A Homotopy Algorithm for the Combined H^2/H^∞ Model Reduction Problem

Yuzhen Ge, Emmanuel G. Collins,
Layne T. Watson, and Dennis S. Bernstein

TR 93-15

Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

May 1, 1993

A HOMOTOPY ALGORITHM FOR THE COMBINED H^2/H^∞ MODEL REDUCTION PROBLEM

Yuzhen Ge
Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

Emmanuel G. Collins
Harris Corporation, Government Aerospace Systems Division
Melbourne, FL 32902

Layne T. Watson
Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

Dennis S. Bernstein
Department of Aerospace Engineering
University of Michigan
Ann Arbor, MI 48109–2140

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 two formulations — input normal form; Ly, Bryson, and Cannon's 2×2 block parametrization — are developed and compared here.

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 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

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,Du(t), \\ y(t) &= C\,x(t), \end{aligned} \tag{1}$$

where $t \in [0, \infty)$, $A \in R^{n\times n}$ is asymptotically stable, $B \in R^{n\times m}$, $C \in R^{l\times n}$, $D \in R^{m\times p}$ $(m \le p)$ 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 $(n_m < n)$

$$\begin{aligned} \dot{x}_m(t) &= A_m\,x_m(t) + B_m\,Du(t), \\ y_m(t) &= C_m\,x_m(t), \end{aligned} \tag{2}$$

where $A_m \in R^{n_m\times n_m}$, $B_m \in R^{n_m\times m}$, $C_m \in R^{l\times n_m}$, 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.,

$$\|H(s) - H_m(s)\|_\infty \le \gamma \tag{3}$$

where $H(s) \equiv EC(sI_n - A)^{-1}BD$, $H_m(s) \equiv EC_m(sI_m - A_m)^{-1}B_mD$, $\gamma > 0$ is a given constant, $E \in R^{q\times l}$ $(q \ge l)$ is a given constant matrix; and

(iii) the H^2 model reduction criterion

$$J(A_m, B_m, C_m) \equiv \lim_{t\to\infty} \mathcal{E}\left[(y - y_m)^T R (y - y_m)\right] \tag{4}$$

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

2. The auxiliary minimization problem.

Define

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

$$\tilde{A} \equiv \begin{pmatrix} A & 0 \\ 0 & A_m \end{pmatrix}, \qquad \tilde{B} \equiv \begin{pmatrix} B \\ B_m \end{pmatrix}, \qquad \tilde{C} \equiv (C \quad -C_m), \tag{5}$$

$$\begin{aligned} \tilde{R} &\equiv \tilde{E}^T\tilde{E} = \tilde{C}^T R \tilde{C} = \begin{pmatrix} C^TRC & -C^TRC_m \\ -C_m^TRC & C_m^TRC_m \end{pmatrix}, \\ \tilde{V} &\equiv \tilde{D}\tilde{D}^T = \tilde{B}V\tilde{B}^T = \begin{pmatrix} BVB^T & BVB_m^T \\ B_mVB^T & B_mVB_m^T \end{pmatrix}. \end{aligned} \tag{6}$$

The full order system (1) and the reduced order system (2) can be written as a single augmented system

$$\dot{\tilde{x}}(t) = \tilde{A}\,\tilde{x}(t) + \tilde{D}\,u(t),$$
$$\tilde{y}(t) = \tilde{C}\,\tilde{x}(t). \tag{7}$$

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

$$\begin{aligned} J(A_m, B_m, C_m) &= \lim_{t\to\infty} \mathcal{E}\left[(y - y_m)^T R\,(y - y_m)\right] \\ &= \lim_{t\to\infty} \mathcal{E}(\tilde{y}^T R\,\tilde{y}) = \lim_{t\to\infty} \mathcal{E}(\tilde{x}^T \tilde{R}\,\tilde{x}) = \text{tr}\,(\tilde{Q}\,\tilde{R}), \end{aligned} \tag{8}$$

where $\tilde{Q}$ satisfies

$$\tilde{A}\,\tilde{Q} + \tilde{Q}\,\tilde{A}^T + \tilde{V} = 0. \tag{9}$$

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

$$\mathcal{Q} \quad \text{is symmetric and nonnegative definite} \tag{10}$$

and

$$\tilde{A}\,\mathcal{Q} + \mathcal{Q}\tilde{A}^T + \gamma^{-2}\mathcal{Q}\tilde{R}\mathcal{Q} + \tilde{V} = 0. \tag{11}$$

Then

$$(\tilde{A}, \tilde{D}) \quad \text{is stabilizable} \tag{12}$$

if and only if

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

Furthermore, if (12) holds, then

$$\|H(s) - H_m(s)\|_\infty \le \gamma, \tag{13}$$

$\tilde{Q} \le \mathcal{Q}$ $\quad$ ($\mathcal{Q} - \tilde{Q}$ 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 (11) 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 (10)–(12) 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 $J(A_m, B_m, C_m)$. (A_m, B_m, C_m) is restricted to the set

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

3. A homotopy approach based on the input normal form.

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

$$\begin{aligned} 0 &= A_m + A_m^T + B_m V B_m^T, \\ 0 &= A_m^T \Omega + \Omega A_m + C_m^T R C_m. \end{aligned} \tag{14}$$

In addition,

$$\begin{aligned} (A_m)_{ii} &= -\frac{1}{2}(B_m V B_m^T)_{ii}, \quad \omega_i = \frac{(C_m^T R C_m)_{ii}}{(B_m V B_m^T)_{ii}}, \\ (A_m)_{ij} &= \frac{(C_m^T R C_m)_{ij} - \omega_j (B_m V B_m^T)_{ij}}{\omega_j - \omega_i}, \qquad \text{if } \omega_i \neq \omega_j. \end{aligned} \tag{15}$$

Definition 1. The triple (A_m, B_m, C_m) satisfying (14) or (15) 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 (11), and (A_m, B_m, C_m) is in input normal form, the following Lagrangian is formed:

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

where the symmetric matrices M_c, M_o, and $\mathcal{P} \in R^{\tilde{n}\times\tilde{n}}$ are Lagrange multipliers. $\Omega =$ diag $(\omega_1, \ldots, \omega_{n_m})$ is related to the input normal form constraint. Setting $\partial\mathcal{L}/\partial\mathcal{Q} = 0$ yields

$$0 = (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T\mathcal{P} + \mathcal{P}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}) + \tilde{R}. \tag{16}$$

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

$$\mathcal{Q} = \begin{pmatrix} \mathcal{Q}_1 & \mathcal{Q}_{12} \\ \mathcal{Q}_{12}^T & \mathcal{Q}_2 \end{pmatrix}, \qquad \mathcal{P} = \begin{pmatrix} \mathcal{P}_1 & \mathcal{P}_{12} \\ \mathcal{P}_{12}^T & \mathcal{P}_2 \end{pmatrix} \tag{17}$$

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

$$\mathcal{P}\mathcal{Q} \equiv Z = \begin{pmatrix} Z_1 & Z_{12} \\ Z_{21} & Z_2 \end{pmatrix} \tag{18}$$

where

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

$\partial L/\partial\Omega = 0$ and $\partial L/\partial A_m = 0$ yield $0 = (A_m M_o)_{ii}$, and

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

A straightforward calculation shows

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial B_m} &= 2(\mathcal{P}_{12}^T BV + \mathcal{P}_2 B_m V) + 2M_c B_m V,\\ \frac{\partial \mathcal{L}}{\partial C_m} &= 2\big(RC_m \mathcal{Q}_2 - RC\mathcal{Q}_{12}\big) + 2RC_m M_o\\ &\quad + \gamma^{-2}\Big[-RC(Z_1^T \mathcal{Q}_{12} + Z_{21}^T \mathcal{Q}_2 + \mathcal{Q}_1 Z_{12} + \mathcal{Q}_{12} Z_2)\\ &\qquad + RC_m(\mathcal{Q}_{12}^T Z_{12} + Z_{12}^T \mathcal{Q}_{12} + \mathcal{Q}_2 Z_2 + Z_2^T \mathcal{Q}_2)\Big].\end{aligned} \tag{19}$$

Theorem 2 [4]. The matrices M_c and M_o in (19) satisfy

$$\begin{aligned} M_c &= -(\frac{1}{2}S + \Omega M_o),\\ \left(M_o\right)_{ii} &= -\frac{1}{\left(A_m\right)_{ii}} \sum_{\substack{j=1\\ j\neq i}}^{n_m} \left(A_m\right)_{ij} \left(M_o\right)_{ji},\\ \left(M_o\right)_{ij} &= \frac{(S)_{ij} - (S)_{ji}}{2(\omega_j - \omega_i)}, \quad \text{if } \omega_j \neq \omega_i,\end{aligned} \tag{20}$$

where

$$S = 2\big(\mathcal{P}_{12}^T \mathcal{Q}_{12} + \mathcal{P}_2 \mathcal{Q}_2\big). \tag{21}$$

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

$$\begin{aligned} A(\lambda) &= A_0 + \lambda(A_f - A_0), & R(\lambda) &= R_0 + \lambda(R_f - R_0),\\ B(\lambda) &= B_0 + \lambda(B_f - B_0), & V(\lambda) &= V_0 + \lambda(V_f - V_0),\\ C(\lambda) &= C_0 + \lambda(C_f - C_0), & \gamma(\lambda) &= \gamma_0 + \lambda(\gamma_f - \gamma_0).\end{aligned} \tag{22}$$

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

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

where $\theta \equiv (\,\text{Vec}\,(B_m) \quad \text{Vec}\,(C_m)\,)^T$ denotes the independent variables B_m and C_m, M_o and M_c satisfy (20), and $\mathcal{Q}$ and $\mathcal{P}$ satisfy respectively (11) and (16) with partitioned forms (17). Vec(P) for a matrix $P \in R^{p\times q}$ is the concatenation of its columns:

$$\text{Vec}(P) \equiv (\,P_{\cdot 1} \quad P_{\cdot 2} \quad \dots \quad P_{\cdot q}\,)^T \in \mathbf{R}^{pq}.$$

The homotopy map is defined as

$$\rho(\theta,\lambda) = (\,\text{Vec}\,\big[H_{B_m}(\theta,\lambda)\big] \quad \text{Vec}\,\big[H_{C_m}(\theta,\lambda)\big]\,)^T, \tag{23}$$

and its Jacobian matrix is

$$D\rho(\theta,\lambda) = \big(D_\theta \rho(\theta,\lambda), D_\lambda \rho(\theta,\lambda)\big). \tag{24}$$

Define

$$\hat{H}_{B_m}(\mathcal{P}^{(j)}, M_c^{(j)}) = 2(\mathcal{P}_{12}^{T\,(j)} B + \mathcal{P}_2^{(j)} B_m)V + 2M_c^{(j)} B_m V,$$
$$\hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}, M_o^{(j)}) = 2R(C_m \mathcal{Q}_2^{(j)} - C\mathcal{Q}_{12}^{(j)}) + 2RC_m M_o^{(j)}$$
$$- \gamma^{-2} RC\big(Z_1^{T\,(j)} \mathcal{Q}_{12} + Z_{21}^{T\,(j)} \mathcal{Q}_2 + Z_1^T \mathcal{Q}_{12}^{(j)} + Z_{21}^T \mathcal{Q}_2^{(j)}$$
$$+ \mathcal{Q}_1^{(j)} Z_{12} + \mathcal{Q}_1 Z_{12}^{(j)} + \mathcal{Q}_{12}^{(j)} Z_2 + \mathcal{Q}_{12} Z_2^{(j)}\big)$$
$$+ \gamma^{-2} RC_m\big(Z_{12}^{T\,(j)} \mathcal{Q}_{12} + Z_{12}^T \mathcal{Q}_{12}^{(j)} + \mathcal{Q}_{12}^{T\,(j)} Z_{12} + \mathcal{Q}_{12}^T Z_{12}^{(j)}$$
$$+ \mathcal{Q}_2^{(j)} Z_2 + Z_2^{T\,(j)} \mathcal{Q}_2 + \mathcal{Q}_2 Z_2^{(j)} + Z_2^T \mathcal{Q}_2^{(j)}\big),$$

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

$$\frac{\partial H_{B_m}}{\partial (B_m)_{kl}} = \hat{H}_{B_m}(\mathcal{P}^{(j)}, M_c^{(j)}) + 2(\mathcal{P}_2 + M_c)E^{(k,l)}V,$$
$$\frac{\partial H_{C_m}}{\partial (B_m)_{kl}} = \hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}, M_o^{(j)}), \tag{25}$$

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

$$\frac{\partial H_{B_m}}{\partial (C_m)_{kl}} = \hat{H}_{B_m}(\mathcal{P}^{(j)}, M_c^{(j)}),$$
$$\frac{\partial H_{C_m}}{\partial (C_m)_{kl}} = \hat{H}_{C_m}(\mathcal{Q}^{(j)}, Z^{(j)}, M_o^{(j)}) + 2RE^{(k,l)}(\mathcal{Q}_2 + M_o) \tag{26}$$
$$+ \gamma^{-2} RE^{(k,l)}\big(Z_{12}^T \mathcal{Q}_{12} + \mathcal{Q}_{12}^T Z_{12} + \mathcal{Q}_2^T Z_2 + Z_2^T \mathcal{Q}_2\big),$$

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

$$0 = (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})\mathcal{Q}^{(j)} + \mathcal{Q}^{(j)}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T + \tilde{V}^{(j)}$$
$$+ \tilde{A}^{(j)}\mathcal{Q} + \mathcal{Q}\tilde{A}^{T\,(j)} + \gamma^{-2}\mathcal{Q}\tilde{R}^{(j)}\mathcal{Q},$$
$$0 = (\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R})^T \mathcal{P}^{(j)} + \mathcal{P}^{(j)}(\tilde{A} + \gamma^{-2}\mathcal{Q}\tilde{R}) + \tilde{R}^{(j)} \tag{27}$$
$$+ (\tilde{A}^{(j)} + \gamma^{-2}\mathcal{Q}^{(j)}\tilde{R} + \gamma^{-2}\mathcal{Q}\tilde{R}^{(j)})^T \mathcal{P}$$
$$+ \mathcal{P}(\tilde{A}^{(j)} + \gamma^{-2}\mathcal{Q}^{(j)}\tilde{R} + \gamma^{-2}\mathcal{Q}\tilde{R}^{(j)}).$$

The computation of $D_\lambda\rho(\theta,\lambda)$ is similar to that of $D_\theta\rho(\theta,\lambda)$.

4. Numerical algorithm for input normal form homotopy.

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 $\Lambda = \text{diag}\,(d_1, d_2, \cdots, d_n)$ 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

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

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 $A_b = {}^{n_m}\{ \overbrace{\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}}^{n_m}$,

$$B_b = {}^{n_m}\{ \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad C_b = (\overbrace{C_1}^{n_m} \quad C_2).$$

3) (A_0, B_0, C_0) is chosen as $A_0 = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$, $B_0 = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}$, $C_0 = (C_1 \quad 0)$.

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

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

and $(\bar{A}_m)_0 = A_{11}$ by construction.

5) Transform the initial point $\left((\bar{A}_m)_0,\ (\bar{B}_m)_0,\ (\bar{C}_m)_0\right)$ to input normal form so that the initial reduced order model is $\left((A_m)_0,\ (B_m)_0,\ (C_m)_0\right) = \left(T^{-1}\,(\bar{A}_m)_0,\ T,\quad T^{-1}\,(\bar{B}_m)_0\ ,\ (\bar{C}_m)_0\, T\right)$.

The initial point for the homotopy map is then $(\theta_0, 0)$, where $\theta_0 = \left(\text{Vec } (B_m)_0 \quad \text{Vec } (C_m)_0 \right)^T$. (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 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 (11) and (16).
3) Evaluate S from (21) and M_o and M_c according to (20).
4) Evaluate the homotopy map $\rho(\theta, \lambda)$ in (23) and $D\rho(\theta, \lambda)$ in (24).
5) Predict the next point $Z^{(0)} = \left(\theta^{(0)}, \lambda^{(0)}\right)$ on the homotopy zero curve using, e.g., a Hermite cubic interpolant.
6) For $k := 0, 1, 2, \cdots$ until convergence do

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

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 $c_2 \geq 0$.

1) 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 $\big((A_m)_0, (B_m)_0, (C_m)_0\big)$ 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 (23) 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 (23), while in other cases it can be less efficient.

5. 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 $2{\times}2$ block-diagonal matrix ($2{\times}2$ blocks with an additional $1{\times}1$ 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 $C_m = \big((C_m)_1 \quad (C_m)_2 \quad \cdots \quad (C_m)_r \big)$ where

$$(C_m)_i = \begin{pmatrix} 1 & * & \cdots & * \\ 0 & * & \cdots & * \end{pmatrix}^T .$$

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 (11), and (A_m, B_m, C_m) is in Ly's form, the following Lagrangian is formed:

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

where $\mathcal{P} \in R^{\tilde{n}\times\tilde{n}}$ is a Lagrange multiplier. Setting $\partial\mathcal{L}/\partial\mathcal{Q} = 0$ yields (16). Partition $\mathcal{Q}, \mathcal{P} \in R^{\tilde{n}\times\tilde{n}}$ as in (17) and define $\mathcal{P}\mathcal{Q} = Z$ as in (18). The partial derivatives of $\mathcal{L}$ can be computed as

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

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 (22) and denote them by A, B, C, R, V, and γ respectively in the following. Let

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

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

$$\theta \equiv ((A_m)_{\mathcal{I}} \quad \mathrm{Vec}\,(B_m) \quad \mathrm{Vec}\,(C_m)_{\mathcal{T}\cdot})^T \tag{28}$$

denotes the independent variables, $\mathcal{Q}$ and $\mathcal{P}$ satisfy respectively (11) and (16), $(A_m)_{\mathcal{I}}$ is a vector consisting of those elements in A_m with indices in the set $\mathcal{I}$, i.e.,

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

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

$$\rho(\theta, \lambda) = \left(\left[H_{A_m}(\theta, \lambda)\right]_{\mathcal{I}} \quad \mathrm{Vec}\,\left[H_{B_m}(\theta, \lambda)\right] \quad \mathrm{Vec}\,\left[H_{C_m}(\theta, \lambda)\right]_{\mathcal{T}\cdot} \right)^T,$$

and its Jacobian matrix is $D\rho(\theta, \lambda) = \left(D_\theta \rho(\theta, \lambda), D_\lambda \rho(\theta, \lambda)\right)$.

The computation of $D\rho(\theta, \lambda)$ above is similar to that described for the input normal form in (23)–(27).

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 $\left((A_m)_0, (B_m)_0, (C_m)_0\right)$ 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 $\left((A_m)_0, (B_m)_0, (C_m)_0\right)_b$ in balanced form in the same way as for the input normal form approach.
2) Transform the balanced $\left((A_m)_0, (B_m)_0, (C_m)_0\right)_b$ to Ly's form, and build θ_0 as described in (28).

The homotopy curve tracking computation is the same as described in Section 4.

6. Numerical Results.

The results given here are all from the input normal form homotopy algorithm of Section 4. The homotopy curve tracking was done with HOMPACK [16].

EXAMPLE 1 [15]. The system is given by

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{pmatrix}, \quad B = (1 \quad 1 \quad 1), \quad C = (1 \quad 0 \quad 1).$$

A model of order $n_m = 2$ when $\gamma = 10$, with cost $\mathcal{J} = 0.678376$, is

$$A_m = \begin{pmatrix} -0.117649 & -0.493522 \\ 1.10166 & -0.785869 \end{pmatrix}, \quad B_m = \begin{pmatrix} -0.485076 \\ 1.25369 \end{pmatrix}, \quad C_m = (-0.751632 \quad -0.870253).$$

A model of order $n_m = 2$ when $\gamma = 1.0$, with cost $\mathcal{J} = 0.723313$, is

$$A_m = \begin{pmatrix} -0.112928 & -0.507912 \\ 1.10526 & -0.789927 \end{pmatrix}, \quad B_m = \begin{pmatrix} -0.475243 \\ 1.25692 \end{pmatrix}, \quad C_m = (-0.737429 \quad -0.896272).$$

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*, vol. 15, pp. 31–73, 1971.

[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.*, vol. 25, pp. 362–382, 1987.

[3] R. F. CURTAIN, "L_∞-approximations of complex functions and robust controllers for large flexible space structures," NASA Technical report, PB88–186226, TW–281, 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, pp. 355–356, June 1992.

[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,* vol. 39, pp. 1115–1193, 1984.

[8] W. M. HADDAD AND D. S. BERNSTEIN, "Combined L^2/H^∞ model reduction," *Int. J. Contr.*, vol. 49, pp. 1523–1535, 1989.

[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.*, vol. AC-30, pp. 1201–1211, 1985.

[10] P. T. KABAMBA, "Balanced forms: canonicity and parametrization," *IEEE Trans. Autom. Contr.*, vol. AC-30, pp. 1106–1109, 1985.

[11] U.-L. LY, A. E. BRYSON AND R. H. CANNON, "Design of low-order compensators using parameter optimization," *Automatica*, vol. 21, pp. 315–318, 1985.

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

[13] D. MUSTAFA AND K. GLOVER, "Minimum Entropy H_∞ Control," Springer–Verlag, Berlin, 1991.

[14] R. E. SKELTON AND P. KABAMBA, "Comments on 'Balanced gains and their significance for L^2 model reduction' ," *IEEE Trans. Autom. Contr.*, vol. AC–31, pp. 796–797, 1986.

[15] R. E. SKELTON AND A. YOUSUFF, "Component cost analysis of large scale systems," *Int. J. Contr.*, vol. 37, pp. 285–304, 1983.

[16] L. WATSON, S. C. BILLUPS AND A. P. MORGAN, "HOMPACK: a suite of codes for globally convergent homotopy algorithms," *ACM Trans. Math. Software*, vol. 13, pp. 281–310, 1987.