced gains d_i as done above [14]. This explains why in some cases the above algorithm for choosing the initial points did not lead to a reduced order model with a minimal cost.)

The above method for choosing the initial point will not give a zero value for the homotopy at $\lambda = 0$ unless the initial γ is chosen so that the term $\gamma^{-2}\tilde{Q}\tilde{R}$ is negligible. The initial γ can be chosen as a sufficiently large positive number ($\gamma(0) = \infty$ corresponds to $\rho(\theta_0, 0) = 0$ exactly).

Once the initial point is chosen, the rest of the computation (which is a standard globally convergent probability-one homotopy algorithm [16]) is as follows:

1) Set $\lambda := 0$, $\theta := \theta_0$.
2) Calculate A_m from B_m and C_m, $\tilde{R}$, $\tilde{V}$, and compute $\mathcal{Q}$ and $\mathcal{P}$ according to (3.7) and (4.3).
3) Evaluate S from (4.8) and M_o and M_c according to (4.7).
4) Evaluate the homotopy map $\rho(\theta, \lambda)$ in (4.10) and $D\rho(\theta, \lambda)$ in (4.11).
5) Predict the next point $Z^{(0)} = (\theta^{(0)}, \lambda^{(0)})$ on the homotopy zero curve using, e.g., a Hermite cubic interpolant.
6) For $k := 0, 1, 2, \cdots$ until convergence do

$$Z^{(k+1)} = Z^{(k)} - \left[D\rho(Z^{(k)})\right]^\dagger \rho(Z^{(k)}),$$

   where $\left[D\rho(Z)\right]^\dagger$ is the Moore-Penrose inverse of $D\rho(Z)$. Let $(\theta_1, \lambda_1) = \lim_{k\to\infty} Z^{(k)}$.
7) If $\lambda_1 < 1$, then set $\theta := \theta_1$, $\lambda := \lambda_1$, and go to step 2).
8) If $\lambda_1 \geq 1$, compute the solution $\bar{\theta}$ at $\lambda = 1$. A_m is then obtained from B_m and C_m.

An alternative strategy for choosing an initial point is as follows:

1) Modify A_f to $A_f' = c_1 I + c_2 A_f$, where $c_1 \leq 0$ and $c_2 \geq 0$.
2) Transform (A_f', B_f, C_f) to balanced form and choose (A_0', B_0', C_0') as before.
3) Compute the initial reduced order model $\left((A_m)_0, (B_m)_0, (C_m)_0\right)$ from the triple (A_0', B_0', C_0') as before.

When $c_1 = 0$, $c_2 = 1$, this strategy reduces to the previous one. For some problems, our numerical experiments show that HOMPACK reaches $\lambda > 1$ in fewer steps with $c_1 \neq 0$ than with $c_1 = 0$. A modification to the homotopy map $\rho(\theta, \lambda)$ in (4.10) is

$$\rho_1(\theta, \lambda) = \lambda\rho(\theta, \lambda) + (1 - \lambda)(\theta - \theta_0),$$

where θ_0 denotes the initial value of θ at $\lambda = 0$. For some problems this homotopy map can be more efficient than the one in (4.10), while in other cases it can be less efficient.

6 Homotopy Algorithm Based on Ly's Formulation

Ly, et al. [11] introduced another canonical form also with $n_m m + n_m l$ parameters as in the input normal form formulation. The reduced order model is represented with respect to a basis such that A_m is a 2×2 block-diagonal matrix (2×2 blocks with an additional 1×1 block if n_m is odd) with 2×2 blocks in the form

$$\begin{pmatrix} 0 & 1 \\ * & * \end{pmatrix},$$

B_m is a full matrix, and

$$C_m = \begin{pmatrix} 1 & 0 & 1 & 0 & \cdots \\ * & * & * & * & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ * & * & * & * & \cdots \end{pmatrix}.$$

Observe that the Ly structure has $n_m(m+l)$ unknowns — n_m from A_m, $n_m m$ from B_m, and $n_m(l-1)$ from C_m.

It is assumed that (A_m, B_m, C_m) is in Ly's form. Let $\mathcal{I}$ be the set of indices of those elements of A_m which are parameters, i.e.,

$$\mathcal{I} \equiv \{(2,1), (2,2), \ldots, (n_m, n_m)\}.$$

To optimize $\mathcal{J}(A_m, B_m, C_m)$ over the open set $\mathcal{S}$ under the constraint that symmetric positive definite $\mathcal{Q}$ satisfies (3.7), and (A_m, B_m, C_m) is in Ly's form, the following Lagrangian is formed:

$$\mathcal{L}(A_m, B_m, C_m, \mathcal{P}, \mathcal{Q}) \equiv \text{tr}\left[\mathcal{Q}\tilde{R} + (\tilde{A}\mathcal{Q} + \mathcal{Q}\tilde{A}^T + \gamma^{-2}\mathcal{Q}\tilde{R}\mathcal{Q} + \tilde{V})\mathcal{P}\right],$$

where $\mathcal{P} \in \mathbf{R}^{\tilde{n} \times \tilde{n}}$ is a Lagrange multiplier. Setting $\partial\mathcal{L}/\partial\mathcal{Q} = 0$ yields (4.3). Partition $\mathcal{Q}$, $\mathcal{P} \in \mathbf{R}^{\tilde{n} \times \tilde{n}}$ as in (4.4) and define $\mathcal{P}\mathcal{Q} = Z$ as in (4.5). The partial derivatives of $\mathcal{L}$ can be computed as

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial (A_m)_{ij}} &= 2\left(\mathcal{P}_{12}^T \mathcal{Q}_{12} + \mathcal{P}_2 \mathcal{Q}_{2},\right)_{ij}, \qquad (i,j) \in \mathcal{I} \\
\frac{\partial \mathcal{L}}{\partial B_m} &= 2\left(\mathcal{P}_{12}^T B V + \mathcal{P}_2 B_m V\right), \\
\frac{\partial \mathcal{L}}{\partial (C_m)_{ij}} &= 2\left(R C_m \mathcal{Q}_2 - R C \mathcal{Q}_{12}\right)_{ij} + \\
&\quad + \gamma^{-2}\left[-RC(Z_1^T \mathcal{Q}_{12} + Z_{21}^T \mathcal{Q}_2 + \mathcal{Q}_1 Z_{12} + \mathcal{Q}_{12} Z_2)\right. \\
&\qquad \left. + RC_m(\mathcal{Q}_{12}^T Z_{12} + Z_{12}^T \mathcal{Q}_{12} + \mathcal{Q}_2 Z_2 + Z_2^T \mathcal{Q}_2)\right]_{ij}.
\end{aligned} \tag{6.1}$$

Now the equations to be solved are the partials in (6.1) set to zero. Define A_0, B_0, C_0, R_0, V_0, and γ_0 as in Section 4. Let A_f, B_f, C_f , R_f, V_f, and γ_f denote A, B, C, R, V, and γ in the above and define $A(\lambda)$, $B(\lambda)$, $C(\lambda)$, $R(\lambda)$, $V(\lambda)$, and $\gamma(\lambda)$ as in (4.9) and denote them by A, B, C, R, V, and γ respectively in the following. Let

$$
\begin{aligned}
H_{A_m}(\theta,\lambda) &= \frac{\partial \mathcal{L}}{\partial A_m} = 2\big(\mathcal{P}_{12}^T \mathcal{Q}_{12} + \mathcal{P}_2 \mathcal{Q}_2\big),\\
H_{B_m}(\theta,\lambda) &= \frac{\partial \mathcal{L}}{\partial B_m} = 2\big(\mathcal{P}_{12}^T B + \mathcal{P}_2 B_m\big) V,\\
H_{C_m}(\theta,\lambda) &= \frac{\partial \mathcal{L}}{\partial C_m} = 2R\big(C_m \mathcal{Q}_2 - C\mathcal{Q}_{12}\big)\\
&\qquad + \gamma^{-2}\big[-RC(Z_1^T \mathcal{Q}_{12} + Z_{21}^T \mathcal{Q}_2 + \mathcal{Q}_1 Z_{12} + \mathcal{Q}_{12} Z_2)\\
&\qquad\qquad + RC_m(\mathcal{Q}_{12}^T Z_{12} + Z_{12}^T \mathcal{Q}_{12} + \mathcal{Q}_2 Z_2 + Z_2^T \mathcal{Q}_2)\big],
\end{aligned}
$$

where in H_{A_m} only those elements corresponding to the parameter elements of A_m are of interest and

$$
\theta \equiv \begin{pmatrix} (A_m)_{\mathcal{I}} \\ \text{Vec}\,(B_m) \\ \text{Vec}\,(C_m)_{\mathcal{T}\cdot} \end{pmatrix} \tag{6.2}
$$

denotes the independent variables, $\mathcal{Q}$ and $\mathcal{P}$ satisfy respectively (3.7) and (4.3), $(A_m)_{\mathcal{I}}$ is a vector consisting of those elements in A_m with indices in the set $\mathcal{I}$, i.e.,

$$
(A_m)_{\mathcal{I}} = \big((A_m)_{21}, (A_m)_{22}, \cdots, (A_m)_{n_m n_m}\big)^T,
$$

and $(C_m)_{\mathcal{T}\cdot}$ is the matrix obtained from rows $\mathcal{T} = \{2, \ldots, l\}$ of C_m.

The homotopy map is defined as

$$
\rho(\theta,\lambda) = \begin{pmatrix} \big[H_{A_m}(\theta,\lambda)\big]_{\mathcal{I}} \\ \text{Vec}\,\big[H_{B_m}(\theta,\lambda)\big] \\ \text{Vec}\,\big[H_{C_m}(\theta,\lambda)\big]_{\mathcal{T}\cdot} \end{pmatrix}, \tag{6.3}
$$

and its Jacobian matrix is

$$
D\rho(\theta,\lambda) = \big(D_\theta \rho(\theta,\lambda), D_\lambda \rho(\theta,\lambda)\big).
$$

Define

$$\begin{aligned}
\hat{H}_{A_m}(\mathcal{P}^{(j)}, \mathcal{Q}^{(j)}) &= 2\big(\mathcal{P}_{12}^{T\,(j)}\mathcal{Q}_{12} + \mathcal{P}_{12}^T\mathcal{Q}_{12}^{(j)} + \mathcal{P}_2^{(j)}\mathcal{Q}_2 + \mathcal{P}_2\mathcal{Q}_2^{(j)}\big),\\
\hat{H}_{B_m}(\mathcal{P}^{(j)}) &= 2\big(\mathcal{P}_{12}^{T\,(j)}B + \mathcal{P}_2^{(j)}B_m\big)V,\\
\hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}) &= 2R\big(C_m\mathcal{Q}_2^{(j)} - C\mathcal{Q}_{12}^{(j)}\big)\\
&\quad - \gamma^{-2}RC\big(Z_1^{T\,(j)}\mathcal{Q}_{12} + Z_{21}^{T\,(j)}\mathcal{Q}_2 + Z_1^T\mathcal{Q}_{12}^{(j)} + Z_{21}^T\mathcal{Q}_2^{(j)}\\
&\qquad + \mathcal{Q}_1^{(j)}Z_{12} + \mathcal{Q}_1Z_{12}^{(j)} + \mathcal{Q}_{12}^{(j)}Z_2 + \mathcal{Q}_{12}Z_2^{(j)}\big)\\
&\quad + \gamma^{-2}RC_m\big(Z_{12}^{T\,(j)}\mathcal{Q}_{12} + Z_{12}^T\mathcal{Q}_{12}^{(j)} + \mathcal{Q}_{12}^{T\,(j)}Z_{12} + \mathcal{Q}_{12}^TZ_{12}^{(j)}\\
&\qquad + \mathcal{Q}_2^{(j)}Z_2 + Z_2^{T\,(j)}\mathcal{Q}_2 + \mathcal{Q}_2Z_2^{(j)} + Z_2^T\mathcal{Q}_2^{(j)}\big),
\end{aligned} \tag{6.4}$$

where the superscript (j) means $\partial/\partial\theta_j$. Using the above definitions, we have for $\theta_j = (A_m)_{kl}$, where $(k,l) \in \mathcal{I}$,

$$\begin{aligned}
\frac{\partial H_{A_m}}{\partial(A_m)_{kl}} &= \hat{H}_{A_m}(\mathcal{P}^{(j)}, Z^{(j)}),\\
\frac{\partial H_{B_m}}{\partial(A_m)_{kl}} &= \hat{H}_{B_m}(\mathcal{P}^{(j)}),\\
\frac{\partial H_{C_m}}{\partial(A_m)_{kl}} &= \hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}),
\end{aligned} \tag{6.5}$$

for $\theta_j = (B_m)_{kl}$,

$$\begin{aligned}
\frac{\partial H_{A_m}}{\partial(B_m)_{kl}} &= \hat{H}_{A_m}(\mathcal{P}^{(j)}, \mathcal{Q}^{(j)}),\\
\frac{\partial H_{B_m}}{\partial(B_m)_{kl}} &= \hat{H}_{B_m}(\mathcal{P}^{(j)}) + 2\mathcal{P}_2E^{(k,l)}V,\\
\frac{\partial H_{C_m}}{\partial(B_m)_{kl}} &= \hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}),
\end{aligned} \tag{6.6}$$

and for $\theta_j = (C_m)_{kl}$, where $k > 1$,

$$\begin{aligned}
\frac{\partial H_{A_m}}{\partial(C_m)_{kl}} &= \hat{H}_{A_m}(\mathcal{P}^{(j)}, \mathcal{Q}^{(j)}),\\
\frac{\partial H_{B_m}}{\partial(C_m)_{kl}} &= \hat{H}_{B_m}(\mathcal{P}^{(j)}),\\
\frac{\partial H_{C_m}}{\partial(C_m)_{kl}} &= \hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}) + 2RE^{(k,l)}\mathcal{Q}_2\\
&\quad + \gamma^{-2}RE^{(k,l)}\big(Z_{12}^T\mathcal{Q}_{12} + \mathcal{Q}_{12}^TZ_{12} + \mathcal{Q}_2^TZ_2 + Z_2^T\mathcal{Q}_2\big),
\end{aligned} \tag{6.7}$$

where $\mathcal{P}^{(j)}$ and $\mathcal{Q}^{(j)}$ can be obtained by solving the Lyapunov equation (4.14). Similarly for λ, using a dot to denote $\partial/\partial\lambda$,

$$\begin{aligned}
\frac{\partial H_{A_m}}{\partial\lambda} &= \hat{H}_{A_m}(\dot{\mathcal{P}},\dot{\mathcal{Q}}),\\
\frac{\partial H_{B_m}}{\partial\lambda} &= \hat{H}_{B_m}(\dot{\mathcal{P}}) + 2\mathcal{P}_{12}^T(\dot{B}V + B\dot{V}) + 2\mathcal{P}_2 B_m \dot{V},\\
\frac{\partial H_{C_m}}{\partial\lambda} &= \hat{H}_{C_m}(\dot{\mathcal{Q}},\dot{Z}) - 2(\dot{R}C + R\dot{C})\mathcal{Q}_{12} + 2\dot{R}C_m\mathcal{Q}_2\\
&\quad + \gamma^{-2}\dot{R}\,h_\lambda - 2\gamma^{-3}\dot{\gamma}R\,h_\lambda\\
&\quad - \gamma^{-2}R\dot{C}(Z_1^T\mathcal{Q}_{12} + Z_{21}^T\mathcal{Q}_2 + \mathcal{Q}_1 Z_{12} + \mathcal{Q}_{12}Z_2),
\end{aligned} \tag{6.8}$$

where

$$\begin{aligned}
h_\lambda &= -C(Z_1^T\mathcal{Q}_{12} + Z_{21}^T\mathcal{Q}_2 + \mathcal{Q}_1 Z_{12} + \mathcal{Q}_{12}Z_2)\\
&\quad + C_m(\mathcal{Q}_{12}^T Z_{12} + Z_{12}^T\mathcal{Q}_{12} + \mathcal{Q}_2 Z_2 + Z_2^T\mathcal{Q}_2),
\end{aligned}$$

and $\dot{\mathcal{P}}$ and $\dot{\mathcal{Q}}$ are obtained by solving (4.16).

Choose the initial γ so that γ_0^{-2} is approximately zero. The initial point $(\theta,\lambda) = (\theta_0, 0)$ is chosen so that the triple $((A_m)_0, (B_m)_0, (C_m)_0)$ is in Ly's form and satisfies $\rho(\theta_0, 0) = 0$. This can be done as follows:

1) Obtain the initial reduced order model $((A_m)_0, (B_m)_0, (C_m)_0)_b$ in balanced form in the same way as for the input normal form approach.
2) Transform the balanced $((A_m)_0, (B_m)_0, (C_m)_0)_b$ to Ly's form, and build θ_0 as described in (6.2).

The homotopy curve tracking computation is the same as described in Section 5.

7 Homotopy Algorithm

Homotopy Algorithm Based on Over-Parametrization Formulation

Now suppose all the components of A_m, B_m, C_m are treated as independent unknowns. To optimize $\mathcal{J}(A_m, B_m, C_m)$ over the open set $\mathcal{S}$ under the constraint that symmetric positive definite $\mathcal{Q}$ satisfies (3.7), the following Lagrangian is formed:

$$\mathcal{L}(A_m, B_m, C_m, \mathcal{P}, \mathcal{Q}) \equiv \mathrm{tr}\ \left[\mathcal{Q}\tilde{R} + (\tilde{A}\mathcal{Q} + \mathcal{Q}\tilde{A}^T + \gamma^{-2}\mathcal{Q}\tilde{R}\mathcal{Q} + \tilde{V})\mathcal{P}\right]$$

where $\mathcal{P} \in \mathbf{R}^{\tilde{n}\times\tilde{n}}$ is a Lagrange multiplier. Setting $\partial\mathcal{L}/\partial\mathcal{Q} = 0$ yields (4.3). Partition $\mathcal{Q}$, $\mathcal{P} \in \mathbf{R}^{\tilde{n}\times\tilde{n}}$ as in (4.4) and define $\mathcal{P}\mathcal{Q} = Z$ as in (4.5). A

Choose the initial γ so that γ_0^{-2} is approximately zero. The initial point $(\theta, \lambda) = (\theta_0, 0)$ is chosen so that the triple $\big((A_m)_0, (B_m)_0, (C_m)_0\big)$ is in balanced form and satisfies $\rho(\theta_0, 0) = 0$. This can be done as follows:

1) Obtain the initial reduced order model $\big((A_m)_0, (B_m)_0, (C_m)_0\big)_b$ in balanced form in the same way as for the input normal form approach.
2) Build θ_0 from $\big((A_m)_0, (B_m)_0, (C_m)_0\big)_b$ as described in (7.2).

The homotopy curve tracking computation is the same as described in Section 5.

8 Numerical Results

The following systems are solved by the homotopy algorithms discussed in the previous sections. The homotopy curve tracking was done with HOMPACK [16].

Systems with transfer functions of the form

$$H(s) = \frac{(s-1)^q}{(s+1)^{10}},$$

where $q = 0, \cdots, 4$, are studied. For these systems, controller canonical form realizations are given by

$$A = \begin{pmatrix} -10 & -45 & -120 & -210 & -250 & -210 & -120 & -45 & -10 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

$$B = (1 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0)^T.$$

For $q = 0$, $C = (0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 1)$.
For $q = 1$, $C = (0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; -1 \;\; 1)$.
For $q = 2$, $C = (0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 1 \;\; -2 \;\; 1)$.
For $q = 3$, $C = (0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 1 \;\; -3 \;\; 3 \;\; -1)$.
For $q = 4$, $C = (0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 1 \;\; -4 \;\; 6 \;\; -4 \;\; 1)$.

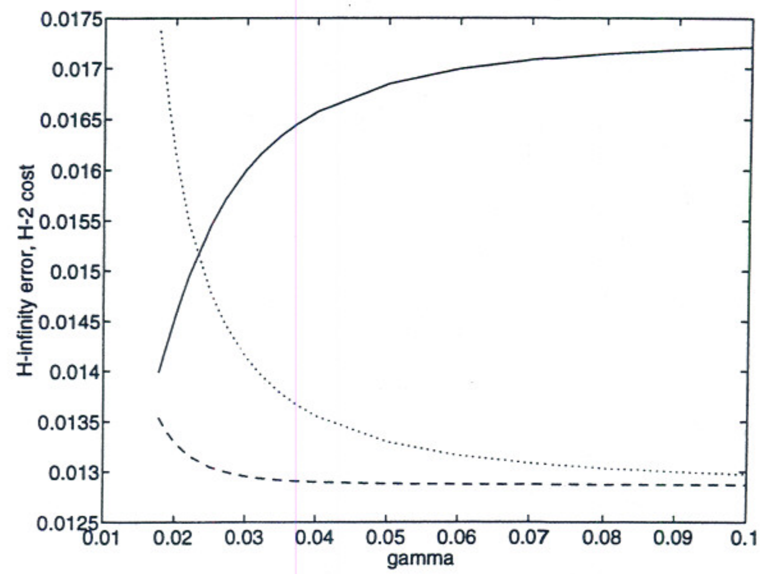

Fig. 1. $\|H(s) - H_m(s)\|_\infty$(solid), $100\,\mathcal{J}$(dotted), and $100\,J$(dashed) versus γ.

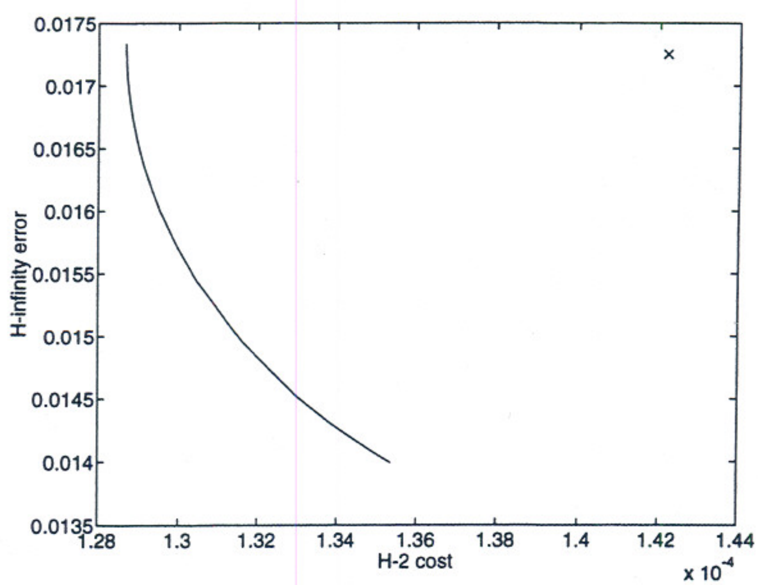

Fig. 2. $\|H(s) - H_m(s)\|_\infty$ versus J for $q = 0$, balanced model at "x".

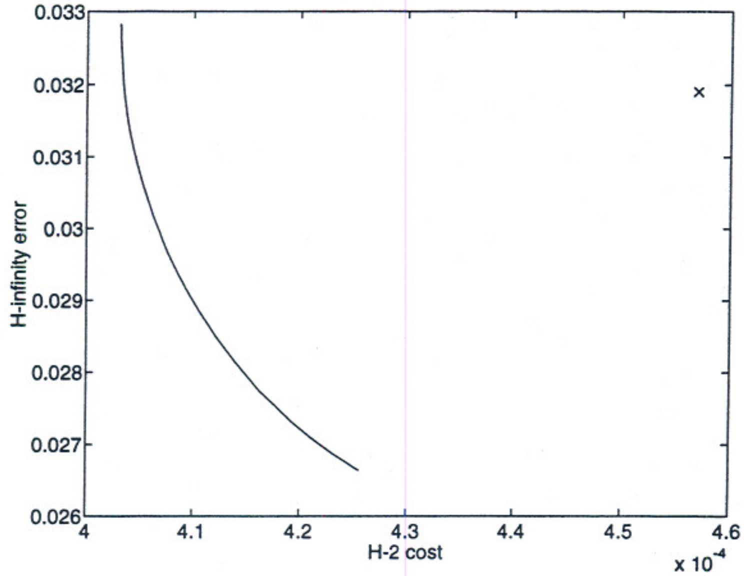

Fig. 3. $\|H(s) - H_m(s)\|_\infty$ versus J for $q = 1$, balanced model at "x".

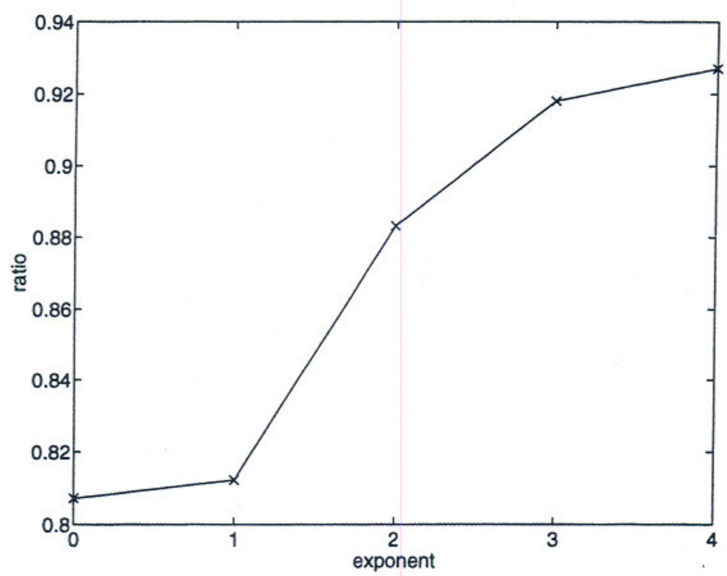

Fig. 4. Ratio of H^∞ error at $\gamma = \gamma_{min}$ to that at $\gamma = \infty$ versus exponent q.

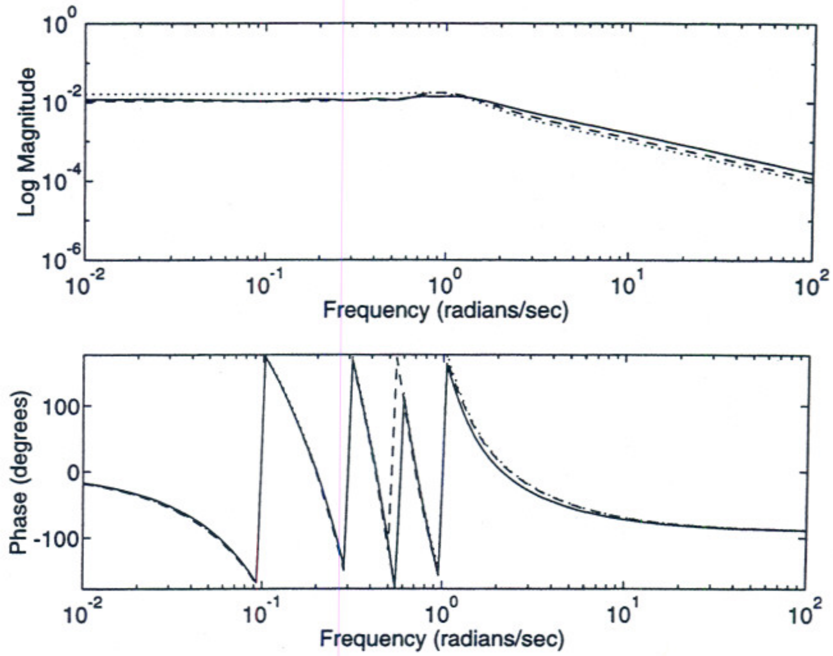

Fig. 5. Bode plots of $H(s) - H_m(s)$ for H^2/H^∞ (solid), H^2(dashed), and balanced (dotted) models with $q = 0$.

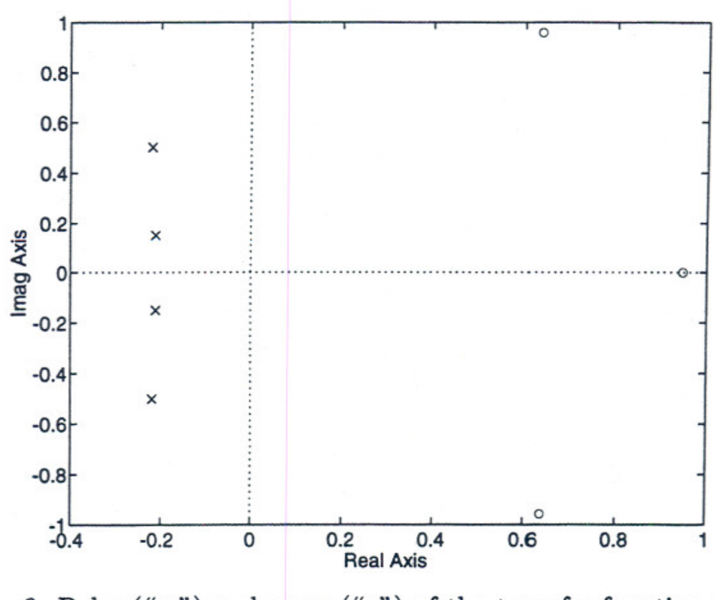

Fig. 6. Poles ("×") and zeros ("o") of the transfer function of the reduced order model for $q = 0$, $n_m = 4$ at $\gamma = 0.0178$.

The H^∞ error $\|H(s) - H_m(s)\|_\infty$ for the balanced reduced model of order 4 and the corresponding Enns-Glover bounds [5] [7] are:

q	H^∞ error	Enns-Glover bound
0	0.017251178	0.021958271
1	0.031901448	0.042197266
2	0.057214882	0.079880388
3	0.098520472	0.14841709
4	0.16125935	0.27001110

For $n_m = 4$ and $q = 0$, solutions of the auxiliary minimization problem are obtained for $\gamma \geq 0.0178$ using the input normal form approach. For $\gamma < \gamma_{min} \equiv 0.0178$, the Riccati equation solver fails and therefore no solution can be found. Let $H_m(s)$ be the transfer function of the reduced order model obtained by minimizing $\mathcal{J}$. In Figure 1, $\|H(s) - H_m(s)\|_\infty$ (solid line), $100\,\mathcal{J}$ (dotted line), and $100\,J$ (dashed line) are plotted against γ. As shown in the figure, as γ decreases, $\|H(s) - H_m(s)\|_\infty$ also decreases while both $\mathcal{J}$ and J increase. As can be seen from the figure, $\mathcal{J}$ is a close bound for J until γ becomes very small. To show the tradeoff between the H^2 cost and the H^∞ error $\|H(s) - H_m(s)\|_\infty$, it is useful to plot $\|H(s) - H_m(s)\|_\infty$ against J (with γ as the parameter of the curve), as shown in Figure 2. In Figure 2, the point marked by "$\times$" corresponds to the balanced reduced model, which has both large H^2 cost and large H^∞ error $\|H(s) - H_m(s)\|_\infty$, relative to the H^2/H^∞ reduced order model. The ratio of H^∞ error at $\gamma = \gamma_{min}$ to that at $\gamma = \infty$ is 0.8071, which indicates that there is about 20% improvement of the reduced order model with $\gamma = \gamma_{min}$ over the reduced order model without the H^∞ constraint. The reduced order models of order 2, 3, 6 were also found, and have qualitative behavior similar to the $n_m = 4$ case.

For $q = 1, \cdots, 4$, the same calculations are carried out. Figure 3 shows similar results to those in Figure 2 for $q = 1$ with an improvement of about 19%. As q increases, the improvement of the optimal reduced order model over the balanced reduced order model decreases. In Figure 4, the ratio of H^∞ error at $\gamma = \gamma_{min}$ to that at $\gamma = \infty$ is plotted against q for $q = 0, \cdots, 4$. The H^∞ norm improvement of the optimal reduced order model with the H^∞ constraint over that without the H^∞ constraint is $1 - ratio$. As q increases, the improvement decreases.

In Figure 5, the Bode plots of $H(s) - H_m(s)$ for the system with $q = 0$, $n_m = 4$, and $\gamma = \gamma_{min} = 0.0178$ are shown. The reduced order model with the H^∞ constraint at $\gamma = \gamma_{min}$ is shown by the solid line; the balanced reduced order model is shown by the dotted line; the reduced order model without the H^∞ constraint is shown by the dashed line. The magnitude plots show that as γ goes to γ_{min}, the H^∞ error becomes increasingly "all pass", that is, flat over a wide frequency range, which indicates H^∞

optimality of the reduced order model. Figure 6 shows the poles and zeros of the transfer function of the reduced order model for the system with $q = 0$ and $n_m = 4$ at $\gamma = \gamma_{min}$.

As another example, consider the system defined by

$$A = \begin{pmatrix} -2 & -8 \\ 0 & -8 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \quad C = B^T.$$

It is easy to verify that the system is balanced and the singular values are all equal to 1, i.e.,

$$A + A^T + BB^T = 0,$$
$$A^T + A + C^T C = 0.$$

The H^∞ error of the balanced reduced model of order 1 is 2 and the Enns-Glover bound is also 2. Optimal reduced models of order 1 are found by the input normal form homotopy approach for $\gamma \geq \gamma_{min} = 1.011$. Figure 7 shows the H^∞ error versus the H^2 cost. The point $\times$ corresponds to the balanced reduced order model. The ratio of the H^∞ error at $\gamma = \gamma_{min}$ to that at $\gamma = \infty$ is 0.6249, which indicates that there is about 37.5% improvement of the reduced order model with $\gamma = \gamma_{min}$ over the reduced order model without the H^∞ constraint, and a 50% improvement over the balanced reduced order model.

In Figure 8, $50\,\|H(s) - H_m(s)\|_\infty$ (solid line), $\mathcal{J}$ (dotted line), and J (dashed line) are plotted against γ. Unlike the previous systems, even for small γ, the actual error $\|H(s) - H_m(s)\|_\infty$ is very close to its bound γ, and $\mathcal{J}$ is a very close bound for the H^2 error J. The Bode plots of $H(s) - H_m(s)$ are shown in Figure 9, where the reduced order model with the H^∞ constraint at $\gamma = \gamma_{min}$ is shown by the solid line; the balanced reduced order model is shown by the dotted line; the reduced order model without the H^∞ constraint is shown by the dashed line. Again the reduced order model for $\gamma = \gamma_{min}$ indicates close to all pass model reduction error. The reduced order model transfer function at $\gamma = \gamma_{min}$ has a single pole at $s = -129.1642$.

9 Conclusion

One of the main conclusions for this study is that the more degrees of freedom that a formulation uses, the more robust is the resulting numerical algorithm. Both the input normal form and Ly form homotopies are very efficient for both the H^2 optimal and the combined H^2/H^∞ model reduction problems. However, they may fail to exist or be very ill conditioned [6]. The over-parametrization formulation solves the ill conditioning issue, but introduces singularity at the solution and may fail for a high dimensional system, which will inevitably have a high order singularity at the solution.

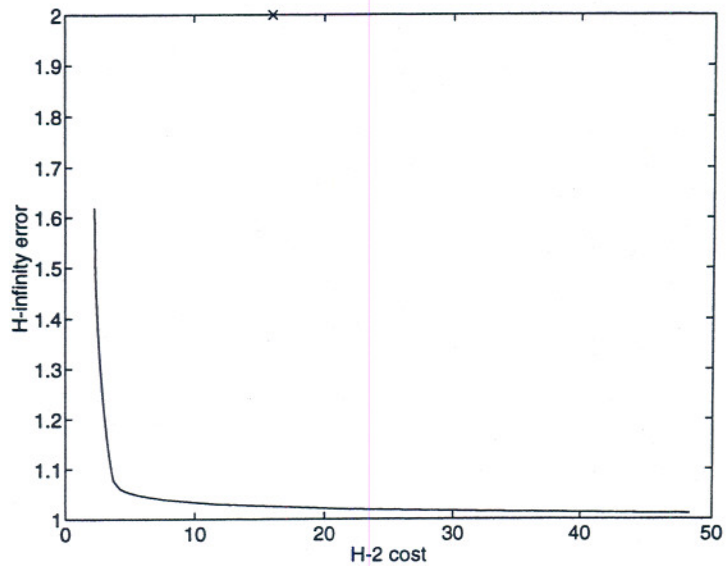

Fig. 7. $\|H(s) - H_m(s)\|_\infty$ versus J.

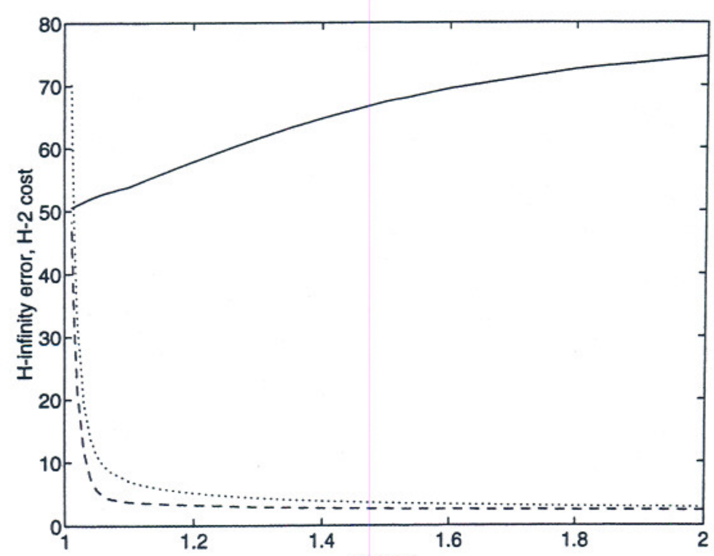

Fig. 8. $50\,\|H(s) - H_m(s)\|_\infty$(solid), $\mathcal{J}$(dotted), and J (dashed) versus γ.

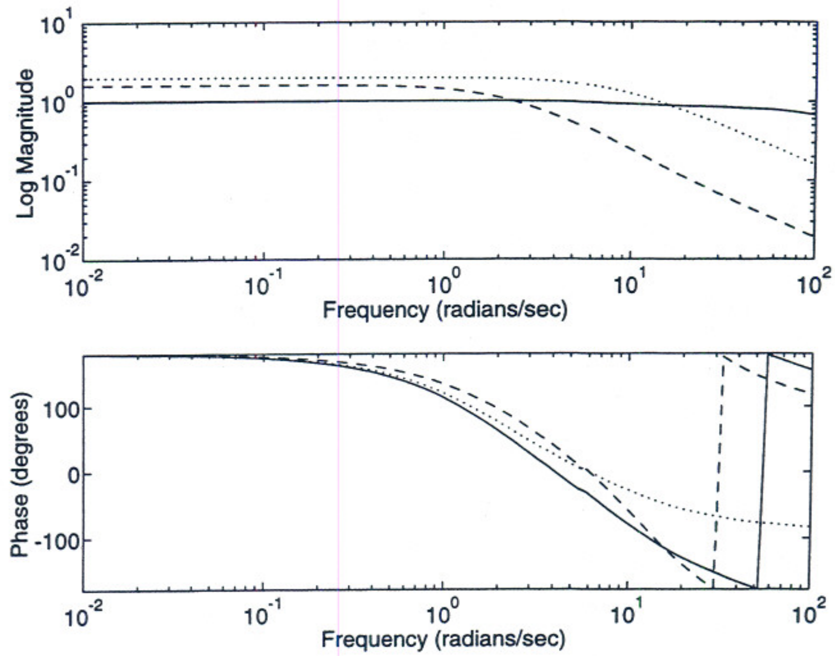

Fig. 9. Bode plots of $H(s) - H_m(s)$ for H^2/H^∞ (solid), H^2(dashed), and balanced (dotted) reduced order model.

Solving the H^2 optimal model order reduction problem may be well worth the effort (compared to simple balancing), as shown by the last example above. The examples also proved the worth of adding the H^∞ constraint, resulting in a difficult combined H^2/H^∞ problem. Finally, globally convergent homotopy methods are a viable approach to the computationally very difficult combined H^2/H^∞ model order reduction problem.

Acknowledgement

We acknowledge Y. Halevi for helpful discussions.

References

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

[2] J.A. Ball and A.C.M. Ran. Optimal Hankel norm model reductions and Wiener-Hopf factorization I: the canonical case, *SIAM J. Contr. Optim.,* **25** (1987), 362–382.

[3] R.F. Curtain. L_∞-approximations of complex functions and robust controllers for large flexible space structures, preprint, 1987.

[4] L.D. Davis, E.G. Collins, Jr., and S.A. Hodel. A parametrization of minimal plants, *Proc. 1992 American Contr. Conf.*, Chicago, IL, June 1992, 355–356.

[5] D.F. Enns. *Model Reduction for Control System Design.* Ph.D. Dissertation, Stanford University, 1984.

[6] Y.Z. Ge, E.G. Collins, L.T. Watson, and L.D. Davis. Minimal parameter homotopies for the L^2 optimal model order reduction problem, Tech. Rep. TR92–36, Dept. Computer Sci., Virginia Polytechnic Inst. & State Univ., Blacksburg, VA, 1992.

[7] K. Glover. All optimal Hankel-Norm approximations of linear multivariable systems and their L^∞-error bounds, *Int. J. Control,* **39** (1984), 1115–1193.

[8] W.M. Haddad and D.S. Bernstein. Combined L^2/H^∞ model reduction, *Int. J. Contr.,* **49** (1989), 1523–1535.

[9] D.C. Hyland and D.S. Bernstein. The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton and Moore, *IEEE Trans. Autom. Contr.,* **AC-30** (1985), 1201–1211.

[10] P.T. Kabamba. Balanced forms: canonicity and parametrization, *IEEE Trans. Autom. Contr.,* **AC-30** (1985), 1106–1109.

[11] U.-L. Ly, A.E. Bryson and R.H. Cannon. Design of low-order compensators using parameter optimization, *Automatica,* **21** (1985), 315–318.

[12] B.C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction, *IEEE Trans. Autom. Contr.,* **AC-26** (1981), 17–32.

[13] D. Mustafa and K. Glover. *Minimum Entropy H_∞ Control.* Berlin: Springer–Verlag, 1991.

[14] R.E. Skelton and P. Kabamba. Comments on 'Balanced gains and their significance for L^2 model reduction,' *IEEE Trans. Autom. Contr.,* **AC–31** (1986), 796–797.

[15] R.E. Skelton and A. Yousuff. Component cost analysis of large scale systems, *Int. J. Contr.,* **37** (1983), 285–304.

[16] L. Watson, S.C. Billups and A.P. Morgan. HOMPACK: a suite of codes for globally convergent homotopy algorithms, *ACM Trans. Math. Software,* **13** (1987), 281–310.

[17] D. Žigić. *Homotopy Methods for Solving the Optimal Projection Equations for the Reduced Order Model Problem.* M.S. thesis, Dept. of Computer Science, Virginia Polytechnic Institute and State Univ., Blacksburg, VA, June 1991.

Departments of Computer Science and Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0106

Departments of Computer Science and Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0106

Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902

Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109

Communicated by John Burns