ELSEVIER Applied Mathematics and Computation 123 (2001) 155–185

APPLIED MATHEMATICS AND COMPUTATION

www.elsevier.com/locate/amc

Probability-one homotopy algorithms for solving the coupled Lyapunov equations arising in reduced-order $\mathrm{H}^2/\mathrm{H}^\infty$ modeling, estimation, and control

Y. Wang [a], D.S. Bernstein [b], L.T. Watson [c,*]

[a] *Analytic Mechanics Associates, 17 Research Drive, Hampton, VA 23666, USA*

[b] *Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109-2118, USA*

[c] *Departments of Computer Science and Mathematics, Virginia Polytechnic Institute and State University, 630 McBryde Hall, Blacksburg, VA 24061-0106, USA*

Abstract

Optimal reduced-order modeling, estimation, and control with respect to combined H^2/H^∞ criteria give rise to coupled Lyapunov and Riccati equations. To develop reliable numerical algorithms for these problems this paper focuses on the coupled Lyapunov equations which appear as a subset of the synthesis equations. In particular, this paper systematically examines the requirements of probability-one homotopy algorithms to guarantee global convergence. Homotopy algorithms for nonlinear systems of equations construct a continuous family of systems and solve the given system by tracking the continuous curve of solutions to the family. The main emphasis is on guaranteeing transversality for several homotopy maps based upon the pseudogramian formulation of the coupled Lyapunov equations and variations based upon canonical forms. These results are essential to the probability-one homotopy approach by guaranteeing good numerical properties in the computational implementation of the homotopy algorithms. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Embedding; Globally convergent homotopy; H^2 optimal model order reduction; Probability-one homotopy

* Corresponding author.
E-mail address: ltw@vtopus.cs.vt.edu (L.T. Watson).

PII: S0096-3003(00)00059-X

1. Introduction

While H^2 optimality constitutes the foundation of modern control theory, the trend during the past decade has been to combine H^2 performance with H^∞ design criteria to bound worst-case performance and to enforce robust stability [2,6,22,30]. This approach has also been applied to problems in model reduction and estimation [3,13,14].

In practice, controllers of low order are desirable, and considerable effort has been devoted to the development of plant and controller reduction methods [1,12,17,20,23,25,32,38,39]. Many of these techniques involve balancing the solutions of Riccati or Lyapunov equations. However, the balancing technique is nonoptimal and may entail instability when used for plant or controller reduction [18].

To achieve both stability and optimality, control gains can be optimized directly with respect to the H^2 criterion while enforcing H^∞ performance. This direct approach gives rise to gradients that can be used for search-based optimization [24] or which can be transformed into coupled Riccati and Lyapunov equations [15,16]. The essential numerical difficulty associated with these equations can be identified with the Lyapunov equations which are coupled by an idempotent matrix determined by the reduced rank pseudogramians [40,41]. The development of numerical techniques for solving these equations is thus of fundamental interest.

The purpose of this paper is to make substantial progress in developing reliable numerical algorithms for solving the coupled Lyapunov equations of reduced order H^2 and H^2/H^∞ synthesis. In particular, the present paper is concerned with the application of homotopy methods for solving the coupled Lyapunov equations arising in H^2 model reduction. In computational practice, homotopy methods are widely used for nonconvex optimization [33,37]. Homotopy methods, in particular, probability-one homotopy methods, have global convergence properties that are often advantageous in comparison to locally convergent methods such as quasi-Newton methods [4,34,35]. Under suitable hypotheses, probability-one homotopy methods are guaranteed to converge globally (from an arbitrary starting point) to a solution of a nonlinear system of equations. The nomenclature "probability-one" is well established in the mathematical literature and reflects the generic, measure theoretic properties of the algorithms rather than stochastic aspects.

The goal of the present paper is to systematically examine the requirements of probability-one homotopy methods to guarantee global convergence. The crucial requirements are (1) transversality and (2) boundedness. As discussed in Section 2, transversality implies the existence of and the ability to track a zero curve of the homotopy map, while boundedness is equivalent to the existence of solutions to the model reduction problem. The existence of optimal reduced-order H^2 models follows from the results in

[32]. The main emphasis in the present paper is on guaranteeing transversality for several homotopy maps based upon the pseudogramian formulation of the optimal projection equations and specialized formulations based upon canonical forms. These results are essential to the probability-one homotopy approach by guaranteeing good numerical properties (explained in [36]) in the computational implementation of the homotopy algorithms. Numerical comparisons with other approaches have been done elsewhere [10,41], and are not the objective of the present paper. Related work is in [7,8,11,21,29,31].

The contents of the paper are as follows. After stating the H^2 model reduction problem in Section 2, we then provide a brief review of probability-one homotopy theory in Section 3. The transversality assumption of probability-one homotopy theory is then proven in Section 4 for several canonical forms. Next, it is shown by example in Section 5 that the boundedness assumption required by probability-one homotopy theory is not always satisfied by the pseudogramian formulation of the optimal projection equations and by some formulations based upon canonical forms. Then it is shown that for a reformulation of the pseudogramian optimal projection equations in complex projective space using homogeneous transformations, the boundedness assumption holds and thus convergence of the homotopy algorithm to a solution (in complex projective space) is guaranteed. The numerical results in [10,41] show that, in practice, it is not necessary to track the homotopy zero curves in complex projective space. Section 6 concludes.

2. H^2 optimal model order reduction

The H^2 optimal model order reduction problem can be formulated as follows: given the nth-order asymptotically stable, controllable and observable linear time-invariant continuous-time system

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{2.1}$$

$$y(t) = Cx(t), \tag{2.2}$$

where $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$, and $C \in \mathbf{R}^{l\times n}$; given $n_m < n$, find an n_mth-order reduced-order model

$$\dot{x}_m(t) = A_m x_m(t) + B_m u(t), \tag{2.3}$$

$$y_m(t) = C_m x_m(t), \tag{2.4}$$

where $A_m \in \mathbf{R}^{n_m\times n_m}$ is asymptotically stable, $B_m \in \mathbf{R}^{n_m\times m}$, $C_m \in \mathbf{R}^{l\times n_m}$, which minimizes the quadratic model-reduction criterion

$$J(A_m, B_m, C_m) \equiv \lim_{t\to\infty} E\Big[(y(t) - y_m(t))^{\mathrm{T}} R(y(t) - y_m(t))\Big], \tag{2.5}$$

where the input $u(t)$ is white noise with positive definite intensity V, and R is a positive definite weighting matrix. Throughout, all positive semidefinite and positive definite matrices are assumed to be symmetric.

To guarantee that J is finite, a solution (A_m, B_m, C_m) is sought in the set $\mathscr{S} = \{(A_m, B_m, C_m)$: A_m is asymptotically stable, (A_m, B_m) is controllable, and (A_m, C_m) is observable$\}$. In this case the quadratic model reduction criterion (2.5) is given by

$$J(A_m, B_m, C_m) = \mathrm{tr}\,[\widetilde{Q}\widetilde{R}], \tag{2.6}$$

where

$$\widetilde{A} \equiv \begin{pmatrix} A & 0 \\ 0 & A_m \end{pmatrix}, \quad \widetilde{B} \equiv \begin{pmatrix} B \\ B_m \end{pmatrix}, \quad \widetilde{C} \equiv (\, C \quad -C_m \,), \quad \widetilde{R} \equiv \widetilde{C}^{\mathrm{T}} R \widetilde{C},$$

and

$$\widetilde{Q} = \int_0^\infty \mathrm{e}^{\widetilde{A}t} \widetilde{B} V \widetilde{B}^{\mathrm{T}} \mathrm{e}^{\widetilde{A}^{\mathrm{T}} t}\,\mathrm{d}t,$$

which is the unique solution of the Lyapunov equation

$$\widetilde{A}\widetilde{Q} + \widetilde{Q}\widetilde{A}^{\mathrm{T}} + \widetilde{B} V \widetilde{B}^{\mathrm{T}} = 0. \tag{2.7a}$$

For future reference define $\widetilde{P}$ by

$$\widetilde{A}^{\mathrm{T}}\widetilde{P} + \widetilde{P}\widetilde{A} + \widetilde{C}^{\mathrm{T}} R \widetilde{C} = 0, \tag{2.7b}$$

and partition $\widetilde{P}$, $\widetilde{Q}$ as

$$\widetilde{P} = \begin{pmatrix} P_1 & P_{12} \\ P_{12}^{\mathrm{T}} & P_2 \end{pmatrix}, \qquad \widetilde{Q} = \begin{pmatrix} Q_1 & Q_{12} \\ Q_{12}^{\mathrm{T}} & Q_2 \end{pmatrix}$$

in conformance with $\widetilde{A}$.

The following theorems and lemmas from [14] and [17] will be needed in Section 4.

Lemma 2.1. *Let positive semidefinite* $\hat{Q}, \hat{P} \in \mathbf{R}^{n\times n}$ *satisfy*

$$\mathrm{rank}\,(\hat{Q}) = \mathrm{rank}\,(\hat{P}) = \mathrm{rank}\,(\hat{Q}\hat{P}) = n_m, \tag{2.8}$$

where $n_m \leqslant n$. *Then there exist nonsingular* $W \in \mathbf{R}^{n\times n}$ *and positive definite diagonal* $\Sigma \in \mathbf{R}^{n_m\times n_m}$ *such that*

$$\hat{Q} = W\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} W^{\mathrm{T}}, \qquad \hat{P} = W^{-\mathrm{T}}\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} W^{-1}.$$

Lemma 2.2. *Let positive semidefinite $\hat{Q}, \hat{P} \in \mathbf{R}^{n\times n}$ satisfy the rank conditions* (2.8), *where $n_m < n$. Then, there exist $G, \Gamma \in \mathbf{R}^{n_m\times n}$ and positive semisimple $M \in \mathbf{R}^{n_m\times n_m}$, unique up to a change of basis in $\mathbf{R}^{n_m}$, such that*

$$\hat{Q}\hat{P} = G^{\mathrm{T}} M \Gamma, \qquad \Gamma G^{\mathrm{T}} = I_{n_m}. \tag{2.9}$$

Theorem 2.3. *Suppose $(A_m, B_m, C_m) \in \mathscr{S}$ solves the optimal model-reduction problem. Then there exist positive semidefinite matrices $\hat{Q}, \hat{P} \in \mathbf{R}^{n\times n}$ satisfying* (2.8) *and such that A_m, B_m, and C_m are given by*

$$A_m = \Gamma A G^{\mathrm{T}}, \quad B_m = \Gamma B, \quad C_m = C G^{\mathrm{T}}, \tag{2.10}$$

and such that, with $\tau \equiv G^{\mathrm{T}}\Gamma$, the following conditions are satisfied:

$$\tau[A\hat{Q} + \hat{Q}A^{\mathrm{T}} + BVB^{\mathrm{T}}] = 0, \tag{2.11}$$

$$[A^{\mathrm{T}}\hat{P} + \hat{P}A + C^{\mathrm{T}}RC]\tau = 0. \tag{2.12}$$

Throughout the paper, necessary equations for a minimum of J are being solved, hence only stationary points of J are being found.

3. Probability-one globally convergent homotopies

A *homotopy* is a continuous map from the interval [0,1] into a function space, where the continuity is with respect to the topology of the function space. Intuitively, a homotopy $\rho(\lambda)$ continuously deforms the function $\rho(0) = g$ into the function $\rho(1) = f$ as λ goes from 0 to 1. In this case, f and g are said to be *homotopic*. Homotopy maps are fundamental tools in topology, and provide a powerful mechanism for defining equivalence classes of functions.

Homotopies provide a mathematical formalism for describing an old procedure in numerical analysis, variously known as continuation, incremental loading, and embedding. The continuation procedure for solving a nonlinear system of equations $f(x) = 0$ starts with a (generally simpler) problem $g(x) = 0$ whose solution x_0 is known. The continuation procedure is to track the set of zeros of

$$\rho(\lambda, x) = \lambda f(x) + (1 - \lambda)g(x) \tag{3.1}$$

as λ is increased monotonically from 0 to 1, starting at the known initial point $(0, x_0)$ satisfying $\rho(0, x_0) = 0$. Each step of this tracking process is done by starting at a point $(\tilde{\lambda}, \tilde{x})$ on the zero set of ρ, fixing some $\Delta\lambda > 0$, and then solving $\rho(\tilde{\lambda} + \Delta\lambda, x) = 0$ for x using a locally convergent iterative procedure, which requires an invertible Jacobian matrix $D_x\rho(\tilde{\lambda} + \Delta\lambda, x)$. The process stops at $\lambda = 1$, since $f(\bar{x}) = \rho(1, \bar{x}) = 0$ gives a zero $\bar{x}$ of $f(x)$. Note that continuation assumes that the zeros of ρ connect the zero x_0 of g to a zero $\bar{x}$ of f, and that the Jacobian matrix $D_x\rho(\lambda, x)$ is invertible along the zero set of ρ; these are strong assumptions, which are frequently not satisfied in practice.

Continuation can fail because the curve γ of zeros of $\rho(\lambda, x)$ emanating from $(0, x_0)$ may (1) have turning points, (2) bifurcate, (3) fail to exist at some λ values, or (4) wander off to infinity without reaching $\lambda = 1$. Turning points and bifurcation correspond to singular $D_x\rho(\lambda, x)$. Generalizations of continuation known as *homotopy methods* attempt to deal with cases (1) and (2), and allow tracking of γ to continue through singularities. In particular, continuation monotonically increases λ, whereas homotopy methods permit λ to both increase and decrease along γ. Homotopy methods can also fail via cases (3) or (4).

The map $\rho(\lambda, x)$ connects the functions $g(x)$ and $f(x)$, hence the use of the word "homotopy." In general the homotopy map $\rho(\lambda, x)$ need not be a simple convex combination of g and f as in (3.1), and can involve λ nonlinearly. Sometimes λ is a physical parameter in the original problem $f(x; \lambda) = 0$, where $\lambda = 1$ is the (nondimensionalized) value of interest, although "artificial parameter" homotopies are generally more computationally efficient than "natural parameter" homotopies $\rho(\lambda, x) = f(x; \lambda)$. An example of an artificial parameter homotopy map is

$$\rho(\lambda, x) = \lambda f(x; \lambda) + (1 - \lambda)(x - a), \tag{3.2}$$

which satisfies $\rho(0, a) = 0$. The name "artificial" reflects the fact that solutions to $\rho(\lambda, x) = 0$ have no physical interpretation for $\lambda < 1$. Note that $\rho(\lambda, x)$ in (3.2) has a unique zero $x = a$ at $\lambda = 0$, regardless of the structure of $f(x; \lambda)$.

All four shortcomings of continuation and homotopy methods have been overcome by probability-one homotopies, proposed in [4]. The supporting theory, based on differential geometry, will be reformulated in less technical jargon here.

Definition 3.1. Let $U \subset \mathbf{R}^q$ be an open set, and let $\rho : U \to \mathbf{R}^p$ be a C^2 map. ρ is said to be *transversal to zero* if either $\rho^{-1}(0) = \emptyset$, or when $\rho^{-1}(0) \neq \emptyset$, the $p \times q$ Jacobian matrix $D\rho$ has full rank on $\rho^{-1}(0)$.

The C^2 requirement is technical, and part of the definition of transversality. The basis for the probability-one homotopy theory is:

Theorem 3.2 (Parametrized Sard's Theorem [4]). *Let $U \subset \mathbf{R}^m$ and $V \subset \mathbf{R}^p$ be open sets, and let $\rho : U \times [0, 1) \times V \to \mathbf{R}^p$ be a C^2 map. If ρ is transversal to zero, then for almost all $a \in U$ the map*

$$\rho_a(\lambda, x) = \rho(a, \lambda, x)$$

is also transversal to zero.

To discuss the import of this theorem, take $U = \mathbf{R}^m$, $V = \mathbf{R}^p$, and suppose that the C^2 map $\rho : \mathbf{R}^m \times [0,1) \times \mathbf{R}^p \to \mathbf{R}^p$ is transversal to zero. A straightforward application of the implicit function theorem yields that for almost all $a \in \mathbf{R}^m$, the zero set of ρ_a consists of smooth, nonintersecting curves which either (1) are closed loops lying entirely in $(0,1) \times \mathbf{R}^p$, (2) have both endpoints in $\{0\} \times \mathbf{R}^p$, (3) have both endpoints in $\{1\} \times \mathbf{R}^p$, (4) are unbounded with one endpoint in either $\{0\} \times \mathbf{R}^p$ or in $\{1\} \times \mathbf{R}^p$, or (5) have one endpoint in $\{0\} \times \mathbf{R}^p$ and the other in $\{1\} \times \mathbf{R}^p$. Furthermore, for almost all $a \in \mathbf{R}^m$, the Jacobian matrix $D\rho_a$ has full rank at *every* point in $\rho_a^{-1}(0)$. The goal is to construct a map ρ_a whose zero set has an endpoint in $\{0\} \times \mathbf{R}^p$, and which rules out (2) and (4). Then (5) obtains, and a zero curve starting at $(0, x_0)$ is *guaranteed* to reach a point $(1, \bar{x})$. All of this holds for almost all $a \in \mathbf{R}^m$, and hence with probability one [4]. Furthermore, since $a \in \mathbf{R}^m$ can be almost any point (and, indirectly, so can the starting point x_0), an algorithm based on tracking the zero curve in (5) is legitimately called *globally convergent*. This discussion is summarized in the following theorem.

Theorem 3.3. *Let $f : \mathbf{R}^p \to \mathbf{R}^p$ be a C^2 map, $\rho : \mathbf{R}^m \times [0,1) \times \mathbf{R}^p \to \mathbf{R}^p$ a C^2 map, and $\rho_a(\lambda, x) = \rho(a, \lambda, x)$. Suppose that*

(1) *ρ is transversal to zero,*

and, for each fixed $a \in \mathbf{R}^m$,

(2) *$\rho_a(0, x) = 0$ has a unique nonsingular solution x_0,*

(3) *$\rho_a(1, x) = f(x) \quad (x \in \mathbf{R}^p)$.*

Then, for almost all $a \in \mathbf{R}^m$, there exists a zero curve γ of ρ_a emanating from $(0, x_0)$, along which the Jacobian matrix $D\rho_a$ has full rank. If, in addition,

(4) *$\rho_a^{-1}(0)$ is bounded,*

then γ reaches a point $(1, \bar{x})$ such that $f(\bar{x}) = 0$. Furthermore, if $Df(\bar{x})$ is invertible, then γ has finite arc length.

Any algorithm for tracking γ from $(0, x_0)$ to $(1, \bar{x})$, based on a homotopy map satisfying the hypotheses of Theorem 3.3, is called a *globally convergent probability-one homotopy algorithm*. Of course the practical numerical details of tracking γ are nontrivial, and have been the subject of twenty years of research in numerical analysis. Production quality software called HOMPACK [36] exists for tracking γ. The distinctions between continuation, homotopy methods, and probability-one homotopy methods are subtle but worth noting. Only the latter are provably globally convergent and (by construction) expressly avoid dealing with singularities numerically, unlike continuation and homotopy methods which must explicitly handle singularities numerically.

The purpose of this paper is to prove or disprove properties (1)–(4) of Theorem 3.3 for some homotopy maps that have been proposed for the H^2 optimal model order reduction problem, and which have been successful in practice. Assumptions (2) and (3) in Theorem 3.3 are usually achieved by the construction of ρ (such as (3.2)), and are straightforward to verify. Although assumption (1) is trivial to verify for some maps, for the H^2 model order reduction homotopies the verification is nontrivial. Assumption (4) is typically very hard to verify, and often is a deep result, since (1)–(4) holding implies the *existence* of a solution to $f(x) = 0$.

Note that (1)–(4) are sufficient, but not necessary, for the existence of a solution to $f(x) = 0$, which is why homotopy maps not satisfying the hypotheses of Theorem 3.3 can still be very successful on practical problems. If (1)–(3) hold and a solution does *not* exist, then (4) must fail, and nonexistence is manifested by γ going off to infinity. Properties (1)–(3) are important because they guarantee good numerical properties along the zero curve γ, which, if bounded, results in a *globally convergent* algorithm. If γ is unbounded, then either the homotopy approach (with this particular ρ) has failed or $f(x) = 0$ has no solution. Furthermore, the goal is not to simply get a map ρ_a which is transversal to zero (small perturbations $\rho_a(\lambda, x) + \epsilon(\lambda, x)$ are transversal to zero), but to construct maps ρ_a which are transversal to zero *throughout a large* neighborhood of some a. This requires starting with a map ρ already transversal to zero.

4. Transversality of homotopies for H^2 optimal model order reduction

This section proves that three homotopies $\rho(a, \lambda, x)$ which have been used in [9,41] for the H^2 optimal model order reduction problem are transversal to zero, the first requirement of Theorem 3.3. An overview and comparison of these homotopy maps is in [10]. The analysis concerns (2.11) and (2.12) where $\hat{Q}$ and $\hat{P}$ are positive semidefinite matrices satisfying (2.8).

4.1. Transversality of homotopies based on decompositions of pseudogramians

Since $\hat{Q}$ and $\hat{P}$ satisfy (2.8), there exists invertible $W \in \mathbf{R}^{n \times n}$ and positive definite diagonal $\Sigma \in \mathbf{R}^{n_m \times n_m}$ such that [17]

$$\hat{Q} = W \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} W^{\mathrm{T}} = W_1 \Sigma W_1^{\mathrm{T}}, \qquad \hat{P} = W^{-\mathrm{T}} \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} W^{-1} = U_1^{\mathrm{T}} \Sigma U_1,$$

where

$$W = (\overbrace{W_1}^{n_m} \quad W_2), \quad W^{-1} = U = {}^{n_m}\{ \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}.$$

Premultiplying (2.11) by U_1 and postmultiplying (2.12) by W_1 yields (recall that $\tau = G^{\mathrm{T}} \Gamma = W_1 U_1$)

$$U_1 A W_1 \Sigma W_1^{\mathrm{T}} + \Sigma W_1^{\mathrm{T}} A^{\mathrm{T}} + U_1 B V B^{\mathrm{T}} = 0, \tag{4.1}$$

$$A^{\mathrm{T}} U_1^{\mathrm{T}} \Sigma + U_1^{\mathrm{T}} \Sigma U_1 A W_1 + C^{\mathrm{T}} R C W_1 = 0. \tag{4.2}$$

A constraint from $W^{-1} = U$ is

$$U_1 W_1 - I = 0. \tag{4.3}$$

The matrix equations (4.1)–(4.3) contain $2nn_m + n_m^2$ scalar equations. However, the only unknowns in (4.1)–(4.3), namely W_1, U_1, and diagonal Σ, contain $2nn_m + n_m$ variables. Hence, some other formulation is necessary in order to make an exact match between the number of equations and the number of unknowns. Following [41], all n_m^2 elements of Σ are considered as unknowns, giving the same number of equations as unknowns. The structure of the problem is such that Σ will turn out to be symmetric, so it can be diagonalized to produce the decomposition of $\hat{Q}$ and $\hat{P}$ described above.

The approach in [41], analyzed next, uses the homotopy map

$$\rho_a(\lambda, x) \equiv \lambda f(x) + (1 - \lambda) g(x; a), \tag{4.4}$$

where the initial problem $\rho_a(0, x) = g(x; a) = 0$ has an easily obtained unique solution and the final problem (4.1)–(4.3) is $\rho_a(1, x) = f(x) = 0$. f and g are displayed in (4.4) simply to point out that the map $\rho_a(\lambda, x)$ can be viewed as a convex combination of two other maps. For notational convenience later when displaying Jacobian matrices the order of the variables is henceforth taken as λ, x, a. Let

$$A(\lambda) = A, \quad B(\lambda) = \lambda B V B^{\mathrm{T}} + (1 - \lambda) B_i, \quad C(\lambda) = \lambda C^{\mathrm{T}} R C + (1 - \lambda) C_i,$$

where $B_i = B(0)$ and $C_i = C(0)$ are matrices defining the initial problem at $\lambda = 0$, and correspond to the parameter vector a in Theorem 3.3. Define

$$\rho_a(\lambda, x) \equiv \rho(\lambda, x, a) \equiv \begin{pmatrix} F_1(\lambda, x, a) \\ F_2(\lambda, x, a) \\ F_3(\lambda, x, a) \end{pmatrix}$$

in (4.4) by

$$F_1(\lambda, x, a) \equiv U_1 A(\lambda) W_1 \Sigma W_1^{\mathrm{T}} + \Sigma W_1^{\mathrm{T}} A^{\mathrm{T}}(\lambda) + U_1 B(\lambda), \tag{4.5}$$

$$F_2(\lambda, x, a) \equiv A^{\mathrm{T}}(\lambda) U_1^{\mathrm{T}} \Sigma + U_1^{\mathrm{T}} \Sigma U_1 A(\lambda) W_1 + C(\lambda) W_1, \tag{4.6}$$

$$F_3(\lambda, x, a) \equiv U_1 W_1 - I, \tag{4.7}$$

where

$$a \equiv \begin{pmatrix} \mathrm{Vec}\,(B_i) \\ \mathrm{Vec}\,(C_i) \end{pmatrix}$$

is the generic parameter vector in Theorem 3.3 and in (4.4),

$$x \equiv \begin{pmatrix} \mathrm{Vec}\,(W_1) \\ \mathrm{Vec}\,(U_1) \\ \mathrm{Vec}\,(\Sigma) \end{pmatrix}$$

denotes the independent variables $W_1 \in \mathbf{R}^{n\times n_m}$, $U_1 \in \mathbf{R}^{n_m\times n}$, $\Sigma \in \mathbf{R}^{n_m\times n_m}$ corresponding to x in Theorem 3.3, and A, B, C, V, R are constants as in Section 2.

The Jacobian matrix of $\rho(\lambda, x, a)$ has $2nn_m + n_m^2$ rows and $2n^2 + 2nn_m + n_m^2 + 1$ columns. Rows 1 through nn_m correspond to (4.5), rows $nn_m + 1$ through $2nn_m$ correspond to (4.6), and rows $2nn_m + 1$ through $2nn_m + n_m^2$ correspond to (4.7). The first column corresponds to the derivatives with respect to λ, columns 2 through $nn_m + 1$ correspond to the derivatives with respect to W_1, columns $nn_m + 2$ through $2nn_m + 1$ correspond to the derivatives with respect to U_1, columns $2nn_m + 2$ through $2nn_m + n_m^2 + 1$ correspond to the derivatives with respect to Σ, columns $2nn_m + n_m^2 + 2$ through $2nn_m + n_m^2 + n^2 + 1$ correspond to the derivatives with respect to B_i, and columns $2nn_m + n_m^2 + n^2 + 2$ through $2nn_m + n_m^2 + 2n^2 + 1$ correspond to the derivatives with respect to C_i:

$$D\rho(\lambda, x, a) = (\, D_\lambda\rho \quad D_{W_1}\rho \quad D_{U_1}\rho \quad D_\Sigma\rho \quad D_{B_i}\rho \quad D_{C_i}\rho\,). \tag{4.8}$$

Since $F_3(\lambda, x, a)$ does not depend upon λ, B_i, and C_i, it follows that

$$D_\lambda F_3(\lambda, x, a) = 0,$$
$$D_{B_i} F_3(\lambda, x, a) = 0,$$
$$D_{C_i} F_3(\lambda, x, a) = 0,$$

and similarly

$$D_{C_i} F_1(\lambda, x, a) = D_{B_i} F_2(\lambda, x, a) = 0.$$

Thus

$$\begin{aligned} D\rho(\lambda, x, a) = D\rho(\lambda, W_1, U_1, \Sigma, B_i, C_i) &= \begin{pmatrix} D_\lambda F_1 & D_x F_1 & D_a F_1 \\ D_\lambda F_2 & D_x F_2 & D_a F_2 \\ 0 & D_x F_3 & 0 \end{pmatrix} \\ &= \begin{pmatrix} D_\lambda F_1 & D_{W_1} F_1 & D_{U_1} F_1 & D_\Sigma F_1 & D_{B_i} F_1 & 0 \\ D_\lambda F_2 & D_{W_1} F_2 & D_{U_1} F_2 & D_\Sigma F_2 & 0 & D_{C_i} F_2 \\ 0 & D_{W_1} F_3 & D_{U_1} F_3 & D_\Sigma F_3 & 0 & 0 \end{pmatrix} \end{aligned} \tag{4.9}$$

The following lemma will be used in the proof of Theorem 4.2:

Lemma 4.1. *Let $X \in \mathbf{R}^{p\times q}$ and $A \in \mathbf{R}^{n\times m}$, $B \in \mathbf{R}^{m\times l}$ be differentiable with respect to x_{ij} for $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant q$. Then*

$$\frac{\partial}{\partial x_{ij}(AB)} = \left(\frac{\partial}{\partial x_{ij}A}\right)B + A\left(\frac{\partial}{\partial x_{ij}B}\right),$$

and for constant M, interpreting the derivative D_X as $D_{\mathrm{Vec}(X)}$,

$$D_X(MX) = I \otimes M, \quad D_X(XM) = M^{\mathrm{T}} \otimes I.$$

The proof of Lemma 4.1 is straightforward calculus.

Theorem 4.2. *The homotopy map given by* (4.5)–(4.7) *is transversal to zero* (*for* $0 \leqslant \lambda < 1$).

Proof. To prove that $D\rho(\lambda, x, a)$ given in (4.9) has full rank, i.e.,

$$\operatorname{rank}(D\rho(\lambda, x, a)) = 2nn_m + n_m^2,$$

it suffices to prove that

$$\operatorname{rank}(D_x F_3) = \operatorname{rank}(\, D_{W_1}F_3 \quad D_{U_1}F_3 \quad D_\Sigma F_3\,) = n_m^2, \tag{4.10}$$

$$\operatorname{rank}(D_a F_1) = \operatorname{rank}(\, D_{B_i}F_1 \quad 0\,) = nn_m, \tag{4.11}$$

$$\operatorname{rank}(D_a F_2) = \operatorname{rank}(\, 0 \quad D_{C_i}F_2\,) = nn_m. \tag{4.12}$$

The meaning of expressions like $D_\Sigma F_3$ is ambiguous until some ordering is specified for the components of the matrices Σ and F_3. Hereafter, whichever ordering is notationally convenient is used. If unspecified, the standard ordering by columns (Vec) is assumed.

Using Lemma 4.1, ordering U_1 and F_3 by rows,

$$D_{U_1}F_3(\lambda, x, a) = D_{U_1}(U_1 W_1) = I_{n_m} \otimes W_1^{\mathrm{T}}, \tag{4.13}$$

and ordering W_1 and F_3 by columns,

$$D_{W_1}F_3(\lambda, x, a) = D_{W_1}(U_1 W_1) = I_{n_m} \otimes U_1. \tag{4.14}$$

Since $U_1 W_1 = I$, by Sylvester's inequality,

$$\operatorname{rank}(U_1) = \operatorname{rank}(W_1) = n_m,$$

and therefore

$$\operatorname{rank}(D_x F_3) = \operatorname{rank}(D_{U_1}F_3) = \operatorname{rank}(D_{W_1}F_3) = n_m^2,$$

which is (4.10).

Using Lemma 4.1, ordering B_i and F_1 by columns yields

$$\begin{aligned} D_{B_i}F_1(\lambda, x, a) &= D_{B_i}(U_1 B(\lambda)) = (1-\lambda)D_{B_i}(U_1 B_i) \\ &= (1-\lambda)I_n \otimes U_1, \end{aligned} \tag{4.15}$$

and using (4.15) for $\lambda < 1$ yields

$$\operatorname{rank}(D_{B_i}F_1) = nn_m.$$

Similarly, ordering C_i and F_2 by rows,

$$D_{C_i}F_2(\lambda, x, a) = D_{C_i}(C(\lambda)W_1) = (1-\lambda)D_{C_i}(C_i W_1) = (1-\lambda)I_n \otimes W_1^{\mathrm{T}}, \tag{4.16}$$

so for $\lambda < 1$

$$\operatorname{rank}(D_{C_i}F_2) = nn_m.$$

This completes the proof of (4.10)–(4.12), and the proof that the homotopy map (4.5)–(4.7) is transversal to zero for all $0 \leqslant \lambda < 1$. □

Remark 4.2.1. One can use more variables in the parameter vector a, e.g., $A(\lambda) = \lambda A + (1-\lambda)A_i$, without affecting the full rank properties.

4.2. Transversality of homotopies based on input normal form

The following theorem from [19] is needed to present the homotopy method for the input normal form.

Theorem 4.3. *Suppose $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$ is asymptotically stable and minimal. Then there exist a similarity transformation U and a positive definite matrix $\Omega = \operatorname{diag}(\omega_1, \cdots, \omega_{n_m})$ 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*

$$\begin{aligned} &A_m + A_m^{\mathrm{T}} + B_m V B_m^{\mathrm{T}} = 0, \\ &A_m^{\mathrm{T}}\Omega + \Omega A_m + C_m^{\mathrm{T}} R C_m = 0. \end{aligned} \tag{4.17}$$

In addition, if the ω_i are distinct,

$$\begin{aligned} &(A_m)_{ii} = -\frac{1}{2}(B_m V B_m^{\mathrm{T}})_{ii}, \\ &\omega_i = \frac{(C_m^{\mathrm{T}} R C_m)_{ii}}{(B_m V B_m^{\mathrm{T}})_{ii}}, \\ &(A_m)_{ij} = \frac{(C_m^{\mathrm{T}} R C_m)_{ij} - \omega_j (B_m V B_m^{\mathrm{T}})_{ij}}{\omega_j - \omega_i}. \end{aligned} \tag{4.18}$$

Definition 4.3.1. The triple (A_m, B_m, C_m) satisfying (4.17) and (4.18) is said to be in *input normal form*.

The utility of the input normal form (4.17) and (4.18) lies in using B_m and C_m as the independent variables, and then being able to recover A_m uniquely from B_m and C_m. The number of variables in B_m and C_m is $n_m(m+l)$, the minimum number of variables possible to describe any reduced order model, and thus the input normal form parameterization is referred to as a "minimal parameterization". If $\omega_i = \omega_j$ for some $i \neq j$, then, regardless of (4.17) holding, (4.18) fails to permit the unique recovery of A_m.

Under the assumption that the solution (A_m, B_m, C_m) being sought exists in input normal form, the only independent variables are B_m and C_m, and in this case the domain is

$$\{(A_m, B_m, C_m)\colon A_m \text{ is asymptotically stable},\\ (A_m, B_m, C_m) \text{ is minimal and in input normal form}\}.$$

Now for (A_m, B_m, C_m) in input normal form, the cost function can be written as

$$J(A_m, B_m, C_m) = \operatorname{tr}(\tilde{Q}_I \tilde{R}_I), \tag{4.19}$$

where $\tilde{Q}_I$ is a symmetric and positive definite matrix satisfying

$$\tilde{A}_I \tilde{Q}_I + \tilde{Q}_I \tilde{A}_I^{\mathrm{T}} + \tilde{V}_I = 0, \tag{4.20}$$

and

$$\tilde{A}_I = \begin{pmatrix} A & 0 \\ 0 & A_m, \end{pmatrix}, \qquad \tilde{R}_I = \begin{pmatrix} C^{\mathrm{T}} R C & -C^{\mathrm{T}} R C_m \\ -C_m^{\mathrm{T}} R C & C_m^{\mathrm{T}} R C_m \end{pmatrix},$$

$$\tilde{V}_I = \begin{pmatrix} B V B^{\mathrm{T}} & B V B_m^{\mathrm{T}} \\ B_m V B^{\mathrm{T}} & B_m V B_m^{\mathrm{T}} \end{pmatrix}. \tag{4.21}$$

$\tilde{Q}_I$ can be written as

$$\tilde{Q}_I = \begin{pmatrix} \bar{Q}_1 & \bar{Q}_{12} \\ \bar{Q}_{12}^{\mathrm{T}} & \bar{Q}_2 \end{pmatrix}, \tag{4.22}$$

where $\bar{Q}_1 \in \mathbf{R}^{n\times n}$, $\bar{Q}_{12} \in \mathbf{R}^{n\times n_m}$, and $\bar{Q}_2 \in \mathbf{R}^{n_m\times n_m}$.

Minimizing (4.19) under the constraints (4.17) and (4.20) leads to the Lagrangian

$$\begin{aligned} L(A_m, B_m, C_m, \Omega, \tilde{Q}_I, M_c, M_o, \tilde{P}_I) = \operatorname{tr}\Big[&\tilde{Q}_I \tilde{R}_I + (A_m + A_m^{\mathrm{T}} + B_m V B_m^{\mathrm{T}}) M_c \\ &+ (A_m^{\mathrm{T}} \Omega + \Omega A_m + C_m^{\mathrm{T}} R C_m) M_o \\ &+ (\tilde{A}_I \tilde{Q}_I + \tilde{Q}_I \tilde{A}_I^{\mathrm{T}} + \tilde{V}_I) \tilde{P}_I\Big], \end{aligned}$$

where the symmetric matrices M_o, M_c, and $\tilde{P}_I$ are Lagrange multipliers.

Setting $\partial L/\partial \tilde{Q}_I = 0$ gives an equation for $\tilde{P}_I$ similar to (4.20) for $\tilde{P}$,

$$\tilde{A}_I^{\mathrm{T}} \tilde{P}_I + \tilde{P}_I \tilde{A}_I + \tilde{R}_I = 0, \tag{4.23}$$

where $\tilde{P}_I$ is symmetric positive definite and can be partitioned similarly to $\tilde{Q}_I$ as

$$\tilde{P}_I = \begin{pmatrix} \bar{P}_1 & \bar{P}_{12} \\ \bar{P}_{12}^{\mathrm{T}} & \bar{P}_2 \end{pmatrix}. \tag{4.24}$$

The matrices M_c and M_o satisfy [5]

$$M_c = -\left(\frac{1}{2} S + \Omega M_o\right), \tag{4.25}$$

$$(M_o)_{ii} = -\frac{1}{(A_m)_{ii}} \sum_{\substack{j=1 \\ j \neq i}}^{n_m} (A_m)_{ij} (M_o)_{ji}, \tag{4.26}$$

$$(M_o)_{ij} = \frac{(S)_{ij} - (S)_{ji}}{2(\omega_j - \omega_i)}, \qquad \text{if } \omega_j \neq \omega_i, \tag{4.27}$$

where $S = 2(\bar{P}_{12}^{\mathrm{T}} \bar{Q}_{12} + \bar{P}_2 \bar{Q}_2)$.

Setting $\partial L/\partial B_m = 0$ and $\partial L/\partial C_m = 0$ gives

$$2(\bar{P}_{12}^{\mathrm{T}} B + \bar{P}_2 B_m) V + 2 M_c B_m V = 0, \tag{4.28}$$
$$2R(C_m \bar{Q}_2 - C \bar{Q}_{12}) + 2 R C_m M_o = 0. \tag{4.29}$$

Observe that $\tilde{P}_I$ through (4.23) and $\tilde{Q}_I$ through (4.20) depend on B_m and C_m as does A_m through (4.18). Similarly M_c through (4.25) and M_0 through (4.26) and (4.27) depend on B_m and C_m. Thus everything in (4.28) and (4.29) is a function of B_m and C_m. Use the homotopy map structure of (4.4) and let

$$B(\lambda) = \lambda B + (1-\lambda) B_i, \qquad C(\lambda) = \lambda C + (1-\lambda) C_i,$$

where B_i and C_i are matrices defining the initial problem at $\lambda = 0$, and correspond to the parameter vector a in Theorem 3.3. The structure of the homotopy map $\rho(\lambda, x, a)$ for the input normal form is now

$$F_1(\lambda, x, a) = (\bar{P}_{12}^{\mathrm{T}} B(\lambda) + \bar{P}_2 B_m) V + M_c B_m V, \tag{4.30}$$
$$F_2(\lambda, x, a) = R(C_m \bar{Q}_2 - C(\lambda) \bar{Q}_{12}) + R C_m M_o, \tag{4.31}$$

where

$$a \equiv \begin{pmatrix} \mathrm{Vec}\,(B_i) \\ \mathrm{Vec}\,(C_i) \end{pmatrix}$$

denotes the parameter variables $B_i \in \mathbf{R}^{n \times m}$, $C_i \in \mathbf{R}^{l \times n}$,

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

denotes the independent variables B_m and C_m corresponding to x in Theorem 3.3, and A, B, C, V, and R are constants as in Section 2.

The Jacobian matrix of $\rho(\lambda, x, a)$ has $n_m m + n_m l$ rows and $(n_m + n)\times(m + l) + 1$ columns. Since $F_1(\lambda, x, a)$ does not involve C_i and $F_2(\lambda, x, a)$ does not involve B_i

$$D_{C_i}F_1(\lambda, x, a) = 0, \qquad D_{B_i}F_2(\lambda, x, a) = 0.$$

The Jacobian matrix is

$$D\rho(\lambda, x, a) = \begin{pmatrix} D_\lambda F_1 & D_{B_m}F_1 & D_{C_m}F_1 & D_{B_i}F_1 & 0 \\ D_\lambda F_2 & D_{B_m}F_2 & D_{C_m}F_2 & 0 & D_{C_i}F_2 \end{pmatrix}. \tag{4.32}$$

The following lemma will be used in the proof of Theorem 4.5.

Lemma 4.4. *Let* $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{A}_I$, $\tilde{B}_I$, $\tilde{C}_I$, $\tilde{P}$, $\tilde{Q}$, $\tilde{R}$, $\tilde{P}_I$, $\tilde{Q}_I$, $\tilde{R}_I$, Ω *and* U *be defined as above. Then*

$$\bar{Q}_1 = Q_1, \qquad \bar{P}_1 = P_1, \tag{4.33}$$

$$\bar{Q}_{12} = Q_{12}U^{-\mathrm{T}}, \quad \bar{P}_{12} = P_{12}U, \tag{4.34}$$

$$\bar{Q}_2 = I, \qquad \bar{P}_2 = \Omega, \tag{4.35}$$

$$Q_2 = UU^{\mathrm{T}}, \qquad P_2 = U^{-\mathrm{T}}\Omega U^{-1}. \tag{4.36}$$

In addition, P_{12}, Q_{12}, $\bar{P}_{12}$, *and* $\bar{Q}_{12}$ *have full column rank*.

Proof. Equations (4.20) and (4.23) can be written in the form

$$\begin{pmatrix} A & 0 \\ 0 & A_m \end{pmatrix}\begin{pmatrix} \bar{Q}_1 & \bar{Q}_{12} \\ \bar{Q}_{12}^{\mathrm{T}} & \bar{Q}_2 \end{pmatrix} + \begin{pmatrix} \bar{Q}_1 & \bar{Q}_{12} \\ \bar{Q}_{12}^{\mathrm{T}} & \bar{Q}_2 \end{pmatrix}\begin{pmatrix} A^{\mathrm{T}} & 0 \\ 0 & A_m^{\mathrm{T}} \end{pmatrix} + \begin{pmatrix} BVB^{\mathrm{T}} & BVB_m^{\mathrm{T}} \\ B_mVB^{\mathrm{T}} & B_mVB_m^{\mathrm{T}} \end{pmatrix} = 0,$$

$$\begin{pmatrix} A^{\mathrm{T}} & 0 \\ 0 & A_m^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} \bar{P}_1 & \bar{P}_{12} \\ \bar{P}_{12}^{\mathrm{T}} & \bar{P}_2 \end{pmatrix} + \begin{pmatrix} \bar{P}_1 & \bar{P}_{12} \\ \bar{P}_{12}^{\mathrm{T}} & \bar{P}_2 \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & A_m \end{pmatrix} + \begin{pmatrix} C^{\mathrm{T}}RC & -C^{\mathrm{T}}RC_m \\ -C_m^{\mathrm{T}}RC & C_m^{\mathrm{T}}RC_m \end{pmatrix} = 0.$$

Expanding these equations yields

$$A\bar{Q}_1 + \bar{Q}_1A^{\mathrm{T}} + BVB^{\mathrm{T}} = 0, \tag{4.37}$$

$$A\bar{Q}_{12} + \bar{Q}_{12}A_m^{\mathrm{T}} + BVB_m^{\mathrm{T}} = 0, \tag{4.38}$$

$$A_m\bar{Q}_2 + \bar{Q}_2A_m^{\mathrm{T}} + B_mVB_m^{\mathrm{T}} = 0, \tag{4.39}$$

$$A^{\mathrm{T}}\bar{P}_1 + \bar{P}_1A + C^{\mathrm{T}}RC = 0, \tag{4.40}$$

$$A^{\mathrm{T}}\bar{P}_{12} + \bar{P}_{12}A_m - C^{\mathrm{T}}RC_m = 0, \tag{4.41}$$

$$A_m^{\mathrm{T}}P_2 + \bar{P}_2A_m + C_m^{\mathrm{T}}RC_m = 0. \tag{4.42}$$

Comparing (2.7a) with (4.37), and (2.7b) with (4.40) yields (4.33).

If the definitions $A_m = U^{-1}\bar{A}_mU$, $B_m = U^{-1}\bar{B}_m$, and $C_m = \bar{C}_mU$ in Theorem 4.3 are substituted into (4.17) then (4.17) becomes

$$\bar{A}_mUU^{\mathrm{T}} + UU^{\mathrm{T}}\bar{A}_m^{\mathrm{T}} + \bar{B}_mV\bar{B}_m^{\mathrm{T}} = 0, \tag{4.43}$$

$$\bar{A}_m^{\mathrm{T}}U^{-\mathrm{T}}\Omega U^{-1} + U^{-\mathrm{T}}\Omega U^{-1}\bar{A}_m + \bar{C}_m^{\mathrm{T}}R\bar{C}_m = 0. \tag{4.44}$$

Comparing (2.7a) and (2.7b) with (4.43) and (4.44) yields (4.36).

If $A_m = U^{-1}\bar{A}_mU$, $B_m = U^{-1}\bar{B}_m$, and $C_m = \bar{C}_mU$ are substituted into (4.38) and (4.41) and the resulting equations are compared with (2.7a) and (2.7b), then (4.34) follows. Comparing (4.17) and (4.18) with (4.39) and (4.42) yields (4.35).

Finally, since Q_2 and P_2 are nonsingular, from Section 6 in [10] it follows that Q_{12} and P_{12} have full column rank. Since U is nonsingular, from (4.34) it follows that $\bar{Q}_{12}$ and $\bar{P}_{12}$ also have full rank. □

Theorem 4.5. *Let $\tilde{P}_I$ and $\tilde{Q}_I$ be defined as above. Then $D\rho(\lambda, x, a)$ given by* (4.32) *has full column rank for $0 \leqslant \lambda < 1$, i.e., the homotopy map* (4.30) *and* (4.31) *is transversal to zero for $0 \leqslant \lambda < 1$.*

Proof. To prove $D\rho(\lambda, x, a)$ given by (4.32) has full column rank, i.e.,

$$\operatorname{rank}(D\rho(\lambda, x, a)) = n_m m + n_m l,$$

it suffices to prove that

$$\operatorname{rank}(D_aF_1) = \operatorname{rank}(D_{B_i}F_1) = n_m m, \tag{4.45}$$

$$\operatorname{rank}(D_aF_2) = \operatorname{rank}(D_{C_i}F_2) = n_m l. \tag{4.46}$$

Since V and R are constant symmetric positive definite matrices, without loss of generality set $V = I$ in (4.30) and $R = I$ in (4.31). Using Lemma 4.1 to compute $D_{B_i}F_1(\lambda, x, a)$, ordering B_i and F_1 by columns,

$$\begin{aligned} D_{B_i}F_1(\lambda, x, a) &= D_{B_i}(\bar{P}_{12}^{\mathrm{T}}B(\lambda)) = (1-\lambda)D_{B_i}(\bar{P}_{12}^{\mathrm{T}}B_i) \\ &= (1-\lambda)I_m \otimes \bar{P}_{12}^{\mathrm{T}}. \end{aligned} \tag{4.47}$$

Ordering C_i and F_2 by rows gives

$$\begin{aligned} D_{C_i}F_2(\lambda, x, a) &= D_{C_i}(-C(\lambda)\bar{Q}_{12}) = (\lambda-1)D_{C_i}(C_i\bar{Q}_{12}) \\ &= (\lambda-1)I_l \otimes \bar{Q}_{12}^{\mathrm{T}}. \end{aligned} \tag{4.48}$$

Now finally, using Lemma 4.4, (4.47), and (4.48), the rank statements of (4.45) and (4.46) follow.

Thus the homotopy map (4.30) and (4.31) for the input normal form parameterization of (A_m, B_m, C_m) for the H^2 model order reduction problem is transversal to zero. □

4.3. Transversality of homotopies based on Ly's formulation

In Ly's formulation [24], 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 $C_m = \left((C_m)_1 \quad (C_m)_2 \quad \cdots \quad (C_m)_i \quad \cdots \quad (C_m)_r \right)$, where

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

$$(C_m)_r = \begin{pmatrix} 1 & * & \cdots & * \end{pmatrix}^{\mathrm{T}}, \quad \text{if } n_m \text{ is odd.}$$

Let $\mathscr{S}$ be the set of indices of those elements of A_m which are independent variables, i.e., $\mathscr{S} \equiv \{(2,1),\ (2,2),\ \ldots,\ (2i,2i-1),\ (2i,2i),\ \ldots,\ (n_m,n_m)\}$. To minimize the cost function $J(A_m,B_m,C_m)$, consider the Lagrangian

$$L(A_m,B_m,C_m,\tilde{Q}) = \operatorname{tr}[\tilde{Q}\tilde{R} + (\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{V})\tilde{P}], \tag{4.49}$$

where the symmetric matrix $\tilde{P}$ is a Lagrange multiplier, $\tilde{Q}$ satisfies (4.20), $\tilde{A}$, $\tilde{R}$, and $\tilde{V}$ are defined in Section 4.2. Setting $\partial L/\partial\tilde{Q} = 0$ gives (4.23); $\tilde{Q}$ and $\tilde{P}$ are symmetric positive definite and can be partitioned as in (4.22) and (4.24). A straightforward calculation shows

$$\begin{aligned}
\frac{\partial L}{\partial (A_m)_{ij}} &= 2(P_{12}^{\mathrm{T}}Q_{12} + P_2Q_2)_{ij}, \quad (i,j) \in \mathscr{S}, \\
\frac{\partial L}{\partial B_m} &= 2(P_{12}^{\mathrm{T}}B + P_2B_m)V, \\
\frac{\partial L}{\partial (C_m)_{ij}} &= 2\frac{\partial}{\partial (C_m)_{ij}}[\operatorname{tr}(-Q_{12}^{\mathrm{T}}C^{\mathrm{T}}RC_m) + \operatorname{tr}(Q_2C_m^{\mathrm{T}}RC_m)] \\
&= 2R(C_mQ_2 - CQ_{12})_{ij}, \quad i > 1.
\end{aligned} \tag{4.50}$$

Let

$$A(\lambda) = A, \quad B(\lambda) = \lambda B + (1-\lambda)B_i, \quad C(\lambda) = \lambda C + (1-\lambda)C_i,$$

where B_i and C_i play the same role as in Section 4.1. Let

$$\begin{aligned} H_{A_m}(\lambda,x) &= \frac{1}{2}\frac{\partial L}{\partial A_m} = (P_{12}^{\mathrm{T}}Q_{12} + P_2Q_2), \\ H_{B_m}(\lambda,x,B_i) &= \frac{1}{2}\frac{\partial L}{\partial B_m} = (P_{12}^{\mathrm{T}}B(\lambda) + P_2B_m)V, \\ H_{C_m}(\lambda,x,C_i) &= \frac{1}{2}\frac{\partial L}{\partial C_m} = R(C_mQ_2 - C(\lambda)Q_{12}), \end{aligned} \tag{4.51}$$

where in H_{A_m} only those elements corresponding to the independent variables of A_m are nonzero and

$$x \equiv \begin{pmatrix} (A_m)_{\mathscr{S}} \\ \mathrm{Vec}\,(B_m) \\ \mathrm{Vec}\,(C_m)_{\mathscr{T}}. \end{pmatrix} \tag{4.52}$$

denotes the independent variables, $(A_m)_{\mathscr{S}}$ is a vector consisting of those elements in A_m with indices in the set $\mathscr{S}$, i.e.,

$$(A_m)_{\mathscr{S}} = ((A_m)_{21}, (A_m)_{22}, \cdots, (A_m)_{n_m n_m})^{\mathrm{T}},$$

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

The homotopy map $\rho(\lambda,x,a)$ for Ly's formulation is now defined as

$$F_1(\lambda,x,a) = [H_{A_m}(\lambda,x)]_{\mathscr{S}}, \tag{4.53}$$

$$F_2(\lambda,x,a) = \mathrm{Vec}\,[H_{B_m}(\lambda,x,B_i)], \tag{4.54}$$

$$F_3(\lambda,x,a) = \mathrm{Vec}\,[H_{C_m}(\lambda,x,C_i)]_{\mathscr{T}}., \tag{4.55}$$

where again the subscripts $\mathscr{S}$ and $\mathscr{T}$ select the appropriate matrix elements, and

$$a \equiv \begin{pmatrix} \mathrm{Vec}\,(B_i) \\ \mathrm{Vec}\,(C_i) \end{pmatrix} \tag{4.56}$$

denotes the parameter variables. As discussed in Section 4.2, without loss of generality set $V = I$ in (4.54) and $R = I$ in (4.55).

The Jacobian matrix $D\rho(\lambda,x,a)$ of $\rho(\lambda,x,a)$ is

$$\begin{pmatrix} D_\lambda F_1 & D_x F_1 & 0 & 0 \\ D_\lambda F_2 & D_x F_2 & D_{B_i} F_2 & 0 \\ D_\lambda F_3 & D_x F_3 & 0 & D_{C_i} F_3 \end{pmatrix}. \tag{4.57}$$

Lemma 4.6. *Suppose* $\mathrm{rank}\,(D_xF_1) = n_m$. *Then the Jacobian matrix* (4.57) *has full column rank for all* $0 \leqslant \lambda < 1$, *i.e., the homotopy map* (4.53)–(4.55) *is transversal to zero for all* $0 \leqslant \lambda < 1$.

Proof. A similar proof to that in Section 4.2 yields

$$\operatorname{rank}(D_{B_i}F_2) = mn_m \quad \text{for} \quad \lambda \neq 1. \tag{4.58}$$

Ordering C_i and F_3 by rows gives

$$D_{C_i}F_3(\lambda,\theta,a) = D_{C_i}(-C(\lambda)Q_{12})_{\mathscr{T}.} = (\lambda-1)D_{C_i}(C_iQ_{12})_{\mathscr{T}.} = (\lambda-1)D_{C_i}[(C_i)_{\mathscr{T}.}Q_{12}]$$

$$= (\lambda-1)\overbrace{\begin{pmatrix} 0 & Q_{12}^{\mathrm{T}} & 0 & \dots & 0 \\ 0 & 0 & Q_{12}^{\mathrm{T}} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & Q_{12}^{\mathrm{T}} \end{pmatrix}}^{l \text{ times}}, \tag{4.59}$$

and then as before

$$\operatorname{rank}(D_{C_i}F_3) = (l-1)n_m \quad \text{for} \quad \lambda \neq 1. \tag{4.60}$$

Note that

$$\operatorname{rank}(D_xF_1) = n_m,$$

which completes the proof. □

Note that there are only n_m components in F_1 but $(l+m)n_m + 1$ independent variables in x and λ. As $l + m \gg 1$ usually in real problems which have been considered previously [10], all Jacobian matrices of F_1 in those problems satisfied the full rank condition. Since each of Q_{12}, P_{12}, Q_2, and P_2 are implicit functions of x and $A(\lambda)$, and one cannot give explicit expressions for D_xF_1 or $D_{A_i}F_1$ as in (4.59) for $D_{C_i}F_3$ (which show clearly the rank conditions), it was necessary to *assume* that $\operatorname{rank}(D_xF_1) = n_m$ in Lemma 4.6. To guarantee the full rank of $D\rho$ without this assumption, instead of using (4.53), let $x = (\eta, \zeta)$, $\eta \in \mathbf{E}^{n_m}$,

$$F_1(\lambda, x, a) = \lambda[H_{A_m}(\lambda, x)]_{\mathscr{S}} + (1-\lambda)(\eta - \eta_0), \tag{4.61}$$

with n_m independent parameter variables in η_0, which gives

$$D_{\eta_0}F_1 = (1-\lambda)I_{n_m} \quad \text{for} \quad \lambda \neq 1. \tag{4.62}$$

Combining (4.58), (4.60), and (4.62) completes the proof that the map (4.54), (4.55), and (4.61) is transversal to zero. Note that the homotopy construction in (4.61) is a theoretical convenience, and in practice the choice (4.53) has been entirely satisfactory.

5. Boundedness of $\rho_a^{-1}(0)$ for H^2 optimal model order reduction problem

5.1. Counterexample for optimal projection homotopies

The zero set $\rho_a^{-1}(0)$ of a given homotopy map based on the optimal projection equations (4.1)–(4.3) is not always bounded, as shown by the following two-dimensional example. The system [20] is given by

$$A = \begin{pmatrix} -0.25 & -0.4 \\ -0.4 & -0.72 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 1.2 \end{pmatrix}, \quad C = (\,1 \quad 1.2\,). \tag{5.1}$$

For the system (2.1)–(2.4) defined by 5.1 and $n_m = 1$, the solution set of the optimal projection equations (4.1)–(4.3) contains an isolated solution and a one-dimensional manifold of solutions.

The isolated solution of this system is

$$A_m = (-0.838521), \quad B_m = (1.537575), \quad C_m = (1.537575),$$

which was obtained by both POLSYS from HOMPACK [36] and by a homotopy approach [41]. The one-dimensional manifold of solutions can be derived directly from equations (4.1)–(4.3) as follows.

Let

$$W_1 = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \quad U_1 = (u_1, u_2), \quad \Sigma = \sigma, \quad V = I, \quad R = I.$$

The optimal projection equations (4.1)–(4.3) for this problem can be written as

$$\begin{aligned}
0 = &- 0.25w_1^2u_1\sigma - 0.4w_1w_2u_1\sigma - 0.4w_1^2u_2\sigma - 0.72w_1w_2u_2\sigma - 0.25w_1\sigma \\
&- 0.4w_2\sigma + u_1 + 1.2u_2, \\
0 = &- 0.25w_1w_2u_1\sigma - 0.4w_2^2u_1\sigma - 0.4w_1w_2u_2\sigma - 0.72w_2^2u_2\sigma - 0.4w_1\sigma \\
&- 0.72w_2\sigma + 1.2u_1 + 1.44u_2, \\
0 = &- 0.25w_1u_1^2\sigma - 0.4w_2u_1^2\sigma - 0.4w_1u_1u_2\sigma - 0.72w_2u_1u_2\sigma - 0.25u_1\sigma \\
&- 0.4u_2\sigma + w_1 + 1.2w_2, \\
0 = &- 0.25w_1u_1u_2\sigma - 0.4w_2u_1u_2\sigma - 0.4w_1u_2^2\sigma - 0.72w_2u_2^2\sigma - 0.4u_1\sigma \\
&- 0.72u_2\sigma + 1.2w_1 + 1.44w_2, \\
0 = &\, w_1u_1 + w_2u_2 - 1.
\end{aligned} \tag{5.2}$$

The triple (A_m, B_m, C_m) is given by

$$\begin{aligned}
A_m = \Gamma A G^{\mathrm{T}} &= (\,u_1 \quad u_2\,)\begin{pmatrix} -0.25 & -0.4 \\ -0.4 & -0.72 \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\
&= w_1(-0.25u_1 - 0.4u_2) + w_2(-0.4u_1 - 0.72u_2),
\end{aligned}$$

$$B_m = \Gamma B = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1.2 \end{pmatrix} = u_1 + 1.2 u_2, \tag{5.3}$$

$$C_m = C G^{\mathrm{T}} = \begin{pmatrix} 1 & 1.2 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = w_1 + 1.2 w_2,$$

where $\Gamma = U_1$ and $G = W_1^{\mathrm{T}}$.

The zero set of (5.2) contains

$$\left\{ (W_1, U_1, \Sigma) \colon w_1 = -1.2 w_2, \quad u_1 = -1.2 u_2, \quad u_2 = \frac{1}{2.44 w_2}, \quad \sigma = 0 \right\},$$

which is unbounded. Every point in this set corresponds to the same triple (A_m, B_m, C_m):

$$A_m = -.0491803, \quad B_m = 0, \quad C_m = 0.$$

The homotopy map based on the optimal projection equations is

$$\begin{aligned} & U_1 A(\lambda) W_1 \Sigma W_1^{\mathrm{T}} + \Sigma W_1^{\mathrm{T}} A^{\mathrm{T}}(\lambda) + U_1 B V B^{\mathrm{T}} = 0, \\ & A^{\mathrm{T}}(\lambda) U_1^{\mathrm{T}} \Sigma + U_1^{\mathrm{T}} \Sigma U_1 A(\lambda) W_1 + C^{\mathrm{T}} R C W_1 = 0, \\ & U_1 W_1 - I = 0, \end{aligned} \tag{5.4}$$

where $A(\lambda) = \lambda A + (1 - \lambda) D$, and D is a part of the parameter vector a in Theorem 3.3. The zero set $\rho_a^{-1}(0)$ of this homotopy map for the system (5.1) includes the subset

$$\begin{aligned} & \Big\{ (\lambda, W_1, U_1, \Sigma) \colon 0 \leqslant \lambda < 1, \quad w_1 = -1.2 w_2, \quad u_1 = -1.2 u_2, \\ & \quad u_2 = \frac{1}{2.44 w_2}, \quad \sigma = 0 \Big\}, \end{aligned} \tag{5.5}$$

which is unbounded. This example shows that the zero set $\rho_a^{-1}(0)$ of a homotopy map can be unbounded and yet some zero curves may still converge to isolated solutions.

Note that, in practice, the algorithm in [41] always maintains $\operatorname{rank}(\Sigma) = n_m$, where $n_m = 1$ in the above example. Solutions with $\Sigma = 0$ in the above example never come into play. Boundedness of $\rho_a^{-1}(0)$ for the optimal projection equations (4.1)–(4.3) can indeed be guaranteed with more sophisticated mathematics, a slightly different homotopy map from the one used in practice, and complex arithmetic for the curve tracking. This is pursued in Section 5.3.

5.2. Simplification and example for input normal form homotopy

The following corollary is needed to simplify the homotopy map based on the input normal form formulation for the H^2 optimal model order reduction problem.

Corollary 5.1. *Let $\tilde{A}_I, \tilde{R}_I, \tilde{V}_I$ be defined as in Section* 4.2, *partitioned as in* (4.21), *let A_m be stable, and $\tilde{Q}_I$ satisfy* (4.20). *To minimize* (4.19) *under the constraints* (4.17) *and* (4.20), *the following two Lagrangians are equivalent*:

$$\begin{aligned} L_1(A_m, B_m, C_m, \Omega, \tilde{Q}_I, M_c, M_o, \tilde{P}_I) = \operatorname{tr}\Big[&\tilde{Q}_I\tilde{R}_I + (A_m + A_m^{\mathrm{T}} + B_m V B_m^{\mathrm{T}})M_c \\ &+ (A_m^{\mathrm{T}}\Omega + \Omega A_m + C_m^{\mathrm{T}} R C_m)M_o \\ &+ (\tilde{A}_I\tilde{Q}_I + \tilde{Q}_I\tilde{A}_I^{\mathrm{T}} + \tilde{V}_I)\tilde{P}_I\Big], \end{aligned} \tag{5.6}$$

where the symmetric matrices M_o, M_c, and $\tilde{P}_I$ are Lagrange multipliers introduced in Section 4.2, *and*

$$L_2(A_m, B_m, C_m, \tilde{Q}_I, \tilde{P}_I) = \operatorname{tr}[\tilde{Q}_I\tilde{R}_I + (\tilde{A}_I\tilde{Q}_I + \tilde{Q}_I\tilde{A}_I^{\mathrm{T}} + \tilde{V}_I)\tilde{P}_I], \tag{5.7}$$

where $\tilde{Q}_I$ is restricted to the form

$$\tilde{Q}_I = \begin{pmatrix} \bar{Q}_1 & \bar{Q}_{12} \\ \bar{Q}_{12}^{\mathrm{T}} & I_{n_m} \end{pmatrix},$$

the Lagrange multiplier $\tilde{P}_I$ is restricted to the form

$$\tilde{P}_I = \begin{pmatrix} \bar{P}_1 & \bar{P}_{12} \\ \bar{P}_{12}^{\mathrm{T}} & \Omega \end{pmatrix},$$

and $\Omega = \operatorname{diag}(\omega_1, \cdots, \omega_{n_m})$ is a positive definite matrix.

Proof. The proof is straightforward. Setting $\partial L/\partial\tilde{Q}_I = 0$ gives the same equation

$$\tilde{A}_I^{\mathrm{T}}\tilde{P}_I + \tilde{P}_I\tilde{A}_I + \tilde{R}_I = 0 \tag{5.8}$$

in both cases. Expanding (4.20) and (5.8) yields the equations for $\bar{Q}_2$ and $\bar{P}_2$. In the first case

$$A_m\bar{Q}_2 + \bar{Q}_2A_m^{\mathrm{T}} + B_m V B_m^{\mathrm{T}} = 0, \quad A_m^{\mathrm{T}}\bar{P}_2 + \bar{P}_2A_m + C_m^{\mathrm{T}} R C_m = 0.$$

Since the constraints (4.17) and (4.20) should be satisfied and A_m is stable, it follows that at a constrained minimum

$$\bar{Q}_2 = I_{n_m}, \quad \bar{P}_2 = \Omega. \qquad \square$$

The partial derivatives $\partial L_2/\partial B_m$ and $\partial L_2/\partial C_m$ of L_2 can be computed as

$$\frac{\partial L_2}{\partial B_m} = 2(\bar{P}_{12}^{\mathrm{T}} B + \Omega B_m)V, \qquad \frac{\partial L_2}{\partial C_m} = 2R(C_m - C\bar{Q}_{12}).$$

The corresponding homotopy map (4.30) and (4.31) is now simplified as

$$\rho(\lambda, x, a) = \begin{pmatrix} \mathrm{Vec}\,(H_{B_m}(\lambda, x, a)) \\ \mathrm{Vec}\,(H_{C_m}(\lambda, x, a)) \end{pmatrix},$$

where

$$H_{B_m}(\lambda, x, a) = (\bar{P}_{12}^{\mathrm{T}} B(\lambda) + \Omega B_m)V, \qquad H_{C_m}(\lambda, x, a) = R(C_m - C(\lambda)\bar{Q}_{12}).$$

The zero set $\rho_a^{-1}(0)$ of a homotopy map based on the input normal form formulation given by [9] is not always bounded, as shown by the following two-dimensional example.

The system is given by

$$A = \begin{pmatrix} -0.895116 & 0.612237 \\ 0.612237 & -0.447393 \end{pmatrix}, \quad B = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad C = (\,-2 \quad 1\,). \tag{5.9}$$

According to [9], the initial point and the triple $(A(\lambda), B(\lambda), C(\lambda))$ are chosen as follows:

(1) Transform the given triple (A, B, C) to balanced form (A_b, B_b, C_b), such that $A_b = T^{-1}AT$, $B_b = T^{-1}B$, and $C_b = CT$ satisfy

$$0 = A_b\Lambda + \Lambda A_b^{\mathrm{T}} + B_b V B_b^{\mathrm{T}}, \qquad 0 = A_b^{\mathrm{T}}\Lambda + \Lambda A_b + C_b^{\mathrm{T}} R C_b,$$

with a positive definite matrix $\Lambda = \mathrm{diag}\,(d_1, d_2, \cdots, d_n)$, $d_i \geqslant d_{i+1}$.

The balanced form of (5.9) is

$$A_b = \begin{pmatrix} -0.25297 & -0.5 \\ -0.5 & -1.0896 \end{pmatrix}, \quad B_b = \begin{pmatrix} -1.232 \\ -1.866 \end{pmatrix},$$
$$C_b = (\,-1.232 \quad -1.866\,),$$

with

$$T = \begin{pmatrix} 0.866 & 0.5 \\ 0.5 & -0.866 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 3 & 0 \\ 0 & 1.5978 \end{pmatrix}.$$

(2) For $n_m = 1$, the parameterization $(A(\lambda), B(\lambda), C(\lambda))$ is chosen as

$$A(\lambda) = \lambda A + (1-\lambda)A_i = \begin{pmatrix} a_1(\lambda) & a_2(\lambda) \\ a_2(\lambda) & a_3(\lambda) \end{pmatrix}$$
$$= \begin{pmatrix} -0.6422\lambda - 0.25297 & 0.612237\lambda \\ 0.612237\lambda & 0.6431\lambda - 1.0896 \end{pmatrix},$$

$$B(\lambda) = \lambda B + (1-\lambda)B_i = \begin{pmatrix} b_1(\lambda) \\ b_2(\lambda) \end{pmatrix} = \begin{pmatrix} -1.232 - 0.768\lambda \\ \lambda \end{pmatrix},$$

$$\begin{aligned} C(\lambda) &= \lambda C + (1-\lambda)C_i = (\, c_1(\lambda) \quad c_2(\lambda)\,) = (\, -1.232 - 0.768\lambda \quad \lambda\,) \\ &= B^{\mathrm{T}}(\lambda). \end{aligned}$$

where

$$A_i = \begin{pmatrix} -0.25297 & 0 \\ 0 & -1.0896 \end{pmatrix}, \quad B_i = \begin{pmatrix} -1.232 \\ 0 \end{pmatrix}, \quad C_i = (\, -1.232 \quad 0\,).$$

For brevity, $a_1(\lambda)$, $a_2(\lambda)$, $a_3(\lambda)$, $b_1(\lambda)$, $b_2(\lambda)$, $c_1(\lambda)$, and $c_2(\lambda)$ will be denoted by a_1, a_2, a_3, b_1, b_2, c_1, and c_2, respectively, in the following. As discussed in Section 4.2, without loss of generality, set $V = I$ and $R = I$.

For any $0 < \lambda < 1$, $B_m \in \mathbf{R}, B_m \neq 0$, let

$$A_m = \frac{-B_m^2}{2}, \quad C_m = -\sqrt{\Omega}B_m, \quad \bar{P}_2 \equiv \Omega = \left[\frac{M(b_1 - b_2)b_1}{a_1 + A_m - Ma_2}\right]^2,$$

$$M = \frac{a_2 b_1 - b_2(b_1 - A_m)}{b_1(a_3 + A_m) - b_2 a_2},$$

$$(\bar{P}_{12})_{12} = \frac{C_m(a_2 b_1 - a_1 b_2 - A_m b_2)}{a_2^2 - (a_1 + A_m)(a_3 + A_m)}, \quad (\bar{P}_{12})_{11} = \frac{b_2 C_m - (\bar{P}_{12})_{12}(a_3 + A_m)}{a_2},$$

$$(\bar{Q}_{12})_{11} = \frac{(\bar{P}_{12})_{11}}{\sqrt{\Omega}}, \quad (\bar{Q}_{12})_{12} = \frac{(\bar{P}_{12})_{12}}{\sqrt{\Omega}}.$$

Then

$$\begin{aligned} &\rho(\lambda, x, a) = 0, \quad \tilde{A}_I(\lambda)\tilde{Q}_I + \tilde{Q}_I\tilde{A}_I^{\mathrm{T}}(\lambda) + \tilde{V}_I(\lambda) = 0, \\ &\tilde{A}_I^{\mathrm{T}}(\lambda)\tilde{P}_I + \tilde{P}_I\tilde{A}_I(\lambda) + \tilde{R}_I(\lambda) = 0 \end{aligned}$$

are satisfied. The zero set $\rho_a^{-1}(0)$ of this homotopy map includes

$$\left\{(\lambda, B_m, C_m)\colon 0 < \lambda < 1, C_m = -\sqrt{\Omega}B_m\right\}. \tag{5.10}$$

Clearly, (5.10) is unbounded. If $B_m \neq 0$, then A_m is stable, (A_m, B_m) is controllable, and (A_m, C_m) is observable.

5.3. Homogeneous transformation to avoid solutions at infinity

As shown by the examples in Sections 5.1 and 5.2, the polynomial systems (4.1)–(4.3) or (4.30) and (4.31) may have solutions at infinity, and $\rho_a^{-1}(0)$ contains paths that diverge to infinity as λ approaches 1. Solutions at infinity can be avoided via the following transformation [26–28], which will be used in Section 5.4.

Let $f(z) = 0$ be a polynomial system of N equations in N unknowns, where $z \in \mathbf{C}^N$, and define $f'(z')$ as the homogenization of $f(z)$:

$$f'_j(z') = z_0^{d_j} f_j(z_1/z_0, \ldots, z_N/z_0), \quad j = 1, \ldots, N, \tag{5.11}$$

where $d_j = \deg(f_j)$. $f'(z') = 0$ is a system of N homogeneous equations in $N + 1$ unknowns.

Note that, if $f'(z^0) = 0$, then $f'(cz^0) = 0$ for any complex scalar c. Therefore, we may take "solutions" of $f'(z') = 0$ to be (complex) lines through the origin in $\mathbf{C}^{N+1}$. The set of these lines is called complex projective space, denoted by $\mathbf{P}^N$, a smooth compact N-complex-dimensional manifold. It is natural to view $\mathbf{P}^N$ as a disjoint union of points $[(z_0, \ldots, z_N)]$ with $z_0 \neq 0$ and the "points at infinity", the points $[(z_0, \ldots, z_N)]$ with $z_0 = 0$. The solutions of $f'(z') = 0$ in $\mathbf{P}^N$ are identified with the solutions and solutions at infinity of $f(z) = 0$ as follows.

First, the solutions to $f(z) = 0$ can be identified with the solutions to $f'(z') = 0$ with $z_0 \neq 0$. Explicitly, if $L \in \mathbf{P}^N$ is a solution to $f'(z') = 0$, and $z' \in L$, with $z' = (z_0, \ldots, z_N)$ and $z_0 \neq 0$, then $z = (z_1/z_0, z_2/z_0, \ldots, z_N/z_0)$ is a solution to $f(z) = 0$. On the other hand, if $z \in \mathbf{C}^N$ is a solution to $f(z) = 0$, then the line through $z' = (1, z)$ is a solution to $f'(z') = 0$ with $z_0 = 1 \neq 0$. A "solution to $f(z) = 0$ at infinity" is simply a solution to $f'(z') = 0$ (in $\mathbf{P}^N$) generated by z' with $z_0 = 0$.

Define a homotopy map (in $\mathbf{P}^N$)

$$h(z', \lambda) = (1 - \lambda)\gamma g'(z') + \lambda f'(z'), \tag{5.12}$$

where g' is a homogeneous system of N polynomials in $N + 1$ variables, and γ is a randomly chosen complex number. Intuitively, let g' be chosen so that its homogeneous structure matches that of f'. Precisely, let $S \in \mathbf{P}^N$ be the set of common solutions of $f'(z') = 0$ and $g'(z') = 0$. Then for each $s \in S$ the following conditions must hold. For $s \in S$ let K denote the full connected component of solutions of $g'(z') = 0$ with $s \in K$.

If s is a geometrically isolated solution of $g'(z') = 0$, assume that: (a) s is also a geometrically isolated solution of $f'(z') = 0$; (b) the multiplicity of s as a solution of $g'(z') = 0$ is less than or equal to the multiplicity of s as a solution of $f'(z') = 0$.

If s is not a geometrically isolated solution of $g'(z') = 0$, assume that: (a) K is contained in S; (b) K is the full solution component of $f'(z') = 0$ containing s; (c) K is a smooth manifold; (d) at each point $z^0 \in K$ the rank of $\nabla g'(z^0)$ is the codimension of K.

Let S' denote the solution set of $g'(z') = 0$ in $\mathbf{P}^N - S$. Under these assumptions, the basic result is the following theorem.

Theorem 5.2. [26,28] *Assume the points in S' are nonsingular solutions of $g'(z') = 0$. For any positive r and for all but a finite number of angles θ, if $\gamma = r\mathrm{e}^{\mathrm{i}\theta}$, then $h^{-1}(0) \cap ((\mathbf{P}^N - S) \times [0,1))$ consists of smooth paths and every geometrically isolated solution of $f'(z') = 0$ not in S has a path in $(\mathbf{P}^N - S) \times [0,1)$ converging to it.*

Let

$$L(z') = \sum_{i=0}^{N} b_i z_i,$$

where $b_i \neq 0$ for some i.

$$U_L = \left\{[z'] \in \mathbf{P}^N \mid L(z') \neq 0\right\}$$

is the Euclidean coordinate patch on $\mathbf{P}^N$ defined by L. Note that U_L, which is an open dense submanifold of $\mathbf{P}^N$, can be identified with $\mathbf{C}^N$ via

$$[(z_0, \ldots, z_N)] \to \frac{1}{L(z')}(z_0, \ldots, z_{i-1}, z_{i+1}, \ldots, z_N),$$

where $b_i \neq 0$.

The following theorem from [28] shows how to keep the homotopy process in complex Euclidean space, even though the basic theorem is formulated in $\mathbf{P}^N$.

Theorem 5.3. *Assume the points in S' are nonsingular solutions of $g'(z') = 0$. Then*

$$\overline{h^{-1}(0) \cap ((\mathbf{P}^N - S) \times [0,1))} \subset U_L \times [0,1],$$

for almost all U_L and all but a finite number of angles θ.

For computations, the coordinate patch U_L is realized via a projective transformation as follows. With homogeneous h in the variables z_i for $i = 0$ to N, let

$$z_0 = \sum_{i=1}^{N} \beta_i z_i + \beta_0, \tag{5.13}$$

where the β_i are constants and $\beta_i \neq 0$ for all i. The projective transformation of h is the system H of N polynomials in the N variables z_i for $i = 1$ to N where $H_j = h_j$, with (5.13) defining z_0 in terms of the other variables. By Theorem 5.3, the homotopy paths, including end points, are completely represented in $\mathbf{C}^N$ via H. The finite solutions of $f(x) = 0$ are recovered via $z_i \leftarrow z_i/z_0$ for $i = 1$ to N. If

$z_0 = 0$, then the solution is at infinity. This concludes the background discussion of polynomial system theory.

5.4. Homogeneous transformation of optimal projection homotopies

In this section the homogeneous transformation introduced in Section 5.3 is used to prevent unbounded zero sets for optimal projection homotopies. Consider the polynomial system given by (4.1)–(4.3) and the corresponding optimal projection homotopies defined in Section 4.1. The start system at $\lambda = 0$ is taken as

$$\begin{aligned} &U_1 A(0) W_1 \Sigma W_1^{\mathrm{T}} + \Sigma W_1^{\mathrm{T}} A(0)^{\mathrm{T}} + U_1 B(0) = 0, \\ &A(0)^{\mathrm{T}} U_1^{\mathrm{T}} \Sigma + U_1^{\mathrm{T}} \Sigma U_1 A(0) W_1 + C(0) W_1 = 0, \\ &U_1 W_1 - I_{n_m} = 0, \end{aligned} \tag{5.14}$$

where $A(0) = D = A - \epsilon I_n$, ϵ is a constant, $A(\lambda) = \lambda A + (1-\lambda) D$. The target system (at $\lambda = 1$) is (4.1)–(4.3).

According to Section 4.3, the homogenization of the target system (4.1)–(4.3) can be taken as

$$\begin{aligned} &U_1' A W_1' \Sigma' W_1'^{\mathrm{T}} + z_0^2 \Sigma' W_1'^{\mathrm{T}} A^{\mathrm{T}} + z_0^3 U_1' B V B^{\mathrm{T}} = 0, \\ &z_0^2 A^{\mathrm{T}} U_1'^{\mathrm{T}} \Sigma' + U_1'^{\mathrm{T}} \Sigma' U_1' A W_1' + z_0^3 C^{\mathrm{T}} R C W_1' = 0, \\ &U_1' W_1' - z_0^2 I_{n_m} = 0, \end{aligned} \tag{5.15}$$

where

$$z = (\mathrm{vec}(U_1), \mathrm{vec}(W_1), \mathrm{vec}(\Sigma)),$$

$$U_1'(z_0, \ldots, z_N) = z_0 U_1(z_1/z_0, \ldots, z_N/z_0),$$

$$W_1'(z_0, \ldots, z_N) = z_0 W_1(z_1/z_0, \ldots, z_N/z_0),$$

$$\Sigma'(z_0, \ldots, z_N) = z_0 \Sigma(z_1/z_0, \ldots, z_N/z_0).$$

The corresponding homogenization of the start system is

$$\begin{aligned} &U_1' D W_1' \Sigma' W_1'^{\mathrm{T}} + z_0^2 \Sigma' W_1'^{\mathrm{T}} D^{\mathrm{T}} + z_0^3 U_1' B_i = 0, \\ &z_0^2 D^{\mathrm{T}} U_1'^{\mathrm{T}} \Sigma' + U_1'^{\mathrm{T}} \Sigma' U_1' D W_1' + z_0^3 C_i W_1' = 0, \\ &U_1' W_1' - z_0^2 I_{n_m} = 0, \end{aligned} \tag{5.16}$$

where $B_i = B(0)$ and $C_i = C(0)$.

Theorem 5.4. *If B_i, C_i, and ϵ can be chosen such that* (5.15) *and* (5.16) *have no common $z_0 \neq 0$, $\Sigma' \neq 0$ solutions, and all $z_0 \neq 0$, $\Sigma' \neq 0$ solutions of* (5.16) *are* *nonsingular, then every geometrically isolated solution of* (5.15) *has a path in* $\mathbf{P}^N$ *converging to it.*

Proof. If $\epsilon = 0$, (5.15) and (5.16) have the same $z_0 = 0$ solution set (corresponding to solutions of (4.1)–(4.3) at infinity). Since B_i and C_i can be chosen such that (5.15) and (5.16) have no common $z_0 \neq 0$, $\Sigma' \neq 0$ solutions and all $z_0 \neq 0$, $\Sigma' \neq 0$ solutions of (5.16) are nonsingular, then all the conditions of Theorems 5.2 and 5.3 hold. For each point in S', the associated path in $H^{-1}(0)$ can be tracked from $\lambda = 0$ to $\lambda = 1$. This will yield the full list of geometrically isolated solutions to $H(z, 1) = 0$. No paths diverge to infinity.

If $\epsilon \neq 0$, $B(\lambda) = BVB^{\mathrm{T}}$, and $C(\lambda) = C^{\mathrm{T}}RC$ for $0 \leq \lambda \leq 1$ as in [41], using the fact $U_1' W_1' = 0$ (when $z_0 = 0$), it is clear that the $z_0 = 0$ solution set of (5.16) is the same as that of (5.15). Similarly, (5.15) and (5.16) have the same $z_0 \neq 0$ solutions when $\Sigma' = 0$. Note that this case corresponds to the counterexample of Section 5.1. Take S be all the $z_0 = 0$ solutions and any solutions corresponding to $z_0 \neq 0$ and $\Sigma' = 0$. Now ϵ can be chosen such that (5.15) and (5.16) have no other common solutions and all other $z_0 \neq 0$ solutions of (5.16) are nonsingular. Then the technical assumptions of Theorem 5.2 can clearly be met for the common solution set S. Thus Theorem 5.2 and 5.3 hold for the start system (5.16) in this case ($\epsilon \neq 0$) also. □

The import of this result is that the real solutions of (4.1)–(4.3), which satisfy the rank condition

$$\operatorname{rank}(W_1) = \operatorname{rank}(U_1) = \operatorname{rank}(\Sigma) = n_m,$$

if they exist, must be connected to the solutions of (5.16) in $\mathbf{P}^N - S$. Technically, this is guaranteed only with a complex multiplier γ in (5.16), and only if complex arithmetic is used and the homotopy curve tracking is done in $\mathbf{P}^N$. However, all this has never been necessary in practice [41]. Furthermore, observe that the solution set (5.15) includes all solutions with rank $\Sigma' \leqslant n_m$, and thus one is guaranteed of finding a reduced order model of order *no greater than* n_m. Since (5.15) represents the optimal projection equations (4.1)–(4.3) for some stable $A(\lambda)$ for every λ, $0 \leq \lambda \leq 1$, it is clear why real arithmetic suffices generically. Generically, the real solutions are isolated, have constant rank, and vary smoothly with respect to λ [26].

Finally, for the target system (5.15), it is always possible to take the starting homogeneous system as

$$p_j z_j^4 - q_j z_0^4 = 0, \quad j = 1, \ldots, N, \tag{5.17}$$

where p_j and q_j are positive constants such that (5.17) has no common solution with (5.15). Since all solutions to (5.17) are nonsingular, all conditions of Theorem 5.2 and 5.3 are satisfied. The drawback is that the starting system

(5.17) is totally unrelated to (5.15), requires complex arithmetic, and may take more steps to converge.

6. Conclusions

Probability-one homotopy methods were considered for the problem of H^2 model reduction. The crucial requirement of transversality was verified for several homotopy maps including the pseudogramian formulation of the optimal projection equations as well as variations based upon canonical forms. These results guarantee good numerical properties in the computational implementation of probability-one homotopy algorithms. Counterexamples to the boundedness requirement of probability-one homotopy theory were provided for the pseudogramian formulation of the optimal projection equations and for some formulations based upon canonical forms. Since a solution may not exist in any particular canonical form, these results are sharp for canonical forms, where unboundedness corresponds to nonexistence of solutions. However, for a reformulation of the pseudogramian optimal projection equations in complex projective space using homogeneous transformations, the boundedness assumption holds and thus global convergence of the homotopy algorithm to a solution (in complex projective space) is guaranteed. Both the genericity of real solutions and considerable computational experience [41] indicate that real-valued homotopies are effective in practice and thus it is not necessary to track the homotopy zero curves in complex projective space.

Acknowledgements

The work of Wang and Watson was supported in part by Department of Energy grant DE-FG05-88ER25068 and Air Force Office of Scientific Research grant F49620-92-J-0236; that of Bernstein was supported in part by Air Force Office of Scientific Research grant F49620-95-1-0019.

References

[1] L. Baratchart, M. Cardelli, M. Olivi, Identification and rational L^2 approximation: a gradient algorithm, Automatica 27 (1991) 413–418.

[2] D.S. Bernstein, W.M. Haddad, LQG control with an H_∞ performance bound: a Riccati equation approach, IEEE Trans. Autom. Control AC-34 (1989) 293–305.

[3] D.S. Bernstein, W.M. Haddad, Steady-state Kalman filtering with an H_∞ error bound, Syst. Control Lett. 12 (1989) 9–16.

[4] S.N. Chow, J. Mallet-Paret, J.A. Yorke, Finding zeros of maps: homotopy methods that are constructive with probability one, Math. Comput. 32 (1978) 887–899.

[5] L.D. Davis, E.G. Collins Jr., S.A. Hodel, A parameterization of minimal plants, in: Proceedings of the 1992 American Contr. Conference, Chicago, IL, June 1992, pp. 355–356.
[6] J.C. Doyle, K. Zhou, B. Bodenheimer, Mixed H_2 and H_∞ performance objectives, II. Optimal control, IEEE Trans. Autom. Control 39 (1994) 1575–1587.
[7] P.A. Furhmann, R. Ober, A functional approach to LQG balancing, Int. J. Control 57 (1993) 627–714.
[8] F.R. Gantmacher, The Theory of Matrices, Chelsea, New York, 1960.
[9] Y. Ge, E.G. Collins Jr., L.T. Watson, L.D. Davis, An input normal form homotopy for the L^2 optimal model order reduction problem, IEEE Trans. Autom. Control 39 (1994) 1302–1305.
[10] Y. Ge, L.T. Watson, E.G. Collins Jr., A comparison of homotopies for alternative formulations of the L^2 optimal model order reduction problem, J. Comput. Appl. Math. 69 (1996) 215–241.
[11] Y. Ge, L.T. Watson, E.G. Collins Jr., D.S. Bernstein, Globally convergent homotopy algorithms for the combined H^2/H_∞ infty model reduction problem, J. Math. Syst. Estim. Control 7 (1997) 129–155.
[12] K. Glover, All optimal Hankel-norm approximations of linear multivariable systems and their L_∞ -error bounds, Int. J. Control 39 (1984) 1115–1195.
[13] K.M. Grigoriadis, Optimal H_∞ model reduction via linear matrix inequalities: continuous- and discrete-time cases, Syst Control Lett. 26 (1995) 321–333.
[14] W.M. Haddad, D.S. Bernstein, Combined L^2/H^∞ model reduction, Int. J. Control 49 (1989) 1523–1535.
[15] W.M. Haddad, D.S. Bernstein, Generalized Riccati equations for the full- and reduced-order mixed-norm H^2/H^∞ standard problem, Syst. Control Lett. 14 (1990) 185–197.
[16] D.C. Hyland, D.S. Bernstein, The optimal projection equations for fixed-order dynamic compensation, IEEE Trans. Autom. Control AC-29 (1984) 1034–1037.
[17] D.C. Hyland, D.S. Bernstein, The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton and Moore, IEEE Trans. Autom. Control AC-30 (1985) 1201–1211.
[18] D.C. Hyland, S. Richter, On direct versus indirect methods for reduced-order controller design, IEEE Trans. Autom. Control 35 (1990) 377–379.
[19] P.T. Kabamba, Balanced forms: canonicity and parameterization, IEEE Trans. Autom. Control AC-30 (1985) 1106–1109.
[20] P.T. Kabamba, Balanced gains and their significance for L^2 model reduction, IEEE Trans. Autom. Control AC-30 (1985) 690–693.
[21] T. Kailath, Linear Systems, Englewood Cliffs, NJ, 1980.
[22] P.P. Khargonekar, M.A. Rotea, Mixed H_2/H_∞ control: a convex optimization approach, IEEE Trans. Autom. Control 36 (1991) 824–837.
[23] W. Krajewski, A. Lepschy, G.A. Main, U. Viaro, Optimality conditions in multivariable L^2 model reduction, J. Franklin Inst. 330 (1993) 431–439.
[24] U.-L. Ly, A.E. Bryson, R.H. Cannon, Design of low-order compensators using parameter optimization, Automatica 21 (1985) 315–318.
[25] B.C. Moore, Principal component analysis in linear systems, IEEE Trans. Autom. Control AC-26 (1981) 17–32.
[26] A.P. Morgan, A.J. Sommese, Coefficient-parameter polynomial continuation, Appl. Math. Comput. 29 (1989) 123–160 [Erratum 51 (1992) 207].
[27] A.P. Morgan, A.J. Sommese, A homotopy for solving general polynomial systems that respects m-homogeneous structures, Appl. Math. Comput. 24 (1987) 101–113.
[28] A.P. Morgan, A.J. Sommese, Computing all solutions to polynomial systems using homotopy continuation, Appl. Math. Comput. 24 (1987) 115–138 [Erratum 51 (1992) 209].
[29] R. Ober, Balanced parameterization of classes of linear systems, SIAM J. Control Optim. 2 (1991) 1251–1287.

[30] M.A. Peters, A.A. Stoorvogel, Mixed H_2/H_∞ control in a stochastic framework, Linear Algebra Appl. 205–206 (1994) 971–996.
[31] C.R. Rao, S.K. Mitra, Generalized Inverse of Matrices and its Applications, Wiley, New York, 1971.
[32] J.T. Spanos, M.H. Milman, D.L. Mingori, Optimal model reduction and frequency-weighted extension, AIAA paper 90-3345-CP, 1990, pp. 271–284.
[33] L.T. Watson, Globally convergent homotopy algorithms for nonlinear systems of equations, Nonlinear Dyn. 1 (1990) 143–191.
[34] L.T. Watson, Globally convergent homotopy methods: a tutorial, Appl. Math. Comput. 31BK (1989) 369–396.
[35] L.T. Watson, Numerical linear algebra aspects of globally convergent homotopy methods, SIAM Rev. 28 (1986) 529–545.
[36] L.T. Watson, S.C. Billups, A.P. Morgan, HOMPACK: a suite of codes for globally convergent homotopy algorithms, ACM Trans. Math. Software 13 (1987) 281–310.
[37] L.T. Watson, R.T. Haftka, Modern homotopy methods in optimization, Comput. Meth. Appl. Mech. Eng. 74 (1989) 289–305.
[38] D.A. Wilson, Optimum solution of model-reduction problem, Proc. IEE 117 (1970) 1161–1165.
[39] K. Zhou, Frequency-weighted L^∞ norm and optimal Hankel norm model reduction, IEEE Trans. Autom. Control 40 (1995) 1687–1699.
[40] D. Žigić, L.T. Watson, C.A. Beattie, Contragredient transformations applied to the optimal projection equations, Linear Algebra Appl. 188–189 (1993) 665–676.
[41] D. Žigić, L.T. Watson, E.G. Collins Jr., D.S. Bernstein, Homotopy approaches to the H^2 reduced order model problem, J. Math. Syst. Estim. Control 3 (1993) 173–205.