A HOMOTOPY ALGORITHM FOR DIGITAL OPTIMAL PROJECTION CONTROL

GASD-HADOC

Emmanuel G. Collins, Jr., Stephen Richter, and Lawrence D. Davis
Harris Corporation
Government Aerospace Systems Division
MS 22/4847
Melbourne, FL 32902

Prepared for:
NASA/MSFC
Marshall Space Flight Center
Huntsville, AL 35812

Contract NAS8-38575

Table of Contents

APPENDICES

1. Introduction and Nomenclature

The linear-quadratic-gaussian (LQG) compensator [1–3] has been developed to facilitate the design of control laws for multi-input multi-output (MIMO) systems. An LQG compensator minimizes a quadratic performance index and (under mild conditions) is guaranteed to yield an internally stable closed-loop system. Unfortunately, however, the minimal dimension of an LQG compensator is almost always equal to the dimension of the plant and can thus often violate practical implementation constraints on controller order. This deficiency is especially highlighted when considering control-design for high-order systems such as flexible space structures. Hence, a very relevant area of research is the development of methodologies that will enable the design of optimal controllers whose dimension is less than that of the design plant (i.e., reduced-order controllers).

Two main approaches have been developed to tackle the reduced-order design problem. The first approach attempts to develop approximations to the optimal reduced-order controller by reducing the dimension of an LQG controller [4–11]. These methods are attractive because they require relatively little computation and should be used if possible. Unfortunately, they tend to yield controllers that either destabilize the system or have poor performance as the requested controller dimension is decreased and/or the requested authority level is increased. Hence, if used in isolation, these methods do not yield a reliable methodology for reduced-order design.

The second approach attempts to directly synthesize an optimal, reduced-order controller by a numerical optimization scheme [12–25]. Almost all of these schemes are parameter optimization approaches; that is, they represent the controller by some parameter vector and attempt to find the vector that optimizes the cost functional. Unfortunately, most of these schemes have only local convergence properties and hence have the potential of failure if the initial controller is not "close" to the optimal controller. One exception is the homotopy algorithm described in [20,25]. A *homotopy* allows an initial controller to be deformed gradually into the desired optimal controller by following a homotopy path. These schemes are particularly useful because they have *global convergence properties.* Hence, this algorithm does not require the initial controller to be near the optimal controller. The algorithm is based on solving a set of "optimal projection" equations [26,27] that are a characterization of the necessary conditions for optimal reduced-order control. Unfortunately, the algorithm has sublinear convergence properties and the convergence slows at higher authority levels and may fail.

This volume describes a new homotopy algorithm for discrete-time systems which has been implemented in MATLAB. The homotopy algorithm is based on a parameter optimization formulation. This algorithm shares the global convergence properties of the homotopy algorithm of [20,25] but potentially has quadratic or superlinear convergence rates. The results reported here may offer the foundation for a reliable approach to optimal, reduced-order controller design.

Nomenclature

$Y \geq Z$	$Y - Z$ is nonnegative definite
$Y > Z$	$Y - Z$ is positive definite
$z_{ij}, Z_{i,j}$ or $Z_{(i,j)}$	(i,j) element of matrix Z
I_r	$r \times r$ identity matrix
$Z^\#$	the group generalized inverse of the square matrix Z satisfying rank Z = rank Z^2 [28,29]
$Z^{\frac{1}{2}}$	the (unique) nonnegative definite square root of Z $(Z^{\frac{1}{2}}Z^{\frac{1}{2}} = Z)$, $Z = Z^{\mathrm{T}} \geq 0$
tr$Z$	trace of square matrix Z
$\lVert Z \rVert_A$	absolute norm of matrix Z $(\lVert Z \rVert_A = \max_{i,j} \mid z_{ij} \mid)$
vec$(\cdot)$	the invertible linear operator defined such that $\mathrm{vec}(s) \triangleq [s_1^{\mathrm{T}}\ s_2^{\mathrm{T}} \cdots s_q^{\mathrm{T}}]^{\mathrm{T}},\ S \in \mathbb{R}^{p \times q}$ where $s_j \in \mathbb{R}^p$ denotes the j^{th} column of S.
$e_m^{(i)}$	the m-dimensional column vector whose i^{th} element equals one and whose additional elements are zeros.
$E_{m \times n}^{(i,j)}$	the $m \times n$ matrix whose (i,j) element equals one and whose additional elements are zero $(= e_m^{(i)} e_m^{(j)^{\mathrm{T}}})$.
$\mathcal{E}_n^{(i)}$	the $m \times m$ matrix whose i^{th} row has all unity elements and whose additional rows are zero.
$\bar{Z}$	for the square matrix Z, $\bar{Z}$ is the identically dimensioned matrix defined by $\bar{z}_{ij} = z_{ii}$.
$Y * Z$	Hadamard product of Y and $Z([y_{ij}z_{ij}])$ (Y and Z must have identical dimensions.)
Y/Z	matrix whose (i,j) element is y_{ij}/z_{ij} (Y and Z must have identical dimensions.) ⟨MATLAB notation⟩
$x = A\backslash b$	x is the least squares solution to $Ax = b$
N_m	$m \times m$ matrix having unity elements (i.e., $N_{m,ij} = 1$)
$[x^{\mathrm{T}}]_{\mathrm{row}-i}$	matrix whose i^{th} row is given by the row vector

	x^{T} and whose additional rows are zero. (The size of the matrix is understood from the context)
$[x]_{\mathrm{col}-j}$	matrix whose j^{th} column is given by the column vector x and whose additional columns are zero. (The size of the matrix is understood from the context.)
$Z(k,:)$	k^{th} row of the matrix Z (MATLAB notation)
$Z(:,k)$	k^{th} column of the matrix Z (MATLAB notation)
$\underline{XYZ}(k,:)$	k^{th} row of the matrix XYZ
$\underline{XYZ}(:,k)$	k^{th} column of the matrix XYZ
$\begin{bmatrix} Z_{11} & Z_{12} \\ SYM & Z_{22} \end{bmatrix}$	partitioned symmetric matrix whose (1,1), (2,2) and (1,2) matrix partitions are as given.

References

1. M. Athans, "The Role and Use of the Stochastic Linear-Quadratic-Gaussian Problem in Control System Design,"*IEEE Trans. Autom. Contr.,* Vol. AC-16, pp. 529-552, Dec. 1971.

2. B. D. O. Anderson and J.B. Moore, Linear Optimal Control, Prentice-Hall, Englewood Cliffs, N.J., 1971.

3. H. Kwakernaak and R. Sivan, Linear Optimal Control Systems, John Wiley and Sons, New York, 1972.

4. A. Yousuff and R.E. Skelton, "A Note on Balanced Controller Reduction," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 254-257, October 1984.

5. A. Yousuff and R.E. Skelton, "Controller Reduction by Component Cost Analysis," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 520-530, June 1984.

6. D.C. Hyland, "Comparison of Various Controller-Reduction Methods: Suboptimal Versus Optimal Projection," *Proc. AIAA Dynamics Specialists Conf.*, pp. 381-389, Palm Spring, CA, May 1984.

7. A. Yousuff and R.E. Skelton, "An Optimal Controller Reduction by Covariance Equivalent Realizations," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 56-60, Jan. 1986.

8. Y. Liu and B. D. O. Anderson, "Controller Reduction Via Stable Factorization and Balancing," *Int. J. Control,* Vol. 44, pp. 507-531, 1986.

9. B. D. O. Anderson and Y. Liu, "Controller Reduction: Concepts and Approaches," *Proc. 1987 American Control Conference,* Minnesota, pp. 1–9, 1987.

10. C. De Villemagne and R.E. Skelton, "Controller Reduction Using Canonical Interactions," *IEEE Trans. Autom. Contr.*, Vol. 33, pp. 740-750, Aug. 1988.

11. Y. Liu, B. D. O. Anderson and U-L. Ly, "Coprime Factorization Controller Reduction with Bezout Identity Induced Frequency Weighting," *Automatica*, Vol. 26, pp. 233-249, 1990.

12. W. S. Levine, T. L. Johnson, and M. Athans, "Optimal Limited State Variable Feedback Controllers for Linear Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-16, pp. 785-793, Dec. 1971.

13. G. D. Martin and A. E. Bryson, Jr., "Attitude Control of a Flexible Spacecraft," *AIAA J. Guid. Contr.,* Vol. 3, Jan. 1980.

14. V. Mukhopadhyay, J. R. Newsom, and I. Abel, "Reduced-Order Optimal Control Law Synthesis for Flutter Suppression," *AIAA J. Guid. Contr. Dyn.*, Vol. 4, pp. 389-395, October 1982.

15. P. Fleming, "A Non-Linear Programming Approach to the Computer-Aided Design of Regulators Using a Linear-Quadratic Formulation," *Int. J. Contr.*, Vol. 42, pp. 257-268, 1985.

16. J. R. Newsom and V. Mukhopadhyay, "A Multiloop Robust Controller Design Study Using Singular Value Gradients," *AIAA J. Guid. Contr. Dyn.*, Vol. 8, pp. 514-519, Aug. 1985.

17. 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.

18. V. Mukhopadhyay, "Stability Robustness Improvement Using Constrained Optimization Techniques," *AIAA J. Guid. Contr. Dyn.*, Vol. 10, pp. 172-177, October 1987.

19. H. T. Toivonen and P. M. Makila, "Newton's Method for Solving Parametric Linear Quadratic Control Problems," *Int. J. Contr.*, Vol. 46, pp. 897-911, 1987.

20. S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality," *Proc. Amer. Contr. Conf.*, pp. 1527-1531, Minneapolis, MN, June 1987.

21. P. M. Makila and H. T. Toivonen, "Computational Methods for Parametric LQ Problems–A Survey," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 658-671, Aug. 1987.

22. F. S. Kramer and A. J. Calise, "Fixed-Order Dynamic Compensation for Multivariable Linear Systems," *AIAA J. Guid. Contr. Dyn.*, Vol. 11, pp. 80-85, Jan. 1988.

23. V. Mukhopadhyay, "Digital Robust Control Law Synthesis Using Constrained Optimization," *AIAA J. Guid. Contr. Dyn.*, Vol. 12, pp. 175-181, October 1989.

24. A. J. Calise and J. V. R. Prasad, "Approximate Loop Transfer Recovery Method for Designing Fixed-Order Compensators," *AIAA J. Guid. Contr. Dyn.*, Vol. 13, pp. 297-302, October 1990.

25. S. Richter and E. G. Collins, Jr., "A Homotopy Algorithm for Reduced-Order Controller Design Using the Optimal Projection Equations," *Proc. IEEE Conf. Dec. Contr.*, pp. 932-933, Tampa, FL, December 1989.

26. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation,"*IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034-1037, Nov. 1984.

27. D. S. Bernstein, L. D. Davis and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order, Discrete-Time Modelling, Estimation and Control," *AIAA J. Guid. Contr. Dyn.*, Vol. 9, pp. 288-293, 1986.

28. A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley and Sons, New York, 1974.

29. S. L. Campbell and C. D. Meyer, Jr. *Generalized Inverses of Linear Transformations*, Pitman, San Francisco, 1979.

2. Optimal Reduced-Order Discrete-Time Dynamic Compensation

Consider the discrete-time system

$$x(k+1) = Ax(k) + Bu(k) + w_1(k) \tag{2.1a}$$

$$y(k) = Cx(k) + Du(k) + w_2(k) \tag{2.1b}$$

where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $y \in \mathbb{R}^{n_y}$, $w_1 \in \mathbb{R}^{n_x}$ is a white noise disturbance with covariance $V_1 \geq 0$, $w_2 \in \mathbb{R}^{n_y}$ is white observation noise with covariance $V_2 > 0$, and w_1 and w_2 have cross covariance $V_{12} \in \mathbb{R}^{n_x \times n_y}$. If $D = 0$, we desire to design a fixed-order dynamic compensator,

$$x_c(k+1) = A_c x_c(k) + B_c y(k) \tag{2.2a}$$

$$u(k) = -C_c x_c(k) - D_c y(k), \tag{2.2b}$$

or if $D \neq 0$, we desire to design a fixed-order dynamic compensator

$$x_c(k+1) = A_c x_c(k) + B_c y(k) \tag{2.3a}$$

$$u(k) = -C_c x_c(k) \tag{2.3b}$$

which minimizes the steady-state performance criterion

$$\mathcal{J}(A_c, B_c, C_c, D_c) = \lim_{k \to \infty} E[x^{\mathrm{T}}(k) R_1 x(k) + 2x^{\mathrm{T}}(k) R_{12} u(k) + u^{\mathrm{T}}(k) R_2 u(k)] \tag{2.4}$$

where $x_c \in \mathbb{R}^{n_c}$, $n_c \leq n_x$, $R_1 = R_1^{\mathrm{T}} \geq 0$, and $R_2 = R_2^{\mathrm{T}} \geq 0$. We will call this problem the optimal reduced-order dynamic compensation problem for discrete-time systems.

The closed-loop system corresponding to (2.1) and (2.2) or (2.1) and (2.3) can be expressed as

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{w}(k) \tag{2.5}$$

where

$$\tilde{x}(k) = \begin{bmatrix} x(k) \\ x_c(k) \end{bmatrix}, \quad \tilde{w}(k) = \begin{bmatrix} w_1(k) - BD_c w_2(k) \\ B_c w_2(k) \end{bmatrix} \tag{2.6}$$

$$\tilde{A} = \begin{bmatrix} A - BD_cC & -BC_c \\ B_cC & A_c - B_cDC_c \end{bmatrix} \tag{2.7}$$

and it understood that either D of D_c is identically zero in (2.6) and (2.7). In addition, the cost (2.4) can be expressed as

$$\mathcal{J}(A_c, B_c, C_c, D_c) = \lim_{k \to \infty} E[\tilde{x}^{\mathrm{T}}(k)\tilde{R}\tilde{x}(k)] + E[w_2^{\mathrm{T}}(k) D_c^{\mathrm{T}} R_2 D_c w_2(k)] \tag{2.8}$$

where

$$\tilde{R} = \begin{bmatrix} \tilde{R}_{11} & \tilde{R}_{12} \\ \tilde{R}_{12}^{\mathrm{T}} & \tilde{R}_{22} \end{bmatrix} \tag{2.9}$$

and

$$\tilde{R}_{11} \triangleq R_1 - C^{\mathrm{T}} D_c^{\mathrm{T}} R_{12}^{\mathrm{T}} - R_{12} D_c C + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c C \tag{2.10a}$$

$$\tilde{R}_{12} \triangleq -R_{12} C_c + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 C_c \tag{2.10b}$$

$$\tilde{R}_{22} \triangleq C_c^{\mathrm{T}} R_2 C_c. \tag{2.10c}$$

To guarantee that the cost J is finite and independent of initial conditions we restrict our attention to the set of stabilizing compensators,

$$\mathcal{S}_c \triangleq \{(A_c, B_c, C_c, D_c): \quad \tilde{A} \text{ is asymptotically stable}\}. \tag{2.11}$$

Assume $(A_c, B_c, C_c, D_c) \in \mathcal{S}_c$ and define $\tilde{Q}$ and $\tilde{P}$ to be the closed-loop steady state covariance and its dual, i.e.,

$$\tilde{Q} = \tilde{A}\tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{V} \tag{2.12}$$

$$\tilde{P} = \tilde{A}^{\mathrm{T}}\tilde{P}\tilde{A} + \tilde{R} \tag{2.13}$$

where

$$\tilde{V} = \begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{12}^{\mathrm{T}} & \tilde{V}_{22} \end{bmatrix} \tag{2.14}$$

and

$$\tilde{V}_{11} \triangleq V_1 - B D_c V_{12}^{\mathrm{T}} - V_{12} D_c^{\mathrm{T}} B^{\mathrm{T}} + B D_c V_2 D_c^{\mathrm{T}} B^{\mathrm{T}} \tag{2.15a}$$

$$\tilde{V}_{12} \triangleq V_{12} B_c^{\mathrm{T}} - B D_c V_2 B_c^{\mathrm{T}} \tag{2.15b}$$

$$\tilde{V}_{22} \triangleq B_c V_2 B_c^{\mathrm{T}}. \tag{2.15c}$$

Then, the cost can be expressed as

$$\mathcal{J}(A_c, B_c, C_c, D_c) = \mathrm{tr}(\tilde{Q}\tilde{R}) + \mathrm{tr}(D_c^{\mathrm{T}} R_2 D_c V_2). \tag{2.16}$$

Also, $\tilde{P}$ and $\tilde{Q}$ can be expressed in the partitioned forms

$$\tilde{P} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{12}^{\mathrm{T}} & \tilde{P}_{22} \end{bmatrix}, \quad \tilde{P}_{11} \in \mathbb{R}^{n_x \times n_x} \tag{2.17}$$

$$\tilde{Q} = \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}^{\mathrm{T}} & \tilde{Q}_{22} \end{bmatrix}, \quad \tilde{Q}_{11} \in \mathbb{R}^{n_x \times n_x}. \tag{2.18}$$

Notice that $\tilde{Q}_{11}$ is the covariance of the plant states, $\tilde{Q}_{22}$ is the covariance of the compensator states and $\tilde{Q}_{22}$ is the cross-covariance of the plant and controller states.

Since the value of J is independent of the internal realization of the compensator, in what follows we will further restrict our attention to minimal compensators. Hence, we define the admissable set,

$$\mathcal{S}_c^+ = \{(A_c, B_c, C_c, D_c) \in \mathcal{S} : \quad (A_c, B_c) \text{ is controllable}, (A_c, C_c) \text{ is observable}\}. \tag{2.19}$$

Note that $\mathcal{S}_c^+$ is an open set.

Optimal Projection theory can be used to characterize all admissable extremals of the optimal reduced-order dynamic compensation problem for discrete-time systems. Before presenting the main theorems we present an important Lemma and some definitions which are useful in stating the main results of optimal projection theory. The lemma also gives many properties of the optimal projection solution (see Theorem 2.1).

Lemma 2.1 [1]. Suppose $\hat{Q} \in \mathbb{R}^{n_x \times n_x}$ and $\hat{P} \in \mathbb{R}^{n_x \times n_x}$ are symmetric and nonnegative definite and rank $\hat{Q}\hat{P} = n_c$. Then, the following statements hold:

(i) $\hat{Q}\hat{P}$ is diagonalizable and has nonnegative eigenvalues.

(ii) The $n_x \times n_x$ matrix

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#} \tag{2.20}$$

is idempotent, i.e., $\tau^2 = \tau$ (τ is an oblique projection) and

$$\text{rank} \;\; \tau = n_c. \tag{2.21}$$

Thus, if τ is given by (2.18), then there exists a nonsingular matrix $W \in \mathbb{R}^{n_x \times n_x}$ such that

$$\tau = W \begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix} W^{-1}. \tag{2.22}$$

(iii) There exists $G, \Gamma \in \mathbb{R}^{n_c \times n_x}$ and nonsingular $M \in \mathbb{R}^{n_c \times n_c}$ such that

$$\hat{Q}\hat{P} = G^{\mathrm{T}} M \Gamma \tag{2.23}$$

$$\Gamma G^{\mathrm{T}} = I_{n_c}. \tag{2.24}$$

(iv) If G, Γ and M satisfy property (iii) then

$$\text{rank} \;\; G = \text{rank} \;\; \Gamma = \text{rank} \;\; M = n_c \tag{2.25}$$

$$(\hat{Q}\hat{P})^{\#} = G^{\mathrm{T}} M^{-1} \Gamma \tag{2.26}$$

$$\tau = G^{\mathrm{T}} \Gamma \tag{2.27}$$

$$\tau G^{\mathrm{T}} = G^{\mathrm{T}}, \quad \Gamma \tau = \tau. \tag{2.28}$$

(v) The matrices G, Γ and M satisfying property (iii) are unique except for a change of basis in $\mathbb{R}^{n_c}$, i.e., if G', Γ' and M' also satisfy property (iv), then there exists nonsingular $T_c \in \mathbb{R}^{n_c \times n_c}$ such that $G' = T_c^{\mathrm{T}} G$, $\Gamma' = T_c^{-1} \Gamma$, $M' = T_c^{-1} M T_c$. Furthermore, all such M are diagonalizable with positive eigenvalues.

(vi) Finally, if rank $\hat{Q}$ = rank $\hat{P}$ = rank $\hat{Q}\hat{P}$ = n_c, there exists a nonsingular transformation $W \in \mathbb{R}_{n_x \times n_x}$ such that

$$\hat{P} = W^{-\mathrm{T}} \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} W^{-1} \tag{2.29}$$

$$\hat{Q} = W \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} W^{\mathrm{T}} \tag{2.30}$$

where $\Omega \in \mathbb{R}^{n_c \times n_c}$ is diagonal and nonsingular. In addition,

$$\hat{P} = \tau^{\mathrm{T}} \hat{P} = \hat{P} \tau = \tau^{\mathrm{T}} \hat{P} \tau \tag{2.31}$$

$$\hat{Q} = \tau \hat{Q} = \hat{Q} \tau^{\mathrm{T}} = \tau \hat{Q} \tau^{\mathrm{T}}. \tag{2.32}$$

Remark 2.1. The transformation W in statement (vii) meets the requirements of statement (ii).

Definition 2.1. A triple (G, M, Γ) satisfying property (iii) of Lemma 2.1 is a *projective factorization* of $\hat{Q}\hat{P}$.

To optimize (2.8) subject to the constraint (2.12) we form the Lagrangian

$$\mathcal{L}(A_c, B_c, C_c, D_c, \tilde{P}, \tilde{Q}) \triangleq \mathrm{tr}[\tilde{Q}\tilde{R} + \tilde{P}(\tilde{A}\tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{V} - \tilde{Q}) + D_c^{\mathrm{T}} R_2 D_c V_2] \tag{2.33}$$

where $\tilde{P} \in \mathbb{R}^{(n_x+n_c)\times(n_x+n_c)}$ is the Lagrange multiplier. The stationary conditions are then given by

$$\frac{\partial \mathcal{L}}{\partial \tilde{P}} = 0, \quad \frac{\partial \mathcal{L}}{\partial \tilde{Q}} = 0 \tag{2.34}$$

$$\frac{\partial \mathcal{L}}{\partial A_c} = 0, \quad \frac{\partial \mathcal{L}}{\partial B_c} = 0, \quad \frac{\partial \mathcal{L}}{\partial C_c} = 0, \quad \frac{\partial \mathcal{L}}{\partial D_c} = 0. \tag{2.35}$$

Definition 2.2. A compensator (A_c, B_c, C_c, D_c) is an *extremal* of the optimal reduced-order dynamic compensation problem for discrete-time systems if it satisfies the stationary conditions (2.32).

Definition 2.3. A compensator (A_c, B_c, C_c, D_c) is an *admissible extremal* of the optimal reduced-order dynamic compensation problem for discrete-time systems if it is an extremal and is also in $\mathcal{S}_c^+$.

Remark 2.2. The optimal (admissible) reduced-order dynamic compensator for discrete-time systems (if it exists) can be found by computing all admissible extremals.

We can now state in the form of two theorems the basic result of Optimal Projection theory, namely a set of necessary conditions which characterize admissible extremals of the optimal fixed-order dynamic compensation problem. For convenience define

$$P_a \triangleq B^{\mathrm{T}}PA + R_{12}^{\mathrm{T}}, \quad Q_a \triangleq AQC^{\mathrm{T}} + V_{12} \tag{2.36a,b}$$

$$R_{2,a} \triangleq R_2 + B^{\mathrm{T}}PB, \quad V_{2,a} \triangleq V_2 + CQC^{\mathrm{T}}. \tag{2.37a,b}$$

Theorem 2.1 [3]. Suppose $D = 0$ and (A_c, B_c, C_c, D_c) is an admissible extremal of the optimal reduced-order dynamic compensation problem for discrete-time systems. Then, there exist nonnegative-definite matrices $P, Q, \hat{P}$ and $\hat{Q}$ such that A_c, B_c, C_c and D_c are given by

$$A_c = \Gamma(A - BR_{2,a}^{-1}P_a - Q_aV_{2,a}^{-1} + Q_aV_{2,a}^{-1}DR_{2,a}^{-1}P_a - BD_cC)G^{\mathrm{T}} \tag{2.38}$$

$$B_c = \Gamma(Q_aV_{2,a}^{-1} + BD_c) \tag{2.39}$$

$$C_c = (R_{2,a}^{-1}P_a + D_cC)G^{\mathrm{T}} \tag{2.40}$$

$$D_c = R_{2,a}^{-1}(B^{\mathrm{T}}PAQC^{\mathrm{T}} + R_{12}^{\mathrm{T}}QC^{\mathrm{T}} + B^{\mathrm{T}}PV_{12}V_{2,a}^{-1} \tag{2.41}$$

for some projective factorization (G, M, Γ) of $\hat{Q}, \hat{P}$ such that the following conditions are satisfied:

$$\begin{aligned} P =& A^{\mathrm{T}}PA + R_1 - P_a^{\mathrm{T}}R_{2,a}^{-1}P_a \\ &+ \tau_\perp^{\mathrm{T}}[(A - Q_aV_{2,a}^{-1}C)^{\mathrm{T}}\hat{P}(A - Q_aV_{2,a}^{-1}C) + (P_a + R_{2,a}D_cC)^{\mathrm{T}}R_{2,a}^{-1}(P_a + R_{2,a}D_cC)\tau_\perp \end{aligned} \tag{2.42}$$

$$\begin{aligned} Q =& AQA^{\mathrm{T}} + V_1 - Q_aV_{2,a}^{-1}Q_a^{\mathrm{T}} \\ &+ \tau_\perp[(A - BR_{2,a}^{-1}P_a)\hat{Q}(A - BR_{2,a}^{-1}P_a)^{\mathrm{T}} + (Q_a + BD_cV_{2,a})V_{2,a}^{-1}(Q_a + BD_cV_{2,a})^{\mathrm{T}}]\tau_\perp^{\mathrm{T}} \end{aligned} \tag{2.43}$$

$$\begin{aligned} \hat{P} =& \tau^{\mathrm{T}}(A - Q_aV_{2,a}^{-1}C)^{\mathrm{T}}\hat{P}(A - Q_aV_{2,a}^{-1}C)\tau \\ &+ \tau^{\mathrm{T}}(P_a + R_{2,a}D_cC)^{\mathrm{T}}R_{2,a}^{-1}(P_a + R_{2,a}D_cC)\tau \end{aligned} \tag{2.44}$$

$$\begin{aligned} \hat{Q} =& \tau(A - BR_{2,a}^{-1}P_a)\hat{Q}(A - BR_{2,a}^{-1}P_a)^{\mathrm{T}}\tau^{\mathrm{T}} \\ &+ \tau(Q_a + BD_cV_{2,a})V_{2,a}^{-1}(Q_a + BD_cV_{2,a})^{\mathrm{T}}\tau^{\mathrm{T}} \end{aligned} \tag{2.45}$$

$$\text{rank } \hat{P} = \text{rank } \hat{Q} = \text{rank } \hat{Q}\hat{P} = n_c \tag{2.46}$$

$$\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^{\#} \tag{2.47}$$

Theorem 2.2 [3]. Suppose $D \neq 0$ and (A_c, B_c, C_c) is an admissible extremal of the optimal reduced-order dynamic compensation problem for discrete-time systems. Then, there exist nonnegative-definite matrices $P, Q, \hat{P}$ and $\hat{Q}$ such that A_c, B_c and C_c are given by (2.38)–(2.40) with $D_c = 0$ for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$ such that conditions (2.42)–(2.47) are satisfied with $D_c = 0$.

Remark 2.3. Theorem 2.1 is a modification of earlier results [2,4]. The primary difference is that the $\hat{P}$ and $\hat{Q}$ in Theorem 2.1 satisfy the rank conditions (2.46), which parallels the corresponding continuous-time theory [4,5], whereas the $\hat{P}$ and $\hat{Q}$ in [2] and [4] do not satisfy these rank conditions.

The following corollary characterizes the optimal, full-order, discrete-time controller.

Corollary 2.1. If $n_c = n_x$, then one can choose $\tau = \Gamma = G = I_{n_x}$, such that $\tau_\perp = 0$ and (2.38)–(2.46) reduce to

$$A_c = A - BC_c - B_cC - BD_cC \tag{2.48}$$

$$B_c = Q_aV_{2,a}^{-1} - BD_c \tag{2.49}$$

$$C_c = R_{2,a}^{-1}P_a - D_cC \tag{2.50}$$

$$D_c = R_{2,a}^{-1}(B^{\mathrm{T}}PAQC^{\mathrm{T}} + R_{12}^{\mathrm{T}}QC^{\mathrm{T}} + B^{\mathrm{T}}PV_{12})V_{2,a}^{-1} \tag{2.51}$$

where

$$P = A^{\mathrm{T}}PA + R_1 - P_a^{\mathrm{T}}R_{2,a}^{-1}P_a \tag{2.52}$$

$$Q = AQA^{\mathrm{T}} + V_1 - Q_aV_{2,a}^{-1}Q_a^{\mathrm{T}} \tag{2.53}$$

$$\begin{aligned}\hat{P} = {} & (A - Q_aV_{2,a}^{-1}C)^{\mathrm{T}}\hat{P}(A - Q_aV_{2,a}^{-1}C) \\ & + (P_a + R_{2,a}D_cC)^{\mathrm{T}}R_{2,a}^{-1}(P_a + R_{2,a}D_cC)\end{aligned} \tag{2.54}$$

$$\begin{aligned}\hat{Q} = {} & (A - BR_{2,a}^{-1}P_a)\hat{Q}(A - BR_{2,a}^{-1}P_a)^{\mathrm{T}} \\ & + (Q_a + BD_cV_{2,a})V_{2,a}^{-1}(Q_a + BD_cV_{2,a})^{\mathrm{T}}\end{aligned} \tag{2.55}$$

$$\text{rank } \hat{P} = \text{rank } \hat{Q} = \text{rank } \hat{Q}\hat{P} = n_x. \tag{2.56}$$

Remark 2.4. Condition (2.56) requires that the LQG controller (A_c, B_c, C_c, D_c) have minimal order n_x. Also, $\hat{P}$ and $\hat{Q}$ are not needed to compute the controller but are the closed-loop grammians to be used in balanced controller reduction.

Remark 2.5. Notice that in the full-order case (i.e., $n_c = n_x$), without loss of generality one can choose $\tau = G = \Gamma = I_{n_x}$ and (2.42) and (2.43) reduce to the standard regulator and observer Riccati equations and (2.38)–(2.41) yield the usual LQG expressions. It can be shown that (2.44)–(2.46) are equivalent to the requirement that the controller (A_c, B_c, C_c) be minimal.

Theorem 2.4 [6]. Suppose there exists nonnegative definite matrices $Q, P, \hat{Q}$ and $\hat{P}$ satisfying (2.40)-(2.45) and A_c, B_c, C_c and D_c satisfy (2.36)–(2.39). Then, the compensator (A_c, B_c, C_c) is an extremal of the optimal fixed-order dynamic compensation problem. Furthermore the following are equivalent:

(i) $\tilde{A}$ is stable

(ii) $(\tilde{A}, \tilde{V}^{\frac{1}{2}})$ is stabilizable

(iii) $(\tilde{A}, \tilde{R}^{\frac{1}{2}})$ is detectable.

In addition,

$$(A_c, B_c) \text{ is controllable} \iff A_c + B_c C G^{\mathrm{T}} \text{ is stable} \tag{2.57}$$

$$(A_c, C_c) \text{ is observable} \iff A_c + \Gamma B C_c \text{ is stable.} \tag{2.58}$$

In the homotopy algorithms to be subsequently defined the optimal projection equations (2.42)–(2.45) due to their relationship to standard LQG equations can be used to give insights into the development of initializing controllers. However, the homotopy algorithms will be based directly on the gradient of the cost functional.

References

1. D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance via Fixed-Order Dynamic Compensation with Guaranteed Cost Bounds," *Math. Contr. Sig. Sys.*, Vol. 3, pp. 139–163, 1990.

2. W. M. Haddad, *Robust Optimal Projection Control-System Synthesis*, Ph.D. Dissertation, Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL, October 1987.

3. W. M. Haddad, private communication, Nov. 1988.

4. D. S. Bernstein, L. D. Davis, and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order, Discrete-Time Modelling, Estimation and Control," *J. Guid. Contr. Dyn.*, Vol. 9, pp. 288–293, 1986.

5. D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed–Order Dynamic Compensation," *IEEE Trans. Autom. Contr.*, Vol. AC–29, pp. 1034-1037, 1984.

6. D. S. Bernstein, private communication, Nov. 1990.

3. Review of Homotopy Methods

A "homotopy" is a continuous deformation of one function into another. Over the past several years, homotopy or continuation methods (whose mathematical basis is algebraic topology and differential topology [1]) have received significant attention in the mathematics literature and have been applied successfully to several important problems [2–7]. Recently, the engineering literature has also begun to recognize the utility of these methods for engineering applications (see e.g. [8–10]). The purpose of this section is to provide a very brief description of homotopy methods for finding the solutions of nonlinear algebraic equations.

The reader is referred to [7,8,11,12] for additional details.

The basic problem is as follows. Given set Θ and Φ contained in $\mathbb{R}^n$ and a mapping $F : \Theta \rightarrow \Phi$, find solutions to

$$F(\theta) = 0. \tag{3.1}$$

Homotopy methods embed the problem (3.1) in a larger problem. In particular let $H : \Theta \times [0,1] \rightarrow \mathbb{R}^n$ be such that:

1) $H(\theta, 1) = F(\theta).$ $\quad$ **(3.2)**

2) There exists at least one known $\theta_0 \in \mathbb{R}^n$ which is a solution to $H(\cdot, 0) = 0$, i.e.,

$$H(\theta_0, 0) = 0. \tag{3.3}$$

3) There exists a continuous curve $(\theta(\lambda), \lambda)$ in $\mathbb{R}^n \times [0,1]$ such that

$$H(\theta(\lambda), \lambda) = 0 \text{ for } \lambda \in [0,1] \tag{3.4}$$

with

$$(\theta(0), 0) = (\theta_0, 0). \tag{3.5}$$

4) The space $\Theta \times [0,1]$ has a differential structure so that the curve $(\theta(\lambda), \lambda)$ is differentiable.

A homotopy algorithm then constructs a procedure to compute the actual curve σ such that the initial solution $\theta(0)$ is transformed to a desired solution $\theta(1)$ satisfying

$$0 = H(\theta(1), 1) = F(\theta(1)). \tag{3.6}$$

Differentiating $H(\theta(\lambda), \lambda) = 0$ with respect to λ to obtain *Davidenko's differential equation*

$$\frac{\partial H}{\partial \theta}\frac{d\theta}{d\lambda} + \frac{\partial H}{\partial \lambda} = 0. \tag{3.7}$$

Together with $\theta(0) = \theta_0$, (3.7) defines an initial value problem which by numerical integration from 0 to 1 yields the desired solution $u(1)$. Some numerical integration schemes are described in [11,12].

References

1. N. G. Lloyd, *Degree Theory*, Cambridge University Press, London, 1978.

2. J. H. Avila, "The Feasibility of Continuation Methods for Nonlinear Equations," *SIAM J. Numer. Anal.*, Vol. 11, pp. 104–144, 1974.

3. H. Wacker, "Continuation Methods," Academic Press, New York, 1978.

4. J. C. Alexander and J. A. Yorke, "The Homotopy Continuation Method: Numerically Implementable Topological Procedures," *Trans. Amer. Math. Soc.*, Vol. 242, pp. 271–284, 1978.

5. C. B. Garcia and W. I. Zangwill, *Pathways to Solutions, Fixed Points and Equilibria,* Prentice-Hall, Englewood Cliffs, NJ, 1981.

6. B. C. Eaves, F. J. Gould, J. O. Peoitgen and M. J. Todd, *Homotopy Methods and Global Convergence,* Plenum Press, New York, 1983.

7. L. T. Watson, "Numerical Linear Algebra Aspects of Globally Convergent Homotopy Methods," *SIAM Rev.*, Vol. 28, pp. 529–545, 1986.

8. S. L. Richter and R. A. DeCarlo, "Continuation Methods: Theory and Applications," *IEEE Trans. Autom. Contr.*, Vol. CAS–30, pp. 347–352, 1983.

9. P. T. Kabamba, R. W. Longman and S. Jian-Guo, "A Homotopy Approach to the Feedback Stabilization of Linear Systems," *J. Guid. Control Dynamics,* Vol. 10, pp. 422-432, 1987.

10. Y. S. Shin, R. T. Raftka, L. T. Watson and R. H. Plautt, "Tracking Structural Optima as a Function of Available Resources by a Homotopy Method," *Computer Methods in Applied Mechanics and Engineering,* Vol. 70, pp. 151–164, 1988.

11. L. T. Watson, "ALGORITHM 652 HOMPACK: A Suite of Codes for Globally Convergent Homotopy Algorithms," *ACM Trans. Math. Software*, Vol. 13, pp. 281–310, Sept. 1987.

12. L. T. Watson, "Numerical Linear Algebra Aspects of Globally Convergent Homotopy Methods," *SIAM Review,* Vol. 28, pp. 529–545, Dec. 1986.

4. The Homotopy Map and It's Jacobian

If we define

$$\theta \triangleq \begin{bmatrix} \mathrm{vec}(A_c) \\ \mathrm{vec}(B_c) \\ \mathrm{vec}(C_c) \\ \mathrm{vec}(D_c) \end{bmatrix}, \tag{4.1}$$

then the cost functional of Section 2 can be expressed as $\mathcal{J}(\theta)$. The homotopy defined in this section is based on finding θ satisfying

$$0 = f(\theta) \triangleq \frac{\partial \mathcal{J}}{\partial \theta}(\theta). \tag{4.2}$$

It is useful to recognize that

$$\nabla \mathcal{J}(\theta)^{\mathrm{T}} \triangleq \frac{\partial \mathcal{J}}{\partial \theta} = \begin{bmatrix} \mathrm{vec}\frac{\partial \mathcal{J}}{\partial A_c} \\ \mathrm{vec}\frac{\partial \mathcal{J}}{\partial B_c} \\ \mathrm{vec}\frac{\partial \mathcal{J}}{\partial C_c} \\ \mathrm{vec}\frac{\partial \mathcal{J}}{\partial D_c} \end{bmatrix}. \tag{4.3}$$

Expressions for the partial derivatives $\frac{\partial \mathcal{J}}{\partial A_c}, \frac{\partial \mathcal{J}}{\partial B_c}, \frac{\partial \mathcal{J}}{\partial C_c}$, and $\frac{\partial \mathcal{J}}{\partial D_c}$ are derived in Appendix A. Here, we cite only the final results. First, we assume that $\tilde{P}$, $\tilde{Q}$ and $\tilde{Z}$ satisfy

$$\tilde{P} = \tilde{A}^{\mathrm{T}} \tilde{P} \tilde{A} + \tilde{R} \tag{4.4}$$

$$\tilde{Q} = \tilde{A} \tilde{Q} \tilde{A}^{\mathrm{T}} + \tilde{V} \tag{4.5}$$

$$\tilde{Z} = \tilde{Q} \tilde{A}^{\mathrm{T}} \tilde{P} \tag{4.6}$$

and note that $\tilde{P}, \tilde{Q}$ and $\tilde{Z}$ have the partitioned forms

$$\tilde{P} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{12}^{\mathrm{T}} & \tilde{P}_{22} \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}^{\mathrm{T}} & \tilde{Q}_{22} \end{bmatrix}, \quad \tilde{Z} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{21} & \tilde{Z}_{22} \end{bmatrix}, \tag{4.7}$$

where the (1.1) and (2.2) blocks of each matrix are respectively $n \times n$ and $n_c \times n_c$. With this in mind, the cost derivatives are given by

$$\frac{\partial \mathcal{J}}{\partial A_c} = 2\tilde{Z}_{22}^{\mathrm{T}} \tag{4.8}$$

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial B_c} &= 2(\tilde{P}_{12}^{\mathrm{T}} V_{12} - \tilde{P}_{12}^{\mathrm{T}} B D_c V_2 + \tilde{P}_{22} B_c V_2 \\ &\qquad + \tilde{Z}_{12}^{\mathrm{T}} C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} D^{\mathrm{T}}) \end{aligned} \tag{4.9}$$

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial C_c} &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{12} + R_2 D_c C \tilde{Q}_{12} + R_2 C_c \tilde{Q}_{22} \\ &\qquad + B^{\mathrm{T}} \tilde{Z}_{21}^{\mathrm{T}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{\mathrm{T}}) \end{aligned} \tag{4.10}$$

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial D_c} &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{11} C^{\mathrm{T}} + R_2 D_c C \tilde{Q}_{11} C^{\mathrm{T}} + R_2 C_c \tilde{Q}_{12}^{\mathrm{T}} C^{\mathrm{T}} \\ &\qquad - B^{\mathrm{T}} \tilde{P}_{11} V_{12} + B^{\mathrm{T}} \tilde{P}_{11} B D_c V_2 - B^{\mathrm{T}} \tilde{P}_{12} B_c V_2 \\ &\qquad\quad - B^{\mathrm{T}} \tilde{Z}_{11}^{\mathrm{T}} C^{\mathrm{T}} + R_2 D_c V_2). \end{aligned} \tag{4.11}$$

Definition of the homotopy map $h(\theta,\lambda)$

To define the homotopy map we assume that the plant matrices (A,B,C,D), the cost weighting matrix (R_1,R_2,R_{12}) and the disturbance matrices (V_1,V_2,V_{12}) are functions of the Homotopy parameter $\lambda \in [0,1]$. In particular, it is assumed that

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} + \lambda\left(\begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix}\right), \tag{4.12}$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^{\mathrm{T}}(\lambda) & R_2(\lambda) \end{bmatrix} = L_R(\lambda)L_R^{\mathrm{T}}(\lambda) \tag{4.13a}$$

where

$$L_R(\lambda) = L_{R,0} + \lambda(L_{R,f} - L_{R,0}) \tag{4.13b}$$

and $L_{R,0}$ and $L_{R,f}$ satisfy

$$L_{R,0}L_{R,0}^{\mathrm{T}} = \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} \tag{4.13c}$$

$$L_{R,f}L_{R,f}^{\mathrm{T}} = \begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^{\mathrm{T}} & R_{2,f} \end{bmatrix}, \tag{4.13d}$$

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^{\mathrm{T}}(\lambda) & V_2^{\mathrm{T}}(\lambda) \end{bmatrix} = L_V(\lambda)L_V^{\mathrm{T}}(\lambda) \tag{4.14a}$$

where

$$L_V(\lambda) = L_{V,0} + \lambda(L_{V,f} - L_{V,0}) \tag{4.14b}$$

and $L_{V,0}$ and $L_{V,f}$ satisfy

$$L_{V_0}L_{V,0}^{\mathrm{T}} = \begin{bmatrix} V_{1,0} & V_{12,0} \\ V_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix} \tag{4.14c}$$

$$L_{V_f}L_{V,f}^{\mathrm{T}} = \begin{bmatrix} V_{1,f} & V_{12,f} \\ V_{12,f}^{\mathrm{T}} & V_{2,0} \end{bmatrix}. \tag{4.14d}$$

Note that (4.12)-(4.14) imply that $A(0) = A_0$ and $A(1) = A_f$, $B(0) = B_0$ and $B(1) = B_f$, etc ... and it is understood that $A_f, B_f, \ldots$ were referred to previously simply as $A, B, \ldots$. The change in notation is simply for convenience.

The homotopy map $h(\theta,\lambda)$ is defined by

$$h(\theta,\lambda) = \begin{bmatrix} \mathrm{vec}(H_{A_c}(\theta,\lambda)) \\ \mathrm{vec}(H_{B_c}(\theta,\lambda)) \\ \mathrm{vec}(H_{C_c}(\theta,\lambda)) \\ \mathrm{vec}(H_{D_c}(\theta,\lambda)) \end{bmatrix} \tag{4.15}$$

where

$$H_{A_c}(\theta,\lambda) = 2\tilde{Z}_{22}^{\mathrm{T}} \tag{4.16}$$

$$\begin{aligned} H_{B_c}(\theta,\lambda) = 2(&\tilde{P}_{12}^{\mathrm{T}} V_{12} - \tilde{P}_{12}^{\mathrm{T}} B D_c V_2 + \tilde{P}_{22} B_c V_2 \\ &+ \tilde{Z}_{12}^{\mathrm{T}} C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} D^{\mathrm{T}} \end{aligned} \tag{4.17}$$

$$\begin{aligned} H_{C_c}(\theta,\lambda) = 2(&-R_{12}^{\mathrm{T}} \tilde{Q}_{12} + R_2 D_c C \tilde{Q}_{12} + R_2 C_c \tilde{Q}_{22} \\ &- B^{\mathrm{T}} \tilde{Z}_{21}^{\mathrm{T}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{\mathrm{T}} \end{aligned} \tag{4.18}$$

$$\begin{aligned} H_{D_c}(\theta,\lambda) = 2(&-R_{12}^{\mathrm{T}} \tilde{Q}_{11} C^{\mathrm{T}} + R_2 D_c C \tilde{Q}_{11} C^{\mathrm{T}} + R_2 C_c \tilde{Q}_{12}^{\mathrm{T}} C^{\mathrm{T}} \\ &- B^{\mathrm{T}} \tilde{P}_{11} V_{12} + B^{\mathrm{T}} \tilde{P}_{11} B D_c V_2 - B^{\mathrm{T}} \tilde{P}_{12} B_c V_2 \\ &\quad - B^{\mathrm{T}} \tilde{Z}_{11}^{\mathrm{T}} C^{\mathrm{T}} + R_2 D_c V_2). \end{aligned} \tag{4.19}$$

The Jacobian of the homotopy map

We now consider that computation of $\nabla h(\theta,\lambda)^{\mathrm{T}}$, the Jacobian of $h(\theta,\lambda)$. Note that

$$\nabla h(\theta,\lambda)^{\mathrm{T}} = \begin{bmatrix} \frac{\partial h}{\partial \theta} & \frac{\partial h}{\partial \lambda} \end{bmatrix}. \tag{4.20}$$

Recalling that θ is defined by (4.1), such that for some integers k and ℓ, θ_j is given by

$$\theta_j = a_{c,k_\ell},\ \theta_j = b_{c,k_\ell},\ \theta_j = c_{c,k_\ell},\ \text{or } \theta_j = d_{c,k_\ell}. \tag{4.21}$$

It follows from (4.6) that $\frac{\partial h}{\partial \theta}$ is of the form

$$\frac{\partial h}{\partial \theta} = \begin{bmatrix} \ldots \mathrm{vec}(\frac{\partial}{\partial a_{c,k_\ell}} H_{A_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial b_{c,k_\ell}} H_{A_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial c_{c,k_\ell}} H_{A_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial d_{c,k_\ell}} H_{A_c}) \\ \ldots \mathrm{vec}(\frac{\partial}{\partial a_{c,k_\ell}} H_{B_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial b_{c,k_\ell}} H_{B_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial c_{c,k_\ell}} H_{B_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial d_{c,k_\ell}} H_{B_c}) \\ \ldots \mathrm{vec}(\frac{\partial}{\partial a_{c,k_\ell}} H_{C_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial b_{c,k_\ell}} H_{C_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial c_{c,k_\ell}} H_{C_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial d_{c,k_\ell}} H_{C_c}) \\ \ldots \mathrm{vec}(\frac{\partial}{\partial a_{c,k_\ell}} H_{D_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial b_{c,k_\ell}} H_{D_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial c_{c,k_\ell}} H_{D_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial d_{c,k_\ell}} H_{D_c}) \end{bmatrix}. \tag{4.22}$$

and $\frac{\partial h}{\partial \lambda}$ can be expressed as

$$\frac{\partial h}{\partial \lambda} = \begin{bmatrix} \mathrm{vec}(\frac{\partial}{\partial \lambda} H_{A_c}) \\ \mathrm{vec}(\frac{\partial}{\partial \lambda} H_{B_c}) \\ \mathrm{vec}(\frac{\partial}{\partial \lambda} H_{C_c}) \\ \mathrm{vec}(\frac{\partial}{\partial \lambda} H_{D_c}) \end{bmatrix}. \tag{4.23}$$

Below, we develop explicit expression for the derivative terms appearing on the right hand sides of (4.22) and (4.23). We use the notation

$$M^{(j)} \triangleq \frac{\partial M}{\partial \theta_j} \tag{4.24}$$

$$\dot{M} \triangleq \frac{\partial M}{\partial \lambda}. \tag{4.25}$$

Differentiating (4.4)–(4.6) with respect to θ_j yields

$$\tilde{P}^{(j)} = \tilde{A}^{\mathrm{T}}\tilde{P}^{(j)}\tilde{A} + ({\tilde{A}^{(j)}}^{\mathrm{T}}\tilde{P}\tilde{A} + \tilde{A}^{\mathrm{T}}\tilde{P}\tilde{A}^{(j)} + \tilde{R}^{(j)}) \tag{4.26}$$

$$\tilde{Q}^{(j)} = \tilde{A}\tilde{Q}^{(j)}\tilde{A}^{\mathrm{T}} + (\tilde{A}^{(j)}\tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{A}\tilde{Q}{\tilde{A}^{(j)}}^{\mathrm{T}} + \tilde{V}^{(j)}) \tag{4.27}$$

$$\tilde{Z}^{(j)} = \tilde{Q}^{(j)}\tilde{A}^{\mathrm{T}}\tilde{P} + \tilde{Q}{\tilde{A}^{(j)}}^{\mathrm{T}}\tilde{P} + \tilde{Q}\tilde{A}^{\mathrm{T}}\tilde{P}^{(j)} \tag{4.28}$$

where expressions for the derivatives $\tilde{A}^{(j)}, \tilde{R}^{(j)}$ and $\tilde{V}^{(j)}$ are given by (B.20)-(B.28) of Appendix B. Similarly, differentiating (4.4)–(4.6) with respect to λ yields

$$\dot{\tilde{P}} = \tilde{A}^{\mathrm{T}}\dot{\tilde{P}}\tilde{A} + (\dot{\tilde{A}}^{\mathrm{T}}\tilde{P}\tilde{A} + \tilde{A}^{\mathrm{T}}\tilde{P}\dot{\tilde{A}} + \dot{\tilde{R}}) \tag{4.29}$$

$$\dot{\tilde{Q}} = \tilde{A}\dot{\tilde{Q}}\tilde{A}^{\mathrm{T}} + (\dot{\tilde{A}}\tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{A}\tilde{Q}\dot{\tilde{A}} + \dot{\tilde{V}}) \tag{4.30}$$

$$\dot{\tilde{Z}} = \dot{\tilde{Q}}\tilde{A}^{\mathrm{T}}\tilde{P} + \tilde{Q}\dot{\tilde{A}}^{\mathrm{T}}\tilde{P} + \tilde{Q}\tilde{A}\dot{\tilde{P}} \tag{4.31}$$

where expressions for $\dot{\tilde{A}}, \dot{\tilde{R}}$ and $\dot{\tilde{V}}$ are given by (B.29)–(B.33) of Appendix B.

Before presenting the desired derivative expressions we define

$$H'_{A_c}(\tilde{Z}^{(j)}) = 2{\tilde{Z}_{22}^{(j)}}^{\mathrm{T}} \tag{4.32}$$

$$\begin{aligned} H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}) = 2(&{\tilde{P}_{12}^{(j)}}^{\mathrm{T}}V_{12} - {\tilde{P}_{12}^{(j)}}^{\mathrm{T}}BD_cV_2 + \tilde{P}_{22}^{(j)}B_cV_2 \\ &+ {\tilde{Z}_{12}^{(j)}}^{\mathrm{T}}C^{\mathrm{T}} - {\tilde{Z}_{22}^{(j)}}^{\mathrm{T}}C_c^{\mathrm{T}}D^{\mathrm{T}}) \end{aligned} \tag{4.33}$$

$$\begin{aligned} H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}) = 2(&-R_{12}^{\mathrm{T}}\tilde{Q}_{12}^{(j)} + R_2D_cC\tilde{Q}_{12}^{(j)} + R_2C_c\tilde{Q}_{22}^{(j)} \\ &- B^{\mathrm{T}}{\tilde{Z}_{21}^{(j)}}^{\mathrm{T}} - D^{\mathrm{T}}B_c^{\mathrm{T}}{\tilde{Z}_{22}^{(j)}}^{\mathrm{T}}) \end{aligned} \tag{4.34}$$

$$\begin{aligned} H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) = 2(&-R_{12}^{\mathrm{T}}\tilde{Q}_{11}^{(j)}C^{\mathrm{T}} + R_2D_cC\tilde{Q}_{11}^{(j)}C^{\mathrm{T}} + R_2C_c{\tilde{Q}_{12}^{(j)}}^{\mathrm{T}}C^{\mathrm{T}}) \\ &- B^{\mathrm{T}}{\tilde{Z}_{11}^{(j)}}^{\mathrm{T}}C^{\mathrm{T}}) \end{aligned} \tag{4.35}$$

Notice that the right hand sides of (4.32)–(4.35) are essentially identical in form to the right hand sides of (4.16)–(4.19). The difference is that $\tilde{P}, \tilde{Q}$, and $\tilde{Z}$ have been replaced by $\tilde{P}^{(j)}, \tilde{Q}^{(j)}$ and $\tilde{Z}^{(j)}$ and the last term ($2R_2D_cV_2$) in (4.19) has no counterpart in (4.35).

Derivatives with Respect to $b_{c,k\ell}$

Differentiating (4.16)–(4.19) with respect to $b_{c,k\ell}(=\theta_j)$ gives the following.

$$\frac{\partial H_{A_c}}{\partial b_{c,k\ell}} = H'_{A_c}(\tilde{Z}^{(j)}) \tag{4.36}$$

$$\frac{\partial H_{B_c}}{\partial b_{c,k\ell}} = H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}) + 2\tilde{P}_{22}E_{n_c \times n_y}^{(k,\ell)}V_2 \tag{4.37}$$

$$\frac{\partial H_{C_c}}{\partial B_{c,k\ell}} = H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}) - 2D^{\mathrm{T}} E^{(\ell,k)}_{n_y \times n_c} \tilde{Z}^{\mathrm{T}}_{22} \tag{4.38}$$

$$\frac{\partial H_{D_c}}{\partial B_{c,k\ell}} = H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) - 2B^{\mathrm{T}} \tilde{P}_{12} E^{(k,\ell)}_{n_c \times n_y} V_2. \tag{4.39}$$

Derivatives with Respect to $c_{c,k\ell}$

Differentiating (4.7)–(4.9) with respect to $\theta_j (= c_{c,k\ell})$ gives the following.

$$\frac{\partial H_{A_c}}{\partial c_{c,k\ell}} = H'_{A_c}(\tilde{Z}^{(j)}) \tag{4.40}$$

$$\frac{\partial H_{B_c}}{\partial c_{c,k\ell}} = H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}) - 2\tilde{Z}^{\mathrm{T}}_{22} E^{(\ell,k)}_{n_c \times n_u} D^{\mathrm{T}} - 2(\Omega^{(j)} Z \Omega + \Omega Z \Omega^{(j)}) B_c \tag{4.41}$$

$$\frac{\partial H_{C_c}}{\partial c_{c,k\ell}} = H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}) + 2R_2 E^{(k,\ell)}_{n_u \times n_c} \tilde{Q}_{22} \tag{4.42}$$

$$\frac{\partial H_{D_c}}{\partial c_{c,k\ell}} = H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) + 2R_2 E^{(k,\ell)}_{n_u \times n_c} \tilde{Q}^{\mathrm{T}}_{12} C^{\mathrm{T}}. \tag{4.43}$$

Derivatives with respect to $d_{c,k\ell}$

Differentiating (4.7)–(4.9) with respect to $d_{c,k\ell}$ gives the following.

$$\frac{\partial H_{A_c}}{\partial d_{c,k\ell}} = H'_{A_c}(\tilde{Z}^{(j)}) \tag{4.44}$$

$$\frac{\partial H_{B_c}}{\partial d_{c,k\ell}} = H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}) - 2\tilde{P}^{\mathrm{T}}_{12} B E^{(k,\ell)}_{n_u \times n_y} V_2 \tag{4.45}$$

$$\frac{\partial H_{C_c}}{\partial d_{c,k\ell}} = H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}) + 2R_2 E^{(k,\ell)}_{n_u \times n_y} C \tilde{Q}_{12} \tag{4.46}$$

$$\begin{aligned}\frac{\partial H_{D_c}}{\partial d_{c,k\ell}} = {} & H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) + 2(R_2 E^{(k,\ell)}_{n_u \times n_y} C \tilde{Q}_{11} C^{\mathrm{T}} + B^{\mathrm{T}} \tilde{P}_{11} B E^{(k,\ell)}_{n_u \times n_y} V_2 \\ & + R_2 E^{(k,\ell)}_{n_c \times n_y} V_2).\end{aligned} \tag{4.47}$$

Derivatives with Respect to λ

Differentiating (4.7)–(4.9) with respect to λ gives

$$\frac{\partial H_{A_c}}{\partial \lambda} = H'_{A_c}(\dot{\tilde{Z}}) \tag{4.48}$$

$$\begin{aligned}\frac{\partial H_{B_c}}{\partial \lambda} = {} & H'_{B_c}(\dot{\tilde{P}}, \dot{\tilde{Z}}) \\ & + 2(\tilde{P}^{\mathrm{T}}_{21} \dot{V}_{12} - \tilde{P}^{\mathrm{T}}_{12} \dot{B} D_c V_2 - \tilde{P}^{\mathrm{T}}_{12} B D_c \dot{V}_2 + \tilde{P}_{22} B_c \dot{V}_2) \\ & + \tilde{Z}^{\mathrm{T}}_{12} \dot{C}^{\mathrm{T}} - \tilde{Z}^{\mathrm{T}}_{22} C^{\mathrm{T}}_c \dot{D}^{\mathrm{T}})\end{aligned} \tag{4.49}$$

$$\frac{\partial H_{C_c}}{\partial \lambda} = H'_{C_c}(\dot{\tilde{Q}}, \dot{\tilde{Z}})$$

$$
\begin{aligned}
&+ 2(-\dot{R}_{12}^{\mathrm{T}}\tilde{Q}_{12} + \dot{R}_2 D_c C\tilde{Q}_{12} + R_2 D_c \dot{C}\tilde{Q}_{12} + \dot{R}_2 C_c \tilde{Q}_{22} \\
&+ \dot{B}^{\mathrm{T}}\tilde{Z}_{21}^{\mathrm{T}} - \dot{D}^{\mathrm{T}} B_c^{\mathrm{T}}\tilde{Z}_{22}^{\mathrm{T}}) \qquad (4.50)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial H_{D_c}}{\partial\lambda} = {} & H'_{D_c}(\dot{\tilde{P}},\dot{\tilde{Q}},\dot{\tilde{Z}}) \\
&+ 2(-\dot{R}_{12}^{\mathrm{T}}\tilde{Q}_{11}C^{\mathrm{T}} - R_{12}^{\mathrm{T}}\tilde{Q}_{11}\dot{C}^{\mathrm{T}} \\
&\qquad + \dot{R}_2 D_c C\tilde{Q}_{11}C^{\mathrm{T}} + R_2 D_c\dot{C}\tilde{Q}_{11}C^{\mathrm{T}} + R_2 D_c C\tilde{Q}_{11}\dot{C}^{\mathrm{T}} \\
&\qquad\qquad + \dot{R}_2 C_c C\tilde{Q}_{12}^{\mathrm{T}}C^{\mathrm{T}} + R_2 C_c\tilde{Q}_{12}^{\mathrm{T}}\dot{C}^{\mathrm{T}}) \\
&+ 2(-\dot{B}^{\mathrm{T}}\tilde{P}_{11}V_{12} - B^{\mathrm{T}}\tilde{P}_{11}\dot{V}_{12} \\
&\qquad + \dot{B}^{\mathrm{T}}\tilde{P}_{11}BD_cV_2 + B^{\mathrm{T}}\tilde{P}_{11}\dot{B}D_cV_2 + B^{\mathrm{T}}\tilde{P}_{11}BD_c\dot{V}_2 \\
&\qquad\qquad - \dot{B}^{\mathrm{T}}\tilde{P}_{12}B_cV_2 - B^{\mathrm{T}}\tilde{P}_{12}B_c\dot{V}_2) \\
&- 2\dot{B}^{\mathrm{T}}\tilde{Z}_{11}^{\mathrm{T}}C^{\mathrm{T}} - B^{\mathrm{T}}\tilde{Z}_{11}\dot{C}^{\mathrm{T}} \\
&+ 2(\dot{R}_2 D_c V_2 + R_2 D_c\dot{V}_2) \qquad (4.51)
\end{aligned}
$$

where from (4.12)–(4.14)

$$
\begin{bmatrix} \dot{A} & \dot{B} \\ \dot{C} & \dot{D} \end{bmatrix} = \begin{bmatrix} A_f - A_0 & B_f - B_0 \\ C_f - C_0 & C_f - C_0 \end{bmatrix} \qquad (4.52)
$$

$$
\begin{bmatrix} \dot{R}_1 & \dot{R}_{12} \\ \dot{R}_{12}^{\mathrm{T}} & \dot{R}_2 \end{bmatrix} = \dot{L}_R\dot{L}_R^{\mathrm{T}} \qquad (4.53a)
$$

where

$$
\dot{L}_R = L_{R,f} - L_{R,0} \qquad (4.53b)
$$

$$
\begin{bmatrix} \dot{V}_1 & \dot{V}_{12} \\ \dot{V}_{12}^{\mathrm{T}} & \dot{V}_2 \end{bmatrix} = \dot{L}_V\dot{L}_V^{\mathrm{T}} \qquad (4.54a)
$$

where

$$
L_V = L_{V,f} - L_{V,0}. \qquad (4.54b)
$$

The homotopy Jacobian can now be computed using (4.20) with (4.22) and (4.23) and (4.32)–(4.51). Note that the primary computations involve the computation of the solutions of the Lyapunov equations (4.26), (4.27), (4.29) and (4.30). Significant computational savings can be made by solving these Lyapunov equations in a basis in which the $\tilde{A}$ matrix is diagonal (or nearly diagonal). This requires transforming the corresponding forcing terms into this basis. But it is seen by (B.20)–(B.33) of Appendix B that these forcing terms are very sparse. Hence this transformation does not have to be expensive. In addition, it is required that the solutions of the Lyapunov equations be transformed into their original basis before substituting into the expressions (4.32)–(4.51). A close examination of these expressions shows that for problems in which $n_y << n_x, n_u << n_x$

and/or $n_c << n_x$ significant computational savings can be made by not actually performing the matrix multiplies to transform the solutions into their original basis until after substituting the transformations into (4.32)–(4.51). Appendix H gives the details of efficient computation of H_θ for the corresponding continuous-time problem. A nearly identical procedure has been implemented for the discrete-time problem considered in this report.

5. Reduction of the Dimension of the Controller Parameter Vector (θ)

The homotopy function $H(\theta, \lambda)$, described earlier, was defined to solve the H_2 optimal reduced-order dynamic compensation problem for discrete-time systems. The vector θ was defined such that it contained each of the elements of the controller matrices, A_c, B_c and C_c. However, for computational efficiency it is desired that θ be as small as possible. Hence, we desire to represent the controller matrix with the fewest parameters possible (i.e., we desire θ to have the smallest dimension possible). The minimal number of parameters $p_{\min}$ with which a compensator can be represented is given by [1,2]

$$p_{\min} = n_c(n_u + n_y) \tag{5.1}$$

One canonical form which allows representation of a controller with a minimal number of parameters is the modal form described in [3]. This form will be called here the Second-Order Polynomial (SP) form. For this parameterization a triple (A_c, B_c, C_c) has the following structure.

$$A_c = \text{block- diag}\{A_{c,1}, A_{c,2} \dots, A_{c,r}\} \tag{5.2}$$

where $A_{c,i} \in \mathbb{R}^{2 \times 2}$ for $i \in \{1, 2, \dots, r\}$ and each $A_{c,i}$ (with the exception of $A_{c,r}$ if the row dimension of A_c is odd) has the form

$$A_{c,i} = \begin{bmatrix} 0 & 1 \\ a_{c,i}^{(1)} & a_{c,i}^{(2)} \end{bmatrix} \tag{5.3}$$

to allow for either a complex conjugate set of poles or two real poles. B_c is completely full and

$$C_c = [C_{c,1},\ C_{c,2}, \dots, C_{c,r}] \tag{5.4}$$

where $C_{c,i}$ has the form

$$C_{c,r} = \begin{bmatrix} 1 & 0 \\ * & * \\ \vdots & \vdots \\ * & * \end{bmatrix}. \tag{5.5}$$

The controller canonical form described in [4,5] also allows representation of a controller with a minimal number of parameters. For single-input, single-output (SISO) systems in controller canonical form the A_c matrix is a companion matrix. In particular, A_c has the form

$$A_c = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & & 0 \\ 0 & 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ * & * & * & * & \cdots & * \end{bmatrix}. \tag{5.6}$$

In addition,

$$B_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{5.7}$$

and C_c is completely full. A dual form of the controller canonical form is the observable canonical form [5].

It is also possible to represent the controller in a basis where the number of free parameters p satisfies

$$p_{\min} < p < p_{\max} \triangleq n_c(n_c + n_u + n_y). \tag{5.8}$$

One such basis is the tridiagonal basis [7–11] in which the controller state matrix is constrained to have nonzero elements only on the diagonal, the super-diagonal and the sub-diagonal. That is,

$$A_c = \begin{bmatrix} * & * & & & \\ * & * & * & & 0 \\ & * & * & & \ddots \\ & & \ddots & \ddots & \\ & & & * & * \end{bmatrix} \tag{5.9}$$

B_c and C_c are completely full. For this form the number of free parameters is given by

$$p = p_{\min} + (3n_c - 2)$$

A common feature of each of the above bases is that they are described by simply constraining certain elements of the controller (or plant) matrices to constant values (e.g., 1 or 0) while allowing the remaining parameters to have arbitrary values (A_c, B_c, C_c). Hence, the corresponding parameter vector (θ_p), gradient vector $(J_{\theta,p})$ and Hessian matrix $(H_{\theta,p})$ are given by

$$\theta_p = \Gamma\theta \tag{5.10}$$

$$J_{\theta,p} = \Gamma J_\theta \tag{5.11}$$

$$H_{\theta,p} = \Gamma H_\theta \Gamma^{\mathrm{T}} \tag{5.12}$$

where Γ is an elemental matrix (i.e., each row has only one nonzero element and this element has unity value). It should be noted here that $H_{\theta,p}$ can be computed more efficiently than shown in (4.64). Since it is not necessary to construct the large Hessian H_θ to compute the smaller Hessian $H_{\theta,p}$.

References

1. D. G. Denery, "An Identification Algorithm That Is Insensitive to Initial Parameter Estimates," *AIAA Journal*, Vol. 9, pp. 371-377, October 1971.

2. G. D. Martin and A. E. Bryson, Jr., "Attitude Control of a Flexible Spacecraft," *J. Guid. Contr.*, Vol. 3, pp. 37-41, Jan. 1980.

3. 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.

4. F. S. Kramer and A. J. Calise, "Fixed-Order Dynamic Compensation for Multivariable Linear Systems," *J. Guid. Contr. Dyn.*, Vol. 11, pp. 80–85, Jan. 1988.

5. A. J. Calise and J. V. R. Prasad, "Approximate Loop Transfer Recovery Method for Designing Fixed-Order Compensators," *J. Guid. Contr. Dyn.*, Vol. 13, pp. 297-302, October 1990.

6. L. D. Davis, E. G. Collins, Jr., and A. S. Hodel, "A Parameterization of Minimal Plants," *Proc. Amer. Contr. Conf.*, 1992, to appear.

7. C. Strachey and J. G. F. Francis, "The Reduction of a Matrix to Codiagonal Form by Elimination," *Comput. J.*, Vol. 4, p. 168, 1961.

8. C. D. LaBudde, "The Reduction of an Arbitrary Real Square Matrix to Tridiagonal Form Using Similarity Transformations," *Math. Comp.*, Vol. 17, pp. 443-447, 1963.

9. B. N. Parlett, "A Note on LaBudde's Algorithm," *Math. Comp.*, Vol. 19, pp. 505-506, 1964.

10. G. A. Geist, "Reduction of A General Matrix to Tridiagonal Form," *SIAM J. Matrix Anal. Appl.*, Vol. 12, pp. 362-373, April 1992.

11. B. N. Parlett, "Reduction to Tridiagonal Form and Minimal Realizations," *SIAM J. Matrix Anal. Appl.*, April 1992, to appear.

12. D. W. Decker and C. T. Kelley, "Newton's Method at Singular Points II," *SIAM J. Numer. Anal.*, Vol. 17, pp. 465-471, June 1980.

13. A. Griewank and M. R. Osborne, "Newton's Method for Singular Problems When the Dimension of the Null Space is ¿1," *SIAM J. Numer. Anal.*, Vol. 18, pp. 145-149, Feb. 1981.

14. D. W. Decker, H. B. Keller and C. T. Kelley, "Convergence Rates for Newton's Method at Singular Points," *SIAM J. Numer. Anal.*, Vol. 20, pp. 296-314, April 1983.

15. A. Griewank and M. R. Osborne, "Analysis of Newton's Method at Irregular Singularities," *SIAM J. Numer. Anal.*, Vol. 20, pp. 747–773, Aug. 1983.

16. D. W. Decker and C. T. Kelley, "Sublinear Convergence of the Chord Method at Singular Points," *Numer. Math.*, Vol. 42, pp. 147–154, 1983.

17. C. T. Kelley and R. Suresh, "A New Acceleration Method for Newton's Method at Singular Points," *SIAM J. Numer. Anal.*, Vol. 20, pp. 1001-1009, Oct. 1983.

18. D. W. Decker and C. T. Kelley, "Broyden's Method for a Class of Problems Having Singular Jacobian at the Root," *SIAM J. Numer. Anal.*, Vol. 22, pp. 566-574, Oct. 1985.

19. D. W. Decker and C. T. Kelley, "Expanded Convergence Domains for Newton's Method at Nearly Singular Roots," *SIAM J. Sci. Stat. Comput.*, Vol. 6, pp. 951-966, Oct. 1985.

20. A. Griewank, "On Solving Nonlinear Equations with Simple Singularities or Nearly Singular Solutions," *SIAM Review*, Vol. 27, pp. 537-563, Dec. 1985.

21. C. T. Kelley, "A Shamanskii-Like Acceleration Scheme for Nonlinear Equations at Singular Roots," *Math. Comput.*, Vol. 47, pp. 609-623, Oct. 1986.

6. Overview of the Homotopy Algorithm

This section describes the general logic and features of the homotopy algorithm for H_2 optimal reduced-order control. It is assumed that the designer has supplied a set of system matrices, $S_f = (A_f, B_f, C_f, D_f, R_{1,f}, R_{2,f}, V_{1,f}, V_{2,f}, V_{12,f})$ describing the optimization problem whose solution is desired. In addition, it is assumed that the designer has chosen an initial set of related system matrices $S_0 = (A_0, B_0, C_0, D_0, R_{1,0}, R_{2,0}, V_{1,0}, V_{2,0}, V_{12,0})$ that has an easily obtained optimal controller $(A_{c,0}, B_{c,0}, C_{c,0}, D_{c,0})$ of order n_c. The initial system S_0 can be chosen to correspond to a low-authority control problem as described in Appendix I since if R_{11}, or $V_{1,0}$ are of the appropriate structure and A_0 is stable, the corresponding LQG controller is nearly nonminimal and can hence be reduced to a nearly optimal n_c^{th}order compensator using, for example, balanced controller reduction [1]. The reader is referred to Appendix I for additional details.

Below, we present an outline of the homotopy algorithm. This algorithm describes a predictor/corrector numerical integration scheme. There are several options to be chosen initially. These options are enumerated before presenting the actual algorithm. Notice that each option corresponds to a particular flag being assigned some integer value.

Controller Basis Options:

$basis = 0$. No basis (i.e., all elements of the controller matrices are considered free.)

$basis = 1$. Tridiagonal Basis.

$basis = 2$. Second-Order Polynomial Form.

$basis = 3$. Controller Canonical Form.

Note that for $basis = 0$ or 1, $p > p_{\min}$ while for $basis = 2$ or 3, $p = p_{\min}$.

Prediction Scheme Options:

Here we use the notation that λ_0, λ_{-1}, and λ_1 represent the values of λ at respectively the current point on the homotopy curve, the previous point and the next point. Also, $\theta_p{}' = d\theta_p/d\lambda$ and is the solution of Davidenko's differential equation (4.7), rewritten here as

$$H_{\theta,p}\theta'_p(\lambda) + H_\lambda = 0. \tag{6.1}$$

If $p = p_{\min}$, $H_{\theta,p}$ is generally invertible, then $\theta'_p(\lambda)$ is given exactly by

$$\theta'_p(\lambda) = -H_{\theta,p}^{-1}H_\lambda. \tag{6.2}$$

If $p > p_{\min}$, then $H_{\theta,p}$ generally has rank $p_{\min}$ and $\theta_p'(\lambda)$ is approximated by the least squares solution of (6.2) or

$$\theta_p' = V \begin{bmatrix} \Sigma_0^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^{\mathrm{T}} \tag{6.3}$$

where it is assumed the $H_{\theta,p}$ has the singular value decomposition

$$H_{\theta,p} = U \begin{bmatrix} \Sigma_0 & 0 \\ 0 & 0 \end{bmatrix} V^{\mathrm{T}}, \quad \Sigma_0 \in \mathrm{I\!R}^{p_{\min} \times p_{\min}}. \tag{6.4}$$

Note that for $p = p_{min}$ (6.3) and (6.4) are equivalent.

pred = 0. No prediction. This option assumes that $\theta_p(\lambda_1) = \theta_p(\lambda_0)$.

pred = 1. Linear prediction. This option assumes predicts $\theta_p(\lambda_1)$ using only $\theta_p(\lambda_0)$ and $\theta_p'(\lambda_0)$. In particular,

$$\theta_p(\lambda_1) = \theta_p(\lambda_0) + (\lambda_1 - \lambda_0)\theta_p'(\lambda_0) \tag{6.5}$$

pred = 2. Cubic spline prediction. This option predicts $\theta_p(\lambda_1)$ using $\theta_p(\lambda_0)$, $\theta_p'(\lambda_0)$, $\theta_p(\lambda_{-1})$ and $\theta_p'(\lambda_{-1})$. In particular,

$$\theta_p(\lambda_1) = a_0 + a_1\lambda_1 + a_2{\lambda_1}^2 + a_3{\lambda_1}^3 \tag{6.6}$$

where a_0, a_1, a_2 and a_3 are computed by solving

$$[a_0 \;\; a_1 \;\; a_2 \;\; a_3] \begin{bmatrix} 1 & 0 & 1 & 0 \\ \lambda_{-1} & 1 & \lambda_0 & 1 \\ \lambda_{-1}^2 & 2\lambda_{-1} & \lambda_0^2 & 2\lambda_0 \\ \lambda_{-1}^3 & 3\lambda_{-1}^2 & \lambda_0^3 & 3\lambda_0^2 \end{bmatrix} = \begin{bmatrix} \theta_p(\lambda_{-1}) \\ \theta_p'(\lambda_{-1}) \\ \theta_p(\lambda_0) \\ \theta_p'(\lambda_0) \end{bmatrix} \tag{6.7}$$

Note that if this option is chosen, then at the initial algorithm prediction step $\theta_p(\lambda_{-1})$ and $\theta_p'(\lambda_{-1})$ are not available, in which case linear prediction is used.

Correction Options:

Here we assume that the homotopy parameter has a fixed value λ_0. The vector θ_p represents the current approximation of the parameter vector at $\lambda = \lambda_0$. Each of the options corresponds to updating θ_p using the formula

$$\theta_p \leftarrow \theta_p + \Delta\theta_p \tag{6.8}$$

where

$$\Delta\theta_p = -G_{\theta,p}J_{\theta,p} \tag{6.9}$$

for some choice of $G_{\theta,p}$.

corr = 1. Newton correction. In this option, if $p = p_{\min}$,

$$G_{\theta,p} = H_{\theta,p}{}^{-1} \tag{6.10}$$

while if $p > p_{\min}$,

$$G_{\theta,p} = V(\Sigma^2 + \alpha^2 I)^{-1}\Sigma U^{\mathrm{T}} \tag{6.11}$$

where α is some (small) scalar and (U, V, Σ) denote the singular value decomposition of $H_{\theta,p}$ such that

$$H_{\theta,p} = U\Sigma V^{\mathrm{T}}. \tag{6.12}$$

It can be shown that if $G_{\theta,p}$ is given by (6.11), then $\Delta\theta_p$ minimizes

$$\frac{1}{2}[||H_{\theta,p}\Delta\theta_p + \theta_p||^2 + \alpha^2||\Delta\theta_p||^2]. \tag{6.13}$$

Hence, $\Delta\theta_p$ is essentially a "Newton correction" that is relatively insensitive to singularities or near singularities in the Hessian, $H_{\theta,p}$.

corr = 2. Quasi-Newton correction. In this option, $G_{\theta,p}$ denotes the estimate of $H_{\theta,p}^{-1}$ using only gradient and cost information. For the algorithm presented here the BFGS inverse Hessian update is used [2].

Outline of the Homotopy Algorithm

Step 1. If *basis* ≥ 1, then transform the initial controller $(A_{c,0}, B_{c,0}, C_{c,0})$ to the chosen basis and let θ_0, p be the corresponding vector of free parameters.

Step 2. Initialize *loop* = 0, $\lambda = 0$, $\Delta\lambda \in (0, 1]$, $S = S_0$, $\theta_p = \theta_{0,p}$ and compute the cost J and the cost gradient $J_{\theta,p}$ corresponding to S and the controller described by θ_p.

Step 3. Let loop = loop+1. If loop = 1, then go to Step 5. Else, continue.

Step 4. Advance the homotopy parameter and predict the corresponding parameter vector θ as follows.

4a. Let $\lambda_0 = \lambda$

4b. Let $\lambda = \lambda_0 + \Delta\lambda$.

4c. If *pred* ≥ 1, then compute $\theta_p'(\lambda_0)$.

4d. Predict $\theta_p(\lambda)$ by using the option defined by *pred*.

4e. If the normalized gradient $J_{\theta,p}\|G_{\theta,p}\|/\Delta\theta_p$ satisfies some preassigned tolerance, then continue. Else, reduce $\Delta\lambda$ and go to Step 4b.

Step 5. Correct the current approximation θ_p to the optimization problem defined by S using the option defined by *corr* until the normalized gradient,

$$\frac{J_{\theta,p}\|G_{\theta,p}\|}{\Delta\theta_p} \tag{6.14}$$

satisfies some preassigned tolerance.

Step 6. If $\lambda = 1$, then stop. Else, go to Step 3.

The above algorithm assumes monotonicity of the solution curve as a function of the homotopy parameter λ. However, it is not difficult to modify the algorithm so that the variable parameter is the arc length as discussed in [3,4] since this modification would still only require the computation of H_θ and H_λ. The modified algorithm would not require monotonicity of the solution curve. However, so far in our computational experience the solution curve has always been monotonic.

Note that if $p = p_{\min}$ and $corr = 1$, then the corrections of Step 5 correspond to Newton corrections. Hence if the prediction tolerance used in Step 4 is sufficiently small, then, entering Step 5, θ_p will be close enough to the optimal value θ_p^* so that the quadratic convergence properties of Newton's method [2] can be realized. In practice, this quadratic convergence property is not always realized due to numerical ill-conditioning associated with the minimal parameterization of the controller. This ill-conditioning is illustrated and discussed further (in the context of continuous-time systems) in Appendix H.

References

1. Yousuff, A. and Skelton, R.E., 1984, A note on balanced controller reduction. *IEEE Transactions on Automatic Control*, **29**, 254-257.

2. Fletcher, R., 1987, *Practical Methods of Optimization: Second Edition*, John Wiley and Sons, New York.

3. Watson, L. T., 1986, Numerical linear algebra aspects of globally convergent homotopy methods. *SIAM Review,* **28**, 529–545.

4. Watson, L. T., 1987, ALGORITHM 652 HOMPACK: A suite of codes for globally convergent homotopy algorithms. *ACM Transactions on Mathematical Software*, **13**, 281–310.

7. A Design Example and Some Rules of Thumb

This section illustrates the design of a reduced order compensator for an axial beam with four disks attached as shown in Figure 7.1. This example has been considered in several previous publications [1-7] and has become a standard benchmark example. The section closes with some general rules of thumb that will aid the control designer in most efficiently utilizing this algorithm.

The basic control objective for the four disk problem is to control the angular displacement at the location of disk 1 using a torque input at the location of disk 3. It is also assumed that a torque disturbance enters the system at the location of disk 3. An 8th order discretized model of the fourdisk plant with nominal performance weights and disturbance covariances is generated by the function **diskmod**.

The design philosophy adopted here is that the scaling **q2** of the nominal control weight R_2 and the nominal sensor noise covariance V_2 are simply design knobs used to determine the controller authority. The system costs are computed assuming that $V_2=0$. This general philosophy is actually motivated by insights into LQG theory. However, it will suffice here to simply note that this philosophy was used successfully on two hardware experiments involving control design and implementation [10-13].

It is now assumed that we are in the MATLAB environment. In what follows the reader is walked through the design process for a 4th order controller. The command sequences are presented after the prompt "≫" and after the commands some of the resultant output is displayed. Explanatory text is interspersed to clarify the motivation of the command sequences and the interpretation of some of the output.

We begin by using **diskmod** to generate the design plant and nominal weights and covariance.

```
≫ diskmod
discretizing a, b, and v1
```

The following variables are now loaded into memory.

```
≫ who

Your variables are:
     a           c           r1          r2          v1          v2
     b           d           r12         ts          v12
```

We choose to display numerical data using the following format.

```
≫ format short e
```

We now begin the search for an authority level that will give us a nearly optimal controller by balanced controller reduction. We commence this process by choosing the initial scaling **q20** of $\mathbf{R}_2$ and $\mathbf{V}_2$ as follows.

```
≫ q20 = .1;
```

We use **dlqg** to design an LQG controller and then check the eigenvalues of the product **phat*qhat** to see if their is any gap between the 4th and 5th eigenvalues (ordered in descending order of magnitude). Note that the warning after the call to **dlqg** in this case is not important.

```
≫[ac,bc,cc,dc,costslqg,phat,qhat] = ...
            dlqd(a,b,c,d,r1,q20*r2,r12,v1,q20*v2,v12,1);
Warning:  Q is not symmetric and positive semi-definite
≫ -sort(-eig(phat*qhat))
ans =
   1.3554e+01
   1.2377e+00
   8.0067e-01
   1.3682e-01
   1.0451e-01
   1.7751e-02
   1.0585e-02
   4.9872e-03
```

Note that their is no gap between the 4th and 5th eigenvalues indicating that balanced controller reduction will probably not yield a nearly optimal reduced-order controller. However to verify this we will actually compute a 4th order controller using balanced controller reduction and compare it's cost with that of the LQG compensator, which is contained in the vector **costslqg**.

```
≫[ac,bc,cc,dc] = balcred(ac,bc,cc,dc,phat,qhat,4);
≫costs=dqcosts(ac,bc,cc,dc,a,b,c,d,r1,q20*r2,r12,
                    v1,q20*v2,v12);

   ≫ costslqg

   costslqg =
          2.1047e-02    1.6027e-02             0        4.2135e-04    3.7495e-02

   ≫ costs

   costs =
          2.5838e-02    2.0623e-02             0        4.2135e-04    4.6882e-02
```

The total cost of the LQG compensator is 3.7495e-02 while the cost of the reduced order controller is 4.6882e-02. The vast differences in these costs is another indication that the reduced order controller is not nearly optimal. We will now repeat the above process with a higher value of **q20**, i.e. the LQG controller is of lower authority.

```
   ≫ q20 = 10;
   ≫[ac,bc,cc,dc,costslqg,phat,qhat] = ...
          dlqg(a,b,c,d,r1,q20*r2,r12,v1,q20*v2,v12,1);
   Warning:  Q is not symmetric and positive semi-definite
   ≫ -sort(-eig(phat*qhat))
   ans =
      4.1835e+02
      2.1594e+00
      5.6033e-01
      4.3915e-01
      4.6232e-02
      3.7616e-02
      4.4658e-03
      4.4134e-03
   ≫ [ac,bc,cc,dc] = balcred(ac,bc,cc,dc,phat,qhat,4);
   ≫costs=dqcosts(ac,bc,cc,dc,a,b,c,d,r1,q20*r2,r12,v1,
                 q20*v2,v12);
   ≫ costslqg

   costslqg =
          2.0046e-01    4.1697e-01             0        2.7474e-03    6.2018e-01
   ≫ costs
   costs =
            2.0137e-01    4.1748e-01         0        2.7474e-03    6.2160e-01
```

This time there is an order of magnitude gap between the 4th and 5th eigenvalues of **phat*qhat** and the costs of the LQG and reduced-order controllers are nearly identical. This indicates that the 4th order balanced controller is nearly optimal. This deduction could also be made by generating a performance curve for the LQG controller (by varying **q20**) and superimposing it with the performance curve for the corresponding 4th order balanced controllers as shown in Figure 7.2. If for a given **q20** the two controllers have essentially the same state and actuation costs then the 4th order balanced controller is probably nearly optimal.

If the 4th order balanced controller is nearly optimal then by using a few Newton corrections (say, 1 to 6) we should be able to converge to the optimal controller (practically, the controller that satisfies a small tolerance on the normalized norm of the cost gradient). This is verified below. Function **nwtpar** is used to intialize the algorithm parameters to their default values while **nwtprint** is used to display these default parameters.

```
≫ par = nwtpar;


≫ nwtprint(par);
 1. Will print intermediate results.
 2. gradient prediction tolerance = 1.00000e-05
 3. gradient correction tolerance = 2.00000e-08
 4. gradient final tolerance = 2.00000e-08
 5. minimum homotopy step size = 1.00000e-06
 6. maximum number of corrections allowed = 10.000000
 7. Will use Hessian from last correction for prediction.
 8. Will not use line search.
 9. Will let program run automatically.
10. initial step size = 1.000000
11. No basis is assumed for the controller.
```

At this time the user has the option of changing any of the default parameters. However, we will be content with them. The default parameters will also be printed by **dnwthom**. In the following call to **dnwthom** the initial and final system parameters are identical so that the algorithm will only perform correction loops.

```
≫ [ac,bc,cc,dc,val,par] = dnwthom(ac,bc,cc,dc, ...
              a,b,c,d,r1,q20*r2,r12,v1,q20*v2,v12, ...
              a,b,c,d,r1,q20*r2,r12,v1,q20*v2,v12,par);
 1. Will print intermediate results.
 2. gradient prediction tolerance = 1.00000e-05
 3. gradient correction tolerance = 2.00000e-08
 4. gradient final tolerance = 2.00000e-08
 5. minimum homotopy step size = 1.00000e-06
 6. maximum number of corrections allowed = 10.000000
 7. Will use Hessian from last correction for prediction.
 8. Will not use line search.
 9. Will let program run automatically.
10. initial step size = 1.000000
11. No basis is assumed for the controller.
Computing Initial Hessian...
Inverting Hessian...
***** INITIAL PARAMETERS *****
                    normalized   normalized   normalized
          cost      cost0-cost    gradient    gradient      del-theta
--------------------------------------------------------------------------------
      6.21598e-01   0.00000e+00   7.35784e-07   5.54898e-010.00000e+00
```

The algorithm is still in process but we note here that the initial normalized gradient value of 7.36e-07 is fairly small. The general rule is that values ≤ to about 2.0e-08 are very small.

```
********** CORRECTING **********
**** lambda = 1.0000e+00 ****
---------- Correction Iteration 1 ----------
Computing Hessian...
Inverting Hessian...
** correcting:  i = 1.000000, lambda = 1.0000e+00 **
                    normalized   normalized                 normalized
          cost      cost0-cost    gradient    gradient      del-theta
--------------------------------------------------------------------------------
      6.21598e-01   0.00000e+00   7.35784e-07   5.54898e-01   0.00000e+00
      6.21515e-01   1.33389e-04   7.34790e-08   5.54259e-02   1.92813e-03
```

```
The normalized gradient correction tolerance is:  2.00000e-08
```

With the algorithm still in progress we note that the top line denotes the initial convergence parameters before the first correction while the bottom line denotes the value of the convergence parameters after the correction. It is seen that both the cost and gradient were improved (i.e., decreased by the first correction iteration). However the normalized gradient is still not below its maximum tolerance of 2.0e-8.

```
--------- Correction Iteration 2 ---------

Computing Hessian...

Inverting Hessian...


** correcting:  i = 2.000000, lambda = 1.0000e+00 **
                      normalized    normalized                    normalized
           cost       cost0-cost    gradient      gradientdel-theta
----------------------------------------------------------------------------
       6.21515e-01   1.33389e-04   7.34790e-08   5.54259e-02   1.92813e-03
       6.21515e-01   5.92847e-07   2.33195e-10   1.77478e-04   1.86134e-04

The normalized gradient correction tolerance is:  2.00000e-08

doubling step size to 0

**** Exiting DNWTHOM with lambda=1.  ****
```

The correction of the initial 4th order balanced controller converged in 2 iterations. This clearly indicates that the balanced controller was nearly optimal. The controller (ac,bc,cc,dc) is now the optimal 4th order controller for the scale factor **q20**.

We now set **q2**=1 and use the **dnwthom** to deform the initial controller into a higher authority controller. We show some of the beginning output of **dnwthom**.

```
≫ [ac,bc,cc,dc,val,par] = dnwthom(ac,bc,cc,dc, ...
          a,b,c,d,r1,q20*r2,r12,v1,q20*v2,v12, ...
          a,b,c,d,r1,q20*r2,r12,v1,q20*v2,v12,par);
          .
          .
          .

***** INITIAL PARAMETERS *****
                      normalized    normalized                    normalized
           cost       cost0-cost    gradient      gradient        del-theta
----------------------------------------------------------------------------
       6.21515e-01   0.00000e+00   2.58673e-10   1.73424e-04   0.00000e+00
```

```
********** PREDICTING **********

**** lambda = 0.0000e+00 ****

** dlambda = 1.0000e+00 **

Computing Pseudo-Inverse of Hessian!

number of sing. vals. retained = 9.000000

** predicting:  lambda = 1.00000e+00, dlambda = 1.00000e+00, **
                   normalized   normalized                normalized
           cost    cost0-cost   gradient      gradient    del-theta
-------------------------------------------------------------------------
       6.21515e-01  0.00000e+00  2.58673e-10  1.73424e-04  0.00000e+00
       1.72524e-01  2.60247e+00  8.79235e-05  5.89473e+01  1.95698e-02

The normalized gradient prediction tolerance is:  1.00000e-05

!!adjusting step size!!

dlambda = 5.0000e-01

              .

              .

              .

** predicting:  lambda = 1.25000e-01, dlambda = 1.25000e-01, **
                   normalized   normalized                normalized
           cost    cost0-cost   gradient      gradient    del-theta
-------------------------------------------------------------------------
       4.79497e-01  2.96182e-01  1.46945e-05  9.85177e+00  4.89245e-03
       5.48829e-01  1.32438e-01  4.47514e-06  3.00030e+00  2.44622e-03

The normalized gradient prediction tolerance is:  1.00000e-05

********** CORRECTING **********

**** lambda = 1.2500e-01 ****

--------- Correction Iteration 1 ---------

Computing Hessian...

Inverting Hessian...

condH(1) = 3.36501e+07 [Hessian condition number]

condH(2) = 9.61511e-01 [free parameter singularity]

condH(3) = 2.05027e+00 [dthetap ratio]

** correcting:  i = 1.000000, lambda = 1.2500e-01 **
                   normalized   normalized                normalized
           cost    cost0-cost   gradient      gradient    del-theta
-------------------------------------------------------------------------
       5.48829e-01  0.00000e+00  4.47514e-06  3.00030e+00  0.00000e+00
       5.48807e-01  4.01362e-05  1.10053e-07  5.74753e-02  1.54763e-04
```

```
.

.

.

**** Exiting DNWTHOM with lambda=1.  ****
```

We now use **valprint** to check the performance parameters recorded in the vector **val**.

```
≫valprint(val)

 1. final homotopy parameter value = 1.00000e+00

 2. total # of Hessian calculations = 33.000000

 3. minimum # of corrections for fixed lambda = 1.000000

 4. maximum # of corrections for fixed lambda = 3.000000

 5. minimum homotopy step size = 3.12500e-02

 6. maximum homotopy step size = 1.25000e-01
```

The costs plotted in Figure 7.2 are computed using **dqcosts** as follows. Note that in the input arguments **v2** is set to zero.

```
≫costs = dqcosts(ac,bc,cc,dc,a,b,c,d,r1,r2,r12,v1,0,v12);

≫costs

costs =
       5.5053e-02    7.8950e-02    0             0             1.3400e-01
```

The optimal controller is listed below.

```
≫ ac

ac =
       9.6632e-01    4.6790e-02    -1.1916e-02   -5.3926e-03
       -3.0758e-02   9.6053e-01    7.4261e-03    -4.1006e-04
       2.9294e-03    -8.6959e-03   9.9335e-01    8.7504e-02
       1.3349e-03    5.5798e-03    -8.8247e-02   9.8980e-01

≫ bc

bc =
       -2.9757e-02
       9.2577e-02
       -3.3036e-02
       -2.8086e-02

≫ cc

cc =
       -8.6600e-02   6.4435e-02    1.6983e-02    -2.9795e-02
```

```
≫ dc
dc =
        3.4606e-02
```

Some Rules of Thumb

1. Choose the initial weights $(R_{1,0}, R_{2,0}, R_{12,0}, V_{1,0}, V_{2,0}, V_{12,0})$ and the final weights $(R_{1,f}, R_{2,f}, R_{12,f}, V_{1,f}, V_{2,f}, V_{12,f})$ so that along the homotopy path the regulator and estimator poles have the same order magnitude. That is avoid situations where the estimator poles are very fast while the regulator poles are slow or vice versa. The algorithm will converge in these cases but the convergence tends to be slow.

2. Our experience indicates that no constraints on the controller basis appear to yield better numerical robustness than constraining the basis to tridiagonal form or some other basis. When attempting a 6th order controller for the four disk problem constraining the controller basis to tridiagonal form yielded very poor numerical robustness. However, when the controller basis was left unconstrained the algorithm performed very well. This phenomena is discussed further in Appendix H.

3. For better control don't take huge steps between the initial and final system parameters. For example don't try to go from very low control authority to very high control authority all at once. Take "reasonable size" increments. You may want to adjust the tolerances along the way to increase the algorithm efficiency.

References

1. R. H. Cannon and D. E. Rosenthal, "Experiments in Control of Flexible Structures with Non-colocated Sensors and Actuators," *AIAA J. Guid. Contr. Dyn.*, Vol. 7, pp. 546-553, 1984.

2. D. Enns, *Model Reduction for Control System Design*, Ph.D. Dissertation, Stanford Univ., Palo Alto, CA, 1984.

3. B. D. O. Anderson and Y. Liu, "Controller Reduction: Concepts and Approaches," Proc. Amer. Contr. Conf., Minnesota, pp. 1-9, June 1987.

4. S. W. Greeley and D. C. Hyland, "Reduced-Order Compensation: Linear Quadratic Reduction Versus Optimal Projection," *AIAA J. Guid. Contr. Dyn.*, Vol. 11, pp. 328-335, 1988.

5. C. DeVillemagne and R. E. Skelton, "Controller Reduction Using Canonical Interactions," *IEEE Trans. Autom. Contr.*, Vol. 33, pp. 740-750, 1988.

6. Y. Liu, B. D. O. Anderson and U-L Ly, "Coprime Factorization Controller Reduction with Bezout Identity Induced Frequency Weighting," Automatica, Vol. 26, pp. 233-249, 1990.

7. D. C. Hyland and S. Richter, "On Direct Versus Indirect Methods for Reduced-Order Controller Design," *IEEE Trans. Autom. Contr.*, Vol. 35, pp. 377-379, 1990.

8. E. G. Collins, Jr., D. J. Phillips and D. C. Hyland, "Design and Implementation of Robust Decentralized Control Laws for the ACES Structure at Marshall Space Flight Center," NASA Contractor Report 4310, Langley Research Center, October 1990.

9. E. G. Collins, Jr., D. J. Phillips and D. C. Hyland, "Robust Decentralized Control Laws for the ACES Structure," *Contr. Sys. Mag.*, pp. 62-70, April 1991.

10. E. G. Collins, Jr., J. A. King, D. J. Phillips and D. C. Hyland, "High Performance, Accelerometer-Based Control of the Minimast Structure at Langley Research Center," NASA Contractor Report 4377, May 1991.

11. E. G. Collins, Jr., J. A. King, D. J. Phillips and D. C. Hyland, "High Performance, Accelerometer-Based Control of the Minimast Structure," *AIAA J. Guid. Contr. Dyn.*, to appear.

8. HADOC Toolbox Reference

8.1 Commands Grouped by Function

Initialization Routines	
balcred	balanced controller reduction
dlqg	discrete linear quadratic gaussian design

Homotopy Algorithm	
dnwthom	discrete Newton homotopy algorithm
ntwpar	set default parameters for dnwthom
nwtprint	print parameters for dnwthom
valprint	print algorithm run-time statistics

Controller Bases	
ccf	convert to controllable canonical form (valid only for SISO controllers)
rnormal	convert to real normal modal form (a special case of tridiagonal form)
rpf	convert to real (or second-order) polynomial form

Costs	
dqcosts	discrete costs

Closed-Loop Matrices	
cla	construct state matrix
clr	construct performance weight
clv	construct noise intensity or covariance
dclp	construct discrete observability grammian
dclq	construct discrete controllability grammian
dclz	construct discrete Z matrix

Utility Routines	
beam	provide plant, noise statistics and perform weights for a simply-supported beam
bodeplot	plot magnitude and phase on same screen
c2dv	discretize disturbance intensity
dlyap2	discrete Lyapunov solution using diagonal basis
dstable	determine discrete stability
eigpq	ordered eigenvalues of product PQ
to180	converts phase vector to lie in interval [- 180,180]

8.2 Command Descriptions

Purpose:

Compute a reduced-order controller using balanced controller reduction.

Synopsis:

[ac,bc,cc] = balcred(ac0,bc0,cc0,pgram,qgram,nc)

Description:

Computes a controller (A_c, B_c, C_c, D_c) of order n_c given an initial controller of greater dimension $(A_{c,0}, B_{c,0}, C_{c,0}, D_{c,0})$ and the corresponding observability and controllability grammians $(P_{gram}$ and $Q_{gram})$.

See also:

dlqg

Purpose:

Compute a continuous-time or discrete-time representation (including noise statistics and performance weights) of a beam example.

Synopsis:

[a,b,c,d,r1,r2,r12,v1,v2,v12] = beam(nmodes,h)

Description:

Computes a continuous-time or discrete-time representation of the beam example described in the following reference:

D.S. Bernstein, L.D. Davis and D.C. Hyland, "The Optimal Projection Equations for Reduced-Order, Discrete-Time Modelling, Estimation and Control," *AIAA J. Guid. Contr. Dyn.*, Vol. 9, pp. 288-293, 1986.

Purpose:

Plot magnitude and phase information on same screen.

Synopsis:

bodeplot(fhz,mag,phase,titlename,axes)

Description:

Plots magnitude and phase on subplots that appear on the same screen. If **axes** is present, it is the 2×4 matrix of axis limits for the magnitude and phase (i.e., **axes** = [**axismag**; **axisphase**].

Purpose:

Transform a single-input, multi-output system to controllable canonical form.

Synopsis:

```
[a,b,c,T,Tinv] = ccf(a0,b0,c0)
```

Description:

Transforms a single-input, multi-output plant to controllable canonical form and also **returns** the transformation matrix and its inverse.

Purpose:

Construct closed-loop matrices.

Synopsis:

acl = cla(a,b,c,d,ac,bc,cc,dc)

rcl = clr(r1,r2,r12,c,cc,dc)

vcl = clv(v1,v2,v12,b,bc,dc)

Description:

Function **cla** computes the closed-loop state matrix using (2.7). Function **clr** computes the closed-loop performance weight using (2.9)-(2.10). Function **clv** computes the closed-loop disturbance intensity or covariance using (2.14)-(2.15).

See also:

dclp, dclq

Purpose:

Discretize a continuous-time disturbance intensity matrix.

Synopsis:

vd = c2dv(v,a,h)

Description:

Converts a continuous-time disturbance intensity V into an equivalent discrete-time covariance (assuming a zero-order hold with sample period h) using

$$V_d = \int_0^h exp(A_s)V\mathrm{exp}(A_s^{\mathrm{T}})\mathrm{ds}.$$

dclp, dclq, dclz dclp, dclq, dclz

Purpose:

Compute the discrete, closed-loop grammians and Z matrix.

Synopsis:

pcl = dclp(acl,rcl)

qcl = dclq(acl,vcl)

zcl = dclz(pcl,qcl,acl)

Description:

Function **dclp** returns the closed-loop discrete observability grammian satisfying

$$P_{c\ell} = A_{c\ell}^{\mathrm{T}} P_{cl} A_{cl} + R_{cl}.$$

Function **dclq** returns the closed-loop discrete controllability grammian satisfying

$$Q_{c\ell} = A_{c\ell} Q_{c\ell} A_{c\ell}^{\mathrm{T}} + V_{c\ell}.$$

Function **dclz** requires the outputs of **dclp** and **dclq** to return the closed-loop discrete Z matrix satisfying

$$Z_{c\ell} = Q_{c\ell} A_{c\ell}^{\mathrm{T}} P_{c\ell}.$$

See also:

cla, clr, clv

Purpose:

Solve the discrete-time Lyapunov equation by transforming to the modal basis.

Synopsis:

q = dlyap2(a,v)

Description:

Computes the solution Q to the discrete-time Lyapunov equation

$$Q = AQA^{\mathrm{T}} + V$$

by transforming to the complex modal basis. If the input A is a column vector, then the system is assumed to be in the diagonal basis and the eigenvalues are the elements of A.

dnwthom — dnwthom

Purpose:

Compute an optimal discrete-time controller using the Newton homotopy algorithm.

Synopsis:

```
[ac,bc,cc,dc,par,val] =  dnwthom(ac0,bc0,cc0,dc0, ...
                         a0,b0,c0,d0,r10,r20,v10,v20,v12, ...
                         af,bf,cf,df,r1r,r2f,v1f,v2f,v12,par)
```

Description:

Computes an optimal discrete-time controller using the Newton homotopy algorithm described in Section 6. On input, the vector **par** contains the variable algorithm parameters whose default values are set using function **nwtpar** as follows:

par = nwtpar.

See the **nwtpar** reference pages for a detailed description of the elements of par. On output, **val** is a vector containing descriptions of important run-time parameters. In particular,

val(1) = value of homotopy parameter on return

val(2) = total number of Hessian calculations

val(3) = min # of corrections for fixed lambda

val(4) = max # of corrections for fixed lambda

val(5) = minimum homotopy step size

val(6) = maximum homotopy step size.

val(7) = number of mega-flops required for run

val(8) = number of seconds required for run.

See also:

nwtpar, nwtprint

Purpose:

Compute each of the quadratic costs for a given discrete-time system.

Synopsis:

[costs, p22, q22] = dqcosts(ac,bc,cc,dc,a,b,c,d,r1,r2,r12,v1,v2,v12)

Description:

Computes the quadratic costs for the given discrete-time system. On return costs is a 5th order vector whose elements have the following values:

costs(1) = state cost ($x^{\mathrm{T}} R_1 x$)

costs(2) = input cost ($u^{\mathrm{T}} R_2 u$)

costs(3) = cross cost ($2x^{\mathrm{T}} R_{12} u$)

costs(4) = feedthrough cost (tr $D_c^{\mathrm{T}} R_2 D_c V_2$)

costs(5) = total cost (sum of the above).

The matrices **p22** and **q22** are respectively equal to the (2,2) blocks of the closed-loop observability grammian ($P_{c\ell}$) and controllability grammian ($Q_{c\ell}$). If the controller is an LQG controller, then **p22** = $\hat{P}$ and **q22** = $\hat{Q}$.

See also:

dpcl, dqcl, dcost

Purpose:

Determine the discrete-time stablility of a matrix.

Synopsis:

sflag = dstable(a)

Description:

Determines the stability of the matrix A in the discrete-time sense (i.e., are the eigenvalues of the matrix in the closed unit circle). On return, **sflag = 1** if A is stable and **sflag=0** if A is unstable.

Purpose:

Return the ordered eigenvalues of the product of two input matrices P and Q.

Synopsis:

eigpq(P, Q)

Description:

Computes and prints the ordered eigenvalues of the product of two input matrices P and Q. If the matrices are controller grammians (i.e., $\hat{P}$ and $\hat{Q}$) the ordering can be used to determine the order of a reduced-order controller.

See also:

balcred

Purpose:

Set the default parameters for the Newton homotopy algorithms.

Synopsis:

par = nwtpar

Description:

Sets the variable algorithm parameters for the homotopy algorithms for optimal, discrete-time, reduced-order controller design. A description of each of these parameters is given in the following table.

See also:

nwtprint, dnwthom

ALGORITHM PARAMETERS

No.	Function	Default	Description
1	Print Option	1	Controls amount of input during the execution of the homotopy algorithm
2	Prediction Tolerance	1.0e+3	Maximum allowable gradient norm for prediction step.
3	Correction Tolerance	1.0e–4	Maximum allowable gradient norm for intermediate correction loops.
4	Final Tolerance	1.0e–4	Maximum allowable gradient norm for the final correction loops.
5	Minimum Step Size	1.0e–6	Minimum allowable step size of the homotopy parameter.
6	Maximum Corrections	10	Maximum number of correction loops for a fixed value of the homotopy. parameter
7	Hessian for Prediction	0	0 uses Hessian from last correction step for prediction. 1 computes a new Hessian. parameter
8	Line Search	0	0 does not use line search unless cost is not decreased. 1 always uses line search.
9	Automatic Run	1	0 lets program run interactively. 1 lets program run automatically
10	Step Size	.01	On input, initial step size On output, last step size used
11	Controller Basis	1	1-no basis 2-tridiagonal form 3-second-order polynomial form 4-controllability canonical form
12	Prediction Option	1	0-no prediction 1-linear prediction 2-circular arc prediction 3-cubic spline prediction

Purpose:

Print the variable homotopy parameters.

Synopsis:

par = nwtprint(par)

Description:

Prints the information contained in the vector **par** that describes the variable algorithm parameters for the Newton homotopy algorithms.

See also:

nwtpar, dnwthom

Purpose:

Convert a plant to real normal modal form or a standard alternative. These forms are special cases of the tridiagonal form.

Synopsis:

[at,bt,ct,dt] = rnormal(a,b,c,d)

or

[at,bt,ct,dt] = rnormal(a,b,c,d, 'modal')

Description:

The first call with four input arguments converts a plant to real normal modal form, i.e., **the** transform of state matrix, **at** has 2×2 blocks of the form

$$\begin{bmatrix} -\nu_i & \omega_i \\ -\omega_i & -\nu_i \end{bmatrix}.$$

If 'modal' is input as a fifth input argument, that **at** has 2×2 blocks of the form

$$\begin{bmatrix} 0 & 1 \\ -(\sigma^2 + \omega^2) & 2\sigma \end{bmatrix}.$$

See also:

trimats

rpfx ——— rpfx

Purpose:

Transform a system to second-order polynomial form.

Synopsis:

[A,B,C,T,Tinv] = rpfx(A0,B0,C0)

Description:

Transforms a system to second-order polynomial form and also returns the transformation matrix and its inverse.

Purpose:

Transform a phase vector to lie in the interval [-180,180].

Synopsis:

phase180 = to180(phase)

Description:

Transforms a phase vector to lie in the interval [-180,180].

Purpose:

Print the run-time homotopy parameters.

Synopsis:

valprint(val)

Description:

Prints the information contained in the vector val that describes important run-time information for the Newton homotopy algorithms.

See also:

dnwthom

Appendix A: Cost Derivatives

In this appendix we consider the cost function $\mathcal{J}(A_c, B_c, C_c)$ defined by (2.4) or, equivalently, (2.16). We derive expressions for $\frac{\partial \mathcal{J}}{\partial A_c}, \frac{\partial \mathcal{J}}{\partial B_c}, \frac{\partial \mathcal{J}}{\partial C_c}$ and $\frac{\partial \mathcal{J}}{\partial D_c}$.

Let $\mathcal{L}$ denote the Lagrangian defined by (2.33) which is rewritten here as

$$\mathcal{L}(A_c, B_c, C_c, D_c, \tilde{P}, \tilde{Q}) \triangleq \mathrm{tr}[\tilde{Q}\tilde{R} + \tilde{P}(\tilde{A}\tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{V} - \tilde{Q}) + D_c^{\mathrm{T}} R_2 D_c V_2]. \tag{A.2}$$

where

$$\tilde{Q} = \tilde{A}\tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{V} \tag{A.3}$$

Then,

$$\frac{\partial \mathcal{J}}{\partial A_c} = \frac{\partial \mathcal{L}}{\partial A_c}, \quad \frac{\partial \mathcal{J}}{\partial B_c} = \frac{\partial \mathcal{J}}{\partial B_c}, \quad \frac{\partial \mathcal{J}}{\partial C_c} = \frac{\partial \mathcal{J}}{\partial C_c}, \quad \frac{\partial \mathcal{J}}{\partial D_c} = \frac{\partial \mathcal{J}}{\partial D_c} \tag{A.4a,b,c,d}$$

subject to the constraint

$$0 = \frac{\partial \mathcal{L}}{\partial \tilde{Q}} \tag{A.5}$$

or, equivalently,

$$\tilde{P} = \tilde{A}^{\mathrm{T}}\tilde{P}\tilde{A} + \tilde{R}. \tag{A.6}$$

Now, let ϕ denote an element of A_c, B_c, C_c or D_c. Then,

$$\frac{\partial \mathcal{J}}{\partial \phi} = \frac{\partial \mathcal{L}}{\partial \phi} = \mathrm{tr}\left[\tilde{Q}\frac{\partial \tilde{R}}{\partial \phi} + \tilde{P}(\frac{\partial \tilde{A}}{\partial \phi}\tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{A}\tilde{Q}\frac{\partial \tilde{A}^{\mathrm{T}}}{\partial \phi} + \frac{\partial \tilde{V}}{\partial \phi})\right] + \frac{\partial}{\partial \phi}\mathrm{tr}(D_c{}^{RT} R_2 D_c V_2) \tag{A.7}$$

or equivalently

$$\frac{\partial \mathcal{J}}{\partial \phi} = \mathrm{tr}(\tilde{Q}\frac{\partial \tilde{R}}{\partial \phi} + \tilde{P}\frac{\partial \tilde{V}}{\partial \phi} + 2\frac{\partial \tilde{A}}{\partial \phi}\tilde{Z}) + \frac{\partial}{\partial \phi}\mathrm{tr}(D_c^{\mathrm{T}} R_2 D_c V_2) \tag{A.8}$$

where

$$\tilde{Z} \triangleq \tilde{Q}\tilde{A}^{\mathrm{T}}\tilde{P}. \tag{A.9}$$

It follows from (A.8) that

$$\frac{\partial \mathcal{J}}{\partial A_c} = \frac{\partial K}{\partial A_c}, \quad \frac{\partial \mathcal{J}}{\partial B_c} = \frac{\partial K}{\partial B_c}, \quad \frac{\partial \mathcal{J}}{\partial C_c} = \frac{\partial K}{\partial C_c}, \quad \frac{\partial \mathcal{J}}{\partial D_c} = \frac{\partial K}{\partial D_c} \tag{A.10}$$

where

$$\begin{aligned} K(A_c, B_c, C_c, D_c, \tilde{P}, \tilde{Q}, \tilde{Z}) \triangleq & \mathrm{tr}[\tilde{Q}\tilde{R}(C_c, D_c) + \tilde{P}\tilde{V}(B_c, D_c) \\ & + 2\tilde{A}(A_c, B_c, C_c, D_c)\tilde{Z}] + \mathrm{tr}(D_c^{\mathrm{T}} R_2 D_c V_2) \end{aligned} \tag{A.11}$$

and from (2.7), (2.9)-(2.10), and (2.14)-(2.15)

$$\tilde{A}(B_c, C_c, D_c) = \begin{bmatrix} A - BD_cC & -BC_c \\ B_cC & A_c - B_cDC_c \end{bmatrix}, \tag{A.12}$$

$$\tilde{R}(B_c, C_c, D_c) = \begin{bmatrix} \tilde{R}_{11} & \tilde{R}_{12} \\ \tilde{R}_{12}^{\mathrm{T}} & \tilde{R}_{22} \end{bmatrix}, \tag{A.13}$$

where

$$\tilde{R}_{11} = R_1 - C^{\mathrm{T}} D_c^{\mathrm{T}} R_{12}^{\mathrm{T}} - R_{12} D_c C + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c C \tag{A.14a}$$

$$\tilde{R}_{12} = -R_{12} C_c + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 C_c \tag{A.14b}$$

$$\tilde{R}_{22} = C_c^{\mathrm{T}} R_2 C_c, \tag{A.14c}$$

and

$$\tilde{V}(B_c, C_c, D_c) = \begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{12}^{\mathrm{T}} & \tilde{V}_{22} \end{bmatrix} \tag{A.15}$$

where

$$\tilde{V}_{11} = V_1 - BD_cV_{12}^{\mathrm{T}} - V_{12} D_c^{\mathrm{T}} B^{\mathrm{T}} + BD_cV_2D_c^{\mathrm{T}} B^{\mathrm{T}} \tag{A.16a}$$

$$\tilde{V}_{12} = V_{12} B_c^{\mathrm{T}} - BD_cV_2B_c^{\mathrm{T}} \tag{A.16b}$$

$$\tilde{V}_{22} = B_cV_2B_c^{\mathrm{T}}. \tag{A.16c}$$

The desired derivative expressions will be derived using (A.10). The development proceeds by considering each of the four terms in the right hand side of (A.1) and differentiating these terms with respect to A_c, B_c, C_c and D_c. It is assumed that $\tilde{P}, \tilde{Q}$ and $\tilde{Z}$ are partitioned conformably with $\tilde{A}, \tilde{R}$ and $\tilde{V}$ such that

$$\tilde{P} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{12}^{\mathrm{T}} & \tilde{P}_{22} \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}^{\mathrm{T}} & \tilde{Q}_{22} \end{bmatrix}, \quad \tilde{Z} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{21} & \tilde{Z}_{22} \end{bmatrix}. \tag{A.17}$$

$\underline{\text{tr}\, \tilde{Q}\tilde{R}}$

$$\mathrm{tr}\tilde{Q}\tilde{R} = \mathrm{tr}(\tilde{Q}_{11}\tilde{R}_{11} + 2\tilde{Q}_{12}\tilde{R}_{12}^{\mathrm{T}} + \tilde{Q}_{22}\tilde{R}_{22}). \tag{A.18}$$

Using (A.14), we may then write

$$\begin{aligned} \mathrm{tr}\tilde{Q}\tilde{R} = {} & \mathrm{tr}(\tilde{Q}_{11}R_1 - 2\tilde{Q}_{11}C^{\mathrm{T}} D_c^{\mathrm{T}} R_{12}^{\mathrm{T}} + \tilde{Q}_{11}C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c C) \\ & +2\mathrm{tr}(-\tilde{Q}_{12}C^{\mathrm{T}} R_{12}^{\mathrm{T}} + \tilde{Q}_{12}C_c^{\mathrm{T}} R_2 D_c C) \\ & +\mathrm{tr}(\tilde{Q}_{22}C_c^{\mathrm{T}} R_2 C_c). \end{aligned} \tag{A.19}$$

Differentiating (A.19) gives

$$\frac{\partial}{\partial A_c}\mathrm{tr}\tilde{Q}\tilde{R} = 0 \tag{A.20}$$

$$\frac{\partial}{\partial B_c}\mathrm{tr}\tilde{Q}\tilde{R} = 0 \tag{A.21}$$

$$\frac{\partial}{\partial C_c}\mathrm{tr}\tilde{Q}\tilde{R} = 2(-R_{12}^{\mathrm{T}}\tilde{Q}_{12} + R_2 D_c C\tilde{Q}_{12} + R_2 C_c\tilde{Q}_{22}) \tag{A.22}$$

$$\frac{\partial}{\partial D_c}\mathrm{tr}\tilde{Q}\tilde{R} = 2(-R_{12}^{\mathrm{T}}\tilde{Q}_{11}C^{\mathrm{T}} + R_2 D_c C\tilde{Q}_{11}C^{\mathrm{T}} + R_2 C_c\tilde{Q}_{12}^{\mathrm{T}}C^{\mathrm{T}}). \tag{A.23}$$

tr $\tilde{P}\tilde{V}$

$$\mathrm{tr}\ \tilde{P}\tilde{V} = \mathrm{tr}(\tilde{P}_{11}\tilde{V}_{11} + 2\tilde{P}_{12}\tilde{V}_{12}^{\mathrm{T}} + \tilde{P}_{22}\tilde{V}_{22}). \tag{A.24}$$

Using (A.16), we may then write

$$\begin{aligned}\mathrm{tr}\ \tilde{P}\tilde{V} = {} & \mathrm{tr}(\tilde{P}_{11}V_1 - 2\tilde{P}_{11}BD_cV_{12}^{\mathrm{T}} + \tilde{P}_{11}BD_cV_2D_c^{\mathrm{T}}B^{\mathrm{T}}) \\ & +2\mathrm{tr}(\tilde{P}_{12}B_cV_{12}^{\mathrm{T}} - \tilde{P}_{12}B_cV_2D_c^{\mathrm{T}}B^{\mathrm{T}}) \\ & +\mathrm{tr}(\tilde{P}_{22}B_cV_2B_c^{\mathrm{T}}).\end{aligned} \tag{A.25}$$

Differentiating (A.25) gives

$$\frac{\partial}{\partial A_c}\mathrm{tr}\tilde{P}\tilde{V} = 0 \tag{A.26}$$

$$\frac{\partial}{\partial B_c}\mathrm{tr}\tilde{P}\tilde{V} = 2(\tilde{P}_{12}^{\mathrm{T}}V_{12} - \tilde{P}_{12}^{\mathrm{T}}BD_cV_2 + \tilde{P}_{22}B_cV_2) \tag{A.27}$$

$$\frac{\partial}{\partial C_c}\mathrm{tr}\tilde{P}\tilde{V} = 0 \tag{A.28}$$

$$\frac{\partial}{\partial D_c}\mathrm{tr}\tilde{P}\tilde{V} = 2(-B^{\mathrm{T}}\tilde{P}_{11}V_{12} + B^{\mathrm{T}}\tilde{P}_{11}BD_cV_2 - B^{\mathrm{T}}\tilde{P}_{12}B_cV_2). \tag{A.29}$$

tr $\tilde{A}\tilde{Z}$

$$\mathrm{tr}\ \tilde{A}\tilde{Z} = \mathrm{tr}[(A - BD_cC)\tilde{Z}_{11} - BC_c\tilde{Z}_{21} + B_cC\tilde{Z}_{12} + (A_c - B_cDC_c)\tilde{Z}_{22}] \tag{A.30}$$

Differentiating (A.30) gives

$$\frac{\partial}{\partial A_c}\mathrm{tr}\tilde{A}\tilde{Z} = 2\tilde{Z}_{22}^{\mathrm{T}}. \tag{A.31}$$

$$\frac{\partial}{\partial B_c}\mathrm{tr}\tilde{A}\tilde{Z} = 2(\tilde{Z}_{12}^{\mathrm{T}}C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}}C_c^{\mathrm{T}}D^{\mathrm{T}}) \tag{A.32}$$

$$\frac{\partial}{\partial C_c}\mathrm{tr}\tilde{A}\tilde{Z} = -2(B^{\mathrm{T}}\tilde{Z}_{21}^{\mathrm{T}} + D^{\mathrm{T}}B_c^{\mathrm{T}}\tilde{Z}_{22}^{\mathrm{T}}) \tag{A.33}$$

$$\frac{\partial}{\partial D_c}\mathrm{tr}\tilde{A}\tilde{Z} = -2B^{\mathrm{T}}\tilde{Z}_{11}^{\mathrm{T}}C^{\mathrm{T}}. \tag{A.34}$$

tr $D_c^{\mathrm{T}} R_2 D_c V_2$

$$\frac{\partial}{\partial A_c} \mathrm{tr} D_c^{\mathrm{T}} R_2 D_c V_2 = 0 \tag{A.35}$$

$$\frac{\partial}{\partial B_c} \mathrm{tr} D_c^{\mathrm{T}} R_2 D_c V_2 = 0 \tag{A.36}$$

$$\frac{\partial}{\partial C_c} \mathrm{tr} D_c^{\mathrm{T}} R_2 D_c V_2 = 0 \tag{A.37}$$

$$\frac{\partial}{\partial D_c} \mathrm{tr} D_c^{\mathrm{T}} R_2 D_c V_2 = 2 R_2 D_c V_2. \tag{A.38}$$

$\mathcal{J}(A_c, B_c, C_c D_c)$

It follows from (A.10) and (A.11) with (A.20)–(A.23), (A.25)–(A.29), (A.31)–(A.34) and (A.35)–(A.38) that

$$\frac{\partial \mathcal{J}}{\partial A_c} = \tilde{Z}_{22}^{\mathrm{T}} \tag{A.39}$$

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial B_c} &= 2(\tilde{P}_{12}^{\mathrm{T}} V_{12} - \tilde{P}_{12}^{\mathrm{T}} B D_c V_2 + \tilde{P}_{22} B_c V_2 \\ &\quad + \tilde{Z}_{12}^{\mathrm{T}} C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} D^{\mathrm{T}}) \end{aligned} \tag{A.40}$$

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial C_c} &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{12} + R_2 D_c C \tilde{Q}_{12} + R_2 C_c \tilde{Q}_{22} \\ &\quad - B^{\mathrm{T}} \tilde{Z}_{21}^{\mathrm{T}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{\mathrm{T}}) \end{aligned} \tag{A.41}$$

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial D_c} &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{11} C^{\mathrm{T}} + R_2 D_c C \tilde{Q}_{11} + R_2 C_c \tilde{Q}_{12}^{\mathrm{T}} C^{\mathrm{T}} \\ &\quad + B^{\mathrm{T}} \tilde{P}_{11} V_{12} + B^{\mathrm{T}} P_{11} B D_c V_2 - B^{\mathrm{T}} \tilde{P}_{12} B_c V_2 \\ &\qquad - B^{\mathrm{T}} \tilde{Z}_{11}^{\mathrm{T}} C^{\mathrm{T}} + R_2 D_c V_2). \end{aligned} \tag{A.42}$$

Appendix B: Closed-Loop Matrix Derivatives

In this appendix we present explicit expressions for the derivatives $\frac{\partial\tilde{A}}{\partial\theta_j}$, $\frac{\partial\tilde{R}}{\partial\theta_j}$, $\frac{\partial\tilde{V}}{\partial\theta_j}$, $\frac{\partial\tilde{A}}{\partial\lambda}$, $\frac{\partial\tilde{R}}{\partial\lambda}$, and $\frac{\partial\tilde{V}}{\partial\lambda}$ where

$$\theta = \begin{bmatrix} \text{vec}(A_c) \\ \text{vec}(B_c) \\ \text{vec}(C_c) \\ \text{vec}(D_c) \end{bmatrix} \tag{B.1}$$

$$\tilde{A} = \begin{bmatrix} A - BD_cC & -BC_c \\ B_cC & A_c - B_cDC_c \end{bmatrix}, \tag{B.2}$$

$$\tilde{R} = \begin{bmatrix} \tilde{R}_{11} & \tilde{R}_{12} \\ \tilde{R}_{12}^{\mathrm{T}} & \tilde{R}_{22} \end{bmatrix}, \tag{B.3}$$

where

$$\tilde{R}_{11} = R_1 - C^{\mathrm{T}}D_c^{\mathrm{T}}R_{12}^{\mathrm{T}} - R_{12}D_cC + C^{\mathrm{T}}D_c^{\mathrm{T}}R_2D_cC \tag{B.4a}$$

$$\tilde{R}_{12} = -R_{12}C_c + C^{\mathrm{T}}D_c^{\mathrm{T}}R_2C_c \tag{B.4b}$$

$$\tilde{R}_{22} = C_c^{\mathrm{T}}R_2C_c, \tag{B.4c}$$

and

$$\tilde{V} = \begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{12}^{\mathrm{T}} & \tilde{V}_{22} \end{bmatrix} \tag{B.5}$$

where

$$\tilde{V}_{11} = V_1 - BD_cV_{12}^{\mathrm{T}} - V_{12}D_c^{\mathrm{T}}B^{\mathrm{T}} + BD_cV_2D_c^{\mathrm{T}}B^{\mathrm{T}} \tag{B.6a}$$

$$\tilde{V}_{12} = V_{12}B_c^{\mathrm{T}} - BD_cV_2B_c^{\mathrm{T}} \tag{B.6b}$$

$$\tilde{V}_{22} = B_c^{\mathrm{T}}V_2B_c^{\mathrm{T}}. \tag{B.6c}$$

It is assumed that the plant matrices (A, B, C, D), the cost weighting matrices (R_1, R_{12}, R_2) and the disturbance matrices (V_1, V_{12}, V_2) are the following functions of λ.

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} + \lambda\left(\begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} \right) \tag{B.7}$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^{\mathrm{T}}(\lambda) & R_2(\lambda) \end{bmatrix} = L_R(\lambda)L_R^{\mathrm{T}}(\lambda) \tag{B.8a}$$

where

$$L_R(\lambda) = L_{R,0} + \lambda(L_{R,f} - L_{R,0}) \tag{B.8b}$$

and $L_{R,0}$ and $L_{R,f}$ satisfy

$$L_{R,0}L_{R,0}^{\mathrm{T}} = \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} \tag{B.8c}$$

$$L_{R,f}L_{R,f}^{\mathrm{T}} = \begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^{\mathrm{T}} & R_{2,f} \end{bmatrix} \tag{B.8d}$$

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^{\mathrm{T}}(\lambda) & V_2^{\mathrm{T}}(\lambda) \end{bmatrix} = L_V(\lambda)L_V^{\mathrm{T}}(\lambda) \tag{B.9a}$$

where

$$L_V(\lambda) = L_{V,0} + \lambda(L_{V,f} - L_{V,0}) \tag{B.9b}$$

and $L_{V,0}$ and $L_{V,f}$ satisfy

$$L_{V_0}L_{V,0}^{\mathrm{T}} = \begin{bmatrix} V_{1,0} & V_{12,0} \\ V_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix} \tag{B.9c}$$

$$L_{V_f}L_{V,f}^{\mathrm{T}} = \begin{bmatrix} V_{1,f} & V_{12,f} \\ V_{12,f}^{\mathrm{T}} & V_{2,0} \end{bmatrix}. \tag{B.9d}$$

Below, we use the notation

$$\dot{M} \triangleq \frac{\partial M}{\partial \lambda}. \tag{B.10}$$

Note that from (B.7)–(B.9)

$$\begin{bmatrix} \dot{A} & \dot{B} \\ \dot{C} & \dot{D} \end{bmatrix} = \begin{bmatrix} A_f - A_0 & B_f - B_0 \\ C_f - C_0 & D_f - D_0 \end{bmatrix} \tag{B.11}$$

$$\begin{bmatrix} \dot{R}_1 & \dot{R}_{12} \\ \dot{R}_{12}^{\mathrm{T}} & \dot{R}_2 \end{bmatrix} = \dot{L}_R\dot{L}_R^{\mathrm{T}} \tag{B.12a}$$

where

$$\dot{L}_R = L_{R,f} - L_{R,0} \tag{B.12b}$$

$$\begin{bmatrix} \dot{V}_1 & \dot{V}_{12} \\ \dot{V}_{12}^{\mathrm{T}} & \dot{V}_2 \end{bmatrix} = \dot{L}_V\dot{L}_V^{\mathrm{T}} \tag{B.12c}$$

where

$$L_V = L_{V,f} - L_{V,0}. \tag{B.12d}$$

The derivations of the expression for $\frac{\partial \dot{A}}{\partial \theta_j}$, $\frac{\partial \dot{R}}{\partial \theta_j}$ and $\frac{\partial \dot{V}}{\partial \theta_j}$ are primarily based on the application of the following derivative formulas. It is assumed that X is an $n \times m$ matrix.

Derivative Formulas

$$\frac{d}{dx_{ij}} XA = [A(j,:)]_{\text{row}-i} \tag{B.14}$$

$$\frac{d}{dx_{ij}} AX = [A(:,i)]_{\text{col}-j} \tag{B.15}$$

$$\frac{d}{dx_{ij}} X^{\mathrm{T}} A = [A(i,:)]_{\text{row}-j} \tag{B.16}$$

$$\frac{d}{dx_{ij}} AX^{\mathrm{T}} = [A(:,j)]_{\text{col}-i} \tag{B.17}$$

$$\frac{d}{dx_{ij}} AXB = A(:,i)B(j,:) \tag{B.18}$$

$$\frac{d}{dx_{ij}} AX^{\mathrm{T}}B = A(:,j)B(i,:) \tag{B.19}$$

$$\underline{\frac{\partial \tilde{A}}{\partial \theta_j}}$$

$$\frac{\partial \tilde{A}}{\partial a_{c,k\ell}} = \begin{bmatrix} 0 & 0 \\ 0 & E^{(k,\ell)}_{n_c \times n_c} \end{bmatrix} \tag{B.20}$$

$$\frac{\partial \tilde{A}}{\partial b_{c,k\ell}} = \begin{bmatrix} 0 & 0 \\ [C(\ell,:)]_{\text{row}-k} & [\underline{DC_c}(\ell,:)]_{\text{row}-k} \end{bmatrix} \tag{B.21}$$

$$\frac{\partial \tilde{A}}{\partial c_{c,k\ell}} = \begin{bmatrix} 0 & -[B(:,k)]_{\text{col}-\ell} \\ 0 & \frac{\partial A_c}{\partial c_{c_{k,\ell}}} - [\underline{B_c D}(:,k)]_{\text{col}-\ell} \end{bmatrix} \tag{B.22}$$

$$\frac{\partial \tilde{A}}{\partial d_{c,k\ell}} = \begin{bmatrix} -B(:,k)C(\ell,:) & 0 \\ 0 & 0 \end{bmatrix} \tag{B.23}$$

where $\frac{\partial A_c}{\partial b_{c,k\ell}}$ and $\frac{\partial A_c}{\partial c_{c,k\ell}}$ are given respectively by (D.2.36) and (D.2.37) of Appendix D.

$$\underline{\frac{\partial \tilde{R}}{\partial \theta_j}}$$

$$\frac{\partial \tilde{R}}{\partial a_{c,k\ell}} = 0 \tag{B.24}$$

$$\frac{\partial \tilde{R}}{\partial b_{c,k\ell}} = 0 \tag{B.25}$$

$$\frac{\partial \tilde{R}}{\partial c_{c,k\ell}} = \begin{bmatrix} 0 & [-R_{12}(:,k) + \underline{C^{\mathrm{T}} D_c^{\mathrm{T}} R_2}(:,k)]_{\mathrm{col}-\ell} \\ \mathrm{SYM} & [\underline{C_c^{\mathrm{T}} R_2}(:,k)]_{\mathrm{col}-\ell} + [\underline{R_2 C_c}(k,:)]_{\mathrm{row}-\ell} \end{bmatrix} \tag{B.26}$$

$$\frac{\partial \tilde{R}}{\partial d_{c,k\ell}} = \begin{bmatrix} \begin{matrix} C^{\mathrm{T}}(:,\ell)[\underline{R_2 D_c C}(k,:) - R_{12}^{\mathrm{T}}(k,:)] \\ +[\underline{C^{\mathrm{T}} D_c^{\mathrm{T}} R_2}(:,k) - R_{12}(:,k)]C(\ell,:) \end{matrix} & C^{\mathrm{T}}(:,\ell)\underline{R_2 C_c}(k,:) \\ \mathrm{SYM} & 0 \end{bmatrix} \tag{B.27}$$

$$\underline{\frac{\partial \tilde{V}}{\partial \theta_j}}$$

$$\frac{\partial \tilde{V}}{\partial a_{c,k\ell}} = 0 \tag{B.28}$$

$$\frac{\partial \tilde{V}}{\partial b_{c,k\ell}} = \begin{bmatrix} 0 & [V_{12}(:,\ell) - \underline{B D_c V_2}(:,\ell)]_{\mathrm{col}-k} \\ \mathrm{SYM} & [\underline{V_2 B_c^{\mathrm{T}}}(\ell,:)]_{\mathrm{row}-k} + [\underline{B_c V_2}(:,\ell)]_{\mathrm{col}-k} \end{bmatrix} \tag{B.29}$$

$$\frac{\partial \tilde{V}}{\partial c_{c,k\ell}} = 0 \tag{B.30}$$

$$\frac{\partial \tilde{V}}{\partial d_{c,k\ell}} = \begin{bmatrix} \begin{matrix} B(:,k)[\underline{V_2 D_c^{\mathrm{T}} B^{\mathrm{T}}}(\ell,:) - V_{12}^{\mathrm{T}}(\ell,:)] \\ +[\underline{B D_c V_2}(:,\ell) - V_{12}(:,\ell)]B^{\mathrm{T}}(k,:) \end{matrix} & -B(:,k)\underline{V_2 B_c^{\mathrm{T}}}(\ell,:) \\ \mathrm{SYM} & 0 \end{bmatrix} \tag{B.31}$$

$$\underline{\dot{\tilde{A}} \triangleq \frac{\partial \tilde{A}}{\partial \lambda}}$$

$$\dot{\tilde{A}} = \begin{bmatrix} \dot{A} - \dot{B} D_c C - B D_c \dot{C} & -\dot{B} C_c \\ B_c \dot{C} & -B_c \dot{D} C_c \end{bmatrix} \tag{B.32}$$

$$\underline{\dot{\tilde{R}} \triangleq \frac{\partial \tilde{R}}{\partial \lambda}}$$

$$\dot{\tilde{R}} = \begin{bmatrix} \dot{\tilde{R}}_{11} & \dot{\tilde{R}}_{12} \\ \dot{\tilde{R}}_{12}^{\mathrm{T}} & \dot{\tilde{R}}_{22} \end{bmatrix} \tag{B.33}$$

where

$$\begin{aligned} \dot{\tilde{R}}_{11} = \dot{R}_1 - \dot{C}^{\mathrm{T}} D_c^{\mathrm{T}} R_{12}^{\mathrm{T}} - C^{\mathrm{T}} D_c^{\mathrm{T}} \dot{R}_{12}^{\mathrm{T}} - \dot{R}_{12} D_c C - R_{12} D_c \dot{C} & \\ + \dot{C}^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c C + C^{\mathrm{T}} D_c^{\mathrm{T}} \dot{R}_2 D_c C + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c \dot{C} & \qquad (\mathrm{B.34}a) \end{aligned}$$

$$\dot{\tilde{R}}_{12} = -\dot{R}_{12} C_c + \dot{C}^{\mathrm{T}} D_c^{\mathrm{T}} R_2 C_c + C^{\mathrm{T}} D_c^{\mathrm{T}} \dot{R}_2 C_c \tag{B.34b}$$

$$\dot{\tilde{R}}_{22} = C_c^{\mathrm{T}} \dot{R}_2 C_c. \tag{B.34c}$$

$$\dot{\tilde{V}} \triangleq \frac{\partial \tilde{V}}{\partial \lambda}$$

$$\dot{\tilde{V}} = \begin{bmatrix} \dot{\tilde{V}}_{11} & \dot{\tilde{V}}_{12} \\ \dot{\tilde{V}}_{12}^{\mathrm{T}} & \dot{\tilde{V}}_{22} \end{bmatrix} \tag{B.35}$$

where

$$\begin{aligned} \dot{\tilde{V}}_{11} &= \dot{V}_1 - \dot{B}D_C V_{12}^{\mathrm{T}} - BD_c\dot{V}_{12}^{\mathrm{T}} - \dot{V}_{12}D_c^{\mathrm{T}}B^{\mathrm{T}} - V_{12}D_c^{\mathrm{T}}\dot{B}^{\mathrm{T}} \\ &\quad + \dot{B}D_cV_2D_cB^{\mathrm{T}} + BD_c\dot{V}_2D_cB^{\mathrm{T}} + BD_cV_2D_c\dot{B}^{\mathrm{T}} \end{aligned} \tag{B.36a}$$

$$\dot{\tilde{V}}_{12} = -\dot{V}_{12}B_c^{\mathrm{T}} - \dot{B}D_c^{\mathrm{T}}V_2B_c^{\mathrm{T}} - BD_c\dot{V}_2B_c^{\mathrm{T}} \tag{B.36b}$$

$$\dot{\tilde{V}}_{22} = B_c\dot{V}_2B_c^{\mathrm{T}}. \tag{B.36c}$$

Appendix C: The Input-Normal Riccati Basis

The homotopy function $H(\boldsymbol{\theta}, \lambda)$ described in Section 4 is defined to solve the optimal reduced-order dynamic compensation problem for discrete-time systems. The elements of the vector $\boldsymbol{\theta}$ include parameters which completely describe the controller (A_c, B_c, C_c, D_c). For computational efficiency it is desired that the vector $\boldsymbol{\theta}$ be as small as possible. Thus, we desire to represent the controller (A_c, B_c, C_c, D_c), assumed to be minimal, with the fewest parameters possible. The results of this section reveal that in a certain basis, which we will denote as the input normal Riccati basis, the controller plant matrix A_c is almost always completely characterized by its input and output matrices B_c and C_c.

Theorem C.1. For every minimal $n_c^{\rm th}$ order plant $(\bar{A}_c, \bar{B}_c, \bar{C}_c, \bar{D}_c)$ there exists a similarity transformation T_I and a positive matrix $\Omega = \operatorname{diag}\{\omega_i\}_{i=1}^{n_c}$ such that $(A_c = T_I^{-1}\bar{A}_c T_I B_c = T_I^{-1}\bar{B}_c,\ C_c = \bar{C}_c T_I,\ D_c = \bar{D}_c)$ satisfies

$$0 = A_c + A_c^{\rm T} + B_c B_c^{\rm T} - C_c^{\rm T} C_c \tag{C.1}$$

$$0 = A_c^{\rm T}\Omega + \Omega A_c + C_c^{\rm T} C_c - \Omega B_c B_c^{\rm T}\Omega. \tag{C.2}$$

In addition,

$$\omega_i = [(C_c^{\rm T} C_c)_{ii}/(B_c B_c^{\rm T})_{ii}]^{1/2} \tag{C.3}$$

$$A_{c,ii} = \frac{1}{2}[(C_c^{\rm T} C_c)_{ii} - (B_c B_c^{\rm T})_{ii}] \tag{C.4}$$

and if $\omega_i \neq \omega_j$ for $i \neq j$, then

$$A_{c,ij} = \frac{\omega_j(1+\omega_i)(B_c B_c^{\rm T})_{ij} - (C_c^{\rm T} C_c)_{ij}(1+\omega_j)}{\omega_i - \omega_j}, \quad i \neq j \tag{C.5}$$

so that A_c is completely and uniquely determined by B_c and C_c.

Proof. The minimality of $(\bar{A}_c, \bar{B}_c, \bar{C}_c)$ insures that there exists unique positive definite solutions Q and P of

$$0 = \bar{A}_c Q + Q\bar{A}_c^{\rm T} + \bar{B}_c \bar{B}_c^{\rm T} - Q\bar{C}_c^{\rm T}\bar{C}_c Q \tag{C.6}$$

$$0 = \bar{A}_c^{\rm T} P + P\bar{A}_c + \bar{C}_c^{\rm T}\bar{C}_c - P\bar{B}_c\bar{B}_c^{\rm T} P. \tag{C.7}$$

It is well known that there exist a transformation T_I such that

$$T_I^{-1} Q T_I^{-\rm T} = I_{n_c}, \quad T_I^{\rm T} P T_I = \Omega \tag{C.8}$$

It follows from (C.6)–(C.8) that (A_c, B_c, C_c) satisfies (C.1) and (C.2).

We now show by construction that (C.3)–(C.5) hold. First recognize that (C.1) and (C.2) are equivalent respectively to

$$0 = A_{c,ij} + A_{c,ji} + (B_c B_c^{\mathrm{T}})_{ij} - (C_c^{\mathrm{T}} C_c)_{ij} \tag{C.9}$$

$$0 = \omega_i A_{c,ij} + \omega_j A_{c,ji} + (C_c^{\mathrm{T}} C_c)_{ij} - \omega_i \omega_j (B_c B_c^{\mathrm{T}})_{ij}. \tag{C.10}$$

Letting $i = j$ in (C.6) gives (C.4) while it follows from (C.7) that for $i = j$

$$0 = 2\omega_i A_{c,ii} + (C_c^{\mathrm{T}} C_c)_{ii} - \omega_i^2 (B_c^{\mathrm{T}} B_c)_{ii}. \tag{C.11}$$

Substituting (C.4) into (C.11) gives

$$0 = -(B_c^{\mathrm{T}} B_c)_{ii} \omega_i^2 + [(C_c^{\mathrm{T}} C_c)_{ii} - (B_c B_c^{\mathrm{T}})_{ii}]\omega_i + (C_c^{\mathrm{T}} C_c)_{ii} \tag{C.12}$$

which has positive solution ω_i given by (C.3).

Multiplying (C.9) by $-\omega_i^2$ and adding the result to (C.10) gives

$$0 = (\omega_i - \omega_j) A_{c,ij} + (C_c^{\mathrm{T}} C_c)_{ij}(1 + \omega_j) - \omega_j (1 + \omega_i)(B_c B_c^{\mathrm{T}})_{ij} \tag{C.13}$$

which implies that if $\omega_i \neq \omega_j$ for $i \neq j$ then $A_{c,ij}$ is given by (C.5). □

Definition C.1. If the minimal order plant (A_c, B_c, C_c, D_c) satisfies (C.1) and (C.2) of Theorem C.1, then the plant is said to be in input normal Riccati form.

Remark C.1. Input normal Riccati form is similar to the input normal form of Moore [1] which is further explored by Kabamba [2].

Now, define

$$\Omega \triangleq \operatorname{diag}\{\omega_i\}_{i=1}^{n_c} \tag{C.14}$$

and $H \in \mathbb{R}^{n_c \times n_c}$ such that

$$h_{ij} \triangleq \begin{cases} (\omega_i = \omega_j)^{-1}, & i \neq j \\ 0, & i = j. \end{cases} \tag{C.15}$$

or equivalently

$$\Omega \triangleq \operatorname{diag}(C_c^{\mathrm{T}} C_c)\operatorname{diag}(B_c B_c^{\mathrm{T}})^{-1} \tag{C.16}$$

$$H \triangleq (N_{n_c} - I_{n_c})/[N_{n_c}\Omega - \Omega N_{n_c} + I_{n_c}] \tag{C.17}$$

where $n_m \in \mathbb{R}^{m \times m}$ defined by

$$N_{m,ij} = 1. \tag{C.18}$$

Then the following remark holds.

Remark C.2. If $\omega_i \neq \omega_j$ for $i \neq j$, then (C.4) and (C.5) are equivalent to

$$A_c = \frac{1}{2}[C_c^\mathrm{T} C_c - B_c B_c^\mathrm{T}] * I_{n_c} + [C_c^\mathrm{T} C_c(I + \Omega) - (I_{n_c} + \Omega)B_c B_c^\mathrm{T} \Omega] * H \tag{C.19}$$

Proposition C.1. Let Λ and Z be in $\mathbb{R}^{n \times n}$ with Λ diagonal. Then,

$$\Lambda Z = \bar{\Lambda} * Z \tag{C.20}$$

and

$$Z\Lambda = Z * \bar{\Lambda}^\mathrm{T}. \tag{C.21}$$

Remark C.3. It follows from Remark C.2 and Proposition C.1 that if $\omega_i \neq \omega_j$ for $i \neq j$, then A_c can be computed by

$$A_c = \frac{1}{2}[C_c^\mathrm{T} C_c - B_c B_c^\mathrm{T}] * I_{n_c} + [C_c^\mathrm{T} C_c + (C_c C_c^\mathrm{T}) * \bar{\Omega}^\mathrm{T} - (B_c B_c^\mathrm{T} + \bar{\Omega} * (B_c B_c^\mathrm{T} + \bar{\Omega} * (B_c B_c^\mathrm{T})) * \bar{\Omega}^\mathrm{T}] * H \tag{C.22}$$

where from (C.17) with (C.20) and (C.21)

$$H = (N_{n_c} - I_{n_c}) / \bar{\Omega}^\mathrm{T} - \bar{\Omega} + I_{n_c}). \tag{C.23}$$

References

1. B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction," *IEEE Trans. Autom. Contr.*, Vol. AC–26, pp. 17–32, Feb. 1981.

2. P. T. Kabamba, "Balanced Forms: Canonicity and Parameterization," *IEEE Trans. Autom. Contr.*, Vol. AC–30, pp. 1106–1109, Nov. 1985.

Appendix D: The Gradient of the Cost Functional for the Input Normal Riccati Basis

In this Appendix it is assumed that the controller (A_c, B_c, C_c, D_c) is in the input normal Riccati form of Appendix C and is hence completely described in terms of B_c, C_c and D_c. We let $\bar{\mathcal{J}}(B_c, C_c, D_c)$ be the restriction of the cost functional $\mathcal{J}(A_c, B_c, C_c, D_c)$ defined by (2.4) or equivalently (2.16) on the set of generic input-normal Riccati triples (A_c, B_c, C_c). Also, define

$$\theta \triangleq \begin{bmatrix} \mathrm{vec}(B_c) \\ \mathrm{vec}(C_c) \\ \mathrm{vec}(D_c) \end{bmatrix}. \tag{D.1}$$

Then, with some abuse of notation we can write the restricted cost functional as $\bar{\mathcal{J}}(\theta)$. The homotopy algorithms to be defined later will be based on finding θ satisfying

$$0 = f(\theta) \triangleq \frac{\partial \bar{\mathcal{J}}}{\partial \theta}(\theta). \tag{D.2}$$

Now, recognize that

$$\nabla \bar{\mathcal{J}}(\theta)^{\mathrm{T}} \triangleq \frac{\partial \bar{\mathcal{J}}}{\partial \theta} = \begin{bmatrix} \mathrm{vec}\frac{\partial \bar{\mathcal{J}}}{\partial B_c} \\ \mathrm{vec}\frac{\partial \bar{\mathcal{J}}}{\partial C_c} \\ \mathrm{vec}\frac{\partial \bar{\mathcal{J}}}{\partial D_c} \end{bmatrix}. \tag{D.3}$$

The next theorem present very useful expressions for $\frac{\partial \bar{\mathcal{J}}}{\partial B_c}$, $\frac{\partial \bar{\mathcal{J}}}{\partial C_c}$, and $\frac{\partial \bar{\mathcal{J}}}{\partial D_c}$. This result is very similar to Theorem 2 of [D.1].

Theorem D.1. The derivatives $\frac{\partial \bar{\mathcal{J}}}{\partial B_c}$, $\frac{\partial \bar{\mathcal{J}}}{\partial C_c}$, and $\frac{\partial \bar{\mathcal{J}}}{\partial D_c}$ are given by

$$\frac{\partial \bar{\mathcal{J}}}{\partial B_c} = \frac{\partial \mathcal{J}}{\partial B_c} + 2(Y - \Omega Z \Omega)B_c \tag{D.4}$$

$$\frac{\partial \bar{\mathcal{J}}}{\partial C_c} = \frac{\partial \mathcal{J}}{\partial C_c} + 2C_c(Z - Y) \tag{D.5}$$

$$\frac{\partial \bar{\mathcal{J}}}{\partial D_c} = \frac{\partial \mathcal{J}}{\partial D_c} \tag{D.6}$$

where $Y \in \mathrm{IR}^{n \times n}$ and $Z \in \mathrm{IR}^{n \times n}$ are symmetric and satisfy

$$0 = \frac{\partial \mathcal{J}}{\partial A_c} + 2(Y + \Omega Z) \tag{D.7}$$

$$0 = [(A_c - B_c B_c^{\mathrm{T}} \Omega) Z]_{ii}, \quad i = 1, 2, \ldots, n_c. \tag{D.8}$$

Proof. Since the triple (A_c, B_c, C_c) is constrained to be in input-normal Riccati form, it satisfies

$$0 = A_c + A_c^{\mathrm{T}} + B_c B_c^{\mathrm{T}} - C_c^{\mathrm{T}} C_c \tag{D.9}$$

$$0 = A_c^{\mathrm{T}} \Omega + \Omega A_c + C_c^{\mathrm{T}} C_c - \Omega B_c B_c^{\mathrm{T}} \Omega. \tag{D.10}$$

Following the proof of Theorem 2 of [D.1] we define the new Lagrangian

$$\begin{aligned}&\bar{\mathcal{L}}(A_c, B_c, C_c, D_c, \tilde{P}, \tilde{Q}, Y, Z)\\ &\quad = \mathcal{L}(A_c, B_c, C_c, D_c, \tilde{P}, \tilde{Q})\\ &\qquad + \mathrm{tr}[Y(A_c + A_c^{\mathrm{T}} - C_c^{\mathrm{T}} C_c + B_c B_c^{\mathrm{T}}) + Z(A_c^{\mathrm{T}} \Omega + \Omega A_c + C_c^{\mathrm{T}} C_c - \Omega B_c B_c^{\mathrm{T}} \Omega)]\end{aligned} \tag{D.11}$$

where Y and Z are $n \times n$ symmetric Lagrange multiplier matrices. Then

$$\frac{\partial \bar{\mathcal{J}}}{\partial B_c} = \frac{\partial \mathcal{J}}{\partial B_c}, \quad \frac{\partial \bar{\mathcal{J}}}{\partial C_c} = \frac{\partial \mathcal{J}}{\partial C_c}, \quad \frac{\partial \bar{\mathcal{J}}}{\partial D_c} = \frac{\partial \mathcal{J}}{\partial D_c} \tag{D.12a,b,c}$$

subject to the constraints

$$0 = \frac{\partial \bar{\mathcal{L}}}{\partial A_c}, \quad 0 = \frac{\partial \bar{\mathcal{L}}}{\partial \Omega}. \tag{D.13a,b}$$

Now,

$$\frac{\partial \bar{\mathcal{L}}}{\partial A_c} = \frac{\partial \mathcal{J}}{\partial A_c} + 2(Y + \Omega Z) \tag{D.14}$$

and

$$\frac{\partial \bar{\mathcal{L}}}{\partial \Omega} = A_c Z + Z A_c^{\mathrm{T}} - Z \Omega B_c B_c^{\mathrm{T}} - B_c B_c^{\mathrm{T}} \Omega Z \tag{D.15}$$

which implies

$$\frac{\partial \bar{\mathcal{L}}}{\partial \omega_i} = 2(A_c Z - B_c B_c^{\mathrm{T}} \Omega Z)_{ii}, \quad i = 1, \ldots, n_c. \tag{D.16}$$

Equations (D.7) and (D.8) follow respectively from (D.13a) with (D.14) and (D.13b) with (D.15). Finally, (D.4)–(D.6) follow respectively from (D.12a,b,c). □

We now state a very important corollary which describes how to efficiently compute Y and Z satisfying (D.7) and (D.8). For convenience, we define

$$\mathcal{L}_{A_c} \triangleq \frac{\partial \mathcal{L}}{\partial A_c} \tag{D.17a}$$

$$F \triangleq A_c - B_c B_c^{\mathrm{T}} \Omega. \tag{D.17b}$$

and rewrite (C.16) here as

$$\Omega \triangleq \mathrm{diag}(C_c^{\mathrm{T}} C_c)\, \mathrm{diag}(B_c B_c^{\mathrm{T}})^{-1}.$$

Corollary D.1. The matrices Y and Z in (D.7) and (D.8) satisfy

$$Y = -(\frac{1}{2}\mathcal{L}_{A_c} + \Omega Z) \tag{D.19}$$

$$z_{ii} = -f_{ii}^{-1} \sum_{\substack{j=1\\ j \neq i}}^{n_c} f_{ij} z_{ji} \tag{D.20}$$

and

$$0 = Z\Omega - \Omega Z + \frac{1}{2}(\mathcal{L}_{A_c} - \mathcal{L}_{A_c}^{\mathrm{T}}) \tag{D.22}$$

which if $\omega_i \neq \omega_j$ for $i \neq j$ is equivalent to

$$z_{ij} = \frac{\frac{1}{2}(\mathcal{L}_{A_c,ij} - \mathcal{L}_{A_c,ji})}{\omega_j - \omega_i}, \quad i \neq j. \tag{D.23}$$

Proof. Equations (D.19) and (D.20) follow respectively from (D.7) and (D.8). Since Y is symmetric

$$0 = Y - Y^{\mathrm{T}}. \tag{D.24}$$

Substituting (D.19) into (D.24) gives (D.22) or equivalently

$$0 = (\omega_j - \omega_i)z_{ij} - \frac{1}{2}(\mathcal{L}_{A_c,ij} - \mathcal{L}_{A_c,ji}), \quad i \neq j \tag{D.25}$$

which if $\omega_i \neq \omega_j$ for $i \neq j$ is equivalent to (D.23). □

Remark D.1. If $\omega_i \neq \omega_j$ for $i \neq j$ (D.20) and (D.21) are equivalent to

$$Z = Z_o + Z_d \tag{D.26}$$

where

$$Z_o \triangleq \frac{1}{2}(\mathcal{L}_{A_c} - \mathcal{L}_{A_c}^{\mathrm{T}}) * H \tag{D.27}$$

$$Z_d \triangleq -\operatorname{diag}(F)^{-1} * (FZ_0) \tag{D.28}$$

and H is given by (C.23), rewritten here as

$$H \triangleq (N_{n_c} - I_{n_c})/[\bar{\Omega}^{\mathrm{T}} - \bar{\Omega} + I_{n_c}]. \tag{D.29}$$

Expressions for the partial derivatives $\frac{\partial \mathcal{J}}{\partial A_c}, \frac{\partial \mathcal{J}}{\partial B_c}, \frac{\partial \mathcal{J}}{\partial C_c}$, and $\frac{\partial \mathcal{J}}{\partial D_c}$ are derived in Appendix A. Here, we cite only the final results. First, we define

$$\tilde{Z} \triangleq \tilde{Q}\tilde{A}^{\mathrm{T}}\tilde{P} \tag{D.30}$$

and note that $\tilde{P}, \tilde{Q}$ and $\tilde{Z}$ have the partitioned forms

$$\tilde{P} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{12}^{\mathrm{T}} & \tilde{P}_{22} \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}^{\mathrm{T}} & \tilde{Q}_{22} \end{bmatrix}, \quad \tilde{Z} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{12}^{\mathrm{T}} & \tilde{Z}_{22} \end{bmatrix}, \tag{D.31}$$

where the (1,1) and (2.2) blocks of each matrix are respectively $n \times n$ and $n_c \times n_c$. With this in mind, the cost derivatives are given by

$$\frac{\partial \mathcal{J}}{\partial A_c} = Q\tilde{Z}_{22}^{\mathrm{T}} \tag{D.32}$$

$$\begin{aligned}
\frac{\partial \mathcal{J}}{\partial B_c} &= 2(\tilde{P}_{12}^{\mathrm{T}} V_{12} - \tilde{P}_{12}^{\mathrm{T}} B D_c V_2 + \tilde{P}_{22} B_c V_2 \\
&\quad + \tilde{Z}_{12}^{\mathrm{T}} C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} D^{\mathrm{T}}) && \text{(D.33)} \\
\frac{\partial \mathcal{J}}{\partial C_c} &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{12} + R_2 D_c C \tilde{Q}_{12} + R_2 C_c \tilde{Q}_{22} \\
&\quad + B^{\mathrm{T}} \tilde{Z}_{21}^{\mathrm{T}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{\mathrm{T}}) && \text{(D.34)} \\
\frac{\partial \mathcal{J}}{\partial D_c} &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{11} C^{\mathrm{T}} + R_2 D_c C \tilde{Q}_{11} C^{\mathrm{T}} + R_2 C_c \tilde{Q}_{12}^{\mathrm{T}} C^{\mathrm{T}} \\
&\quad - B^{\mathrm{T}} \tilde{P}_{11} V_{12} + B^{\mathrm{T}} \tilde{P}_{11} B D_c V_2 - B^{\mathrm{T}} \tilde{P}_{12} B_c V_2 \\
&\qquad - B^{\mathrm{T}} \tilde{Z}_{11}^{\mathrm{T}} C^{\mathrm{T}} + R_2 D_c V_2). && \text{(D.35)}
\end{aligned}$$

References

1. P. T. Kabamba, "Balanced Forms: Cononicity and Parameterization," *IEEE Trans. Autom. Contr.*, Vol. AC-30, pp. 1106–1109, Nov. 1985.

Appendix E: The Homotopy Map and It's Jacobian for the Normal Riccati Basis

As stated in the previous Appendix the objective is to find θ satisfying

$$f(\theta) = 0 \tag{E.1}$$

where

$$f(\theta) \triangleq \nabla \bar{J}(\theta)^{\mathrm{T}} \tag{E.2}$$

and $\bar{J}(\theta)$ denotes the restricted cost functional for the input-normal Riccati basis. In this section we define a homotopy map to accomplish this task and show how to efficiently compute it's Jacobian.

Definition of the homotopy map $h(\theta, \lambda)$

To define the homotopy map we assume that the plant matrices (A, B, C, D), the cost weighting matrix (R_1, R_2, R_{12}) and the disturbance matrices (V_1, V_2, V_{12}) are functions of the Homotopy parameter $\lambda \in [0, 1]$. In particular, it is assumed that

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} + \left(\begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} \right) \tag{E.3}$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^{\mathrm{T}}(\lambda) & R_2(\lambda) \end{bmatrix} = \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} + \lambda \left(\begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^{\mathrm{T}} & R_{2,f} \end{bmatrix} - \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} \right) \tag{E.4}$$

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^{\mathrm{T}}(\lambda) & V_2(\lambda) \end{bmatrix} = \begin{bmatrix} V_{1,0} & V_{12,0} \\ R_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix} + \lambda \left(\begin{bmatrix} V_{1,f} & V_{12,f} \\ R_{12,f}^{\mathrm{T}} & V_{2,f} \end{bmatrix} - \begin{bmatrix} V_{1,0} & V_{12,0} \\ R_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix} \right). \tag{E.5}$$

Note that (E.3)-(E.5) imply that $A(0) = A_0$ and $A(1) = A_f$, $B(0) = B_0$ and $B(1) = B_f$, etc ... and it is understood that $A_f, B_f, \ldots$ were referred to in the previous sections simply as $A, B, \ldots$. The change in notation is simply for convenience.

The homotopy map $h(\theta, \lambda)$ is defined by

$$h(\theta, \lambda) = \begin{bmatrix} \mathrm{vec}\big(H_{B_c}(\theta, \lambda)\big) \\ \mathrm{vec}\big(H_{C_c}(\theta, \lambda)\big) \\ \mathrm{vec}\big(H_{D_c}(\theta, \lambda)\big) \end{bmatrix} \tag{E.6}$$

where

$$\begin{aligned} H_{B_c}(\theta, \lambda) = 2(&\tilde{P}_{12}^{\mathrm{T}} V_{12} - \tilde{P}_{12}^{\mathrm{T}} B D_c V_2 + \tilde{P}_{22} B_c V_2 \\ &+ \tilde{Z}_{12}^{\mathrm{T}} C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} D^{\mathrm{T}} + (Y - \Omega Z \Omega) B_c) \end{aligned} \tag{E.7}$$

$$\begin{aligned} H_{C_c}(\theta, \lambda) = 2(&-R_{12}^{\mathrm{T}} \tilde{Q}_{12} + R_2 D_c C \tilde{Q}_{12} + R_2 C_c \tilde{Q}_{22} \\ &- B^{\mathrm{T}} \tilde{Z}_{21}^{\mathrm{T}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{\mathrm{T}} + C_c(Z - Y)) \end{aligned} \tag{E.8}$$

$$\begin{aligned} H_{D_c}(\theta, \lambda) = 2(&-R_{12}^{\mathrm{T}} \tilde{Q}_{11} C^{\mathrm{T}} + R_2 D_c C \tilde{Q}_{11} C^{\mathrm{T}} + R_2 C_c \tilde{Q}_{12}^{\mathrm{T}} C^{\mathrm{T}} \\ &- B^{\mathrm{T}} \tilde{P}_{11} V_{12} + B^{\mathrm{T}} \tilde{P}_{11} B D_c V_2 - B^{\mathrm{T}} \tilde{P}_{12} B_c V_2 \\ &- B^{\mathrm{T}} \tilde{Z}_{11}^{\mathrm{T}} C^{\mathrm{T}} + R_2 D_c V_2). \end{aligned} \tag{E.9}$$

Here, it is assumed that $\tilde{P}, \tilde{Q}$ and $\tilde{Z}$ satisfy

$$\tilde{P} = \tilde{A}^{\mathrm{T}} \tilde{P} \tilde{A} + \tilde{R} \tag{E.10}$$

$$\tilde{Q} = \tilde{A} \tilde{Q} \tilde{A}^{\mathrm{T}} + \tilde{V}. \tag{E.11}$$

$$\tilde{Z} = \tilde{Q} \tilde{A}^{\mathrm{T}} \tilde{P}. \tag{E.12}$$

In addition, Y and Z are given by

$$Y = -(\frac{1}{2}\mathcal{L}_{A_c} + \Omega Z) \tag{E.13}$$

$$Z = Z_o + Z_d \tag{E.14}$$

where

$$Z_o \triangleq \frac{1}{2}(\mathcal{L}_{A_c} - \mathcal{L}_{A_c}^{\mathrm{T}}) * H \tag{E.15}$$

$$Z_d \triangleq -\operatorname{diag}(F)^{-1} * (F Z_0) \tag{E.16}$$

$$\mathcal{L}_{A_c} \triangleq Q \tilde{Z}_{22}^{\mathrm{T}} \tag{E.17}$$

$$F \triangleq A_c - B_c B_c^{\mathrm{T}} \Omega \tag{E.18}$$

$$\Omega \triangleq \operatorname{diag}(C_c^{\mathrm{T}} C_c) \operatorname{diag}(B_c B_c^{\mathrm{T}})^{-1} \tag{E.19}$$

$$H \triangleq (N_{n_c} - I_{n_c})./[\bar{\Omega}^{\mathrm{T}} - \bar{\Omega} + I_{n_c}]. \tag{E.20}$$

Note that (E.13) and (E.14) are equivalent to

$$0 = \mathcal{L}_{A_c} + 2(Y + \Omega Z) \tag{E.21}$$

$$0 = [F\ Z]_{ii}, \quad i = 1, 2, \ldots, n_c. \tag{E.22}$$

Also, note that it follows from the results of the previous section that

$$h(\theta, 1) = f(\theta)(\triangleq \nabla \bar{J}(\theta)^{\mathrm{T}}). \tag{E.23}$$

Also, note that $h(\theta, \lambda)$ is the transposed gradient for the optimization problem with parameters $(A(\lambda), \ldots, R_1(\lambda), \ldots, V_{12}(\lambda))$.

We now consider that computation of $\nabla h(\theta, \lambda)^{\mathrm{T}}$, the Jacobian of $h(\theta, \lambda)$. Note that

$$\nabla h(\theta, \lambda)^{\mathrm{T}} = \begin{bmatrix} \frac{\partial h}{\partial \theta} & \frac{\partial h}{\partial \lambda} \end{bmatrix}. \tag{E.24}$$

Recalling that θ is defined by (D.1), such that for some integers k and ℓ, θ_j is given by

$$\theta_j = b_{c,k_\ell},\ \theta_j = c_{c,k_\ell},\ \text{or}\ \theta_j = d_{c,k_\ell}. \tag{E.25}$$

It follows from (E.6) that $\frac{\partial h}{\partial \theta}$ is of the form

$$\frac{\partial h}{\partial \theta} = \begin{bmatrix} \ldots \mathrm{vec}(\frac{\partial}{\partial b_{c,k_\ell}} H_{B_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial c_{c,k_\ell}} H_{B_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial d_{c,k_\ell}} H_{B_c}) \\ \ldots \mathrm{vec}(\frac{\partial}{\partial b_{c,k_\ell}} H_{C_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial c_{c,k_\ell}} H_{C_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial d_{c,k_\ell}} H_{C_c}) \\ \ldots \mathrm{vec}(\frac{\partial}{\partial b_{c,k_\ell}} H_{D_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial c_{c,k_\ell}} H_{D_c}) \ldots \mathrm{vec}(\frac{\partial}{\partial d_{c,k_\ell}} H_{D_c}) \end{bmatrix}. \tag{E.26}$$

and $\frac{\partial h}{\partial \lambda}$ can be expressed as

$$\frac{\partial h}{\partial \lambda} = \begin{bmatrix} \mathrm{vec}(\frac{\partial}{\partial \lambda} H_{B_c}) \\ \mathrm{vec}(\frac{\partial}{\partial \lambda} H_{C_c}) \\ \mathrm{vec}(\frac{\partial}{\partial \lambda} H_{D_c}) \end{bmatrix}^{\mathrm{T}}. \tag{E.27}$$

Below, we develop explicit expression for the derivative terms appearing on the right hand sides of (E.26) and (E.27). We use the notation

$$\begin{aligned} M^{(j)} &\triangleq \frac{\partial M}{\partial \theta_j} \\ \dot{M} &\triangleq \frac{\partial M}{\partial \lambda}. \end{aligned} \tag{E.28}$$

Differentiating (E.10)–(E.12) with respect to θ_j yields

$$\tilde{P}^{(j)} = \tilde{A}^{\mathrm{T}} \tilde{P}^{(j)} \tilde{A} + (\tilde{A}^{(j)^{\mathrm{T}}} \tilde{P} \tilde{A} + \tilde{A}^{\mathrm{T}} \tilde{P} \tilde{A}^{(j)} + \tilde{R}^{(j)}) \tag{E.29}$$

$$\tilde{Q}^{(j)} = \tilde{A} \tilde{Q}^{(j)} \tilde{A}^{\mathrm{T}} + (\tilde{A}^{(j)} \tilde{Q} \tilde{A}^{\mathrm{T}} + \tilde{A} \tilde{Q} \tilde{A}^{(j)^{\mathrm{T}}} + \tilde{V}^{(j)}) \tag{E.30}$$

$$\tilde{Z}^{(j)} = \tilde{Q}^{(j)} \tilde{A}^{\mathrm{T}} \tilde{P} + \tilde{Q} \tilde{A}^{(j)^{\mathrm{T}}} \tilde{P} + \tilde{Q} \tilde{A}^{\mathrm{T}} \tilde{P}^{(j)} \tag{E.31}$$

where expressions for the derivatives $\tilde{A}^{(j)}, \tilde{R}^{(j)}$ and $\tilde{V}^{(j)}$ are given by (F.20)-(F.28) of Appendix F. Differentiating (E.21), (E.22), (E.17) and (E.18) with respect to θ_j yields

$$0 = [\mathcal{L}_{A_c}^{(j)} + 2 + \Omega^{(j)} Z] + 2(Y^{(j)} + \Omega Z^{(j)}) \tag{E.32}$$

$$-[F^{(j)} Z]_{ii} = [F Z^{(j)}]_{ii}, \quad i = 1, 2, \ldots, n_c \tag{E.33}$$

where

$$\mathcal{L}_{A_c}^{(j)} \triangleq Q \tilde{Z}_{22}^{(j)^{\mathrm{T}}} \tag{E.34}$$

$$F^{(j)} = A_c^{(j)} - (B_c B_c^{\mathrm{T}})^{(j)} \Omega - B_c B_c^{\mathrm{T}} \Omega^{(j)} \tag{E.35}$$

and the derivatives $A_c^{(j)}, (B_c B_c^{\mathrm{T}})^{(j)}$ and $\Omega^{(j)}$ may be computed using the results of Appendix G. Note that if we define

$$\mathcal{L}' \triangleq \mathcal{L}_{A_c}^{(j)} + 2\Omega^{(j)} Z \tag{E.36}$$

then (E.32) and (E.33) are equivalent to

$$Y^{(j)} = -(\frac{1}{2}\mathcal{L}' + \Omega Z^{(j)}) \tag{E.37}$$

$$Z^{(j)} = Z_o^{(j)} + Z_d^{(j)} \tag{E.38}$$

where

$$Z_0^{(j)} \triangleq \frac{1}{2}(\mathcal{L}' - \mathcal{L}'^{\mathrm{T}}) * H \tag{E.39}$$

$$Z_d^{(j)} \triangleq -\operatorname{diag}(F)^{-1} * (F Z_0^{(j)} + F^{(j)} Z). \tag{E.40}$$

Differentiating (E.10)–(E.12) with respect to λ yields

$$\dot{\bar{P}} = \bar{A}^{\mathrm{T}} \dot{\bar{P}} \bar{A} + (\dot{\bar{A}}^{\mathrm{T}} \bar{P} \bar{A} + \bar{A}^{\mathrm{T}} \bar{P} \dot{\bar{A}} + \dot{\bar{R}}) \tag{E.41}$$

$$\dot{\bar{Q}} = \bar{A} \dot{\bar{Q}} \bar{A}^{\mathrm{T}} + (\dot{\bar{A}} \bar{Q} \bar{A}^{\mathrm{T}} + \bar{A} \bar{Q} \dot{\bar{A}} + \dot{\bar{V}}) \tag{E.42}$$

$$\dot{\bar{Z}} = \dot{\bar{Q}} \bar{A}^{\mathrm{T}} \bar{P} + \bar{Q} \dot{\bar{A}}^{\mathrm{T}} \bar{P} + \bar{Q} \bar{A} \dot{\bar{P}}. \tag{E.43}$$

where expressions for $\dot{\bar{A}}, \dot{\bar{R}}$ and $\dot{\bar{V}}$ are given by (F.29)–(F.33) of Appendix F. Differentiating (E.21), (E.22) and (E.17) with respect to λ yields

$$0 = \dot{\mathcal{L}}_{A_c} + 2(\dot{Y} + \Omega \dot{Z}) \tag{E.44}$$

$$0 = [F\dot{Z}]_{ii}, \quad i = 1, 2, \ldots, n_c \tag{E.45}$$

where

$$\dot{\mathcal{L}}_{A_c} = Q \dot{\bar{Z}}_{22}^{\mathrm{T}}. \tag{E.46}$$

Note that (E.44) and (E.45) are equivalent to

$$\dot{Y} = -(\frac{1}{2}\dot{\mathcal{L}}_{A_c} + \Omega \dot{Z}) \tag{E.47}$$

$$\dot{Z} = \dot{Z}_o + \dot{Z}_d \tag{E.48}$$

where

$$\dot{Z}_o = \frac{1}{2}(\dot{\mathcal{L}}_{A_c} - \dot{\mathcal{L}}_{A_c}^{\mathrm{T}}) * H \tag{E.49}$$

$$\dot{Z}_d = -\operatorname{diag}(F)^{-1} * (F \dot{Z}_o). \tag{E.50}$$

Before presenting the desired derivative expressions we define

$$\begin{aligned} H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) &= 2(\tilde{P}_{12}^{(j)^{\mathrm{T}}} V_{12} - \tilde{P}_{12}^{(j)^{\mathrm{T}}} B D_c V_2 + \tilde{P}_{22}^{(j)} B_c V_2 \\ &\quad + \tilde{Z}_{12}^{(j)^{\mathrm{T}}} C^{\mathrm{T}} - \tilde{Z}_{22}^{(j)^{\mathrm{T}}} C_c^{\mathrm{T}} D^{\mathrm{T}} + (Y^{(j)} - \Omega Z^{(j)} \Omega) B_c) \qquad \text{(E.51)} \\ H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{12}^{(j)} + R_2 D_c C \tilde{Q}_{12}^{(j)} + R_2 C_c \tilde{Q}_{22}^{(j)} \\ &\quad - B^{\mathrm{T}} \tilde{Z}_{21}^{(j)^{\mathrm{T}}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{(j)^{\mathrm{T}}} + C_c(Z^{(j)} - Y^{(j)})) \qquad \text{(E.52)} \\ H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) &= 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{11}^{(j)} C^{\mathrm{T}} + R_2 D_c C \tilde{Q}_{11}^{(j)} C^{\mathrm{T}} + R_2 C_c \tilde{Q}_{12}^{(j)^{\mathrm{T}}} C^{\mathrm{T}}) \\ &\quad - B^{\mathrm{T}} \tilde{Z}_{11}^{(j)^{\mathrm{T}}} C^{\mathrm{T}}) \qquad \text{(E.53)} \end{aligned}$$

Notice that the right hand sides of (E.51)–(E.53) are essentially identical in form to the right hand sides of (E.7)–(E.9). The difference is that $\tilde{P}, \tilde{Q}, \tilde{Z}, Y$ and Z and have been replaced by $\tilde{P}^{(j)}, \tilde{Q}^{(j)}$, $\tilde{Z}^{(j)}, Y^{(j)}$ and $Z^{(j)}$ and the last term $(2R_2 D_c V_2)$ in (E.9) has no counterpart in (E.53).

Derivatives with Respect to $B_{c,k\ell}$

Differentiating (E.7)–(E.9) with respect to $b_{c,k\ell}(= \theta_j)$ gives the following.

$$\begin{aligned} \frac{\partial H_{B_c}}{\partial B_{c,k\ell}} &= H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) + 2[\tilde{P}_{22} E_{n_c \times n_y}^{(k,\ell)} V_2 + (Y - \Omega Z \Omega) E_{k\ell}] \\ &\quad - 2(\Omega^{(j)} Z \Omega + \Omega Z \Omega^{(j)}) B_c \end{aligned} \qquad \text{(E.54)}$$

$$\frac{\partial H_{C_c}}{\partial B_{c,k\ell}} = H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) - 2D^{\mathrm{T}} E_{n_y \times n_c}^{(\ell,k)} \tilde{Z}_{22}^{\mathrm{T}} \qquad \text{(E.55)}$$

$$\frac{\partial H_{D_c}}{\partial B_{c,k\ell}} = H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) - 2B^{\mathrm{T}} \tilde{P}_{12} E_{n_c \times n_y}^{(k,\ell)} V_2. \qquad \text{(E.56)}$$

Derivatives with Respect to $C_{c,k\ell}$

Differentiating (E.7)–(E.9) with respect to $\theta_j(= c_{c,k\ell})$ gives the following.

$$\frac{\partial H_{B_c}}{\partial C_{c,k\ell}} = H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) - 2\tilde{Z}_{22}^{\mathrm{T}} E_{n_c \times n_u}^{(\ell,k)} D^{\mathrm{T}} - 2(\Omega^{(j)} Z \Omega + \Omega Z \Omega^{(j)}) B_c \qquad \text{(E.57)}$$

$$\frac{\partial H_{C_c}}{\partial C_{c,k\ell}} = H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) + 2R_2 E_{n_u \times n_c}^{(k,\ell)} \tilde{Q}_{22} + 2E_{n_u \times n_c}^{(k,\ell)} (Z - Y) \qquad \text{(E.58)}$$

$$\frac{\partial H_{D_c}}{\partial C_{c,k\ell}} = H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) + 2R_2 E_{n_u \times n_c}^{(k,\ell)} \tilde{Q}_{12}^{\mathrm{T}} C^{\mathrm{T}}. \qquad \text{(E.59)}$$

Derivatives with respect to $d_{c,k\ell}$

Differentiating (E.7)–(E.9) with respect to $d_{c,k\ell}$ gives the following.

$$\frac{\partial H_{B_c}}{\partial D_{c,k\ell}} = H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) - 2\tilde{P}_{12}^{\mathrm{T}} B E^{(k,\ell)}_{n_u \times n_y} V_2 \tag{E.60}$$

$$\frac{\partial H_{C_c}}{\partial D_{c,k\ell}} = H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}, Y^{(j)}, Z^{(j)}) + 2R_2 E^{(k,\ell)}_{n_u \times n_y} C \tilde{Q}_{12} \tag{E.61}$$

$$\begin{aligned}\frac{\partial H_{D_c}}{\partial D_{c,k\ell}} = H'_{D_c}(\tilde{P}^{(j)}, \tilde{Q}^{(j)}, \tilde{Z}^{(j)}) &+ 2(R_2 E^{(k,\ell)}_{n_u \times n_y} C \tilde{Q}_{11} C^{\mathrm{T}} + B^{\mathrm{T}} \tilde{P}_{11} B E^{(k,\ell)}_{n_u \times n_y} V_2 \\ &+ R_2 E^{(k,\ell)}_{n_c \times n_y} V_2).\end{aligned} \tag{E.62}$$

Derivatives with Respect to λ

Differentiating (E.7)–(E.9) with respect to λ gives

$$\begin{aligned}\frac{\partial H_{B_c}}{\partial \lambda} = H'_{B_c}&(\dot{\tilde{P}}, \dot{\tilde{Z}}, \dot{Y}, \dot{Z}) \\ &+ 2(\tilde{P}_{21}^{\mathrm{T}} \dot{V}_{12} - \tilde{P}_{12}^{\mathrm{T}} \dot{B} D_c V_2 - \tilde{P}_{12}^{\mathrm{T}} B D_c \dot{V}_2 + \tilde{P}_{22} B_c \dot{V}_2) \\ &+ \tilde{Z}_{12}^{\mathrm{T}} \dot{C}^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} \dot{D}^{\mathrm{T}})\end{aligned} \tag{E.63}$$

$$\begin{aligned}\frac{\partial H_{C_c}}{\partial \lambda} = H'_{C_c}&(\dot{\tilde{Q}}, \dot{\tilde{Z}}, \dot{Y}, \dot{Z}) \\ &+ 2(-\dot{R}_{12}^{\mathrm{T}} \tilde{Q}_{12} + \dot{R}_2 D_c C \tilde{Q}_{12} + R_2 D_c \dot{C} \tilde{Q}_{12} + \dot{R}_2 C_c \tilde{Q}_{22} \\ &+ \dot{B}^{\mathrm{T}} \tilde{Z}_{21}^{\mathrm{T}} - \dot{D}^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{\mathrm{T}})\end{aligned} \tag{E.64}$$

$$\begin{aligned}\frac{\partial H_{D_c}}{\partial \lambda} = H'_{D_c}&(\dot{\tilde{P}}, \dot{\tilde{Q}}, \dot{\tilde{Z}}) \\ &+ 2(-\dot{R}_{12}^{\mathrm{T}} \tilde{Q}_{11} C^{\mathrm{T}} - R_{12}^{\mathrm{T}} \tilde{Q}_{11} \dot{C}^{\mathrm{T}} \\ &\quad + \dot{R}_2 D_c C \tilde{Q}_{11} C^{\mathrm{T}} + R_2 D_c \dot{C} \tilde{Q}_{11} C^{\mathrm{T}} + R_2 D_c C \tilde{Q}_{11} \dot{C}^{\mathrm{T}} \\ &\qquad + \dot{R}_2 C_c C \tilde{Q}_{12}^{\mathrm{T}} C^{\mathrm{T}} + R_2 C_c \tilde{Q}_{12}^{\mathrm{T}} \dot{C}^{\mathrm{T}}) \\ &+ 2(-\dot{B}^{\mathrm{T}} \tilde{P}_{11} V_{12} - B^{\mathrm{T}} \tilde{P}_{11} \dot{V}_{12} \\ &\quad + \dot{B}^{\mathrm{T}} \tilde{P}_{11} B D_c V_2 + B^{\mathrm{T}} \tilde{P}_{11} \dot{B} D_c V_2 + B^{\mathrm{T}} \tilde{P}_{11} B D_c \dot{V}_2 \\ &\qquad - \dot{B}^{\mathrm{T}} \tilde{P}_{12} B_c V_2 - B^{\mathrm{T}} \tilde{P}_{12} B_c \dot{V}_2) \\ &- 2\dot{B}^{\mathrm{T}} \tilde{Z}_{11}^{\mathrm{T}} C^{\mathrm{T}} - B^{\mathrm{T}} \tilde{Z}_{11} \dot{C}^{\mathrm{T}} \\ &+ 2(\dot{R}_2 D_c V_2 + R_2 D_c \dot{V}_2)\end{aligned} \tag{E.65}$$

where

$$\begin{bmatrix} \dot{A} & \dot{B} \\ \dot{C} & \dot{D} \end{bmatrix} = \begin{bmatrix} A_f - A_0 & B_f - B_0 \\ C_f - C_0 & C_f - C_0 \end{bmatrix} \tag{E.66}$$

$$\begin{bmatrix} \dot{R}_1 & \dot{R}_{12} \\ \dot{R}_{12}^{\mathrm{T}} & \dot{R}_2 \end{bmatrix} = \begin{bmatrix} R_{1,f} - R_{1,0} & R_{12,f} - R_{12,0} \\ R_{12,f}^{\mathrm{T}} - R_{12,0}^{\mathrm{T}} & R_{2,f} - R_{2,0} \end{bmatrix} \tag{E.67}$$

$$\begin{bmatrix} \dot{V}_1 & \dot{V}_{12} \\ \dot{V}_{12}^{\mathrm{T}} & \dot{V}_2 \end{bmatrix} = \begin{bmatrix} V_{1,f} - V_{1,0} & V_{12,f} - V_{12,0} \\ V_{12,f}^{\mathrm{T}} - V_{12,0}^{\mathrm{T}} & V_{2,f} - V_{2,0} \end{bmatrix}. \tag{E.68}$$

Appendix F: Closed-Loop Matrix Derivatives for the Input Normalized Riccati Basis

In this appendix we assume that (A_c, B_c, C_c, D_c) is restricted to the input-Normal Riccati basis and present explicit expressions for the derivatives $\frac{\partial \tilde{A}}{\partial \theta_j}$, $\frac{\partial \tilde{R}}{\partial \theta_j}$, $\frac{\partial \tilde{V}}{\partial \theta_j}$, $\frac{\partial \tilde{A}}{\partial \lambda}$, $\frac{\partial \tilde{R}}{\partial \lambda}$, and $\frac{\partial \tilde{V}}{\partial \lambda}$ where

$$\theta = \begin{bmatrix} \text{vec}(B_c) \\ \text{vec}(C_c) \\ \text{vec}(D_c) \end{bmatrix} \tag{F.1}$$

$$\tilde{A} = \begin{bmatrix} A - BD_cC & -BC_c \\ B_cC & A_c - B_cDC_c \end{bmatrix}, \tag{F.2}$$

$$\tilde{R} = \begin{bmatrix} \tilde{R}_{11} & \tilde{R}_{12} \\ \tilde{R}_{12}^{\mathrm{T}} & \tilde{R}_{22} \end{bmatrix}, \tag{F.3}$$

where

$$\tilde{R}_{11} = R_1 - C^{\mathrm{T}} D_c^{\mathrm{T}} R_{12}^{\mathrm{T}} - R_{12} D_c C + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c C \tag{F.4a}$$

$$\tilde{R}_{12} = -R_{12} C_c + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 C_c \tag{F.4b}$$

$$\tilde{R}_{22} = C_c^{\mathrm{T}} R_2 C_c, \tag{F.4c}$$

and

$$\tilde{V} = \begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{12}^{\mathrm{T}} & \tilde{V}_{22} \end{bmatrix} \tag{F.5}$$

where

$$\tilde{V}_{11} = V_1 - BD_c V_{12}^{\mathrm{T}} - V_{12} D_c^{\mathrm{T}} B^{\mathrm{T}} + BD_c V_2 D_c^{\mathrm{T}} B^{\mathrm{T}} \tag{F.6a}$$

$$\tilde{V}_{12} = V_{12} B_c^{\mathrm{T}} - BD_c V_2 B_c^{\mathrm{T}} \tag{F.6b}$$

$$\tilde{V}_{22} = B_c^{\mathrm{T}} V_2 B_c^{\mathrm{T}}. \tag{F.6c}$$

It is assumed that the plant matrices (A, B, C, D), the cost weighting matrices (R_1, R_{12}, R_2) and the disturbance matrices (V_1, V_{12}, V_2) are the following functions of λ.

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} + \lambda\left(\begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} \right) \tag{F.7}$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^{\mathrm{T}}(\lambda) & R_2(\lambda) \end{bmatrix} = \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} + \lambda\left(\begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^{\mathrm{T}} & R_{2,f} \end{bmatrix} - \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} \right) \tag{F.8}$$

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^{\mathrm{T}}(\lambda) & V_2(\lambda) \end{bmatrix} = \begin{bmatrix} V_{1,0} & V_{12,0} \\ V_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix} + \lambda\left(\begin{bmatrix} V_{1,f} & V_{12,f} \\ V_{12,f}^{\mathrm{T}} & V_{2,f} \end{bmatrix} - \begin{bmatrix} V_{1,0} & V_{12,0} \\ V_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix} \right). \tag{F.9}$$

Below, we use the notation

$$\dot{M} \triangleq \frac{\partial M}{\partial \lambda}. \tag{F.10}$$

Note that

$$\begin{bmatrix} \dot{A} & \dot{B} \\ \dot{C} & \dot{D} \end{bmatrix} = \begin{bmatrix} A_f - A_0 & B_f - B_0 \\ C_f - C_0 & D_f - D_0 \end{bmatrix} \tag{F.11}$$

$$\begin{bmatrix} \dot{R}_1 & \dot{R}_{12} \\ \\ \dot{R}_{12}^{\mathrm{T}} & \dot{R}_2 \end{bmatrix} = \begin{bmatrix} R_{1,f} - R_{1,0} & R_{12,f} - R_{12,0} \\ \\ R_{12,f}^{\mathrm{T}} - R_{12,0}^{\mathrm{T}} & R_{2,f} - R_{2,0} \end{bmatrix} \tag{F.12}$$

$$\begin{bmatrix} \dot{V}_1 & \dot{V}_{12} \\ \\ \dot{V}_{12}^{\mathrm{T}} & \dot{V}_2 \end{bmatrix} = \begin{bmatrix} V_{1,f} - V_{1,0} & V_{12,f} - V_{12,0} \\ \\ V_{12,f}^{\mathrm{T}} - V_{12,0}^{\mathrm{T}} & V_{2,f} - V_{2,0} \end{bmatrix}. \tag{F.13}$$

The derivations of the expression for $\frac{\partial \tilde{A}}{\partial \theta_j}$, $\frac{\partial \tilde{R}}{\partial \theta_j}$ and $\frac{\partial \tilde{V}}{\partial \theta_j}$ are primarily based on the application of the following derivative formulas. It is assumed that X is an $n \times m$ matrix.

Derivative Formulas

$$\frac{d}{dx_{ij}} XA = [A(j,:)]_{\mathrm{row}-i} \tag{F.14}$$

$$\frac{d}{dx_{ij}} AX = [A(:,i)]_{\mathrm{col}-j} \tag{F.15}$$

$$\frac{d}{dx_{ij}} X^{\mathrm{T}} A = [A(i,:)]_{\mathrm{row}-j} \tag{F.16}$$

$$\frac{d}{dx_{ij}} AX^{\mathrm{T}} = [A(:,j)]_{\mathrm{col}-i} \tag{F.17}$$

$$\frac{d}{dx_{ij}} AXB = A(:,i)B(j,:) \tag{F.18}$$

$$\frac{d}{dx_{ij}} AX^{\mathrm{T}}B = A(:,j)B(i,:) \tag{F.19}$$

$\frac{\partial \tilde{A}}{\partial \theta_j}$

$$\frac{\partial \tilde{A}}{\partial b_{c,k\ell}} = \begin{bmatrix} 0 & 0 \\ \\ [C(\ell,:)]_{\mathrm{row}-k} & \frac{\partial A_c}{\partial b_{c_{k,\ell}}} - [\underline{DC_c}(\ell,:)]_{\mathrm{row}-k} \end{bmatrix} \tag{F.20}$$

$$\frac{\partial \tilde{A}}{\partial c_{c,k\ell}} = \begin{bmatrix} 0 & -[B(:,k)]_{\mathrm{col}-\ell} \\ \\ 0 & \frac{\partial A_c}{\partial c_{c_{k,\ell}}} - [\underline{B_c D}(:,k)]_{\mathrm{col}-\ell} \end{bmatrix} \tag{F.21}$$

$$\frac{\partial \tilde{A}}{\partial d_{c,k\ell}} = \begin{bmatrix} -B(:,k)C(\ell,:) & 0 \\ 0 & 0 \end{bmatrix} \tag{F.22}$$

where $\frac{\partial A_c}{\partial b_{c,k\ell}}$ and $\frac{\partial A_c}{\partial c_{c,k\ell}}$ are given respectively by (D.2.36) and (D.2.37) of Appendix D.

$$\underline{\frac{\partial \tilde{R}}{\partial \theta_j}}$$

$$\frac{\partial \tilde{R}}{\partial b_{c,k\ell}} = 0 \tag{F.23}$$

$$\frac{\partial \tilde{R}}{\partial c_{c,k\ell}} = \begin{bmatrix} 0 & [-R_{12}(:,k) + \underline{C^{\mathrm{T}} D_c^{\mathrm{T}} R_2}(:,k)]_{\mathrm{col}-\ell} \\ \mathrm{SYM} & [\underline{C_c^{\mathrm{T}} R_2}(:,k)]_{\mathrm{col}-\ell} + [\underline{R_2 C_c}(k,:)]_{\mathrm{row}-\ell} \end{bmatrix} \tag{F.24}$$

$$\frac{\partial \tilde{R}}{\partial d_{c,k\ell}} = \begin{bmatrix} \begin{array}{c} C^{\mathrm{T}}(:,\ell)[\underline{R_2 D_c C}(k,:) - R_{12}^{\mathrm{T}}(k,:)] \\ +[\underline{C^{\mathrm{T}} D_c^{\mathrm{T}} R_2}(:,k) - R_{12}(:,k)]C(\ell,:) \end{array} & C^{\mathrm{T}}(:,\ell)\underline{R_2 C_c}(k,:) \\ \mathrm{SYM} & 0 \end{bmatrix} \tag{F.25}$$

$$\underline{\frac{\partial \tilde{V}}{\partial \theta_j}}$$

$$\frac{\partial \tilde{V}}{\partial b_{c,k\ell}} = \begin{bmatrix} 0 & [V_{12}(:,\ell) - \underline{B D_c V_2}(:,\ell)]_{\mathrm{col}-k} \\ \mathrm{SYM} & [\underline{V_2 B_c^{\mathrm{T}}}(\ell,:)]_{\mathrm{row}-k} + [\underline{B_c V_2}(:,\ell)]_{\mathrm{col}-k} \end{bmatrix} \tag{F.26}$$

$$\frac{\partial \tilde{V}}{\partial c_{c,k\ell}} = 0 \tag{F.27}$$

$$\frac{\partial \tilde{V}}{\partial d_{c,k\ell}} = \begin{bmatrix} \begin{array}{c} B(:,k)[\underline{V_2 D_c^{\mathrm{T}} B^{\mathrm{T}}}(\ell,:) - V_{12}^{\mathrm{T}}(\ell,:)] \\ +[\underline{B D_c V_2}(:,\ell) - V_{12}(:,\ell)]B^{\mathrm{T}}(k,:) \end{array} & -B(:,k)\underline{V_2 B_c^{\mathrm{T}}}(\ell,:) \\ \mathrm{SYM} & 0 \end{bmatrix} \tag{F.28}$$

$$\underline{\dot{\tilde{A}} \triangleq \frac{\partial \tilde{A}}{\partial \lambda}}$$

$$\dot{\tilde{A}} = \begin{bmatrix} \dot{A} - \dot{B} D_c C - B D_c \dot{C} & -\dot{B} C_c \\ B_c \dot{C} & -B_c \dot{D} C_c \end{bmatrix} \tag{F.29}$$

$$\underline{\dot{\tilde{R}} \triangleq \frac{\partial \tilde{R}}{\partial \lambda}}$$

$$\dot{\tilde{R}} = \begin{bmatrix} \dot{\tilde{R}}_{11} & \dot{\tilde{R}}_{12} \\ \dot{\tilde{R}}_{12}^{\mathrm{T}} & \dot{\tilde{R}}_{22} \end{bmatrix} \tag{F.30}$$

where

$$\begin{aligned}\dot{\tilde{R}}_{11} &= \dot{R}_1 - \dot{C}^{\mathrm{T}} D_c^{\mathrm{T}} R_{12}^{\mathrm{T}} - C^{\mathrm{T}} D_c^{\mathrm{T}} \dot{R}_{12}^{\mathrm{T}} - \dot{R}_{12} D_c C - R_{12} D_c \dot{C} \\ &\quad + \dot{C}^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c C + C^{\mathrm{T}} D_c^{\mathrm{T}} \dot{R}_2 D_c C + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2 D_c \dot{C} \qquad \text{(F.31}a\text{)}\end{aligned}$$

$$\dot{\tilde{R}}_{12} = -\dot{R}_{12} C_c + \dot{C}^{\mathrm{T}} D_c^{\mathrm{T}} R_2 C_c + C^{\mathrm{T}} D_c^{\mathrm{T}} \dot{R}_2 C_c \qquad \text{(F.31}b\text{)}$$

$$\dot{\tilde{R}}_{22} = C_c^{\mathrm{T}} \dot{R}_2 C_c. \qquad \text{(F.31}c\text{)}$$

$$\underline{\dot{\tilde{V}} \triangleq \frac{\partial \tilde{V}}{\partial \lambda}}$$

$$\dot{\tilde{V}} = \begin{bmatrix} \dot{\tilde{V}}_{11} & \dot{\tilde{V}}_{12} \\ \dot{\tilde{V}}_{12}^{\mathrm{T}} & \dot{\tilde{V}}_{22} \end{bmatrix} \qquad \text{(F.32)}$$

where

$$\begin{aligned}\dot{\tilde{V}}_{11} &= \dot{V}_1 - \dot{B} D_C V_{12}^{\mathrm{T}} - B D_c \dot{V}_{12}^{\mathrm{T}} - \dot{V}_{12} D_c^{\mathrm{T}} B^{\mathrm{T}} - V_{12} D_c^{\mathrm{T}} \dot{B}^{\mathrm{T}} \\ &\quad + \dot{B} D_c V_2 D_c B^{\mathrm{T}} + B D_c \dot{V}_2 D_c B^{\mathrm{T}} + B D_c V_2 D_c \dot{B}^{\mathrm{T}} \qquad \text{(F.33}a\text{)}\end{aligned}$$

$$\dot{\tilde{V}}_{12} = -\dot{V}_{12} B_c^{\mathrm{T}} - \dot{B} D_c^{\mathrm{T}} V_2 B_c^{\mathrm{T}} - B D_c \dot{V}_2 B_c^{\mathrm{T}} \qquad \text{(F.33}b\text{)}$$

$$\dot{\tilde{V}}_{22} = B_c \dot{V}_2 B_c^{\mathrm{T}}. \qquad \text{(F.33}c\text{)}$$

Appendix G: Derivation of $\frac{\partial A_c}{\partial b_{c,k\ell}}$ and $\frac{\partial A_c}{\partial c_{c,k\ell}}$ for the Input-Normal Riccati Basis

G.1 Problem Statement

In this Appendix we assume that the controller triple (A_c, B_c, C_c) is in the input normalized Riccati basis described in Section 3, such that

$$A_c = \frac{1}{2}[C_c^{\mathrm{T}} C_c - B_c B_c^{\mathrm{T}}] * I_{n_c} + [(C_c^{\mathrm{T}} C_c) * (N_{n_c} + \bar{\Omega}^{\mathrm{T}}) - (N_{n_c} + \bar{\Omega}) * (B_c B_c^{\mathrm{T}}) * \bar{\Omega}^{\mathrm{T}}] * H \quad \text{(G.1.1)}$$

where

$$\Omega \triangleq \operatorname{diag}(C_c^{\mathrm{T}} C_c) \operatorname{diag}(B_c B_c^{\mathrm{T}})^{-1} \quad \text{(G.1.2)}$$

and

$$H = (N_{n_c} - I_{n_c}) / (\bar{\Omega}^{\mathrm{T}} - \bar{\Omega} + I_{n_c}). \quad \text{(G.1.3)}$$

We derive explicit expressions for the derivatives $\frac{\partial A_c}{\partial b_{c,k\ell}}$ and $\frac{\partial A_c}{\partial c_{c,k\ell}}$.

Below, we use the notation

$$F_{B_c} \triangleq B_c B_c^{\mathrm{T}} \quad \text{(G.1.4)}$$

$$F_{C_c} \triangleq C_c^{\mathrm{T}} C_c \quad \text{(G.1.5)}$$

$$M \triangleq N_{n_c} - I_{n_c} \quad \text{(G.1.6)}$$

$$u_{n_c} \triangleq [1, 1, \ldots, 1]^{\mathrm{T}}, \quad u_{n_c} \in \mathbb{R}^{n_c} \quad \text{(G.1.7)}$$

and recognize that

$$\bar{Z} = \operatorname{diag}(Z) u_{n_c}^{\mathrm{T}}, \quad Z \in \mathbb{R}^{n_c \times n_c} \quad \text{(G.1.8)}$$

$$N_{n_c} = u_{n_c} u_{n_c}^{\mathrm{T}} \quad \text{(G.1.9a)}$$

$$\mathcal{E}_{n_c}^{(k)} = e_{n_c}^{(k)} u_{n_c}^{\mathrm{T}} \quad \text{(G.1.9b)}$$

$$\bar{\Omega} = \bar{G} / \bar{F} \quad \text{(G.1.10)}$$

$$H = M / (\bar{\Omega}^{\mathrm{T}} - \bar{\Omega} + I_{n_c}) \quad \text{(G.1.11)}$$

$$A_c = \frac{1}{2} I * [F_{C_c} - F_{B_c}] + [F_{C_c} * (N_{n_c} + \bar{\Omega}^{\mathrm{T}}) - (N_{n_c} + \bar{\Omega}) * F_{B_c} * \bar{\Omega}^{\mathrm{T}}] * H \quad \text{(G.1.12)}$$

$$[F_{C_c} * (N_{n_c} + \bar{\Omega}^{\mathrm{T}}) - (N_{n_c} + \bar{\Omega}) * F_{B_c} * \bar{\Omega}^{\mathrm{T}}] = A_c * M / H. \quad \text{(G.1.13)}$$

The derivations of the expressions for $\frac{\partial A_c}{\partial b_{c,k\ell}}$ and $\frac{\partial A_c}{\partial c_{c,k\ell}}$ use the following identities.

$$E_{m \times n}^{(k,\ell)} = e_m^{(k)} e_n^{(\ell)\mathrm{T}} \quad \text{(G.1.14)}$$

$$e_m^{(k)^{\mathrm{T}}} Z = Z(k,:), \quad Z \in \mathrm{I\!R}^{m \times n} \tag{G.1.15}$$

$$Z e_n^{(\ell)} = Z(:,\ell), \quad Z \in \mathrm{I\!R}^{m \times n} \tag{G.1.16}$$

$$\frac{d}{dx}[N/D(x)] = -[N/D(x)/D(x)] * \frac{\partial D}{\partial x} \tag{G.1.17}$$

$$\frac{d}{dx}[N(x)/D] = \frac{\partial N}{\partial x}/D \tag{G.1.18}$$

$$Z = ab^{\mathrm{T}} \;\Rightarrow\; \bar{Z} = (a * b)u_n, \quad a, b, \in \mathrm{I\!R}^n \tag{G.1.19}$$

$$M * M = M \tag{G.1.20}$$

G.2 Expression for $\frac{\partial A_c}{\partial b_{c,k\ell}}$ and $\frac{\partial A_c}{\partial C_{c,k\ell}}$

Differentiating (G.1.12) with respect to $b_{c,k\ell}$ gives respectively

$$\begin{aligned}\frac{\partial A_C}{\partial b_{c,k\ell}} &= -\frac{1}{2}I_{n_c} * \frac{\partial F}{\partial b_{c,k\ell}}\\ &+ [F_{C_c} * \frac{\partial \bar{\Omega}^{\mathrm{T}}}{\partial b_{c,k\ell}} - \frac{\partial \bar{\Omega}}{\partial b_{c,k\ell}} * F_{B_c} * \bar{\Omega}^{\mathrm{T}} - (N_{n_c} + \bar{\Omega}) * (\frac{\partial F_{B_c}}{\partial b_{c,k\ell}} * \bar{\Omega}^{\mathrm{T}} + F_{B_c} * \frac{\partial \bar{\Omega}^{\mathrm{T}}}{\partial b_{c,k\ell}})] * H\\ &+ (A_c * M/H) * \frac{\partial H}{\partial b_{c,k\ell}} \qquad \text{(G.2.1)}\end{aligned}$$

$$\begin{aligned}\frac{\partial A_C}{\partial c_{c,k\ell}} &= \frac{1}{2}I_{n_c} * \frac{\partial F_{B_c}}{\partial c_{c,k\ell}}\\ &+ [\frac{\partial F_{C_c}}{\partial c_{c,k\ell}} * (N_{n_c} + \bar{\Omega}^{\mathrm{T}}) + F_{C_c} * \frac{\partial \bar{\Omega}^{\mathrm{T}}}{\partial c_{c,k\ell}} - \frac{\partial \bar{\Omega}}{\partial c_{c,k\ell}} * F_{B_c} * \bar{\Omega}^{\mathrm{T}} - (N_{n_c} + \bar{\Omega}) * F_{B_c} * \frac{\partial \bar{\Omega}^{\mathrm{T}}}{\partial c_{c,k\ell}}] * H\\ &+ (A_c * M/H) * \frac{\partial H}{\partial c_{c,k\ell}}. \qquad \text{(G.2.2)}\end{aligned}$$

Below, we develop explicit expressions for the derivatives on the right hand sides of (G.2.1) and (G.2.2).

$$\underline{\frac{\partial F_{B_c}}{\partial b_{c,k\ell}} \text{ and } \frac{\partial F_{C_c}}{\partial c_{c,k\ell}}}$$

Differentiating (G.1.4) and (G.1.5) respectively with respect to $b_{c,k\ell}$ and $c_{k,\ell}$ gives

$$\frac{\partial F_{B_c}}{\partial b_{c,k\ell}} = E^{(k,\ell)}_{n_c \times n_y} B_c^{\mathrm{T}} + B_c E^{(k,\ell)^{\mathrm{T}}}_{n_y \times n_c} \qquad \text{(G.2.3)}$$

$$\frac{\partial F_{C_c}}{\partial c_{c,k\ell}} = E^{(k,\ell)^{\mathrm{T}}}_{n_u \times n_c} C_c + C_c^{\mathrm{T}} E^{(k,\ell)}_{n_u \times n_c} \qquad \text{(G.2.4)}$$

which using (G.1.14) are equivalent to

$$\frac{\partial F_{B_c}}{\partial b_{c,k\ell}} = e^{(k)}_{n_c} e^{(\ell)^{\mathrm{T}}}_{n_y} B_c^{\mathrm{T}} + B_c e^{(\ell)}_{n_y} e^{(k)^{\mathrm{T}}}_{n_c} \qquad \text{(G.2.5)}$$

$$\frac{\partial F_{C_c}}{\partial c_{c,k\ell}} = e^{(\ell)}_{n_u} e^{(k)^{\mathrm{T}}}_{n_u} C_c + C_c^{\mathrm{T}} e^{(k)}_{n_u} e^{(\ell)^{\mathrm{T}}}_{n_c}. \qquad \text{(G.2.6)}$$

From (G.2.5) and (G.2.6) with (G.15) and (G.16) we obtain

$$\boxed{\begin{aligned}\frac{\partial F_{B_c}}{\partial b_{c,k\ell}} &= e^{(k)}_{n_c} B_c(:,\ell)^{\mathrm{T}} + B_c(:,\ell) e^{(k)^{\mathrm{T}}}_{n_c} \qquad \text{(G.2.7)}\\ \frac{\partial F_{C_c}}{\partial c_{c,k\ell}} &= e^{(\ell)}_{n_c} C_c(k,:) + C_c(k,:)^{\mathrm{T}} e^{(\ell)^{\mathrm{T}}}_{n_c} \qquad \text{(G.2.8)}\end{aligned}}$$

$$\underline{\frac{\partial \bar{\Omega}}{\partial b_{c,k\ell}} \text{ and } \frac{\partial \bar{\Omega}}{\partial c_{c,k\ell}}}$$

Differentiating (C.1.10) with respect to $b_{c,k\ell}$ and $c_{c,k\ell}$ using respectively (C.1.17) and (C.1.18) gives

$$\frac{\partial \bar{\Omega}}{\partial b_{c,k\ell}} = -(\bar{F}_{C_c}/\bar{F}_{B_c}/\bar{F}_{B_c}) * \frac{\partial F_{B_c}}{\partial b_{c,k\ell}} \tag{G.2.9}$$

$$\frac{\partial \bar{\Omega}}{\partial c_{c,k\ell}} = \frac{\partial \bar{F}_{C_c}}{\partial c_{c,k\ell}}/\bar{F}_{B_c}. \tag{G.2.10}$$

Also, it follows form (G.2.7) and (G.2.8) with (G.1.9b) that

$$\frac{\partial \bar{F}_{B_c}}{\partial b_{c,k\ell}} = 2b_{c,k\ell}e_{n_c}^{(k)}u_{n_c}^{\mathrm{T}} \tag{G.2.11}$$

$$\frac{\partial \bar{F}_{C_c}}{\partial c_{c,k\ell}} = 2c_{c,k\ell}e_{n_c}^{(\ell)}u_{n_c}^{\mathrm{T}} \tag{G.2.12}$$

or, equivalently,

$$\frac{\partial \bar{F}_{B_c}}{\partial b_{c,k\ell}} = 2b_{c,k\ell}\mathcal{E}_{n_c}^{(k)} \tag{G.2.13}$$

$$\frac{\partial \bar{F}_{C_c}}{\partial c_{c,k\ell}} = 2c_{c,k\ell}\mathcal{E}_{n_c}^{(\ell)}. \tag{G.2.14}$$

Substituting (G.2.13) into (G.2.9) gives

$$\frac{\partial \bar{\Omega}}{\partial b_{c,k\ell}} = -(\bar{\Omega}/\bar{F}_{B_c}) * (2b_{c,k\ell}\mathcal{E}_{n_c}^{(k)}) \tag{G.2.15}$$

or, equivalently,

$$\boxed{\frac{\partial \bar{\Omega}}{\partial b_{c,k\ell}} = -\frac{2b_{c,k\ell}\omega_k}{f_{B_c,kk}}\mathcal{E}_{n_c}^{(k)}. \qquad \text{(G.2.16)}}$$

Substituting (G.2.14) into (G.2.10) gives

$$\frac{\partial \bar{\Omega}}{\partial c_{c,k\ell}} = 2c_{c,k\ell}\mathcal{E}_{n_c}^{(\ell)}/\bar{F}_{B_c} \tag{G.2.17}$$

or, equivalently,

$$\boxed{\frac{\partial \bar{\Omega}}{\partial c_{c,k\ell}} = \frac{2c_{c,k\ell}}{f_{B_c,\ell\ell}}\mathcal{E}_{n_c}^{(\ell)}. \qquad \text{(G.2.18)}}$$

$\dfrac{\partial H}{\partial b_{c,k\ell}}$ and $\dfrac{\partial H}{\partial c_{c,k\ell}}$

Let $x = b_{c,k\ell}$ or $c_{c,k\ell}$. Differentiating (G.1.11) by x yields

$$\frac{\partial H}{\partial x} = -[M/(\bar{\Omega}^{\mathrm{T}} - \bar{\Omega} + I_{n_c})/(\bar{\Omega}^{\mathrm{T}} - \bar{\Omega} + I_{n_c})] * (\frac{\partial \bar{\Omega}^{\mathrm{T}}}{\partial x} - \frac{\partial \bar{\Omega}}{\partial x}). \tag{G.2.19}$$

It follows from (G.1.11) and (G.1.20) that

$$\frac{\partial H}{\partial x} = H * H * (\frac{\partial \bar{\Omega}}{\partial x} - \frac{\partial \bar{\Omega}^{\mathrm{T}}}{\partial x}). \tag{G.2.20}$$

Hence, from (G.2.20) with (G.2.16) and (G.2.18) it follows that

$$\frac{\partial H}{\partial b_{c,k\ell}} = -\frac{2b_{c,k\ell}\omega_k}{f_{B_c,kk}} H * H * (\mathcal{E}_{n_c}^{(k)} - \mathcal{E}_{n_c}^{(k)^{\mathrm{T}}}). \tag{G.2.21}$$

$$\frac{\partial H}{\partial c_{c,k\ell}} = \frac{2c_{c,k\ell}}{f_{B_c,\ell\ell}} H * H * (\mathcal{E}_{n_c}^{(\ell)} - \mathcal{E}_{n_c}^{(\ell)^{\mathrm{T}}}). \tag{G.2.22}$$

Substituting (G.2.7), (G.2.16) and (G.2.21) into (G.2.1) gives

$$\begin{aligned}
\frac{\partial A_c}{\partial b_{c,k\ell}} &= -\frac{1}{2} I_{n_c} * [e_{n_c}^{(k)} B_c(:,\ell)^{\mathrm{T}} + B_c(:,\ell) e_{n_c}^{(k)^{\mathrm{T}}} \\
&\quad + [F_{C_c} * (\frac{-2b_{c,k\ell}\omega_k}{f_{B_c,kk}} \mathcal{E}_{n_c}^{(k)^{\mathrm{T}}}) + (\frac{2b_{c,k\ell}\omega_k}{f_{B_c,kk}} \mathcal{E}_{n_c}^{(k)}) * F_{B_c} * \bar{\Omega}^{\mathrm{T}} \\
&\quad - (N_{n_c} + \bar{\Omega}) * (e_{n_c}^{(k)} B_c(:,\ell)^{\mathrm{T}} + B_c(:,\ell) e_{n_c}^{(k)^{\mathrm{T}}}) * \bar{\Omega}^{\mathrm{T}} \\
&\quad - (N_{n_c} + \bar{\Omega}) * F_{B_c} * (-\frac{2b_{c,k\ell}\omega_k}{f_{B_c,kk}} \mathcal{E}_{n_c}^{(k)^{\mathrm{T}}})] * H \\
&\quad + (A_c * M/H) * [\frac{-2b_{c,k\ell}\omega k}{f_{B_c,kk}} H * H * (\mathcal{E}_{n_c}^{(k)} - \mathcal{E}_{n_c}^{(k)^{\mathrm{T}}})]
\end{aligned} \tag{G.2.23}$$

or, equivalently,

$$\begin{aligned}
\frac{\partial A_c}{\partial b_{c,k\ell}} &= -b_{c,k\ell} E_{n_c \times n_c}^{(k,k)} \\
&\quad + \frac{2b_{c,k\ell}\omega_k}{f_{B_c,kk}} \mathcal{E}_{n_c}^{(k)} * [F_{B_c} * \bar{\Omega}^{\mathrm{T}} * H - A_c * M * H] \\
&\quad + \frac{2b_{c,k\ell}\omega_k}{f_{B_c,kk}} \mathcal{E}_{n_c}^{(k)^{\mathrm{T}}} * [-F_{C_c} * H + (N_{n_c} + \bar{\Omega}) * F_{B_c} * H + A_c * M * H] \\
&\quad - [(N_{n_c} + \bar{\Omega}) * H * \bar{\Omega}^{\mathrm{T}}] * [e_{n_c}^{(k)} B_c(:,\ell)^{\mathrm{T}} + B_c(:,\ell) e_{n_c}^{(k)^{\mathrm{T}}}].
\end{aligned} \tag{G.2.24}$$

Substituting (G.2.8), (G.2.18) and (G.2.22) into (G.2.2) gives

$$\begin{aligned}\frac{\partial A_c}{\partial c_{c,k\ell}} &= \frac{1}{2}I_{n_c} * [e_{n_c}^{(\ell)}C_c(k,:) + C_c(k,:)^{\mathrm{T}}e_{n_c}^{(\ell)^{\mathrm{T}}} \\ &+ [(e_{n_c}^{(\ell)}C_c(k,:) + C_c(k,:)^{\mathrm{T}}e_{n_c}^{(\ell)^{\mathrm{T}}} * (N_{n_c} + \bar{\Omega}^{\mathrm{T}}) \\ &\quad + F_{C_c} * (\frac{2C_{c,k\ell}}{f_{B_c,\ell\ell}}\mathcal{E}_{n_c}^{(\ell)^{\mathrm{T}}}) - (\frac{2C_{c,k\ell}}{f_{B_c,\ell\ell}}\mathcal{E}_{n_c}^{(\ell)^{\mathrm{T}}})] * H \\ &\quad - (N_{n_c} + \bar{\Omega}) * F_{B_c} * (\frac{2c_{c,k\ell}}{f_{B_c,\ell\ell}}\mathcal{E}_{n_c}^{(\ell)^{\mathrm{T}}})] * H \\ &(A_c * M/H) * [(\frac{2c_{c,k\ell}}{f_{B_c,\ell\ell}}H * H * (\mathcal{E}_{n_c}^{(\ell)} - \mathcal{E}_{n_c}^{(\ell)^{\mathrm{T}}})]\end{aligned} \tag{G.2.25}$$

or, equivalently,

$$\begin{aligned}\frac{\partial A_c}{\partial c_{c,k\ell}} &= c_{c,k\ell}E_{n_c \times n_c}^{(\ell,\ell)} \\ &+ \frac{2c_{c,k\ell}}{f_{B_c,\ell\ell}}\mathcal{E}_{n_c}^{(\ell)}[-F_{B_c} * \bar{\Omega}^{\mathrm{T}} * H + A_c * M * H] \\ &+ \frac{2c_{c,k\ell}}{f_{B_c,\ell\ell}}\mathcal{E}_{n_c}^{(\ell)^{\mathrm{T}}}[F_{C_c} * H - (N_{n_c} + \bar{\Omega}) * F_{B_c} * H + A_c * M * H] \\ &+ (N_{n_c} + \bar{\Omega}^{\mathrm{T}}) * H * [e_{n_c}^{(\ell)}C_c(k,:) + C_c(k,:)^{\mathrm{T}}e_{n_c}^{(\ell)^{\mathrm{T}}}].\end{aligned} \tag{G.2.26}$$

Now, define

$$H_{row} \triangleq [(F_{B_c} * \bar{\Omega}^{\mathrm{T}}) - (A_c * M)] * H \tag{G.2.27}$$

$$H_{col} \triangleq \left[((N_{n_c} + \bar{\Omega}) * F_{B_c}) + (A_c * M) - F_{C_c}\right] * H \tag{G.2.28}$$

$$H_{B_c} \triangleq (N_{n_c} + \bar{\Omega}) * \bar{\Omega}^{\mathrm{T}} * H \tag{G.2.29}$$

$$H_{C_c} \triangleq (N_{n_c} + \bar{\Omega}^{\mathrm{T}}) * H. \tag{G.2.30}$$

Then, it follows from (G.2.24) and (G.2.26) that

$$\begin{aligned}\frac{\partial A_c}{\partial b_{c,k\ell}} &= -b_{c,k\ell}E_{n_c \times n_c}^{(k,k)} + \frac{2b_{c,k\ell}}{f_{B_c,kk}}\omega_k[e_{n_c}^{(k)}H_{row}(k,:) + H_{col}(:,k)e_{n_c}^{(k)^{\mathrm{T}}}] \\ &\quad + \hat{H}_{B_c} * [B_c(:,\ell)e_{n_c}^{(k)^{\mathrm{T}}} + e_{n_c}^{(k)}B_c(:,\ell)^{\mathrm{T}}]\end{aligned} \tag{G.2.31}$$

$$\begin{aligned}\frac{\partial A_c}{\partial c_{c,k\ell}} &= c_{c,k\ell}E_{n_c \times n_c}^{(\ell,\ell)} - \frac{2c_{c,k\ell}}{f_{B_c,\ell\ell}}[e_{n_c}^{(\ell)}H_{row}(\ell,:) + H_{col}(:,\ell)e_{n_c}^{(\ell)^{\mathrm{T}}}] \\ &\quad + \hat{H}_{C_c} * [e_{n_c}^{(\ell)}C_c(k,:) + C_c(k,:)^{\mathrm{T}}e_{n_c}^{(\ell)^{\mathrm{T}}}]\end{aligned} \tag{G.2.32}$$

Note that $\frac{\partial A_c}{\partial b_{c,k\ell}}$ only has nonzero entries in the k^{th} row and column, while $\frac{\partial A_c}{\partial c_{c,k\ell}}$ only has nonzero entries in the ℓ^{th} row and column.

Appendix H:

"Design of Reduced-Order H_2 Optimal Controllers
Using a Homotopy Algorithm

October 1992
Revised: March 1993
Revised: September 1993

Design of Reduced-Order, H2 Optimal Controllers Using a Homotopy Algorithm

by

Emmanuel G. Collins, Jr., Larry D. Davis, Stephen Richter
Harris Corporation
Government Aerospace Systems Division
MS 22/4849
Melbourne, Florida 32902
FAX: (407) 727-6007
email: ecollins@x102a.ess.harris.com

Abstract

The minimal dimension of a linear-quadratic-gaussian (LQG) compensator is almost always equal to the dimension of the design plant. This deficiency can lead to implementation problems when considering control design for high-order systems such as flexible structures and has led to the development of methodologies for the design of optimal (or near optimal) controllers whose dimension is less than that of the design plant. This paper presents a new (gradient-based) homotopy algorithm for the design of reduced-order, H_2 optimal controllers. An important result is the development of an *efficient* methodology for computation of the of the cost functional Hessian which is required by the algorithm. The optimal controller is represented by a parameter vector and various parameterizations of the optimal controller are considered to reduce the algorithm dimensionality. The algorithm has been implemented in MATLAB and the results are illustrated using a benchmark, non-collocated flexible structure control problem. It is seen that the choice of a particular controller parameterization often introduces numerical ill-conditioning in the algorithm implementation.

*Supported in part by SANDIA National Laboratories under contract 54-7609 and NASA under contract NAS8-38575.

1. Introduction and Nomenclature

The linear-quadratic-gaussian (LQG) compensator (Athans 1971, Kwakernaak and Sivan 1972) has been developed to facilitate the design of control laws for multi-input multi-output (MIMO) systems. An LQG compensator minimizes a quadratic performance index and (under mild conditions) is guaranteed to yield an internally stable closed-loop system. Unfortunately, however, the minimal dimension of an LQG compensator is almost always equal to the dimension of the plant and can thus often violate practical implementation constraints on controller order. This deficiency is especially highlighted when considering control-design for high-order systems such as flexible space structures. Hence, a very relevant area of research is the development of methodologies that will enable the design of optimal controllers whose dimension is less than that of the design plant (i.e., reduced-order controllers).

Two main approaches have been developed to tackle the reduced-order design problem. The first approach attempts to develop approximations to the optimal reduced-order controller by reducing the dimension of an LQG controller (Yousuff and Skelton 1984a, Yousuff and Skelton 1984b, Anderson and Liu 1989, De Villemagne and Skelton 1988, Liu, Anderson and Ly 1990). These methods are attractive because they require relatively little computation and should be used if possible. Unfortunately, they tend to yield controllers that either destabilize the system or have poor performance as the requested controller dimension is decreased and/or the requested authority level is increased. Hence, if used in isolation, these methods do not yield a reliable methodology for reduced-order design.

The second approach attempts to directly synthesize an optimal, reduced-order controller by a numerical optimization scheme (Levine, Johnson and Athans 1971, Martin and Bryson 1980, Mukhopadhyay, Newsom and Abel 1982, Ly, Brysib and Cannon 1985, Mukhopadhyay 1987, Kuhn and Schmidt 1987, Richter 1987, Makila and Toivonen 1987, Kramer and Calise 1988, Mukhopadhyay 1989, Richter and Collins 1989, Mercadal 1991). Almost all of these schemes are gradient-based parameter optimization approaches; that is, they represent the controller by some parameter vector and attempt to find a vector for which the gradient of the cost functional is zero.

With the exception of Mercadal 1991, all of the previous, gradient-based optimization techniques are descent methods. That is, at each iteration the cost function is decreased. An alternative (Mercadal 1991) is to develop a gradient-based homotopy algorithm that allows an initial controller to be deformed gradually into a desired optimal controller by following a homotopy path. This type of algorithm is distinct from the previous algorithms in that each iteration does not necessarily decrease the cost function. In fact, the cost may actually increase as the homotopy path is traversed.

However, it is quite possible that the shortest path from the initial controller to the desired optimal controller is not a descent path.

Efficient path following requires accurate computation of the Hessian of the cost functional. Hence, this paper develops an *efficient* method for computing the Hessian. An alternative method for computing the Hessian is presented in an earlier publication (Sun 1991). However, to our knowledge, this previous method, based on the results of Sun 1990, does not exploit certain low rank matrices as does the method presented here.

A homotopy algorithm for optimal reduced-order design is described in Richter 1987 and Richter and Collins 1989. This algorithm is based on solving a set of "optimal projection" equations (Hyland and Bernstein 1984, Haddad 1987) that are a characterization of the necessary conditions for optimal reduced-order control. Unfortunately, the algorithm has sublinear convergence properties and the convergence slows at higher control authority levels and may fail. Homotopy algorithms for optimal reduced-order modeling, based on optimal projection equations, are discussed in Zigic et al. 1992 and Zigic et al. 1993. These algorithms are based on more efficient path following techniques but are relatively slow due to the large dimensionality of the algorithm formulation.

This paper describes a homotopy algorithm for the design of reduced-order, H_2 optimal controllers which is not based on the optimal projection equations. The algorithm relies on the first and second derivatives (i.e., the gradient and Hessian) of the cost functional with respect to a parameter vector describing the controller and an efficient methodology for computing the Hessian is developed. To reduce the dimensionality of the algorithm, various parameterizations of the optimal controller are considered. The algorithm has the potential for quadratic convergence rates along the homotopy curve. The results have been implemented in MATLAB and are illustrated using a benchmark, non-collocated flexible structure control problem. It is seen that the choice of a particular controller parameterization often introduces numerical ill-conditioning in the algorithm implementation. The algorithm presented here is similar to that described in Mercadal 1991. However, whereas Mercadal 1991 focuses on theoretical issues related to homotopies and only describes a rudimentary homotopy algorithm, the present paper focusses on numerical algorithmic issues and describes a much more refined and efficient homotopy algorithm.

The paper is organized as follows. Section 2 describes the H_2 optimal reduced order dynamic compensation problem. Section 3 gives a brief overview of homotopy methods. Section 4 then develops a homotopy algorithm for the design of reduced-order H_2 optimal controllers. Section 5 applies the algorithm to a benchmark structural control problem and compares various algorithm options. Finally, Section 6 presents the conclusions.

Nomenclature

$Y \geq Z$	$Y - Z$ is nonnegative definite
$Y > Z$	$Y - Z$ is positive definite
$z_{ij}, Z_{i,j}$ or $Z_{(i,j)}$	(i,j) element of matrix Z
I_r	$r \times r$ identity matrix
$\mathrm{tr} Z$	trace of square matrix Z
$\mathrm{vec}(\cdot)$	the invertible linear operator defined such that $\mathrm{vec}(s) \triangleq [s_1^{\mathrm{T}}\ s_2^{\mathrm{T}} \cdots s_q^{\mathrm{T}}]^{\mathrm{T}},\ S \in \mathbb{R}^{p \times q}$ where $s_j \in \mathbb{R}^p$ denotes the j^{th} column of S.
$e_m^{(i)}$	the m-dimensional column vector whose i^{th} element equals one and whose additional elements are zeros.
$E_{m,n}^{(k,\ell)}$	the $m \times n$ matrix whose (k,ℓ) element equals one and whose additional elements are zeros.
$Z(k,:)$	k^{th} row of the matrix Z (MATLAB notation)
$Z(:,k)$	k^{th} column of the matrix Z (MATLAB notation)

2. H$_2$ Optimal Reduced-Order Dynamic Compensation

Consider the system

$$\dot{x}(t) = Ax(t) + Bu(t) + w_1(t) \tag{2.1a}$$

$$y(t) = Cx(t) + Du(t) + w_2(t) \tag{2.1b}$$

where $x \in \mathbb{R}^{n_x}, u \in \mathbb{R}^{n_u}, y \in \mathbb{R}^{n_y}, w_1 \in \mathbb{R}^{n_x}$ is white disturbance noise with intensity $V_1 \geq 0, w_2 \in \mathbb{R}^{n_y}$ is white observation noise with intensity $V_2 \geq 0$, and w_1 and w_2 have cross correlation $V_{12} \in \mathbb{R}^{n_x \times n_y}$. We desire to design a fixed-order dynamic compensator,

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t) \tag{2.2a}$$

$$u(t) = -C_c x_c(t) \tag{2.2b}$$

which minimizes the steady-state performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \to \infty} E[x^{\mathrm{T}}(t) R_1 x(t) + 2x^{\mathrm{T}}(t) R_{12} u(t) + u^{\mathrm{T}}(t) R_2 u(t)] \tag{2.3}$$

where $x_c \in \mathrm{IR}^{n_c}, n_c \leq n, R_1 = R_1^{\mathrm{T}} \geq 0$ and $R_2 = R_2^{\mathrm{T}} \geq 0$. We will call this problem *the optimal reduced-order dynamic compensation problem.*

The closed-loop system corresponding to (2.1) and (2.2) can be expressed as

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{w}(t) \tag{2.4}$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix}, \quad \tilde{w}(t) \triangleq \begin{bmatrix} w_1(t) \\ B_c w_2(t) \end{bmatrix} \tag{2.5,6,7}$$

$$\tilde{A} \triangleq \begin{bmatrix} A & -BC_c \\ B_c C & A_c - B_c D C_c \end{bmatrix}. \tag{2.8}$$

In addition, the cost (2.3) can be expressed as

$$J(A_c, B_c, C_c) = \lim_{t \to \infty} E[\tilde{x}^{\mathrm{T}}(t)\tilde{R}\tilde{x}(t)] \tag{2.9}$$

where

$$\tilde{R} \triangleq \begin{bmatrix} R_1 & R_{12}C_c \\ C_c^{\mathrm{T}} R_{12}^{\mathrm{T}} & C_c^{\mathrm{T}} R_2 C_c \end{bmatrix}. \tag{2.10}$$

To guarantee that the cost J is finite and independent of initial conditions we restrict our attention to the set of stabilizing compensators, $\mathcal{S}_c \triangleq \{(A_c, B_c, C_c) \colon \tilde{A}$ is asymptotically stable$\}$. Assume $(A_c, B_c, C_c) \in \mathcal{S}_c$ and define $\tilde{Q} \in \mathrm{IR}^{n_x \times n_x}$ and $\tilde{P} \in \mathrm{IR}^{n_x \times n_x}$ to be the closed-loop steady-state covariance and its dual, i.e.,

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{V} \tag{2.11}$$

$$0 = \tilde{A}^{\mathrm{T}}\tilde{P} + \tilde{P}\tilde{A} + \tilde{R} \tag{2.12}$$

where

$$\tilde{V} \triangleq \begin{bmatrix} V_1 & V_{12}B_c^{\mathrm{T}} \\ B_c V_{12}^{\mathrm{T}} & B_c V_2 B_c^{\mathrm{T}} \end{bmatrix}. \tag{2.13}$$

Then, the cost can be expressed as

$$J(A_c, B_c, C_c) = \operatorname{tr} \tilde{Q}\tilde{R} = \operatorname{tr} \tilde{P}\tilde{V}. \tag{2.14}$$

Also, $\tilde{Q}$ and $\tilde{P}$ can be expressed in the partitioned forms

$$\tilde{Q} = \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}^{\mathrm{T}} & \tilde{Q}_{22} \end{bmatrix}, \quad \tilde{Q}_{11} \in \mathrm{IR}^{n_x \times n_x}, \quad \tilde{Q}_{22} \in \mathrm{IR}^{n_c \times n_c} \tag{2.15}$$

$$\tilde{P} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{12}^{\mathrm{T}} & \tilde{P}_{22} \end{bmatrix}, \quad \tilde{P}_{11}\mathrm{IR}^{n_x \times n_x}, \quad \tilde{P}_{22} \in \mathrm{IR}^{n_c \times n_c}. \tag{2.16}$$

Notice that $\tilde{Q}_{11}$ is the covariance of the plant states, $\tilde{Q}_{22}$ is the covariance of the compensator states and $\tilde{Q}_{12}$ is the cross-covariance of the plant and controller states.

Expressions for the partial derivatives $\frac{\partial \mathcal{J}}{\partial A_c}, \frac{\partial \mathcal{J}}{\partial B_c}$, and $\frac{\partial \mathcal{J}}{\partial C_c}$, are given below. First, we define $\tilde{Z}$ satisfying

$$\tilde{Z} = \tilde{Q}\tilde{P} \tag{2.17}$$

and assign $\tilde{Z}$ the partitioned form

$$\tilde{Z} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{21} & \tilde{Z}_{22} \end{bmatrix}, \quad \tilde{Z}_{11} \in \mathbb{R}^{n_s x n_s}, \quad \tilde{Z}_{22} \in \mathbb{R}^{n_c x n_c}, \tag{2.18}$$

The cost derivatives are then given by

$$\frac{\partial J}{\partial A_c} = 2\tilde{Z}_{22}^{\mathrm{T}} \tag{2.19}$$

$$\frac{\partial J}{\partial B_c} = 2(\tilde{P}_{12}^{\mathrm{T}} V_{12} + \tilde{P}_{22} B_c V_2 + \tilde{Z}_{12}^{\mathrm{T}} C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} D^{\mathrm{T}}) \tag{2.20}$$

$$\frac{\partial J}{\partial C_c} = 2(-R_{12}^{\mathrm{T}} \tilde{Q}_{12} + R_2 C_c \tilde{Q}_{22} - B^{\mathrm{T}} \tilde{Z}_{21}^{\mathrm{T}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \tilde{Z}_{22}^{\mathrm{T}}) \tag{2.21}$$

Definition 2.1. A compensator (A_c, B_c, C_c) is an *extremal* of the optimal reduced-order dynamic compensation problem if it satisfies the stationary conditions

$$\frac{\partial J}{\partial A_c} = 0, \quad \frac{\partial J}{\partial B_c} = 0, \quad \frac{\partial J}{\partial C_c} = 0. \tag{2.22}$$

The homotopy algorithm of Section 4 is based on finding extremals of the optimal reduced-order dynamic compensation problem.

3. Homotopy Methods for the Solution of Nonlinear Algebraic Equations

A "homotopy" is a continuous deformation of one function into another. Over the past several years, homotopy or continuation methods (whose mathematical basis is algebraic topology and differential topology (Lloyd 1978)) have received significant attention in the mathematics literature and have been applied successfully to several important problems (Avila 1874, Wacker 1978, Alexander and Yorke 1978, Garcia and Zangwill 1981, Eaves, Gould, Peoitgen, and Todd 1983, Watson, 1986). Recently, the engineering literature has also begun to recognize the utility of these methods for engineering applications (see e.g. Richter and DeCarlo 1983, Richter and DeCarlo 1984, Turner and Chun 1984, Dunyak, Junkins, and Watson 1984, Lefebvre, Richter and DeCarlo 1985, Sebok, Richter, and Decarlo 1986, Horta, Juang and Junkins 1986, Kabamba, Longman and

Jian-Guo 1987, Shin, Raftka, Watson, and Plautt 1988, Rakowska, Haftka, and Watson 1991)). The purpose of this section is to provide a very brief description of homotopy methods for finding the solutions of nonlinear algebraic equations. The reader is referred to (Watson 1986, Richter and DeCarlo 1983, Watson 1987, Watson 1986) for additional details.

The basic problem is as follows. Given sets Θ and Φ contained in $\mathbb{R}^n$ and a mapping $F : \Theta \to \Phi$, find solutions to

$$F(\theta) = 0. \tag{3.1}$$

Homotopy methods embed the problem (3.1) in a larger problem. In particular let $H : \Theta \times [0,1] \to \mathbb{R}^n$ be such that:

1) $H(\theta, 1) = F(\theta)$. (3.2)

2) There exists at least one known $\theta_0 \in \mathbb{R}^n$ which is a solution to $H(\cdot, 0) = 0$, i.e.,

$$H(\theta_0, 0) = 0. \tag{3.3}$$

3) There exists a continuous curve $(\theta(\lambda), \lambda)$ in $\mathbb{R}^n \times [0,1]$ such that

$$H(\theta(\lambda), \lambda) = 0 \text{ for } \lambda \in [0,1] \tag{3.4}$$

with

$$(\theta(0), 0) = (\theta_0, 0). \tag{3.5}$$

4) The space $\Theta \times [0,1]$ has a differential structure so that the curve $(\theta(\lambda), \lambda)$ is differentiable.

A homotopy algorithm then constructs a procedure to compute the actual curve σ such that the initial solution $\theta(0)$ is transformed to a desired solution $\theta(1)$ satisfying

$$0 = H(\theta(1), 1) = F(\theta(1)). \tag{3.6}$$

Differentiating $H(\theta(\lambda), \lambda) = 0$ with respect to λ yields *Davidenko's differential equation*

$$\frac{\partial H}{\partial \theta}\frac{d\theta}{d\lambda} + \frac{\partial H}{\partial \lambda} = 0. \tag{3.7}$$

Together with $\theta(0) = \theta_0$, (3.7) defines an initial value problem which by numerical integration from 0 to 1 yields the desired solution $\theta(1)$. Some numerical integration schemes are described in Watson 1986 and Watson 1987).

4. A Homotopy Algorithm for H$_2$ Optimal Reduced-Order Dynamic Compensation

This section presents a new homotopy algorithm that can be used to design H$_2$ optimal reduced-order dynamic compensators. Particular attention is given to construction of the Jacobian of the homotopy map.

4.1 The Homotopy Map

If we define

$$\theta \triangleq \begin{bmatrix} \mathrm{vec}(A_c) \\ \mathrm{vec}(B_c) \\ \mathrm{vec}(C_c) \end{bmatrix}, \tag{4.1}$$

then the cost functional of Section 2 can be expressed as $J(\theta)$. The homotopy defined in this section is based on finding θ satisfying

$$0 = f(\theta) \triangleq \frac{\partial J}{\partial \theta}(\theta). \tag{4.2}$$

It is useful to recognize that

$$\frac{\partial J}{\partial \theta} = \begin{bmatrix} \mathrm{vec}\frac{\partial J}{\partial A_c} \\ \mathrm{vec}\frac{\partial J}{\partial B_c} \\ \mathrm{vec}\frac{\partial J}{\partial C_c} \end{bmatrix}. \tag{4.3}$$

Expressions for the partial derivatives $\frac{\partial J}{\partial A_c}, \frac{\partial J}{\partial B_c}$ and $\frac{\partial J}{\partial C_c}$ are given by (2.19)-(? 21).

Definition of the homotopy map $H(\theta, \lambda)$

To define the homotopy map we assume that the plant matrices (A, B, C, D), the cost weighting matrices (R_1, R_2, R_{12}) and the disturbance matrices (V_1, V_2, V_{12}) are functions of the homotopy parameter $\lambda \in [0,1]$. In particular, it is assumed that

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} + \lambda\left(\begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} \right), \tag{4.4}$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^{\mathrm{T}}(\lambda) & R_2(\lambda) \end{bmatrix} = L_R(\lambda) L_R^{\mathrm{T}}(\lambda) \tag{4.5}$$

where

$$L_R(\lambda) = L_{R,0} + \lambda(L_{R,f} - L_{R,0}) \tag{4.6}$$

and $L_{R,0}$ and $L_{R,f}$ satisfy

$$L_{R,0} L_{R,0}^{\mathrm{T}} = \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} \tag{4.7}$$

$$L_{R,f} L_{R,f}^{\mathrm{T}} = \begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^{\mathrm{T}} & R_{2,f} \end{bmatrix}, \tag{4.8}$$

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^{\mathrm{T}}(\lambda) & V_2^{\mathrm{T}}(\lambda) \end{bmatrix} = L_V(\lambda)L_V^{\mathrm{T}}(\lambda) \tag{4.9}$$

where

$$L_V(\lambda) = L_{V,0} + \lambda(L_{V,f} - L_{V,0}) \tag{4.10}$$

and $L_{V,0}$ and $L_{V,f}$ satisfy

$$L_{V,0}L_{V,0}^{\mathrm{T}} = \begin{bmatrix} V_{1,0} & V_{12,0} \\ V_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix} \tag{4.11}$$

$$L_{V,f}L_{V,f}^{\mathrm{T}} = \begin{bmatrix} V_{1,f} & V_{12,f} \\ V_{12,f}^{\mathrm{T}} & V_{2,f} \end{bmatrix}. \tag{4.12}$$

Note that (4.4)-(4.12) imply that $A(0) = A_0$ and $A(1) = A_f$, $B(0) = B_0$ and $B(1) = B_f$, etc ... and it is understood that $A_f, B_f, \ldots$ were referred to previously simply as $A, B, \ldots$. The change in notation is simply for convenience.

Let $\tilde{P}(\lambda)$, $\tilde{Q}(\lambda)$ and $\tilde{Z}(\lambda)$ satisfy

$$0 = \tilde{A}(\lambda)^{\mathrm{T}}\tilde{P}(\lambda) + \tilde{P}(\lambda)\tilde{A}(\lambda) + \tilde{R}(\lambda) \tag{4.13}$$

$$0 = \tilde{A}(\lambda)\tilde{Q}(\lambda) + \tilde{Q}(\lambda)\tilde{A}(\lambda)^{\mathrm{T}} + \tilde{V}(\lambda) \tag{4.14}$$

$$\tilde{Z}(\lambda) = \tilde{Q}(\lambda)\tilde{P}(\lambda) \tag{4.15}$$

with partitioned forms

$$\tilde{P}(\lambda) = \begin{bmatrix} \tilde{P}_{11}(\lambda) & \tilde{P}_{12}(\lambda) \\ \tilde{P}_{12}^{\mathrm{T}}(\lambda) & \tilde{P}_{22}(\lambda) \end{bmatrix}, \quad \tilde{Q}(\lambda) = \begin{bmatrix} \tilde{Q}_{11}(\lambda) & \tilde{Q}_{12}(\lambda) \\ \tilde{Q}_{12}^{\mathrm{T}}(\lambda) & \tilde{Q}_{22}(\lambda) \end{bmatrix}, \quad \tilde{Z}(\lambda) = \begin{bmatrix} \tilde{Z}_{11}(\lambda) & \tilde{Z}_{12}(\lambda) \\ \tilde{Z}_{21}(\lambda) & \tilde{Z}_{22}(\lambda) \end{bmatrix} \tag{4.16}$$

where the (1,1) and (2,2) blocks of each matrix are respectively $n_x \times n_x$ and $n_c \times n_c$. The homotopy map $H(\theta, \lambda)$ is defined as the gradient of the cost of the system defined by the homotopy parameter λ. In particular,

$$H(\theta, \lambda) \triangleq \begin{bmatrix} \mathrm{vec}(H_{A_c}(\theta, \lambda)) \\ \mathrm{vec}(H_{B_c}(\theta, \lambda)) \\ \mathrm{vec}(H_{C_c}(\theta, \lambda)) \end{bmatrix} \tag{4.17}$$

where

$$H_{A_c}(\theta, \lambda) = 2\tilde{Z}_{22}^{\mathrm{T}} \tag{4.18}$$

$$H_{B_c}(\theta, \lambda) = 2(\tilde{P}_{12}^{\mathrm{T}}V_{12} + \tilde{P}_{22}B_cV_2 + \tilde{Z}_{12}^{\mathrm{T}}C^{\mathrm{T}} - \tilde{Z}_{22}^{\mathrm{T}}C_c^{\mathrm{T}}D^{\mathrm{T}}) \tag{4.19}$$

$$H_{C_c}(\theta, \lambda) = 2(-R_{12}^{\mathrm{T}}\tilde{Q}_{12} + R_2C_c\tilde{Q}_{22} - B^{\mathrm{T}}\tilde{Z}_{21}^{\mathrm{T}} - D^{\mathrm{T}}B_c^{\mathrm{T}}\tilde{Z}_{22}^{\mathrm{T}}) \tag{4.20}$$

Note that in (4.18)-(4.20) and below the argument λ is omitted for notational convenience.

4.2 The Jacobian of the Homotopy Map

We now consider that computation of $\nabla H(\theta,\lambda)^{\rm T}$, the Jacobian of $H(\theta,\lambda)$. Note that

$$\nabla H(\theta,\lambda)^{\rm T} = [H_\theta \;\; H_\lambda] \tag{4.21}$$

where

$$H_\theta \triangleq \frac{\partial H}{\partial \theta}, \quad H_\lambda \triangleq \frac{\partial H}{\partial \lambda}. \tag{4.22}$$

Since $H(\theta,\lambda)$ is the gradient for the system defined by λ, H_θ is the corresponding Hessian. Recalling that θ is defined by (4.1), such that for some integers k and ℓ, θ_j is given by

$$\theta_j = a_{c,k\ell},\; \theta_j = b_{c,k\ell},\; \theta_j = c_{c,k\ell},\; \text{or}\; \theta_j = d_{c,k\ell}. \tag{4.23}$$

It follows from (4.13) that H_θ is of the form

$$H_\theta = \begin{bmatrix} \ldots \operatorname{vec}(\frac{\partial}{\partial a_{c,k\ell}} H_{A_c}) \ldots \operatorname{vec}(\frac{\partial}{\partial b_{c,k\ell}} H_{A_c}) \ldots \operatorname{vec}(\frac{\partial}{\partial c_{c,k\ell}} H_{A_c}) \\ \ldots \operatorname{vec}(\frac{\partial}{\partial a_{c,k\ell}} H_{B_c}) \ldots \operatorname{vec}(\frac{\partial}{\partial b_{c,k\ell}} H_{B_c}) \ldots \operatorname{vec}(\frac{\partial}{\partial c_{c,k\ell}} H_{B_c}) \\ \ldots \operatorname{vec}(\frac{\partial}{\partial a_{c,k\ell}} H_{C_c}) \ldots \operatorname{vec}(\frac{\partial}{\partial b_{c,k\ell}} H_{C_c}) \ldots \operatorname{vec}(\frac{\partial}{\partial c_{c,k\ell}} H_{C_c}) \end{bmatrix}. \tag{4.24}$$

and H_λ can be expressed as

$$H_\lambda = \begin{bmatrix} \operatorname{vec}(\frac{\partial}{\partial \lambda} H_{A_c}) \\ \operatorname{vec}(\frac{\partial}{\partial \lambda} H_{B_c}) \\ \operatorname{vec}(\frac{\partial}{\partial \lambda} H_{C_c}) \end{bmatrix}. \tag{4.25}$$

Below, we develop explicit expression for the derivative terms appearing on the right hand sides of (4.20) and (4.21). We use the notation

$$M^{(j)} \triangleq \frac{\partial M}{\partial \theta_j} \tag{4.26}$$

$$\dot{M} \triangleq \frac{\partial M}{\partial \lambda}. \tag{4.27}$$

Differentiating (4.13)-(4.15) with respect to θ_j yields

$$0 = \tilde{A}^{\rm T}\tilde{P}^{(j)} + \tilde{P}^{(j)}\tilde{A} + (\tilde{A}^{(j)^{\rm T}}\tilde{P} + \tilde{P}\tilde{A}^{(j)} + \tilde{R}^{(j)}) \tag{4.28}$$

$$0 = \tilde{A}\tilde{Q}^{(j)} + \tilde{Q}^{(j)}\tilde{A}^{\rm T} + (\tilde{A}^{(j)}\tilde{Q} + \tilde{Q}\tilde{A}^{(j)^{\rm T}} + \tilde{V}^{(j)}) \tag{4.29}$$

$$\tilde{Z}^{(j)} = \tilde{Q}^{(j)}\tilde{P} + \tilde{Q}\tilde{P}^{(j)} \tag{4.30}$$

where expressions for the derivatives $\tilde{A}^{(j)}, \tilde{R}^{(j)}$ and $\tilde{V}^{(j)}$ are given by (A.20)-(A.28) of Appendix A. Similarly, differentiating (4.13)-(4.15) with respect to λ yields

$$0 = \tilde{A}^{\rm T}\dot{\tilde{P}} + \dot{\tilde{P}}\tilde{A} + (\dot{\tilde{A}}^{\rm T}\tilde{P} + \tilde{P}\dot{\tilde{A}} + \dot{\tilde{R}}) \tag{4.31}$$

$$0 = \tilde{A}\dot{\tilde{Q}} + \dot{\tilde{Q}}\tilde{A}^{\rm T} + (\dot{\tilde{A}}\tilde{Q} + \tilde{Q}\dot{\tilde{A}}^{\rm T} + \dot{\tilde{V}}) \tag{4.32}$$

$$\dot{\tilde{Z}} = \dot{\tilde{Q}}\tilde{P} + \tilde{Q}\dot{\tilde{P}} \tag{4.33}$$

where expressions for $\dot{A}, \dot{R}$ and $\dot{V}$ are given by (A.29)–(A.33) of Appendix A.

Before presenting the desired derivative expressions we define

$$H'_{A_c}(\bar{Z}^{(j)}) \triangleq 2\bar{Z}_{22}^{(j)^{\mathrm{T}}} \tag{4.34}$$

$$H'_{B_c}(\bar{P}^{(j)}, \bar{Z}^{(j)}) \triangleq 2(\bar{P}_{12}^{(j)^{\mathrm{T}}} V_{12} + \bar{P}_{22}^{(j)} B_c V_2 + \bar{Z}_{12}^{(j)^{\mathrm{T}}} C^{\mathrm{T}} - \bar{Z}_{22}^{(j)^{\mathrm{T}}} C_c^{\mathrm{T}} D^{\mathrm{T}}) \tag{4.35}$$

$$H'_{C_c}(\bar{Q}^{(j)}, \bar{Z}^{(j)}) \triangleq 2(-R_{12}^{\mathrm{T}} \bar{Q}_{12}^{(j)} + R_2 C_c \bar{Q}_{22}^{(j)} - B^{\mathrm{T}} \bar{Z}_{21}^{(j)^{\mathrm{T}}} - D^{\mathrm{T}} B_c^{\mathrm{T}} \bar{Z}_{22}^{(j)^{\mathrm{T}}}) \tag{4.36}$$

Notice that the right hand sides of (4.34)–(4.36) are identical in form to the right hand sides of (4.18)–(4.20). The only difference is that $\bar{P}, \bar{Q}$, and $\bar{Z}$ have been replaced by $\bar{P}^{(j)}, \bar{Q}^{(j)}$ and $\bar{Z}^{(j)}$

Derivatives with respect to $a_{c,k\ell}$

Differentiating (4.18)–(4.20) with respect to $a_{c,k\ell}(=\theta_j)$ gives

$$\frac{\partial H_{A_c}}{\partial a_{c,k\ell}} = H'_{A_c}(\bar{Z}^{(j)}) \tag{4.37}$$

$$\frac{\partial H_{B_c}}{\partial a_{c,k\ell}} = H'_{B_c}(\bar{P}^{(j)}, \bar{Z}^{(j)}) \tag{4.38}$$

$$\frac{\partial H_{C_c}}{\partial a_{c,k\ell}} = H'_{C_c}(\bar{Q}^{(j)}, \bar{Z}^{(j)}). \tag{4.39}$$

Derivatives with respect to $b_{c,k\ell}$

Differentiating (4.18)–(4.20) with respect to $b_{c,k\ell}(=\theta_j)$ gives

$$\frac{\partial H_{A_c}}{\partial b_{c,k\ell}} = H'_{A_c}(\bar{Z}^{(j)}) \tag{4.40}$$

$$\frac{\partial H_{B_c}}{\partial b_{c,k\ell}} = H'_{B_c}(\bar{P}^{(j)}, \bar{Z}^{(j)}) + 2\bar{P}_{22} E_{n_c \times n_y}^{(k,\ell)} V_2 \tag{4.41}$$

$$\frac{\partial H_{C_c}}{\partial b_{c,k\ell}} = H'_{C_c}(\bar{Q}^{(j)}, \bar{Z}^{(j)}) - 2D^{\mathrm{T}} E_{n_y \times n_c}^{(\ell,k)} \bar{Z}_{22}^{\mathrm{T}}. \tag{4.42}$$

Derivatives with Respect to $c_{c,k\ell}$

Differentiating (4.18)–(4.20) with respect to $c_{c,k\ell}(=\theta_j)$ gives the following.

$$\frac{\partial H_{A_c}}{\partial c_{c,k\ell}} = H'_{A_c}(\bar{Z}^{(j)}) \tag{4.43}$$

$$\frac{\partial H_{B_c}}{\partial c_{c,k\ell}} = H'_{B_c}(\bar{P}^{(j)}, \bar{Z}^{(j)}) - 2\bar{Z}_{22}^{\mathrm{T}} E_{n_c \times n_u}^{(\ell,k)} D^{\mathrm{T}} \tag{4.44}$$

$$\frac{\partial H_{C_c}}{\partial c_{c,k\ell}} = H'_{C_c}(\bar{Q}^{(j)}, \bar{Z}^{(j)}) + 2R_2 E_{n_u \times n_c}^{(k,\ell)} \bar{Q}_{22}. \tag{4.45}$$

Derivatives with respect to λ

Differentiating (4.18)–(4.20) with respect to λ gives

$$\frac{\partial H_{A_c}}{\partial \lambda} = H'_{A_c}(\dot{\bar{Z}}) \tag{4.46}$$

$$\frac{\partial H_{B_c}}{\partial \lambda} = H'_{B_c}(\dot{\bar{P}}, \dot{\bar{Z}}) + 2(\bar{P}_{21}^{\mathrm{T}} \dot{V}_{12} + \bar{P}_{22} B_c \dot{V}_2 + \bar{Z}_{12}^{\mathrm{T}} \dot{C}^{\mathrm{T}} - \bar{Z}_{22}^{\mathrm{T}} C_c^{\mathrm{T}} \dot{D}^{\mathrm{T}}) \tag{4.47}$$

$$\frac{\partial H_{C_c}}{\partial \lambda} = H'_{C_c}(\dot{\bar{Q}}, \dot{\bar{Z}}) + 2(-\dot{R}_{12}^{\mathrm{T}} \bar{Q}_{12} + \dot{R}_2 C_c \bar{Q}_{22} + \dot{B}^{\mathrm{T}} \bar{Z}_{21}^{\mathrm{T}} - \dot{D}^{\mathrm{T}} B_c^{\mathrm{T}} \bar{Z}_{22}^{\mathrm{T}}) \tag{4.48}$$

where from (4.4)–(4.12)

$$\begin{bmatrix} \dot{A} & \dot{B} \\ \dot{C} & \dot{D} \end{bmatrix} = \begin{bmatrix} A_f - A_0 & B_f - B_0 \\ C_f - C_0 & C_f - C_0 \end{bmatrix} \tag{4.49}$$

$$\begin{bmatrix} \dot{R}_1 & \dot{R}_{12} \\ \dot{R}_{12}^{\mathrm{T}} & \dot{R}_2 \end{bmatrix} = \dot{L_R} {L_R}^{\mathrm{T}} + L_R \dot{L_R}^{\mathrm{T}} \tag{4.50}$$

where

$$\dot{L_R} = L_{R,f} - L_{R,0} \tag{4.51}$$

$$\begin{bmatrix} \dot{V}_1 & \dot{V}_{12} \\ \dot{V}_{12}^{\mathrm{T}} & \dot{V}_2 \end{bmatrix} = \dot{L_V} {L_V}^{\mathrm{T}} + L_V \dot{L_V}^{\mathrm{T}} \tag{4.52}$$

where

$$L_V = L_{V,f} - L_{V,0}. \tag{4.53}$$

H_θ can now be computed using (4.24) and (4.37)–(4.45).

Note that the calculation of the j^{th} column of H_θ requires the computation of the Lyapunov equations described by (4.28) and (4.29). Significant computational savings can be made by solving these Lyapunov equations in a basis in which the closed-loop state matrix $\tilde{A}$ is nearly diagonal (i.e., a modal form) or nearly block triangular (i.e., a Schur form). This requires transforming the corresponding forcing terms into this basis which can be costly if the dimension of the closed-loop system, $n_{cl}(= n_x + n_c)$ is large. In fact, if the forcing terms are dense, this transformation requires $2n_{cl}^3$ operations. Fortunately, it is seen by (A.20)–(A.28) of Appendix A that these forcing terms are low rank. Hence, these transformations do not have to be expensive and often require only about $2n_{cl}^2$ operations. Computation of the expressions (4.37)–(4.45) requires the solutions of the Lyapunov equations in their original basis. However, it is not efficient to numerically perform this transformation before substituting into (4.37)–(4.45). Instead, symbolic substitution and judicious choice of the order of matrix multiplications can result in significant computational savings. The details of efficient computation of H_θ are presented in Appendix B.

H_λ is computed using (4.25) and (4.46)-(4.48). This requires computation of the Lyapunov equations (4.31) and (4.32). The forcing terms for these Lyapunov equations are not sparse so that computing H_λ in a particular basis requires $2n_{c\ell}^3$ operations to transform the forcing terms. However, the rest of the optimization associated with the computation of H_θ does apply to the computation of H_λ.

4.3 Reduction of the Dimension of the Controller Parameter Vector (θ)

The homotopy function $H(\theta, \lambda)$, described earlier, was defined to solve the H_2 optimal reduced-order dynamic compensation problem. The vector θ was defined such that it contained each of the elements of the controller matrices, A_c, B_c and C_c. However, for computational efficiency it is desired that θ be as small as possible. Hence, we desire to represent the controller matrix with the fewest parameters possible (i.e., we desire θ to have the smallest dimension possible). The minimal number of parameters p_{min} with which a compensator can be represented is given by (Martin and Bryson 1980, Denery 1971)

$$p_{\min} = n_c(n_u + n_y) \tag{4.54}$$

One canonical form which allows representation of a controller with a minimal number of parameters is the modal form described in (Martin and Bryson 1980). This form will be called here the Second-Order Polynomial (SP) form. For this parameterization a triple (A_c, B_c, C_c) has the following structure.

$$A_c = \text{block- diag}\{A_{c,1}, A_{c,2} \ldots, A_{c,r}\} \tag{4.55}$$

where $A_{c,i} \in \mathbb{R}^{2\times 2}$ for $i \in \{1, 2, \ldots, r\}$ and each $A_{c,i}$ (with the exception of $A_{c,r}$ if the row dimension of A_c is odd) has the form

$$A_{c,i} = \begin{bmatrix} 0 & 1 \\ a_{c,i}^{(1)} & a_{c,i}^{(2)} \end{bmatrix} \tag{4.56}$$

to allow for either a complex conjugate set of poles or two real poles. B_c is completely full and

$$C_c = [C_{c,1}, \ C_{c,2}, \ldots, C_{c,r}] \tag{4.60}$$

where $C_{c,i}$ has the form

$$C_{c,r} = \begin{bmatrix} 1 & 0 \\ * & * \\ \vdots & \vdots \\ * & * \end{bmatrix}. \tag{4.57}$$

The controller canonical form described in Kailath 1980 also allows representation of a controller with a minimal number of parameters. For single-input, single-output (SISO) systems in controller canonical form the A_c matrix is a companion matrix. In particular, A_c has the form

$$A_c = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & & 0 \\ 0 & 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ * & * & * & * & \cdots & * \end{bmatrix}. \tag{4.58}$$

In addition,

$$B_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{4.59}$$

and C_c is completely full. A dual form of the controller canonical form is the observable canonical form (Kailath 1980).

It is also possible to represent the controller in a basis where the number of free parameters p satisfies

$$p_{\min} < p < p_{\max} \triangleq n_c(n_c + n_u + n_y). \tag{4.60}$$

One such basis is the tridiagonal basis (Geist 1991, Parlett 1992) in which the controller state matrix is constrained to have nonzero elements only on the diagonal, the super-diagonal and the sub-diagonal. That is,

$$A_c = \begin{bmatrix} * & * & & & \\ * & * & * & & 0 \\ & * & * & & \ddots \\ & & \ddots & \ddots & \\ & & & * & * \end{bmatrix} \tag{4.61}$$

B_c and C_c are completely full. For this form the number of free parameters is given by

$$p = p_{\min} + (3n_c - 2)$$

A common feature of each of the above bases is that they are described by simply constraining certain elements of the controller (or plant) matrices to constant values (e.g., 1 or 0) while allowing the remaining parameters to have arbitrary values (A_c, B_c, C_c). Hence, the corresponding parameter vector (θ_p), gradient vector $(J_{\theta,p})$ and Hessian matrix $(H_{\theta,p})$ are given by

$$\theta_p = \Gamma\theta \tag{4.62}$$

$$J_{\theta,p} = \Gamma J_\theta \tag{4.63}$$

$$H_{\theta,p} = \Gamma H_\theta \Gamma^{\mathrm{T}} \tag{4.64}$$

$$J_{\theta,p} = \Gamma J_{\theta} \tag{4.63}$$

$$H_{\theta,p} = \Gamma H_{\theta} \Gamma^{\mathrm{T}} \tag{4.64}$$

where Γ is an elemental matrix (i.e., each row has only one nonzero element and this element has unity value). It should be noted here that $H_{\theta,p}$ can be computed more efficiently than shown in (4.64). Since it is not necessary to construct the large Hessian H_θ to compute the smaller Hessian $H_{\theta,p}$.

4.4 Overview of the Homotopy Algorithm

This section describes the general logic and features of the homotopy algorithm for H_2 optimal reduced-order control. It is assumed that the designer has supplied a set of system matrices, $S_f = (A_f, B_f, C_f, D_f, R_{1,f}, R_{2,f}, R_{12,f}, V_{1,f}, V_{2,f}, V_{12,f})$ describing the optimization problem whose solution is desired. In addition, it is assumed that the designer has chosen an initial set of related system matrices $S_0 = (A_0, B_0, C_0, D_0, R_{1,0}, R_{2,0}, R_{12,0}, V_{1,0}, V_{2,0}, V_{12,0})$ that has an easily obtained optimal controller $(A_{c,0}, B_{c,0}, C_{c,0})$ of order n_c.

It is always possible to choose the initial system S_0 such that (A_0, B_0, C_0, D_0) in nonminimial with minimal dimension n_c. In this case, it is easy to show that the corresponding LQG compensator has minimal dimension $n_r \leq n_c$ and will usually have minimal dimension $n_r = n_c$. In the latter case, $(A_{c,0}, B_{c,0}, C_{c,0}, D_{c,0})$ is chosen as a minimal realization of the LQG compensator. However, we have seen experimentally that the corresponding homotopy can lead to failure of the homotopy algorithm. Similar observations have been made by Mercadal (Mercadal 1991). In particular, Mercadal has shown that allowing the plant parameters to vary along the homotopy path can lead to the development of destabilizing controllers or path bifurcations.

That the above type of homotopy would cause problems is somewhat intuitive since for a given λ, say $\lambda_1 \in [0,1]$, a controller $(A_c(\lambda_1). B_c(\lambda_1), C_c(\lambda_1))$ that stabilizes the plant $(A(\lambda_1), B(\lambda_1), C(\lambda_1), D(\lambda_1))$ may not stabilize the plant $(A(\lambda_2), B(\lambda_2), C(\lambda_2), D(\lambda_2))$ for $\lambda_2 \neq \lambda_1$. Hence, below we present ways of constructing the initial system S_0 that does *not* require the plant paramaters (A,B,C,D) to vary along the homotopy path. In this case, a controller that stabilizes the plant at λ_1 will also stabilize the plant at $\lambda_2 > \lambda_1$. This argument in itself does *not* ensure that at every step along the homotopy algorithm the controller design remains stabilizing. This is a subject that requires further research. It should be mentioned that another advantage of a homotopy that varies only the performance weights $(R_1, R_2, R_{12}, V_1, V_2, V_{12})$ is that the optimal controller at each point is optimal with respect to the real nominal plant (A_f, B_f, C_f, D_f).

Now, we present three options for constructing S_0 and hence defining the homotopy.

Option 1. One alternative for constructing S_0 is to choose A_0 to be stable (e.g., if A_f is stable, let $A_0 = A_f$ or if A_f is unstable, let $A_0 = A_f - \sigma I$ where σ is sufficiently large to ensure stability of A_0), and let either $R_{1,0}$ or $V_{1,0}$ be zero with all other parameters equal to their final values. In this case $(A_{c,0}, B_{c,0}, C_{c,0})$ is chosen such that it's input-output map is zero, i.e., $C_{c,0}(sI_{n_c} - A_{c,0})^{-1}B_{c,0}$ = 0.

Option 2. Another alternative is to choose A_0 to be stable and as elaborated in (Collins, Haddad, and Ying 1993) and choose either $(R_{1,0}, V_{2,0})$ or $(V_{1,0}, R_{2,0})$ as given below. (Again, all other initial parameters are equal to their final values.)

(i) In a basis in which

$$A_0 = \begin{bmatrix} (A_0)_{11} & 0 \\ (A_0)_{21} & (A_0)_{22} \end{bmatrix}, \quad (A_0)_{11} \in \mathrm{I\!R}^{n_c \times n_c}, \tag{4.65}$$

choose $R_{1,0}$ to be of the form

$$R_{1,0} = \begin{bmatrix} (R_{1,0})_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad (R_{1,0})_{11} \in \mathrm{I\!R}^{n_c \times n_c} \tag{4.66}$$

and for some positive scalar α choose

$$V_{2,0} = \alpha V_{2,f} \tag{4.67}$$

(ii) In a basis in which

$$A_0 = \begin{bmatrix} (A_0)_{11} & (A_0)_{12} \\ 0 & (A_0)_{22} \end{bmatrix}, \quad (A_0)_{11} \in \mathrm{I\!R}^{n_c \times n_c}, \tag{4.68}$$

choose $V_{1,0}$ to be of the form

$$V_{1,0} = \begin{bmatrix} (V_{1,0})_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad (V_{1,0})_{11} \in \mathrm{I\!R}^{n_c \times n_c} \tag{4.69}$$

and for some positive scalar α choose

$$R_2 = \alpha R_{2,f} \tag{4.70}$$

As discussed in (Collins, Haddad, and Ying 1993), α in (4.67) or (4.70) can always be chosen sufficiently large that the corresponding LQG compensator is nearly nonminimal. In this case, a very close approximation to $(A_{c,0}, B_{c,0}, C_{c,0})$ is easily obtained by reducing the LQG compensator to it's (nearly) minimal realization using an appropriate technique such as balanced controller reduction (Yousuff and Skelton 1984). This initialization option can sometimes present a shorter path to optimal solution than the first option given above.

Option 3. A third alternative (which does not require A_0 to be stable) is based on the following experimental observation. The initial system can be chosen to correspond to a low authority control problem, e.g., one can choose

$$R_{2,0} = \alpha R_{2,f}, \quad V_{2,0} = \beta V_{2,f} \tag{4.71}$$

with α and β large and let all other initial system parameters equal their final values. In this case it has been observed that the reduced-order controller $(A_{c,r}, B_{c,r}, C_{c,r})$ obtained by suboptimal reduction of an LQG controller will often yield virtually the same cost as the LQG controller (see, e.g., De Villemagne and Skelton 1988), hence indicating that $(A_{c,r}, B_{c,r}, C_{c,r})$ is nearly optimal. In this case we choose $(A_{c,0}, B_{c,0}, C_{c,0}) = (A_{c,r}, B_{c,r}, C_{c,r})$. It should be noted that these observations are partially (but not fully) explained by the results of (Collins, Haddad, and Ying 1993).

Below, we present an outline of the homotopy algorithm. This algorithm describes a predictor/corrector numerical integration scheme. There are several options to be chosen initially. These options are enumerated before presenting the actual algorithm. Notice that each option corresponds to a particular flag being assigned some integer value.

Controller Basis Options:

$basis = 0$. No basis (i.e., all elements of the controller matrices are considered free.)

$basis = 1$. Tridiagonal Basis.

$basis = 2$. Second-Order Polynomial Form.

$basis = 3$. Controllable Canonical Form.

Note that for $basis = 0$ or 1, $p > p_{\min}$ while for $basis = 2$ or 3, $p = p_{\min}$.

Prediction Scheme Options:

Here we use the notation that λ_0, λ_{-1}, and λ_1 represent the values of λ at respectively the current point on the homotopy curve, the previous point and the next point. Also, $\theta_p{}' = d\theta_p/d\lambda$ and is the solution of Davidenko's differential equation (4.7), rewritten here as

$$H_{\theta,p}\theta_p'(\lambda) + H_\lambda = 0. \tag{4.72}$$

If $p = p_{\min}$, $H_{\theta,p}$ is generally invertible, then $\theta_p'(\lambda)$ is given exactly by

$$\theta_p'(\lambda) = -H_{\theta,p}^{-1}H_\lambda. \tag{4.73}$$

If $p > p_{\min}$, then $H_{\theta,p}$ generally has rank $p_{\min}$ and $\theta_p'(\lambda)$ is approximated by the least squares solution of (4.73) or

$$\theta_p' = -V\begin{bmatrix} \Sigma_0^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^{\mathrm{T}} H_\lambda \tag{4.74}$$

where it is assumed the $H_{\theta,p}$ has the singular value decomposition

$$H_{\theta,p} = U\begin{bmatrix} \Sigma_0 & 0 \\ 0 & 0 \end{bmatrix} V^{\mathrm{T}}, \quad \Sigma_0 \in \mathbb{R}^{p_{\min} \times p_{\min}}. \tag{4.75}$$

Note that for $p = p_{min}$ (4.73) and (4.74) are equivalent.

$pred = 0$. No prediction. This option assumes that $\theta_p(\lambda_1) = \theta_p(\lambda_0)$.

$pred = 1$. Linear prediction. This option assumes predicts $\theta_p(\lambda_1)$ using only $\theta_p(\lambda_0)$ and $\theta_p{}'(\lambda_0)$. In particular,

$$\theta_p(\lambda_1) = \theta_p(\lambda_0) + (\lambda_1 - \lambda_0)\theta_p{}'(\lambda_0) \tag{4.76}$$

$pred = 2$. Cubic spline prediction. This option predicts $\theta_p(\lambda_1)$ using $\theta_p(\lambda_0)$, $\theta_p{}'(\lambda_0)$, $\theta_p(\lambda_{-1})$ and $\theta_p{}'(\lambda_{-1})$. In particular,

$$\theta_p(\lambda_1) = a_0 + a_1\lambda_1 + a_2{\lambda_1}^2 + a_3{\lambda_1}^3 \tag{4.77}$$

where a_0, a_1, a_2 and a_3 are computed by solving

$$[a_0 \;\; a_1 \;\; a_2 \;\; a_3]\begin{bmatrix} 1 & 0 & 1 & 0 \\ \lambda_{-1} & 1 & \lambda_0 & 1 \\ \lambda_{-1}^2 & 2\lambda_{-1} & \lambda_0^2 & 2\lambda_0 \\ \lambda_{-1}^3 & 3\lambda_{-1}^2 & \lambda_0^3 & 3\lambda_0^2 \end{bmatrix} = \begin{bmatrix} \theta_p(\lambda_{-1}) \\ \theta_p'(\lambda_{-1}) \\ \theta_p(\lambda_0) \\ \theta_p'(\lambda_0) \end{bmatrix} \tag{4.78}$$

Note that if this option is chosen, then at the initial algorithm prediction step $\theta_p(\lambda_{-1})$ and $\theta_p'(\lambda_{-1})$ are not available, in which case linear prediction is used.

Correction Options:

Here we assume that the homotopy parameter has a fixed value λ_0. The vector θ_p represents the current approximation of the parameter vector at $\lambda = \lambda_0$. Each of the options corresponds to updating θ_p using the formula

$$\theta_p \leftarrow \theta_p + \Delta\theta_p \tag{4.79}$$

where

$$\Delta\theta_p = -G_{\theta,p}J_{\theta,p} \tag{4.80}$$

for some choice of $G_{\theta,p}$.

corr = 1. Newton correction. In this option, if $p = p_{\min}$,

$$G_{\theta,p} = H_{\theta,p}{}^{-1} \tag{4.81}$$

while if $p > p_{\min}$,

$$G_{\theta,p} = V(\Sigma^2 + \alpha^2 I)^{-1}\Sigma U^{\mathrm{T}} \tag{4.82}$$

where α is some (small) scalar and (U, V, Σ) denote the singular value decomposition of $H_{\theta,p}$ such that

$$H_{\theta,p} = U\Sigma V^{\mathrm{T}}. \tag{4.83}$$

It can be shown that if $G_{\theta,p}$ is given by (4.82), then $\Delta\theta_p$ minimizes

$$\frac{1}{2}[||H_{\theta,p}\Delta\theta_p + \theta_p||^2 + \alpha^2||\Delta\theta_p||^2]. \tag{4.84}$$

Hence, $\Delta\theta_p$ is essentially a "Newton correction" that is relatively insensitive to singularities or near singularities in the Hessian, $H_{\theta,p}$.

corr = 2. Quasi-Newton correction. In this option, $G_{\theta,p}$ denotes the estimate of $H_{\theta,p}^{-1}$ using only gradient and cost information. For the algorithm presented here the BFGS inverse Hessian update is used (Fletcher 1987).

Outline of the Homotopy Algorithm

Step 1. If *basis* ≥ 1, then transform the initial controller $(A_{c,0}, B_{c,0}, C_{c,0})$ to the chosen basis and let θ_0, p be the corresponding vector of free parameters.

Step 2. Initialize *loop* $= 0$, $\lambda = 0$, $\Delta\lambda \in (0,1]$, $S = S_0$, $\theta_p = \theta_{0,p}$ and compute the cost J and the cost gradient $J_{\theta,p}$ corresponding to S and the controller described by θ_p.

Step 3. Let loop = loop+1. If loop = 1, then go to Step 5. Else, continue.

Step 4. Advance the homotopy parameter and predict the corresponding parameter vector θ as follows.

4a. Let $\lambda_0 = \lambda$

4b. Let $\lambda = \lambda_0 + \Delta\lambda$.

4c. If *pred* ≥ 1, then compute $\theta'_p(\lambda_0)$.

4d. Predict $\theta_p(\lambda)$ by using the option defined by *pred*.

4e. If the normalized gradient $J_{\theta,p}\|G_{\theta,p}\|/\|\theta_p\|$ satisfies some preassigned tolerance, then continue. Else, reduce $\Delta\lambda$ and go to Step 4b.

Step 5. Correct the current approximation θ_p to the optimization problem defined by S using the option defined by *corr* until the normalized gradient,

$$\frac{J_{\theta,p}\|G_{\theta,p}\|}{\|\theta_p\|} \tag{4.85}$$

satisfies some preassigned tolerance.

Step 6. If $\lambda = 1$, then stop. Else, go to Step 3.

The above algorithm assumes monotonicity of the solution curve as a function of the homotopy parameter λ. However, it is not difficult to modify the algorithm so that the variable parameter is the arc length as discussed in Watson 1986 and Watson 1987 since this modification would still only require the computation of H_θ and H_λ. The modified algorithm would not require monotonicity of the solution curve. However, so far in our computational experience the solution curve has always been monotonic.

Note that if $p = p_{\min}$ and *corr* $= 1$, then the corrections of Step 5 correspond to Newton corrections. Hence if the prediction tolerance used in Step 4 is sufficiently small, then, entering

Step 5, θ_p will be close enough to the optimal value θ_p^* so that the quadratic convergence properties of Newton's method (Fletcher 1987) can be realized. In practice, this quadratic convergence property is not always realized due to numerical ill-conditioning associated with the minimal parameterization of the controller. This ill-conditioning is illustrated and discussed further below.

5. Illustration of Reduced-Order Design Using a Four Disk Example

This section illustrates the homotopy algorithm of Section 5 by considering control design for an axial beam with four disks attached as shown in Figure 5.1. This example was derived from a laboratory experiment described in (Cannon and Rosenthal 1984) and has been considered in several subsequent publications [Anderson and Liu 1989, De Villemagne and Skelton 1988, Liu, Anderson, and Ly 1990, Hyland and Richter 1990). The basic control objective for the four-disk problem is to control the angular displacement at the location of disk 1 using a torque input at the location of disk 3. It is also assumed that a torque disturbance enters the system at the location of disk 3.

The design philosophy adopted here is that the scaling $q2$ of the nominal control weight $R_{2,0} = 1$ and the nominal sensor noise intensity $V_{2,0} = 1$ are simply design knobs used to determine the control authority. (Hence, $R_2(\lambda) = q_2(\lambda)R_{2,0}$ and $V_2(\lambda) = q_2(\lambda)V_{2,0}$.) The system costs are computed assuming $V_2 = 0$ although $V_2 = 0$ is *not* assumed in the design process. This general philosophy is actually motivated by insights into LQG theory. However, it will suffice here to simply note that this philosophy was used successfully on two hardware experiments involving control design and implementation [Collins, Phillips, and Hyland 1991, Collins, King, Phillips and Hyland 1992). It should be noted that these assumptions do *not* influence the qualitative results described below.

Below, we will compare various algorithm options. In particular, we desire to illustrate the types of convergence that are sometimes achieved when various bases are used to represent the controller, and the speed of the algorithm when various prediction options are used. We will also investigate what type of convergence and speed are achieved when H_θ^{-1}, the inverse of the Hessian of the cost is not computed explicitly but is estimated using a Quasi-Newton method. The comparisons are all based on a MATLAB implementation of the algorithm and the program in each case was run on a 486, 33 MHz PC.

We choose to base the comparison on the design of an 8th order controller (for the 8th order design plant). Of course, we can solve for optimal full-order controllers using Riccati equations but we choose this order controller because experientially we have seen that the higher the order of the controller the more the algorithm struggles when a particular basis is chosen for the controller. Hence, we are essentially basing our comparisons on the worst-case controller order for this particular design model. The controller that is used to initialize the algorithm is the LQG controller corresponding to the choice $q_2 = 1$. The algorithm is used to deform this controller into the higher authority controller corresponding to $q_2 = 0.1$.

Table 5.1 shows a comparison of the algorithm when various bases are chosen for the controller. Linear prediction is used in each case. In fact, it was seen experimentally that if cubic spline prediction was used, the algorithm performance degraded if an over-parameterized controller basis (i.e., tridiagonal basis or no basis) was used. This phenomenon is almost certainly due to the fact that in these cases the tangent vectors $(\theta'_p(\lambda))$ are only estimated using (4.71) and hence are not accurate. As evidenced from Table 5.1, the performance of the controllable canonical form was worse in terms of clock time and minimum and maximum step size. The minimum step size of 7.8e-16 indicates substantial ill-conditioning along the homotopy path. For this example, the second-order polynomial form required the least number of flops although it did require slightly more clock time than the tridiagonal basis. In terms of minimum and maximum step size, the choice of no controller basis was better conditioned than restriction to any of the bases.

Controller Basis	Megaflops	Real Time (sec.)	No. Hessian Calculations	Minimum Step Size	Maximum Step Size
None	1098	1098.2	47	0.01	0.32
Tridiagonal	590	880.7	120	0.0003	0.08
SPF	518	930.4	283	0.0001	0.04
CCF	828	1524.7	461	7.8e−16	0.02

Table 5.1. Comparison of Controller Basis Options

Table 5.2 shows a comparison of the algorithm when the second-order polynomial form was chosen for the controller and various prediction options were used. Notice that in terms of real time, linear prediction required only 17.8% of the time required when no prediction was used. Cubic spline prediction required only 5.6% of the time required when no prediction was used. The ability to predict along the curve described by the changing parameters is one of the practical benefits of formulating an optimization problem formally in terms of a homotopy.

Prediction Option	Megaflops	Real Time (sec.)	No. Hessian Calculations	Minimum Step Size	Maximum Step Size	Correction Tolerance
None	3560	5205.0	1552	6e−15	0.01	10^{-4}
Linear	518	930.4	283	1.5e−4	0.04	10^{-4}
Cubic	160	293.2	86	0.01	0.08	10^{-6}

Table 5.2. Comparison of Prediction Options for SPF Controller Basis

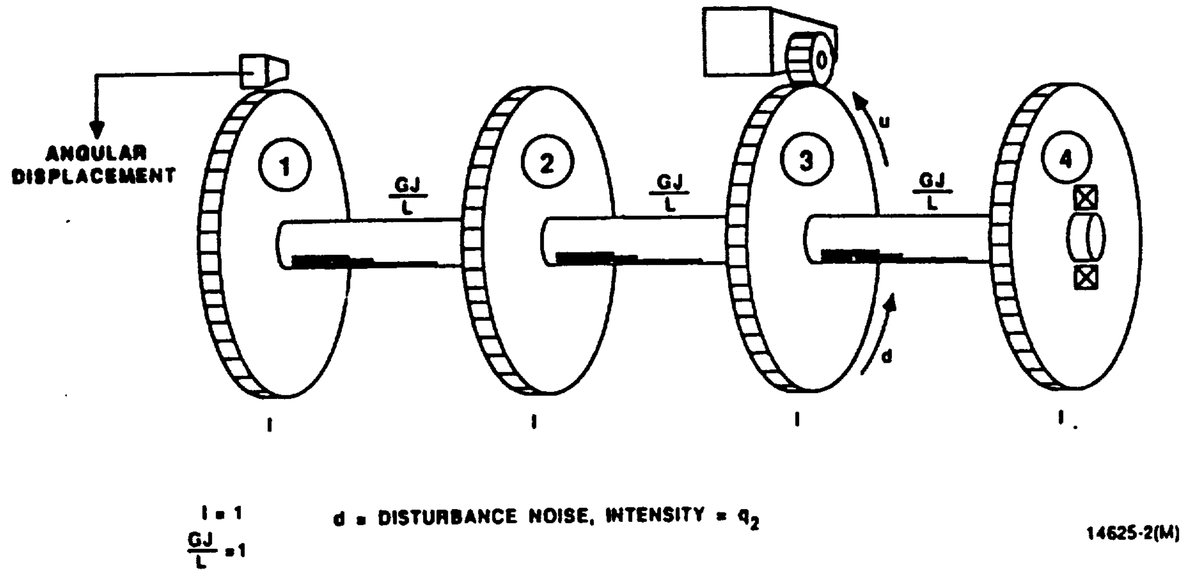

Figure 5.1. The Four Disk Model.

Table 5.3 shows a comparison of the algorithm when the second-order polynomial form was chosen for the controller, H_θ^{-1} was estimated using a Quasi-Newton (in particular BFGS) method and various prediction options were used. The "*" under the Megaflop heading indicates that the MATLAB flop counter overflowed and so the flop data is unavailable. Notice that when the Quasi-Newton method was used, the prediction did not help. This is because of the inaccuracies in the tangent vectors due to the errors in the estimate of the inverse Hessian. Also note that by comparing Table 5.2 with Table 5.3, the behavior of the Quasi-Newton method was substantially worse than the behavior of the algorithm when the Hessian inverse was calculated exactly. In fact the best clock time for the Quasi-Newton method was 27 times slower than the best clock time when the inverse Hessian was calculated exactly.

Prediction Option	Megaflops	Real Time (sec.)	Minimum Step Size	Maximum Step Size
None	*	7960.3	1.0e–14	0.01
Linear	*	8011.4	1.0e–14	0.01
Cubic	*	8902.1	1.0e–14	0.01

Table 5.3. Comparison of Prediction Options for SPF Controller Basis with Quasi-Newton Approximation to Inverse Hessian

Figures 5.2 through 5.4 consider respectively the design of 2nd, 4th and 6th order controllers for authority levels corresponding to $q_2 \in (1, 0.1, 0.01, ...1.0e-6)$ and compare the optimal curves for an LQG controller, a reduced-order controller obtained by balancing and an optimal reduced-order controller. In each case, the optimal reduced-order controller performs substantially better than the balanced controller as the authority level is increased (i.e., q_2 is decreased). At low authority, the cost curves of the balanced and optimal controllers coincided, indicating that the two controllers are probably very similar. In fact the low authority balanced controllers were used to initialize the homotopy algorithm in the design of the optimal reduced order controllers as discussed in Option 3 of Subsection 4.4. Figure 5.5 compares the optimal controllers of various orders. This type of figure can be used in practice to determine the order of the controller to be implemented.

6.0 Conclusions

This paper has presented a new homotopy algorithm for the design of H_2 optimal reduced-order controllers. The example of the previous section illustrated some of the features of the various algorithm options. For the test case considered, the option of estimating the inverse Hessian (H_θ^{-1}) via a Quasi-Newton method performed considerably worse than the option of actually computing the Hessian and inverting it. The results also show the ill-conditioning that can occur when a particular basis is chosen for the controller. For example, the second-order polynomial form was particularly ill-conditioned for the test case. In addition, the tridiagonal basis, which over-parameterizes the compensator, actually outperformed the second-order polynomial form in terms of clock time required.

This ill-conditioning is not new. It is well known that restriction to a particular controller basis can cause numerical ill-conditioning or even instability (Kuhn and Schmidt 1987, Ge, Collins, Watson, and Davis). At least two solutions are possible. One is to have a family of minimal controller bases and have the algorithm switch to the basis that is best conditioned (Kuhn and Schmidt 1987,Ge, Collins, Watson, and Davis). Besides the second-order polynomial form and the controller canonical form mentioned here, another basis that could be included in this family is the input normal Riccati basis of (Davis, Collins, and Hodel 1992). As observed here, one can also use a slightly over parameterized controller basis such as the tridiagonal form. However, even these bases will not always be well-conditioned. One other option is to augment the cost function with a term that includes the squares of the free controller elements (Kuhn and Schmidt 1987). Unfortunately, this alternative requires a cost function that is not well motivated physically. In our opinion, finding practical solutions to ill-conditioning is the fundamental problem in the numerical computation of optimal reduced-order controllers.

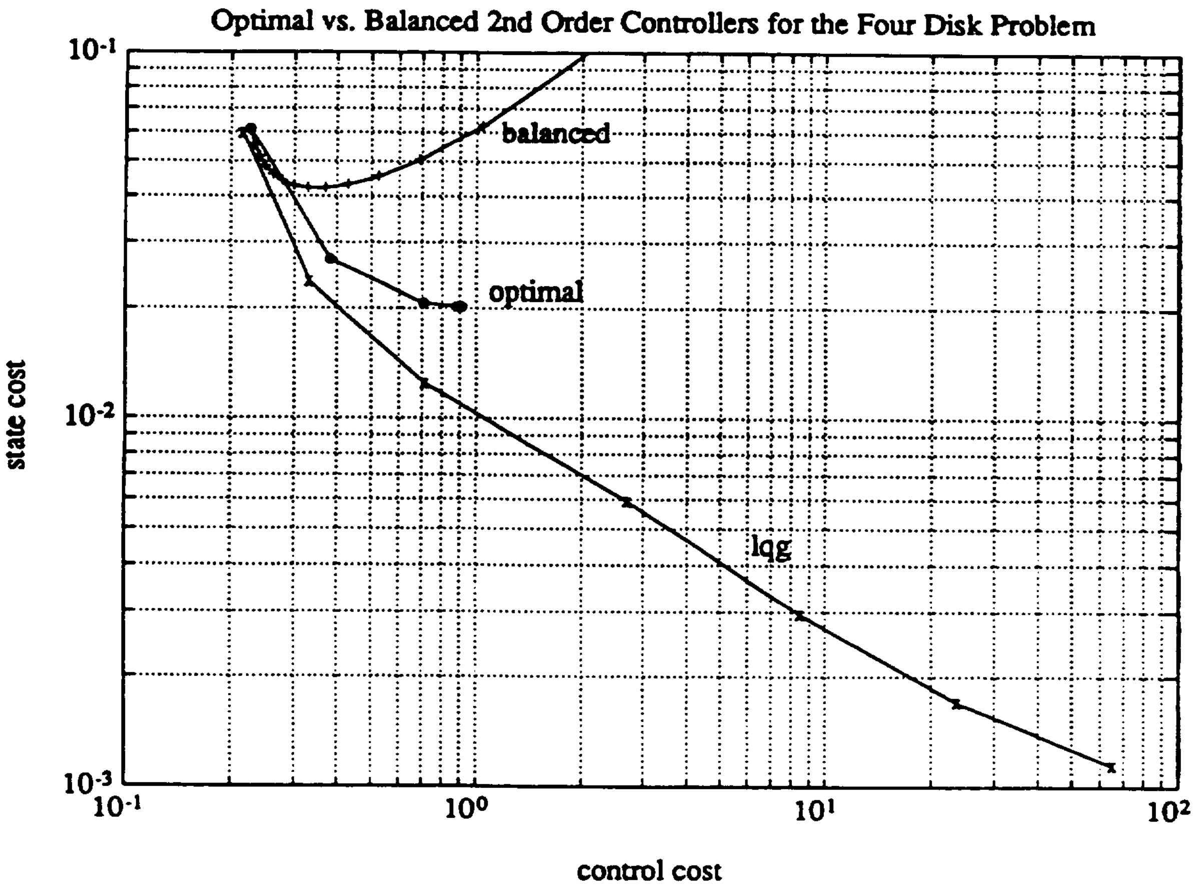

Figure 5.2. Performance Curves for the 2nd Order Controllers.

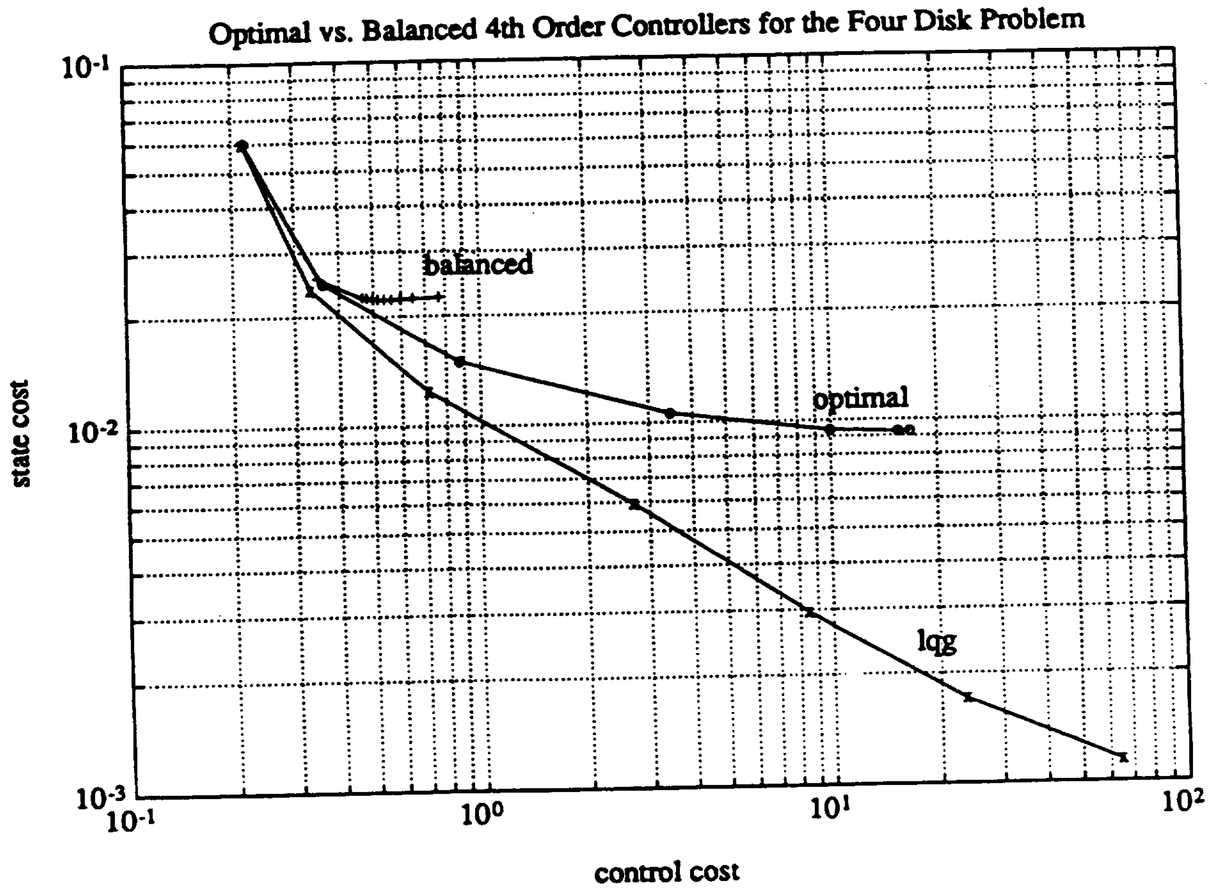

Figure 5.3. Performance Curves for the 4th Order Controllers.

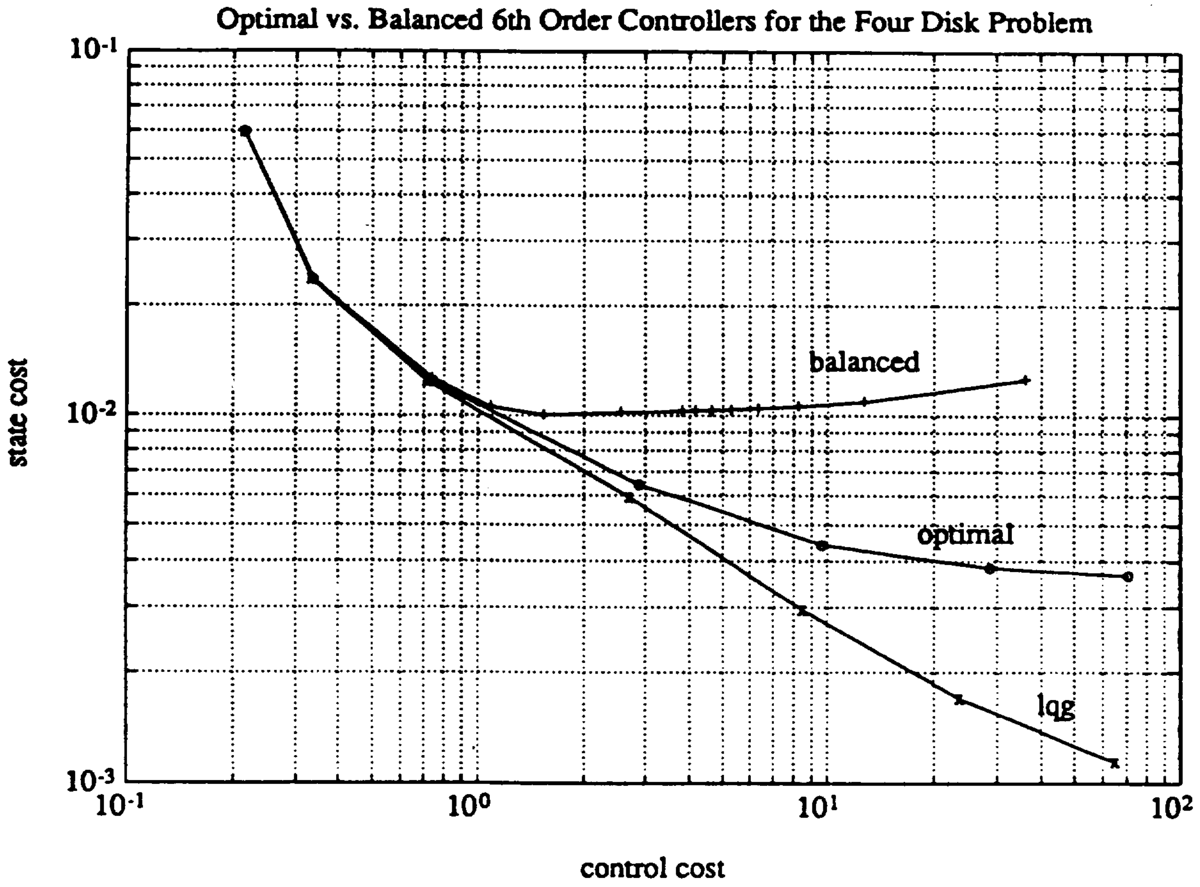

Figure 5.4. Performance Curves for the 6^{th} Order Controllers.

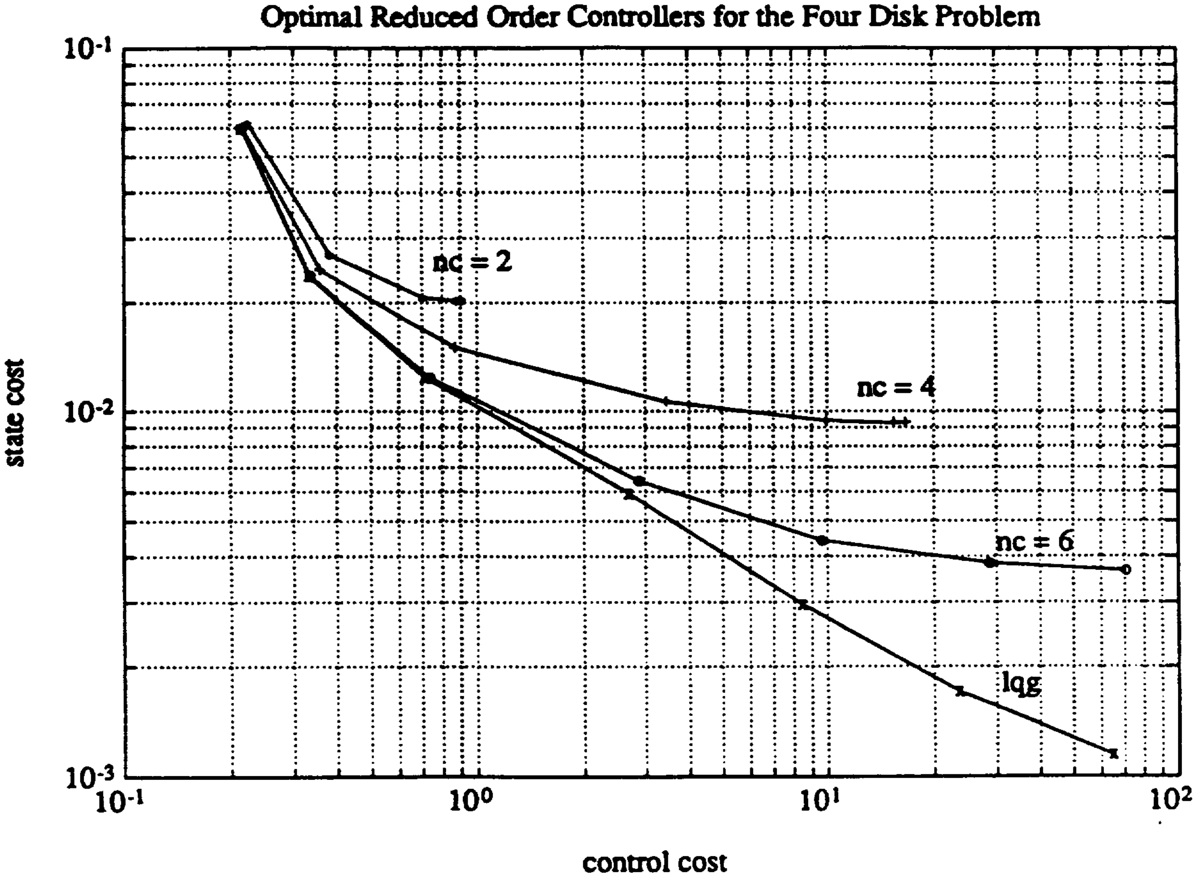

Figure 5.5. Comparison of the Performance Curves for Various Order Controllers

References

Alexander, J. C., and Yorke, J. A., 1978, The homotopy continuation method: numerically implementable topological procedures. *Transactions of the American Mathematical Society,* **242**, 271–284.

Anderson, B. D. O. and Liu, Y., 1989, Controller reduction: concepts and approaches. *IEEE Transactions on Automatic Control,* **34**, 802–812.

Athans, M., 1971, The role and use of the stochastic linear-quadratic-gaussian problem in control system design. *IEEE Transaction on Automatic Control,* **16**, 529-552.

Avila, J. H., 1974, The feasibility of continuation methods for nonlinear equations. *SIAM Journal of Numerical Analysis,* **11**, 104–144.

Cannon, R. H., and Rosenthal, D. E., 1984, Experiments in control of flexible structures with noncolocated sensors and actuators. *AIAA Journal of Guidance, Control and Dynamics,* **7**, 546-553.

Collins, E. G., Jr., Phillips, D. J., and Hyland, D. C., 1991, Robust decentralized control laws for the ACES Structure. *Control Systems Magazine*, 62-70.

Collins, E. G. Jr., King, J. A., Phillips, D. J. and Hyland, D. C., 1992, High performance, accelerometer-based control of the Minimast Structure. *AIAA Journal of Guidance Control and Dynamics*, 885–892.

Collins, E. G. Jr., Haddad, W. M., and Ying, S. S., 1993, Construction of low authority, nearly non-minimal LQG compensators for reduced-order control design. Submitted to *IEEE Transactions on Automatic Control and the 1994 American Control Conference.*

Davis, L. D., Collins, E. G., Jr., and Hodel, A. S., 1992, A parameterization of minimal plants. *Proceedings of the American Control Conference*, 2639-2643, Chicago, Il.

Denery, D. G., 1971, An identification algorithm that is insensitive to initial parameter estimates. *AIAA Journal,* **9**, 371-377.

De Villemagne, C. and Skelton, R.E., 1988, Controller reduction using canonical interactions. *IEEE Transactions on Automatic Control,* **33**, 740-750.

Dunyak, J. P., Junkins, J. L., and Watson, L. T., 1984, Robust nonlinear least squares estimation using the Chow-Yorke homotopy method. *Journal of Guidance, Control and Dynamics,* 7, 752–755.

Eaves, B. C., Gould, F. J., Peoitgen, J. O., and Todd, M. J., 1983, *Homotopy Methods and Global Convergence,* Plenum Press, New York.

Fletcher, R., 1987, *Practical Methods of Optimization: Second Edition*, John Wiley and Sons, New York.

Garcia, C. B., and Zangwill, W. I., 1981, *Pathways to Solutions, Fixed Points and Equilibria,* Prentice-Hall, Englewood Cliffs, NJ.

Ge, Y., Collins, E. G., Jr., Watson, L. T., and Davis, L. D., Minimal parameter homotopies for the L^2 optimal model order reduction problem. submitted to *IEEE Transactions on Automatic Control.*

Geist, G. A., 1991, Reduction of a general matrix to tridiagonal form. *SIAM Journal of Matrix Analysis and Applications,* **12**, 362-373.

Haddad, W. M., 1987, *Robust Optimal Projection Control-System Synthesis,* Ph.D. Thesis, Florida Institute of Technology, Melbourne, FL.

Hyland, D. C. and Bernstein, D. S., 1984, The optimal projection equations for fixed-order dynamic compensation. *IEEE Transactions on Automatic Control,* **29**, 1034-1037.

Hyland, D. C. and Richter, S., 1990, On direct versus indirect methods for reduced-order controller design. *IEEE Transactions on Automatic Control,* **35**, 377-379.

Kabamba, P. T., Longman, R. W., and Jian-Guo, S., 1987, A homotopy approach to the feedback stabilization of linear systems. *AIAA Journal of Guidance, Control and Dynamics,* **10**, 422-432.

Kailath, T., 1980, *Linear Systems*, Prentice-Hall, NJ.

Horta, L. G., Juang, J.-N., and Junkins, Journal L., 1986, A sequential linear optimization approach for controller design. *AIAA Journal of Guidance, Control and Dynamics,* **9**, 699–703.

Kramer, F. S., and Calise, A. Journal, 1988, Fixed-order dynamic compensation for multivariable linear systems. *AIAA Journal of Guidance, Control and Dynamics*, **11**, 80-85.

Kuhn, U. and Schmidt, G., 1987, Fresh look into the design and computation of optimal output feedback controls for linear multivariable systems. *International Journal of Control,* **46**, 75–95.

Kwakernaak, H., and Sivan, R., 1972, Linear optimal control systems. John Wiley and Sons, New York.

Lefebvre, S., Richter, S., DeCarlo, R., 1985, A continuation algorithm for eigenvalue assignment by decentralized constant-output feedback. *International Journal of Control*, **41**, 1273–1292.

Levine, W. S., Johnson, T. L., and Athans, M., 1971, Optimal limited state variable feedback controllers for linear systems. *IEEE Transactions on Automatic Control*, **16**, 785-793.

Liu, Y., Anderson, B. D. O. and Ly, U-L., 1990, Coprime factorization controller reduction with bezout identity induced frequency Weighting. *Automatica*, **26**, 233-249.

Lloyd, N. G., 1978, *Degree Theory*, Cambridge University Press, London.

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

Makila, P. M., Toivonen, H. T., 1987, Computational methods for parametric LQ problems–a survey. *IEEE Transactions on Automatic Control*, **32**, 658-671.

Martin, G. D., and Bryson, A. E., Jr., 1980, Attitude control of a flexible spacecraft. *AIAA Journal of Guidance, Control, and Dynamics* **3**.

Mercadal, M., 1991, Homotopy approach to optimal, linear quadratic, fixed architecture compensation. *Journal of Guidance, Control and Dynamics*, **14**, 1224–1233.

Mukhopadhyay, V., Newsom, Journal R., and Abel, I., 1982, Reduced-order optimal control law synthesis for flutter suppression. *AIAA Journal of Guidance, Control and Dynamics*, **4**, 389-395.

Mukhopadhyay, V., 1987, Stability robustness improvement using constrained optimization techniques. *AIAA Journal of Guidance Control and Dynamics*, **10**, 172-177.

Mukhopadhyay, V., 1989, Digital robust control law synthesis using constrained optimization. *AIAA Journal of Guidance, Control and Dynamics*, **12**, 175-181.

Parlett, B. N., 1992, Reduction to tridiagonal form and minimal realizations. *SIAM Journal of Matrix Analysis and Applications*, 567–593.

Rakowska, J., Haftka, R. T., and Watson, L. T., 1991, "Tracing the Efficient Curve for Multi-

Objective Control-Structure Optimization," *Comput. Systems. Engrg.*, **2**, 461–472.

Richter, S., 1987, A homotopy algorithm for solving the optimal projection equations for fixed-order dynamic compensation: existence, convergence and global optimality. *Proceedings of the American Control Conference,* 1527-1531, Minneapolis, MN.

Richter, S. and Collins, E. G. Jr., 1989, A homotopy algorithm for reduced-order controller design using the optimal projection equations. *Proc. IEEE Conference on Decision and Control,* 932-933.

Richter, S. L., and DeCarlo, R. A., 1983, Continuation methods: theory and applications. *IEEE Transactions on Circuits and Systems*, **30**, 347–352.

Richter, S., and DeCarlo, R., 1984, A homotopy method for eigenvalue assignment using decentralized state feedback. *IEEE Transactions on Automatic Control*, **29**, No. 2, 148–158.

Sebok, D. R., Richter, S. and DeCarlo, R., 1986, Feedback gain optimization in decentralized eigenvalue assignment. *Automatica*, **22**, 433-447.

Shin, Y. S., Haftka, R. T., Watson, L. T., and Plaut, R. H., 1988, Tracking structural optima as a function of available resources by a homotopy method. *Computer Methods in Applied Mechanics and Engineering,* **70**, 151–164.

Sun, J. L., 1990, How to perform differentiations on matrices. *Journal of Guidance, Control and Dynamics*, **31**, 924–925.

Sun, J. L., 1991, Constrained optimal controller design: the next iteration. *Proceedings of the AIAA Guidance, Navigation and Control Conference,* New Orleans, LA, AIAA paper 91-2601.

Turner, Journal D., and Chun, H. M., 1984, Optimal distributed control of a flexible spacecraft during a large-angle maneuver. *Journal of Guidance, Control and Dynamics*, **7**, 257–264.

Wacker, H., 1978, Continuation methods. Academic Press, New York.

Watson, L. T., 1986, Numerical linear algebra aspects of globally convergent homotopy methods. *SIAM Review,* **28**, 529–545.

Watson, L. T., 1987, ALGORITHM 652 HOMPACK: A suite of codes for globally convergent homotopy algorithms. *ACM Transactions on Mathematical Software*, **13**, 281–310.

Yousuff, A. and Skelton, R.E., 1984a, A note on balanced controller reduction. *IEEE Trans-*

actions on Automatic Control, **29**, 254-257.

Yousuff, A., and Skelton, R.E., 1984b, Controller reduction by component cost analysis. *IEEE Transactions on Automatic Control*, **29**, 520–530.

Zigić, D., Watson, L. T., Collins, E. G., Jr., and Bernstein, D. S., 1992, "Homotopy Methods for Solving the Optimal Projection Equations for the H_2 Reduced Order model Problem," *International Journal of Control* , **56**, 173–191.

Zigić, D., Watson, L. T., Collins, E. G., Jr., and Bernstein, D. S., 1993, "Homotopy Approaches to the H_2 Reduced Order Model Problem," *Journal of Mathematical Systems, Estimation, Control*, to appear.

Appendix A: Closed-Loop Matrix Derivatives

In this appendix we present explicit expressions for the derivatives $\frac{\partial \tilde{A}}{\partial \theta_j}$, $\frac{\partial \tilde{R}}{\partial \theta_j}$, $\frac{\partial \tilde{V}}{\partial \theta_j}$, $\frac{\partial \tilde{A}}{\partial \lambda}$, $\frac{\partial \tilde{R}}{\partial \lambda}$, and $\frac{\partial \tilde{V}}{\partial \lambda}$ where

$$\theta = \begin{bmatrix} \mathrm{vec}(A_c) \\ \mathrm{vec}(B_c) \\ \mathrm{vec}(C_c) \end{bmatrix}, \tag{A.1}$$

$$\tilde{A} = \begin{bmatrix} A & -BC_c \\ B_cC & A_c - B_cDC_c \end{bmatrix}, \tag{A.2}$$

$$\tilde{R} = \begin{bmatrix} \tilde{R}_{11} & \tilde{R}_{12} \\ \tilde{R}_{12}^{\mathrm{T}} & \tilde{R}_{22} \end{bmatrix}, \tag{A.3}$$

where

$$\tilde{R}_{11} = R_1 \tag{A.4a}$$

$$\tilde{R}_{12} = -R_{12}C_c \tag{A.4b}$$

$$\tilde{R}_{22} = C_c^{\mathrm{T}} R_2 C_c, \tag{A.4c}$$

and

$$\tilde{V} = \begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{12}^{\mathrm{T}} & \tilde{V}_{22} \end{bmatrix} \tag{A.5}$$

where

$$\tilde{V}_{11} = V_1 \tag{A.6a}$$

$$\tilde{V}_{12} = V_{12}B_c^{\mathrm{T}} \tag{A.6b}$$

$$\tilde{V}_{22} = B_c^{\mathrm{T}} V_2 B_c^{\mathrm{T}}. \tag{A.6c}$$

It is assumed that the plant matrices (A, B, C, D), the cost weighting matrices (R_1, R_{12}, R_2) and the disturbance matrices (V_1, V_{12}, V_2) are the following functions of λ.

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} = +\lambda\left(\begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix}\right) \tag{A.7}$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^{\mathrm{T}}(\lambda) & R_2(\lambda) \end{bmatrix} = L_R(\lambda)L_R^{\mathrm{T}}(\lambda) \tag{A.8a}$$

where

$$L_R(\lambda) = L_{R,0} + \lambda(L_{R,f} - L_{R,0}) \tag{A.8b}$$

and $L_{R,0}$ and $L_{R,f}$ satisfy

$$L_{R,0}L_{R,0}^{\mathrm{T}} = \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix} \tag{A.8c}$$

$$L_{R,f}L_{R,f}^{\mathrm{T}} = \begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^{\mathrm{T}} & R_{2,f} \end{bmatrix} \tag{A.8d}$$

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^{\mathrm{T}}(\lambda) & V_2^{\mathrm{T}}(\lambda) \end{bmatrix} = L_V(\lambda)L_V^{\mathrm{T}}(\lambda) \tag{A.9a}$$

where

$$L_V(\lambda) = L_{V,0} + \lambda(L_{V,f} - L_{V,0}) \tag{A.9b}$$

Below, we use the notation

$$\dot{M} \triangleq \frac{\partial M}{\partial \lambda}. \tag{A.10}$$

Note that from (A.7)-(A.9)

$$\begin{bmatrix} \dot{A} & \dot{B} \\ \dot{C} & \dot{D} \end{bmatrix} = \begin{bmatrix} A_f - A_0 & B_f - B_0 \\ C_f - C_0 & D_f - D_0 \end{bmatrix} \tag{A.11}$$

$$\begin{bmatrix} \dot{R}_1 & \dot{R}_{12} \\ \dot{R}_{12}^{\mathrm{T}} & \dot{R}_2 \end{bmatrix} = \dot{L}_R {L_R}^{\mathrm{T}} + L_R \dot{L}_R^{\mathrm{T}} \tag{A.12a}$$

where

$$\dot{L}_R = L_{R,f} - L_{R,0} \tag{A.12b}$$

$$\begin{bmatrix} \dot{V}_1 & \dot{V}_{12} \\ \dot{V}_{12}^{\mathrm{T}} & \dot{V}_2 \end{bmatrix} = \dot{L}_V {L_V}^{\mathrm{T}} + L_V \dot{L}_V^{\mathrm{T}} \tag{B.12c}$$

where

$$L_V = L_{V,f} - L_{V,0}. \tag{A.12d}$$

The derivations of the expression for $\frac{\partial \dot{A}}{\partial \theta_i}$, $\frac{\partial \dot{R}}{\partial \theta_i}$ and $\frac{\partial \dot{V}}{\partial \theta_i}$ are primarily based on the application of the following derivative formulas. It is assumed that X is an $m \times n$ matrix and A is an $n \times p$ matrix.

Derivative Formulas

$$\frac{d}{dx_{ij}} XA = e_n^{(i)} A(j,:) \tag{A.14}$$

$$\frac{d}{dx_{ij}} AX = A(:,i) e_p^{(j)} \tag{A.15}$$

$$\frac{d}{dx_{ij}} X^{\mathrm{T}} A = e_n^{(j)} A(i,:) \tag{A.16}$$

$$\frac{d}{dx_{ij}} AX^{\mathrm{T}} = A(:,j) e_p^{(i)} \tag{A.17}$$

$$\frac{d}{dx_{ij}} AXB = A(:,i) B(j,:) \tag{A.18}$$

$$\frac{d}{dx_{ij}} AX^{\mathrm{T}} B = A(:,j) B(i,:) \tag{A.19}$$

Derivatives with respect to θ_j for $\theta_j = a_{c,k\ell}$

$$\frac{\partial \tilde{A}}{\partial a_{c,k\ell}} = \begin{bmatrix} 0 & 0 \\ 0 & e_{n_c}^{(\ell)} e_{n_c}^{(k)^{\mathrm{T}}} \end{bmatrix} \tag{A.20}$$

$$\frac{\partial \tilde{R}}{\partial b_{c,k\ell}} = 0 \tag{A.21}$$

$$\frac{\partial \tilde{V}}{\partial b_{c,k\ell}} = 0 \tag{A.22}$$

Derivatives with respect to θ_j for $\theta_j = b_{c,k\ell}$

$$\frac{\partial \tilde{A}}{\partial b_{c,k\ell}} = \begin{bmatrix} 0 & 0 \\ e_{n_c}^{(k)} C(\ell,:) & e_{n_c}^{(k)} D(\ell,:) C_c \end{bmatrix} \tag{A.23}$$

$$\frac{\partial \tilde{R}}{\partial b_{c,k\ell}} = 0 \tag{A.24}$$

$$\frac{\partial \tilde{V}}{\partial b_{c,k\ell}} = \begin{bmatrix} 0 & [V_{12}(:,\ell) - BD_c V_2(:,\ell)] e_{n_c}^{(k)^{\mathrm{T}}} \\ \mathrm{SYM} & e_{n_c}^{(k)} V_2(\ell,:) B_c^{\mathrm{T}} + B_c V_2(:,\ell) e_{n_c}^{(k)^{\mathrm{T}}} \end{bmatrix} \tag{A.25}$$

Derivatives with respect to θ_j for $\theta_j = c_{c,k\ell}$

$$\frac{\partial \bar{A}}{\partial c_{c,k\ell}} = \begin{bmatrix} 0 & -B(:,k){e_{n_c}^{(\ell)}}^{\mathrm{T}} \\ 0 & -B_c D(:,k){e_{n_c}^{(\ell)}}^{\mathrm{T}} \end{bmatrix} \tag{A.26}$$

$$\frac{\partial \bar{R}}{\partial c_{c,k\ell}} = \begin{bmatrix} 0 & [-R_{12}(:,k) + C^{\mathrm{T}} D_c^{\mathrm{T}} R_2(:,k)]{e_{n_c}^{(\ell)}}^{\mathrm{T}} \\ \mathrm{SYM} & C_c^{\mathrm{T}} R_2(:,k){e_{n_c}^{(\ell)}}^{\mathrm{T}} + e_{n_c}^{(\ell)} R_2(k,:)C_c \end{bmatrix} \tag{A.27}$$

$$\frac{\partial \bar{V}}{\partial c_{c,k\ell}} = 0 \tag{A.28}$$

Derivatives with respect to λ

$$\dot{\bar{A}} \triangleq \frac{\partial \bar{A}}{\partial \lambda} = \begin{bmatrix} \dot{A} & -\dot{B}C_c \\ B_c\dot{C} & -B_c\dot{D}C_c \end{bmatrix} \tag{A.29}$$

$$\dot{\bar{R}} \triangleq \frac{\partial \bar{R}}{\partial \lambda} = \begin{bmatrix} \dot{\bar{R}}_{11} & \dot{\bar{R}}_{12} \\ \dot{\bar{R}}_{12}^{\mathrm{T}} & \dot{\bar{R}}_{22} \end{bmatrix} \tag{A.30}$$

where

$$\dot{\bar{R}}_{11} = \dot{R}_1 \tag{A.31a}$$

$$\dot{\bar{R}}_{12} = -\dot{R}_{12}C_c \tag{A.31b}$$

$$\dot{\bar{R}}_{22} = C_c^{\mathrm{T}} \dot{R}_2 C_c. \tag{A.31c}$$

$$\dot{\bar{V}} = \begin{bmatrix} \dot{\bar{V}}_{11} & \dot{\bar{V}}_{12} \\ \dot{\bar{V}}_{12}^{\mathrm{T}} & \dot{\bar{V}}_{22} \end{bmatrix} \tag{A.32}$$

where

$$\dot{\bar{V}}_{11} = \dot{V}_1 \tag{A.33a}$$

$$\dot{\bar{V}}_{12} = -\dot{V}_{12}B_c^{\mathrm{T}} - \dot{B}D_c^{\mathrm{T}}V_2B_c^{\mathrm{T}} - BD_c\dot{V}_2B_c^{\mathrm{T}} \tag{A.36b}$$

$$\dot{\bar{V}}_{22} = B_c\dot{V}_2B_c^{\mathrm{T}}. \tag{A.36c}$$

Appendix B: Efficient Computation of H_θ

In this appendix we show how to efficiently compute H_0, using (4.26) with (4.37)–(4.39), (4.40)–(4.48), (4.31)–(4.33) and (A.20)–(A.28). First, we assume that Ψ transforms $\tilde{A} \in \mathbb{R}^{n_{cl} \times n_{n_{cl}}}$ to either, complex modal form or complex Schur form, such that

$$\Psi^{-1}\tilde{A}\Psi = \Lambda \tag{B.1}$$

where $\Lambda \in \mathbb{C}^{n_{cl} \times n_{cl}}$ is diagonal or upper triangular. The pre- and post-multiplying (4.31) respectively by Ψ^H and Ψ, pre- and post-multiplying (4.32) by respectively Ψ^{-1} and Ψ^{-H}, and pre-post-multiplying (4.33) by T^{-1} and T give

$$0 = \Lambda^* \bar{P}^{(j)} + \bar{P}^{(j)}\Lambda + (\bar{A}^{(j)^{\mathrm{T}}}\bar{P} + \bar{P}\bar{A}^{(j)} + \bar{R}^{(j)}) \tag{B.2}$$

$$0 = \Lambda\bar{Q}^{(j)} + \bar{Q}^{(j)}\Lambda^* + (\bar{A}^{(j)}\bar{Q} + \bar{Q}\bar{A}^{(j)^{\mathrm{T}}} + \bar{V}^{(j)}) \tag{B.3}$$

$$\bar{Z}^{(j)} = \bar{Q}^{(j)}\bar{P} + \bar{Q}\bar{P}^{(j)} \tag{B.4}$$

where

$$\bar{P}^{(j)} = \Psi^H \tilde{P}^{(j)}\Psi \tag{B.5}$$

$$\bar{Q}^{(j)} = \Psi^{-1}\tilde{Q}^{(j)}\Psi^{-H} \tag{B.6}$$

$$\bar{Z}^{(j)} = \Psi^{-1}\tilde{Z}^{(j)}\Psi \tag{B.7}$$

$$\bar{P} = \Psi^H \tilde{P}\Psi \tag{B.8}$$

$$\bar{Q} = \Psi^{-1}\tilde{Q}\Psi^{-H} \tag{B.9}$$

$$\bar{A}^{(j)} = \Psi^{-1}\tilde{A}^{(j)}\Psi \tag{B.10}$$

$$\bar{R}^{(j)} = \Psi^H \tilde{R}^{(j)}\Psi \tag{B.11}$$

$$\bar{V}^{(j)} = \Psi^{-1}\tilde{V}^{(j)}\Psi. \tag{B.12}$$

Next, partition Ψ as

$$\Psi = \begin{bmatrix} \Psi_1 \\ \Psi_2 \end{bmatrix}, \quad \Psi_1 \in \mathbb{R}^{n_x \times n_{cl}}, \quad \Psi_2 \in \mathbb{R}^{n_c \times n_{cl}} \tag{B.13}$$

and partition Ψ^{-1} as

$$\Psi^{-1} = \left[(\Psi^{-1})_1, (\Psi^{-1})_2\right], \quad (\Psi^{-1})_1 \in \mathbb{R}^{n_{cl} \times n_x}, \quad (\Psi^{-1})_2 \in \mathbb{R}^{n_{cl} \times n_c}. \tag{B.14}$$

Also, define

$$\bar{B} \triangleq (\Psi^{-1})_1 B \tag{B.15}$$

$$\bar{C} \triangleq C\Psi_1 \tag{B.16}$$

$$\bar{R}_{12} \triangleq \Psi_1^H R_{12} \tag{B.17}$$

$$\bar{V}_{12} \triangleq (\Psi^{-1})_1 V_{12} \tag{B.18}$$

$$\bar{B}_c \triangleq (\Psi^{-1})_2 B_c \tag{B.19}$$

$$\bar{C}_c \triangleq C_c \Psi_2. \tag{B.20}$$

Now, recall that $M^{(j)} \triangleq \frac{\partial M}{\partial \theta_j}$ where θ_j represents either $a_{c,k\ell}, b_{c,k\ell}, c_{c,k\ell}$. It then follows from (B.10)–(B.12) and (A.20)–(A.28) that $\bar{A}^{(j)}$, $\bar{R}^{(j)}$ and $\bar{V}^{(j)}$ are given as follows.

for $\theta_{\mathrm{j}} = \mathbf{a}_{\mathbf{c},\mathbf{k}\ell}$

$$\bar{A}^{(j)} = \Psi_2^{-1}(:,\ell)\Psi_2(k,:) \tag{B.21}$$

$$\bar{R}^{(j)} = 0 \tag{B.22}$$

$$\bar{V}^{(j)} = 0 \tag{B.23}$$

for $\theta_{\mathrm{j}} = \mathbf{b}_{\mathbf{c},\mathbf{k}\ell}$

$$\bar{A}^{(j)} = \Psi_2^{-1}(:,k)\left[\bar{C}(\ell,:) + D(\ell,:)\bar{C}_c\right] \tag{B.24}$$

$$\bar{R}^{(j)} = 0 \tag{B.25}$$

$$\begin{aligned}\bar{V}^{(j)} = &\left\{\left[\bar{V}_{12}(:,\ell) - \bar{B}D_c V_2(:,\ell) + \bar{B}_c V_2(:,\ell)\right]\left[(\Psi^{-1})_2(:,k)\right]^H\right\} \\ &+ \left\{\left[\bar{V}_{12}(:,\ell) + \bar{B}_c V_2(:,\ell)\right]\left[(\Psi^{-1})_2(:,k)\right]^H\right\}^H.\end{aligned} \tag{B.26}$$

for $\theta_{\mathrm{j}} = \mathbf{c}_{\mathbf{c},\mathbf{k}\ell}$

$$\bar{A}^{(j)} = \left[\bar{B}(:,k) + \bar{B}_c D(:,k)\right]\Psi_2(\ell,:) \tag{B.27}$$

$$\begin{aligned}\bar{R}^{(j)} = &\left\{\left[(\Psi^{-1})_2(\ell,:)\right]^H\left[-\bar{R}_{12}(:,k)^H + R_2(k,:)D_c\bar{C} + R_2(k,:)\bar{C}_c\right]\right\} \\ &+ \left\{\left[(\Psi^{-1})_2(\ell,:)\right]^H\left[-\bar{R}_{12}(:,k)^H + R_2(k,:)\bar{C}_c\right]\right\}^H\end{aligned} \tag{B.28}$$

$$\bar{V}^{(j)} = 0 \tag{B.29}$$

Note that (B.21)-(B.29) allow the transformation of $\bar{A}^{(j)}$, $\bar{R}^{(j)}$ and $\bar{V}^{(j)}$ to the modal of Schur basis to be performance very efficiently.

Now, it follows from (B.5)-(B.7) that

$$\tilde{P}^{(j)} = \Psi^{-H}\bar{P}^{(j)}\Psi^{-1} \tag{B.30}$$

$$\tilde{Q}^{(j)} = \Psi\bar{Q}^{(j)}\Psi^{H} \tag{B.31}$$

$$\tilde{Z}^{(j)} = \Psi\bar{Z}^{(j)}\Psi^{-1} \tag{B.32}$$

or, equivalently,

$$\begin{bmatrix} \tilde{P}_{11}^{(j)} & \tilde{P}_{12}^{(j)} \\ \tilde{P}_{12}^{(j)\mathrm{T}} & \tilde{P}_{22}^{(j)} \end{bmatrix} = \begin{bmatrix} (\Psi^{-1})_1^H\bar{P}^{(j)}(\Psi^{-1})_1 & (\Psi^{-1})_1^H\bar{P}^{(j)}(\Psi^{-1})_2 \\ (\Psi^{-1})_2^H\bar{P}^{(j)}(\Psi^{-1})_1 & (\Psi^{-1})_2^H\bar{P}^{(j)}(\Psi^{-1})_2 \end{bmatrix} \tag{B.33}$$

$$\begin{bmatrix} \tilde{Q}_{11}^{(j)} & \tilde{Q}_{12}^{(j)} \\ \tilde{Q}_{12}^{(j)\mathrm{T}} & \tilde{Q}_{22}^{(j)} \end{bmatrix} = \begin{bmatrix} \Psi_1\bar{Q}\Psi_1^H & \Psi_1\bar{Q}^{(j)}\Psi_2^H \\ \Psi_2\bar{Q}^{(j)}\Psi_1^H & \Psi_2\bar{Q}^{(j)}\Psi_2^H \end{bmatrix} \tag{B.34}$$

$$\begin{bmatrix} \tilde{Z}_{11}^{(j)} & \tilde{Z}_{12}^{(j)} \\ \tilde{Z}_{21}^{(j)} & \tilde{Z}_{22}^{(j)} \end{bmatrix} = \begin{bmatrix} \Psi_1\bar{Z}^{(j)}(\Psi^{-1})_1 & \Psi_1\bar{Z}^{(j)}(\Psi^{-1})_2 \\ \Psi_2\bar{Z}^{(j)}(\Psi^{-1})_1 & \Psi_2\bar{Z}^{(j)}(\Psi^{-1})_2 \end{bmatrix}. \tag{B.35}$$

It follows from (4.37), (B.35), and (B.4) that

$$H'_{A_c}(\tilde{Z}^{(j)}) \triangleq H_{A_c}^{(j)} = \Psi_2\bar{Q}^{(j)}\bar{P}(\Psi^{-1})_2 + \Psi_2^{-1}\bar{Q}\bar{P}^{(j)}\Psi 2 \tag{B.36}$$

It follows from (4.38), (B.33), (B.35), (B.13), (B.14), and (B.16)-(B.18) that

$$H'_{B_c}(\tilde{P}^{(j)}, \tilde{Z}^{(j)}) \triangleq H_{B_c}^{(j)} = 2\left[\Psi_2^H(\bar{P}^{(j)}M_{HBC}) + \bar{P}_2(\bar{Q}^{(j)}C_{HBC})\right] \tag{B.37}$$

where

$$\bar{P}_2 \triangleq \Psi_2^{-H}\bar{P} \tag{B.38}$$

$$M_{HBC} \triangleq \bar{V}_{12} + \bar{B}_cV_2 + \bar{Q}C_{HBC}^H \tag{B.39}$$

$$C_{HBC} \triangleq \bar{C} - D\bar{C}_c. \tag{B.40}$$

Similarly, it follows from (4.39), (B.34), (B.35), (B.13)-(B.15), (B.17) and (B.18) that

$$H'_{C_c}(\tilde{Q}^{(j)}, \tilde{Z}^{(j)}) = H_{C_c}^{(j)} = 2\left[(M_{HCC}\bar{Q}^{(j)})\Psi_2^H - (B_{HCC}^H\bar{P}^{(j)})\bar{Q}_2^H\right] \tag{B.41}$$

where

$$\bar{Q}_2 \triangleq \Psi_2 \bar{Q} \tag{B.42}$$

$$M_{HCC} \triangleq R_2 \bar{C}_c - \bar{R}_{12}^H - B_{HCC}\bar{P} \tag{B.43}$$

$$B_{HCC} \triangleq \bar{B} + \bar{B}_c D. \tag{B.44}$$

Finally, substituting (B.33)–(B.35) into (4.40)–(4.48), using (B.13)–(B.18) and recalling the definitions (B.36), (B.37) and (B.41) gives the following.

Derivations with respect to $a_{c,k\ell}$

$$\frac{\partial H_{A_c}}{\partial a_{c,k\ell}} = H_{A_c}^{(j)} \tag{B.45}$$

$$\frac{\partial H_{B_c}}{\partial a_{c,k\ell}} = H_{B_c}^{(j)} \tag{B.46}$$

$$\frac{\partial H_{C_c}}{\partial a_{c,k\ell}} = H_{C_c}^{(j)} \tag{B.47}$$

Derivations with respect to $b_{c,k\ell}$

$$\frac{\partial H_{A_c}}{\partial b_{c,k\ell}} = H_{A_c}^{(j)} \tag{B.48}$$

$$\frac{\partial H_{B_c}}{\partial b_{c,k\ell}} = H_{B_c}^{(j)} - 2(\bar{P}_2(\Psi^{-1})_2(:,k))V_2(\ell,:) \tag{B.49}$$

$$\frac{\partial H_{C_c}}{\partial b_{c,k\ell}} = H_{C_c}^{(j)} - 2D(\ell,:)^{\mathrm{T}}(\bar{P}_2(k,:)\bar{Q}_2^H) \tag{B.50}$$

Derivations with respect to $c_{c,k\ell}$

$$\frac{\partial H_{A_c}}{\partial c_{c,k\ell}} = H_{A_c}^{(j)} \tag{B.51}$$

$$\frac{\partial H_{B_c}}{\partial c_{c,k\ell}} = H_{B_c}^{(j)} - 2(\bar{P}_2\bar{Q}_2(\ell,:)^H)D(:,k)^{\mathrm{T}} \tag{B.52}$$

$$\frac{\partial H_{C_c}}{\partial c_{c,k\ell}} = H_{C_c}^{(j)} - 2R_2(:,k)(\bar{Q}_2(\ell,:)\Psi_2^H) \tag{B.53}$$

Appendix I:

"Contruction of Low Authority, Nearly Non-Minimal LQG Compensators for Initializing Optimal Reduced-Order Control Design Algorithms"

October 1993

Construction of Low Authority, Nearly Non-Minimal LQG Compensators for Reduced-Order Control Design

by

Emmanuel G. Collins, Jr.
Harris Corporation
Government Aerospace Systems Division
MS 22/4849
Melbourne, FL 32902
(407) 727-6358
FAX: (407) 727-4016
ecollins@x102a.ess.harris.com

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901
(407) 768-8000 Ext. 7241
FAX: (407) 984-8461
haddad@zach.fit.edu

Sidney S. Ying
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901
(407) 768-7063
FAX: (407) 254-7805
ssy@dllws.cca.cr.rockwell.com

Abstract

It has been observed numerically that low authority LQG compensators are often nearly non-minimal. However, to date a rigorous justification for this phenomenon has not yet been established. This paper helps provide the needed theoretical foundation. In particular, it is shown that for both continuous-time and discrete-time stable systems, by proper choice of the structure of the design weights, the corresponding LQG compensator becomes nonminimal as the control authority is decreased. Thus, the results provide a partial explanation of why the suboptimal controller reduction methods tend to work best at low control authority. The results also can be used as rigorous guidelines to efficiently initialize homotopy algorithms for directly synthesizing optimal reduced-order controllers. The restriction to stable systems is not necessarily limiting since the freedom involved in defining a homotopy allows this assumption to always be satisfied.

This research was supported in part by the National Science Foundation under Grants ECS-9109558 and ECS-9350181, the National Aeronautical and Aerospace Administration under Contract NAS8-38575, and the Air Force Office of Scientific Research under Contract F49620-91-C-0019.

1. Introduction

The development of linear-quadratic-gaussian (LQG) theory [1-3] was a major breakthrough in modern control theory since it provides a systematic way to synthesize high performance controllers for nominal models of complex, multi-input multi-output systems. However, one of the well known deficiencies of an LQG compensator is that its minimal dimension is usually equal to the dimension of the design plant. This has led to the development of techniques to directly synthesize optimal reduced-order controllers [4–17] and techniques to synthesize reduced-order approximations of the optimal full-order compensator (i.e., controller reduction methods) [18–23].

The controller reduction methods almost always yield suboptimal (and sometimes destabilizing) reduced-order control laws since an optimal reduced-order controller is not usually a direct function of the parameters used to compute or describe the optimal full-order controller. Nevertheless, these methods are computationally inexpensive and sometimes do yield high performing and even nearly optimal control laws. An observation that holds true about most of these methods is that they tend to work best at low control authority [17, 21, 23]. However, to date no rigorous explanation has been presented to explain this phenomenon.

One of the purposes of this paper is to provide a partial explanation as to why the suboptimal projection methods tend to work at low control authority. The discussion here focuses on stable systems. It is shown that if the state weighting matrix R_1 or disturbance intensity (or covariance for discrete systems) V_1 has a specific structure in a basis in which the A matrix is upper or lower block triangular, respectively, then at low control authority the corresponding LQG compensator is nearly nonminimal and can hence be easily reduced to a nearly optimal reduced-order controller. The conditions presented for R_1 and V_1 often are satisfied or nearly satisfied in practice. Hence, for stable systems the results proved in this paper do offer one explanation of why suboptimal controller reduction methods often provide nearly optimal control laws at low authority. The results can also be used as guidelines for choosing R_1 and V_1 such that suboptimal controller reduction methods yield "good" reduced-order controllers.

Suboptimal controller reduction methods can be used to initialize algorithms for synthesizing optimal reduced-order controllers. Of particular interest are the homotopy algorithms of [11. 15–17] since they are based on allowing the plant and weights defining an optimization problem to vary as a function of the homotopy parameter $\lambda \in [0.1]$. These homotopy algorithms rely on choosing the initial plant and weights so that the corresponding LQG compensator is easily reduced to a nearly optimal reduced-order compensator of the desired dimension. Hence, the results presented here provide some rigorous guidelines for initializing these algorithms. Note that the restriction to stable systems is not necessarily limiting since the freedom involved in defining a homotopy allows this assumption to be satisfied. However, future work will focus on theory that directly applies to unstable systems.

Notation

$\mathbb{R}, \mathbb{R}^{r\times s}, \mathbb{R}^r$	real numbers, $r \times s$ real matrices, $\mathbb{R}^{r\times 1}$
$\mathbb{E}$	expected value
$X \geq 0, X > 0$	matrix X is nonnegative definite, X is positive definite
$0_{r\times s}, 0_r$	$r \times s$ zero matrix, $r \times r$ zero matrix
I_r	$r \times r$ identity matrix
$\mathrm{vec}(\cdot)$	the invertible linear operator defined such that $\mathrm{vec}\ S \triangleq [s_1^{\mathrm{T}}\ s_2^{\mathrm{T}} \cdots s_q^{\mathrm{T}}]^{\mathrm{T}},\ S \in \mathbb{R}^{p\times q}$, where $s_j \in \mathbb{R}^p$ denotes the j^{th} column of S.

2. Low Authority LQG Compensation: Continuous-Time Systems

Consider the n^{th}-order linear time-invariant plant

$$\dot{x}(t) = Ax(t) + Bu(t) + D_1w(t), \tag{2.1a}$$

$$y(t) = Cx(t) + D_2w(t), \tag{2.1b}$$

where (A, B) is stabilizable, (A, C) is detectable, $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^l$, and $w \in \mathbb{R}^d$ is a standard white noise disturbance with intensity I_d and rank $D_2 = l$. The intensities of $D_1w(t)$ and $D_2w(t)$ are thus given, respectively, by $V_1 \triangleq D_1D_1^{\mathrm{T}} \geq 0$, and $V_2 \triangleq D_2D_2^{\mathrm{T}} > 0$. For convenience, we assume that $V_{12} \triangleq D_1D_2^{\mathrm{T}} = 0$, i.e., the plant disturbance and measurement noise are uncorrelated. Then, the LQG compensator

$$\dot{x}_c(t) = A_cx_c(t) + B_cy(t), \tag{2.2a}$$

$$u(t) = -C_cx_c(t), \tag{2.2b}$$

for the plant (2.1) minimizing the steady-state quadratic performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t\to\infty} \frac{1}{t}\mathbb{E} \int_0^t [x^{\mathrm{T}}(s)R_1x(s) + u^{\mathrm{T}}(s)R_2u(s)]ds. \tag{2.3}$$

where $R_1 \geq 0$ and $R_2 > 0$ are the weighting matrices for the controlled states and controller input, respectively, is given by:

$$A_c = A - \Sigma P - Q\bar{\Sigma}, \tag{2.4a}$$

$$B_c = QC^{\mathrm{T}}V_2^{-1}, \quad C_c = R_2^{-1}B^{\mathrm{T}}P, \tag{2.4b,c}$$

where $\Sigma \triangleq BR_2^{-1}B^{\mathrm{T}}$, $\bar{\Sigma} \triangleq C^{\mathrm{T}}V_2^{-1}C$, and P and Q are the unique, nonnegative-definite solutions respectively of

$$0 = A^{\mathrm{T}}P + PA + R_1 - P\Sigma P, \tag{2.5}$$

$$0 = AQ + QA^{\mathrm{T}} + V_1 - Q\bar{\Sigma}Q. \tag{2.6}$$

Furthermore, the "shifted" observability and controllability grammians [18, 24] of the compensator, $\hat{P}$ and $\hat{Q}$, are the unique, nonnegative-definite solutions respectively of

$$0 = (A - Q\bar{\Sigma})^{\mathrm{T}}\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P, \tag{2.7}$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^{\mathrm{T}} + Q\bar{\Sigma}Q. \tag{2.8}$$

Although a cross-weighting term of the form $2x^{\mathrm{T}}(t)R_{12}u(t)$ can also be included in (2.3), we shall not do so here to facilitate the presentation. The magnitudes of R_2 and V_2 relative to the state weighting matrix R_1 and plant disturbance intensity V_1 govern the regulator and estimator authorities, respectively. The selection of R_2 and V_2 such that $\|R_2\| >> \|R_1\|$, or $\|V_2\| >> \|V_1\|$, yields a low authority compensator. It has been observed numerically that low authority LQG compensators are often nearly nonminimal [17, 21]. This section provides a rigorous justification for this observation when the open-loop plant is stable and (A, R_1) or (A, V_1) have a particular structure. In order to prove this result, we first exploit some interesting structural properties of the solutions of the Riccati equations and Lyapunov equations assuming the coefficient matrix A and the constant driving term R_1 have certain partitioned forms.

Lemma 2.1. Suppose

$$A = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad R_1 = \begin{bmatrix} R_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.9a,b,c}$$

where $A_1, R_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $B_1 \in \mathbb{R}^{n_r \times m}$, $R_{1,1} > 0$.

(*i*) If (A, B) and (A_1, B_1) are stabilizable, then the unique, nonnegative-definite solution of the Riccati equation:

$$0 = A^{\mathrm{T}}P + PA + R_1 - PBB^{\mathrm{T}}P. \tag{2.10}$$

is given by

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.11}$$

where the $n_r \times n_r$ matrix P_1 is the unique, positive-definite solution of

$$0 = A_1^{\mathrm{T}} P_1 + P_1 A_1 + R_{1,1} - P_1 B_1 B_1^{\mathrm{T}} P_1. \tag{2.12}$$

(*ii*) If A is asymptotically stable, then the unique, nonnegative-definite solution of the Lyapunov equation:

$$0 = A^{\mathrm{T}} P + PA + R_1, \tag{2.13}$$

is given by

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.14}$$

where the $n_r \times n_r$ matrix P_1 is the unique, positive-definite solution of

$$0 = A_1^{\mathrm{T}} P_1 + P_1 A_1 + R_{1,1}. \tag{2.15}$$

Proof.

(*i*) Since (A, B) is stabilizable and $R_1 \geq 0$, it follows from Theorem 12.2 of [25] that there exists a unique, nonnegative-definite solution of the Riccati equation (2.10). Similarly, the assumptions that (A_1, B_1) is stabilizable and $R_{1,1} > 0$ imply that there exists a positive-definite matrix P_1 satisfying the Riccati equation (2.12). Using (2.12), it follows by construction that (2.11) is the solution of (2.10).

(*ii*) This is a special case of the Riccati equation of property (*i*). □

The following lemma states the dual of Lemma 2.1 if the coefficient matrix A is upper block triangular and V_1 is upper block diagonal.

Lemma 2.2. Suppose

$$A = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \quad C = [C_1 \quad C_2], \quad V_1 = \begin{bmatrix} V_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.16a, b, c}$$

where $A_1, V_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $C_1 \in \mathbb{R}^{l \times n_r}$, $V_{1,1} > 0$.

(*i*) If (A, C) and (A_1, C_1) are detectable, then the unique, nonnegative-definite solution of the Riccati equation:

$$0 = AQ + QA^{\mathrm{T}} + V_1 - QC^{\mathrm{T}} CQ, \tag{2.17}$$

is given by

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.18}$$

where the $n_r \times n_r$ matrix Q_1 is the unique, positive-definite solution of

$$0 = A_1Q_1 + Q_1A_1^{\mathrm{T}} + V_{1,1} - Q_1C_1^{\mathrm{T}}C_1Q_1. \tag{2.19}$$

(*ii*) If A is asymptotically stable, then the unique, nonnegative-definite solution of the Lyapunov equation:

$$0 = AQ + QA^{\mathrm{T}} + V_1, \tag{2.20}$$

is given by

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.21}$$

where the $n_r \times n_r$ matrix Q_1 is the unique, positive-definite solution of

$$0 = A_1Q_1 + Q_1A_1^{\mathrm{T}} + V_{1,1}. \tag{2.22}$$

Proof. The proof is dual to the proof of Lemma 2.1. □

The following theorem shows that, with proper choice of the weighting matrices, a low authority LQG controller for a stable plant is nearly nonminimal. The proof of this theorem relies on the above two lemmas.

Theorem 2.1. Consider the plant given by (2.1).

(*i*) Suppose

$$A = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix}, \quad R_1 = \begin{bmatrix} R_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.23a,b}$$

where $A_1, R_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $R_{1,1} > 0$, and A is asymptotically stable. Let

$$V_2 \triangleq \beta \hat{V}_2 \tag{2.24}$$

where $\hat{V}_2$ is some finite, positive-definite matrix and $\beta \in \mathbb{R}$ is a positive scalar. Then

$$\lim_{\beta \to \infty} \text{rank } (\hat{Q}\hat{P}) \le \lim_{\beta \to \infty} \text{rank } \hat{P} \le n_r, \tag{2.25}$$

where $\hat{Q}$ and $\hat{P}$ are the shifted controllability and observability grammians of the corresponding LQG compensator, satisfying (2.8) and (2.7), respectively. Equivalently, for

$\delta > 0$, there exists N such that for all $\beta > N$, $\lambda_{n_r+1} < \delta\lambda_{n_r}$, where λ_i represents the i^{th} eigenvalue of $\hat{Q}\hat{P}$ and $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_i \geq \lambda_{i+1} ... \geq 0$.

(ii) Suppose

$$A = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \quad V_1 = \begin{bmatrix} V_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.26a, b}$$

where $A_1, V_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $V_{1,1} > 0$, and A is asymptotically stable. Let

$$R_2 \triangleq \alpha\hat{R}_2, \tag{2.27}$$

where $\hat{R}_2$ is some finite, positive-definite matrix and $\alpha \in \mathbb{R}$ is a positive scalar. Then

$$\lim_{\alpha\to\infty} \text{rank } (\hat{Q}\hat{P}) \leq \lim_{\alpha\to\infty} \text{rank } \hat{Q} \leq n_r, \tag{2.28}$$

where $\hat{Q}$ and $\hat{P}$ are the shifted controllability and observability grammians of the corresponding LQG compensator, satisfying (2.8) and (2.7), respectively. Equivalently, for $\delta > 0$, there exists N such that for all $\alpha > N$, $\lambda_{n_r+1} < \delta\lambda_{n_r}$, where λ_i represents the i^{th} eigenvalue of $\hat{Q}\hat{P}$ and $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_i \geq \lambda_{i+1} ... \geq 0$.

Proof.

(i) Partition $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{12}^{\mathrm{T}} & \Sigma_2 \end{bmatrix}$, conformal to A in (2.23). The assumptions (2.23) and that A is asymptotically stable imply that (A, B) and (A_1, B_1) are both stabilizable. Thus, it follows from property (i) of Lemma 2.1 that the unique, nonnegative-definite solution P of the Riccati equation (2.5) has the structure given by (2.11), which implies that

$$P\Sigma P = \begin{bmatrix} P_1\Sigma_1 P_1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{2.29}$$

Thus, noting the special partitioned structures in (2.29) and (2.23), and that A is asymptotically stable, it follows from property (ii) of Lemma 2.1 that there exists

$$\hat{P}_0 \triangleq \begin{bmatrix} \hat{P}_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{2.30}$$

which is the unique. nonnegative-definite solution of

$$0 = A^{\mathrm{T}}\hat{P}_0 + \hat{P}_0 A + P\Sigma P, \tag{2.31}$$

where $n_r \times n_r$ matrix $\hat{P}_1$ is the unique. nonnegative-definite solution of

$$0 = A_1^{\mathrm{T}}\hat{P}_1 + P_1 A + P_1\Sigma_1 P_1.$$

Next, computing (2.31) – (2.7) and using (2.24), yields the following modified Lyapunov equation:

$$0 = A^{\mathrm{T}}\Delta\hat{P} + \Delta\hat{P}A + \beta^{-1}[(C^{\mathrm{T}}\hat{V}_2^{-1}CQ\hat{P}) + (C^{\mathrm{T}}\hat{V}_2^{-1}CQ\hat{P})^{\mathrm{T}}]. \qquad (2.32)$$

where

$$\Delta\hat{P} \triangleq \hat{P}_0 - \hat{P}. \qquad (2.33)$$

Since A is asymptotically stable and Q and $\hat{P}$ satisfy (2.6) and (2.7), respectively, Q and $\hat{P}$ are bounded for all β. Thus, (2.32) implies that $\lim_{\beta\to\infty}\|\Delta\hat{P}\| = 0$. Hence, for $\epsilon > 0$, there exists M such that for all $\beta > M$, $\|\Delta\hat{P}\| < \epsilon$. Using (2.33), it follows that

$$\lim_{\beta\to\infty}\hat{P} = \hat{P}_0 = \begin{bmatrix} \hat{P}_1 & 0 \\ 0 & 0 \end{bmatrix}. \qquad (2.34)$$

Thus, $\lim_{\beta\to\infty}$ rank $(\hat{Q}\hat{P}) \leq \lim_{\beta\to\infty}$ rank $\hat{P} = n_r$, which implies the following inequalities of the eigenvalues of $\hat{Q}\hat{P}$. Suppose λ_i represents an eigenvalue of $\hat{Q}\hat{P}$ and $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_i \geq \lambda_{i+1} \ldots \geq 0$. Then, for $\delta > 0$, there exists N such that for all $\beta > N$, $\lambda_{n_r+1} < \delta\lambda_{n_r}$.

(*ii*) The proof is dual to the proof of (*i*). □

Remark 2.1. Theorem 2.1 provides two ways of weighting matrices selection resulting in a nearly nonminimal, low authority LQG compensator for a stable plant. The first approach starts by transforming the plant A into coordinates such that A has the representation as in equation (2.23a) after transformation. Then select the weighting matrix R_1 with the partitioned form as in (2.23b) and with rank $R_1 = n_r$. By decreasing the authority of the compensator, or, equivalently, increasing $\|V_2\|$ or β, the LQG compensator approaches nonminimality with minimal dimension of n_r. Using a dual approach, with A and V_1 partitioned as in (2.26), by increasing $\|R_2\|$ or α, the resulting LQG compensator approaches nonminimality.

Remark 2.2. Note that if A is in a modal form, then it satisfies both (2.23a) and (2.26a) of Theorem 2.1. In this case, R_1 given by (2.23b), describes a state weighting matrix in which only the states pertaining to selected modes are weighted. Similarly, V_1 given by (2.26b) describes a disturbance that excites only certain modes. It is not uncommon for these conditions to be satisfied or nearly satisfied in practice.

Remark 2.3. The suboptimal controller reduction methods of [18–23] characterize the reduced-order controller by a projection or some other type of reduction of the LQG controller. It has been observed that these suboptimal reduced-order controllers for the low-authority control problem will yield virtually the same cost as the LQG controller. According to Theorem 2.1, for a stable plant and with proper choice of the weighting matrices, the LQG controller for a low authority control problem is nearly nonminimal, which provides a theoretical justification for the above observation.

Remark 2.4. The homotopy algorithms for reduced-order dynamic compensation problems developed in [15–17] are based on allowing the plant and weights defining an optimization problem to vary as functions of the homotopy parameter $\lambda \in [0,1]$. In particular, it is assumed that

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & 0 \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & 0 \end{bmatrix} + \lambda(\begin{bmatrix} A_f & B_f \\ C_f & 0 \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & 0 \end{bmatrix}),$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^{\mathrm{T}}(\lambda) & R_2(\lambda) \end{bmatrix} = L_R(\lambda) L_R^{\mathrm{T}}(\lambda),$$

where

$$L_R(\lambda) = L_{R,0} + \lambda(L_{R,f} - L_{R,0}),$$

and $L_{R,0}$ and $L_{R,f}$ satisfy

$$L_{R,0} L_{R,0}^{\mathrm{T}} = \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^{\mathrm{T}} & R_{2,0} \end{bmatrix}, \quad L_{R,f} L_{R,f}^{\mathrm{T}} = \begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^{\mathrm{T}} & R_{2,f} \end{bmatrix},$$

and

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^{\mathrm{T}}(\lambda) & V_2(\lambda) \end{bmatrix} = L_V(\lambda) L_V^{\mathrm{T}}(\lambda),$$

where

$$L_V(\lambda) = L_{V,0} + \lambda(L_{V,f} - L_{V,0}),$$

and $L_{V,0}$ and $L_{V,f}$ satisfy

$$L_{V,0} L_{V,0}^{\mathrm{T}} = \begin{bmatrix} V_{1,0} & V_{12,0} \\ V_{12,0}^{\mathrm{T}} & V_{2,0} \end{bmatrix}, \quad L_{V,f} L_{V,f}^{\mathrm{T}} = \begin{bmatrix} V_{1,f} & V_{12,f} \\ V_{12,f}^{\mathrm{T}} & V_{2,f} \end{bmatrix}.$$

Note that the above equations imply that $A(0) = A_0$, $B(0) = B_0$, etc ... which are the initial set of system matrices and that $A(1) = A_f$, $B(1) = B_f$, etc ... which are the final and given system matrices, To initialize the homotopy algorithm efficiently, the designer can choose $(A_0, B_0, C_0, R_{1,0}, R_{12,0}, R_{2,0}, V_{1,0}, V_{12,0}, V_{2,0})$, to correspond to a low authority control problem with stable open-loop plant as stated in Theorem 2.1, for which a nearly optimal reduced-order controller may be easily obtained by balanced controller reduction [18] or an alternative suboptimal controller reduction method [19–23].

3. Low Authority LQG Compensation: Discrete-Time Systems

In this section, we consider the discrete-time counterpart of the previous section. In particular, a rigorous justification is provided for a nearly nonminimal low authority discrete-time LQG compensator when the open-loop plant is stable and certain weighting matrix has specific structure. Consider the n^{th}-order linear, discrete time-invariant plant

$$x(k+1) = Ax(k) + Bu(k) + D_1 w(k), \tag{3.1a}$$

$$y(k) = Cx(k) + D_2 w(k), \tag{3.1b}$$

where (A, B) is stabilizable, (A, C) is detectable, $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^l$, and $w \in \mathbb{R}^d$ is a standard white noise disturbance with covariance I_d and rank $D_2 = l$. The covariances of $D_1 w(k)$ and $D_2 w(k)$ are thus given, respectively, by $V_1 \triangleq D_1 D_1^{\mathrm{T}} \geq 0$, and $V_2 \triangleq D_2 D_2^{\mathrm{T}} > 0$. For convenience, once again we assume that $V_{12} \triangleq D_1 D_2^{\mathrm{T}} = 0$. Then, the LQG compensator

$$x_c(k+1) = A_c x_c(k) + B_c y(k), \tag{3.2a}$$

$$u(k) = -C_c x_c(k) - D_c y(k), \tag{3.2b}$$

for the plant (3.1) minimizing the steady-state quadratic performance criterion

$$J(A_c, B_c, C_c, D_c) \triangleq \lim_{k \to \infty} \mathbb{E}[x^{\mathrm{T}}(k) R_1 x(k) + u^{\mathrm{T}}(k) R_2 u(k)], \tag{3.3}$$

where $R_1 \geq 0$ and $R_2 > 0$ are respectively the weighting matrices for the controlled states and controller input, is given by [26]:

$$A_c = A - Q_a V_{2a}^{-1} C - B R_{2a}^{-1} P_a, \tag{3.4a}$$

$$B_c = Q_a V_{2a}^{-1}, \quad C_c = R_{2a}^{-1} P_a, \quad D_c = R_{2a}^{-1} B^{\mathrm{T}} P A Q C^{\mathrm{T}} V_{2a}^{-1}, \tag{3.4b, c, d}$$

where

$$Q_a \triangleq AQC^{\mathrm{T}}, \;\; P_a \triangleq B^{\mathrm{T}} P A, \;\; V_{2a} \triangleq V_2 + CQC^{\mathrm{T}}, \;\; R_{2a} \triangleq R_2 + B^{\mathrm{T}} P B, \tag{3.5, 6, 7, 8}$$

and P and Q are the unique. nonnegative-definite solutions respectively of

$$P = A^{\mathrm{T}} P A + R_1 - P_a^{\mathrm{T}} R_{2a}^{-1} P_a, \tag{3.9}$$

$$Q = AQA^{\mathrm{T}} + V_1 - Q_a V_{2a}^{-1} Q_a^{\mathrm{T}}. \tag{3.10}$$

Furthermore, the "shifted" observability and controllability grammians of the compensator, $\hat{P}$ and $\hat{Q}$, satisfy

$$\hat{P} = (A - Q_a V_{2a}^{-1} C)^{\mathrm{T}} \hat{P} (A - Q_a V_{2a}^{-1} C) + (P_a - R_{2a} D_c C)^{\mathrm{T}} R_{2a}^{-1} (P_a - R_{2a} D_c C), \tag{3.11}$$

$$\hat{Q} = (A - B R_{2a}^{-1} P_a) \hat{Q} (A - B R_{2a}^{-1} P_a)^{\mathrm{T}} + (Q_a - B D_c V_{2a}) V_{2a}^{-1} (Q_a - B D_c V_{2a})^{\mathrm{T}}, \tag{3.12}$$

and $\hat{P}$ and $\hat{Q}$ are nonnegative definite.

As in the continuous-time case a cross-weighting term of the form $2x^{\mathrm{T}}(k) R_{12} u(k)$ can also be included in (3.3), we shall not do so here to facilitate the presentation. Similar to the continuous-time compensation problem, the magnitudes of R_2 and V_2 relatively to R_1 and V_1 govern the regulator and estimator authority, respectively. The following theorem is the discrete-time counterpart of the continuous-time result stated in Theorem 2.1. It provides a rigorous justification for a nearly nonminimal low authority discrete-time LQG compensator when the open-loop plant is stable and R_1 or V_1 has certain structure.

Theorem 3.1. Consider the plant described in (3.1).

(*i*) Suppose

$$A = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix}, \quad R_1 = \begin{bmatrix} R_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{3.13a, b}$$

where $A_1, R_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $R_{1,1} > 0$, and A is asymptotically stable. Let

$$V_2 \triangleq \beta \hat{V}_2, \tag{3.14}$$

where $\hat{V}_2$ is some finite, positive-definite matrix and $\beta \in \mathbb{R}$ is a positive scalar. Then

$$\lim_{\beta \to \infty} \text{rank } (\hat{Q}\hat{P}) \leq \lim_{\beta \to \infty} \text{rank } \hat{P} \leq n_r, \tag{3.15}$$

where $\hat{Q}$ and $\hat{P}$ are the shifted controllability and observability grammians of the corresponding LQG compensator, satisfying (3.11) and (3.12), respectively. Equivalently, for $\delta > 0$, there exists N such that for all $\beta > N$, $\lambda_{n_r+1} < \delta \lambda_{n_r}$, where λ_i represents the i^{th} eigenvalue of $\hat{Q}\hat{P}$ and $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_i \geq \lambda_{i+1} ... \geq 0$.

(*ii*) Suppose

$$A = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \quad V_1 = \begin{bmatrix} V_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \tag{3.16a, b}$$

where $A_1, V_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $V_{1,1} > 0$, and A is asymptotically stable. Let

$$R_2 \triangleq \alpha \hat{R}_2, \tag{3.17}$$

where $\hat{R}_2$ is some finite, positive-definite matrix and $\alpha \in \mathbb{R}$ is a positive scalar. Then

$$\lim_{\alpha \to \infty} \text{ rank } (\hat{Q}\hat{P}) \leq \lim_{\alpha \to \infty} \text{ rank } \hat{Q} \leq n_r, \tag{3.18}$$

where $\hat{Q}$ and $\hat{P}$ are the shifted observability and controllability grammians of the corresponding LQG compensator, satisfying (3.11) and (3.12), respectively. Equivalently, for $\delta > 0$, there exists N such that for all $\alpha > N$, $\lambda_{n_r+1} < \delta\lambda_{n_r}$, where λ_i represents the i^{th} eigenvalue of $\hat{Q}\hat{P}$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_i \geq \lambda_{i+1} \ldots \geq 0$.

Proof. The proof is similar to the proof of Theorem 2.1. □

4. Numerical Illustrative Examples

To illustrate the proper choices of the weighting matrices resulting in a nearly nonminimal, low authority LQG compensator for a stable continuous-time plant, consider a simply supported beam with two collocated sensor/actuator pairs. Assuming the beam has length 2 and that the sensor/actuator pairs are placed at coordinates $a = \frac{55}{172}$, and $b = \frac{46}{43}$, a continuous-time model retaining the first five modes is obtained:

$$\dot{x} = Ax + Bu + D_1 w, \qquad y = Cx + D_2 w,$$

where

$$A = \text{block-diag}(\begin{bmatrix} 0 & 1 \\ -1 & -0.01 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -16 & -0.04 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -81 & -0.09 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -256 & -0.16 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -625 & -0.25 \end{bmatrix}),$$

$$B = \begin{bmatrix} 0 & -0.2174 & 0 & 0.4245 & 0 & -0.6112 & 0 & 0.7686 & 0 & -0.8893 \\ 0 & -0.8439 & 0 & -0.9054 & 0 & -0.1275 & 0 & 0.7686 & 0 & 0.9522 \end{bmatrix}^{\mathrm{T}}, \quad C = B^{\mathrm{T}}.$$

The noise intensities are $V_1 \triangleq D_1 D_1^{\mathrm{T}} = 0.1 I_{10}$ and $V_2 \triangleq D_2 D_2^{\mathrm{T}} = \beta I_2$, and it is assumed that $V_{12} \triangleq D_1 D_2^{\mathrm{T}} = 0$. The design objective is to minimize the continuous-time cost $J = \lim_{t \to \infty} \mathbb{E}[x^{\mathrm{T}} R_1 x + u^{\mathrm{T}} R_2 u]$, where $R_2 = \alpha I_2$. Note that the magnitude of the positive real numbers α and β are the indicators of the controller authority level. For this particular plant, A has the representation as in (2.23a) and (2.26a) with $A_{12} = 0$ and $A_{21} = 0$, respectively. Here, we illustrate the results of property (i) of Theorem 2.1 for the cases of $n_r = 2$ and $n_r = 6$. Setting $\alpha = 0.1$, by selecting the weighting matrix $R_1 = \begin{bmatrix} I_{n_r} & 0 \\ 0 & 0 \end{bmatrix}$, and increasing β (hence, decreasing the compensator authority), the resulting LQG compensator approaches nonminimality with minimal dimension of n_r or, equivalently, $\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})} \rightarrow 0$ where λ_i is the sorted (in descending order) i^{th} eigenvalue of $\hat{Q}\hat{P}$. Figure 1 shows the ratio curve for $n_r = 2$ with $\beta \in (0.01, 0.1, 1, 10, 10^2, 10^3, 10^4, 10^5, 10^6)$. The curve clearly indicates that the ratio decreases as β increases. To illustrate that suboptimal controller reduction methods yield nearly optimal reduced-order compensators for low authority control problems, Figure 1 also shows the norm of the cost gradient of the 2^{nd}-order controller obtained by balancing. The cost gradient is defined as $[(\text{vec}\ \frac{\partial J}{\partial A_c})^{\mathrm{T}} \quad (\text{vec}\ \frac{\partial J}{\partial B_c})^{\mathrm{T}} \quad (\text{vec}\ \frac{\partial J}{\partial C_c})^{\mathrm{T}}\,]^{\mathrm{T}}$. The cost gradient curve indicates the balanced controller approaches the optimal reduced-order compensator as β increases, or as the control authority decreases. Figure 2 shows the eigenvalue ratio of the LQG controller for $n_r = 6$ and the norm of the cost gradient of the corresponding 6^{th}-order balanced controller.

Conversely, if the weighting term R_1 for the above example does not have the structure given by (2.23b), decreasing the controller authority (i.e., increasing β) may not yield a nearly nonminimal LQG compensator. As a result, the norm of the cost gradient of the corresponding 2^{nd}-order balanced controller does not approach zero as the control authority decreases. This is illustrated in Figure 3 for $n_r = 2$ and $R_1 = I_{10}$. Note that for this particular example, at $\beta = 0.01$ the balanced controller destablizes the closed-loop system and hence the norm of the cost gradient becomes infinite.

5. Conclusion

By exploiting structural properties of the solutions of the Riccati equations and Lyapunov equations, this paper shows that for both continuous-time and discrete-time stable systems, if the coefficient matrix A and driving weighting term R_1 (or V_1) have specific structures, the corresponding LQG compensator becomes nonminimal as the control authority is decreased. This result provides a partial explanation of why suboptimal projection methods tend to work best at low authority. This paper also establishes some rigorous guidelines to initialize homotopy algorithms for directly synthesizing optimal reduced-order controllers. In particular, to initialize the homotopy algorithm efficiently the designer can choose the plant and weighting matrices to correspond to a low authority control problem with stable open-loop systems as stated in Theorem 2.1 or 3.1. In this case, a nearly optimal reduced-order controller may be easily obtained using an appropriate suboptimal controller reduction method such as balancing since the resulting LQG controller is nearly non-minimal. These results are clearly illustrated by numerical examples.

Conversely, if the structure of the plant and weighting matrices do not satisfy the conditions specified in Theorem 2.1 or 3.1, the resulting LQG compensator is not necessarily nearly minimal even at low control authority. In this case, reduced-order controllers obtained by suboptimal projection methods may not be nearly optimal even at low authority. This result is illustrated in the last example with a reduced-order controller obtained by balancing.

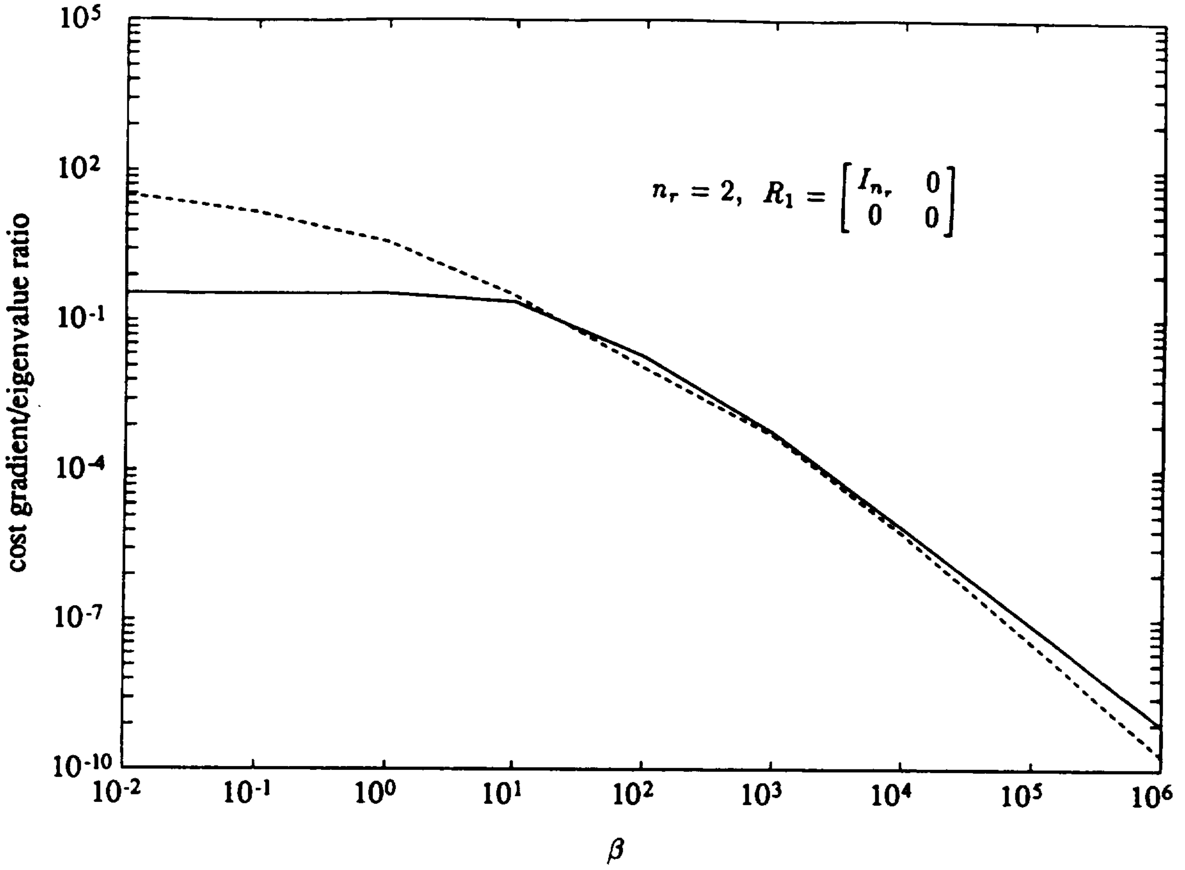

Figure 1. Non-minimality indicator ($\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})}$) of the LQG controller (—) and the norm of the cost gradient of the 2^{nd}-order balanced controller (- - -) vs control authority (β) for $n_r = 2$ and "good structured" R_1

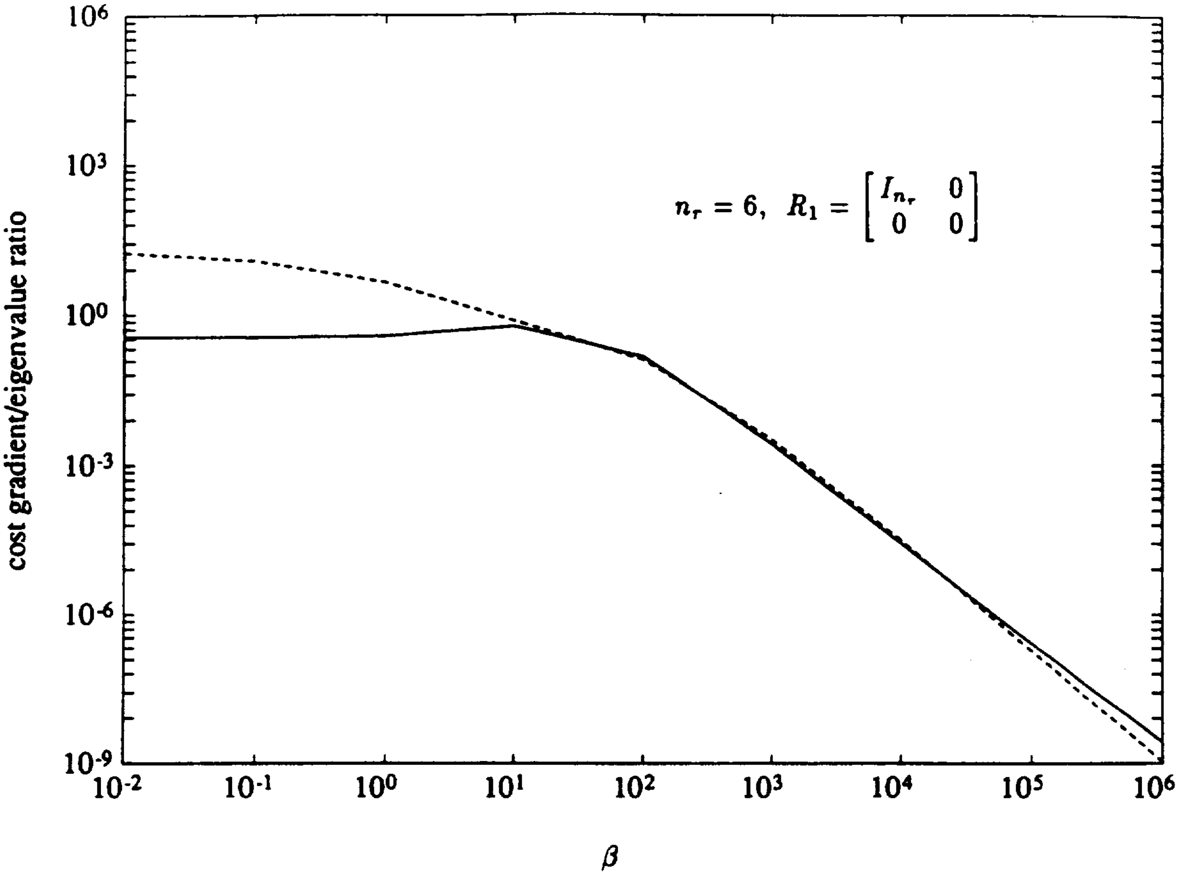

Figure 2. Non-minimality indicator ($\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})}$) of the LQG controller (—) and the norm of the cost gradient of the 6^{th}-order balanced controller (- - -) vs control authority (β) for $n_r = 6$ and "good structured" R_1

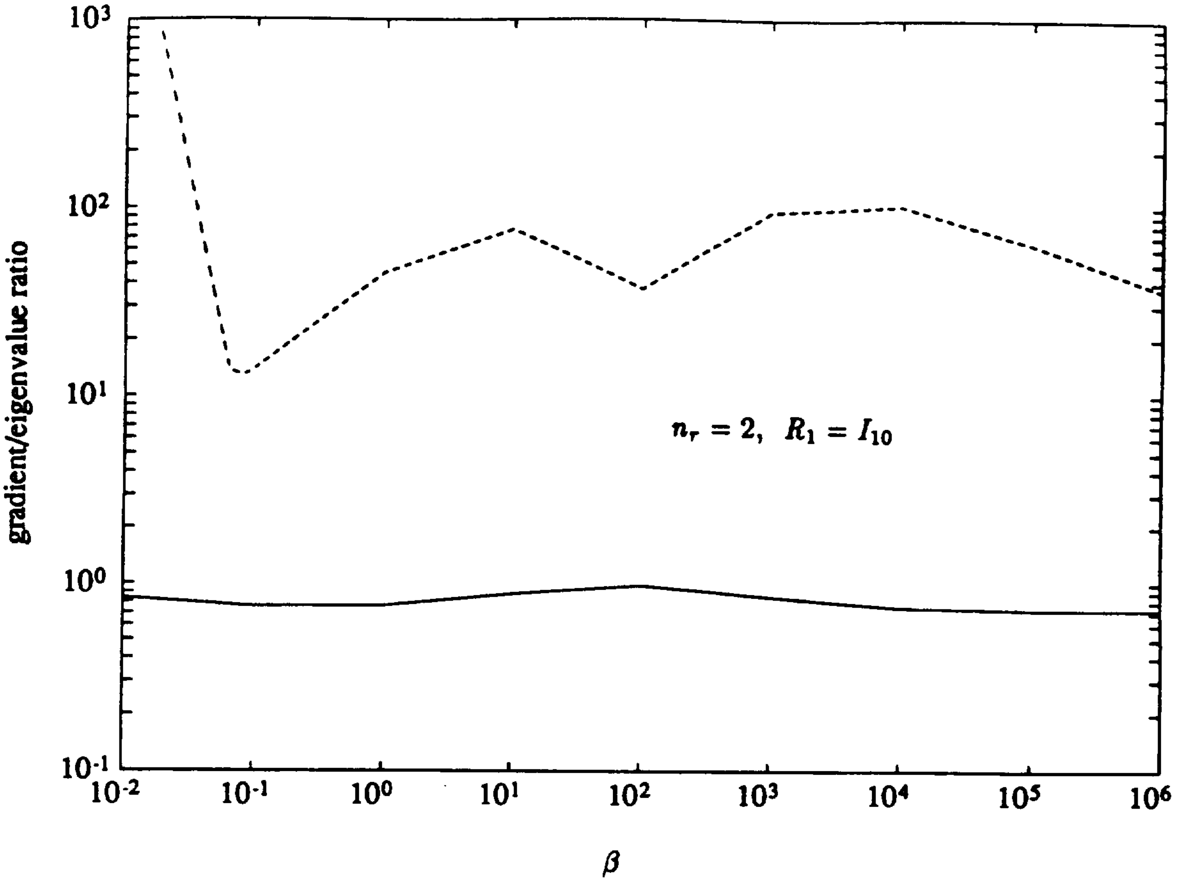

Figure 3. Non-minimality indicator ($\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})}$) of the LQG controller (—) and the norm of the cost gradient of the 2^{nd}-order balanced controller (- - -) vs control authority (β) for $n_r = 2$ and "bad structured" R_1

References

1. M. Athans, "The Role and Use of the Stochastic Linear- Quadratic-Guassian Problem in Control System Design," *IEEE Trans. Autom. Contr.*, Vol. AC-16, pp. 529–552, Dec. 1971.
2. B.D.O. Anderson and J.B. Moore, *Linear Optimal Control*, Prentice-Hall, Englewood Cliffs, N.J., 1971.
3. H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, John Wiley and Sons, New York, 1972.
4. D.C. Hyland and D.S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 1034–1037, 1984.
5. W.S. Levine, T.L. Johnson, and M. Athans, "Optimal Limited State Variable Feedback Controllers for Linear Systems," *IEEE Trans. Autom. Contr.*, Vol. AC-16, pp. 785–793, Dec. 1971.
6. G. D. Martin and A. E. Bryson, Jr., "Attitude Control of a Flexible Spacecraft," *AIAA Journal of Guidance, Control, and Dynamics*, Vol 3, 1980.
7. V. Mukhopadhyay, R. Newsom, and I. Abel, "Reduced-Order Optimal Control Law Synthesis for Flutter Suppression," *AIAA Journal of Guidance, Control and Dynamics*, Vol. 4, pp. 389–395, 1982.
8. 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.
9. V. Mukhopadhyay, "Stability Robustness Improvement using Constrained Optimization Techniques," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 10, pp. 172–177, 1987.
10. U. Kuhn and G. Schmidt, "Fresh Look into the Design and Computation of Optimal Output Feedback Controls for Linear Multivariable Systems," *International Journal of Control*, Vol. 46, pp. 75–95, 1987.
11. S. Richter, "A Homotopy Algorithm for Solving the Optimal Projection Equations for Fixed-Order Dynamic Compensation: Existence, Convergence and Global Optimality," *Proceedings of the American Control Conference*, pp. 1527–1531, Minneapolis, MN, June 1987.
12. P. M. Makila and H. T. Toivonen, "Computational Methods for Parametric LQ Problems–A Survey," *IEEE Transactions on Automatic Control*, Vol. 32, pp 658–671, 1987.
13. F. S. Kramer and A. Calise, "Fixed-Order Dynamic Compensation for Multivariable Linear Systems," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 9, pp. 699–703, 1988.
14. V. Mukhopadhyay, "Digital Robust Control Law Synthesis using Constrained Optimization," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 12, pp. 175–181, 1989.
15. S. Richter and E. G. Collins, Jr., "A Homotopy Algorithm for Reduced-Order Controller Design using the Optimal Projection Equations," *Proc. IEEE Conference on Decision and Control*, pp. 932–933, Tampa, FL, December 1989.
16. M. Mercadal, "Homotopy Approach to Optimal, Linear Quadratic, Fixed Architecture Compensation," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 14, pp. 1224–1233.

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No.

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle

A HOMOTOPY ALGORITHM FOR DIGITAL
OPTIMAL PROJECTION CONTROL

5. Report Date

October 1993

6. Performing Organization Code

7. Author(s)

Emmanuel G. Collins, Jr., Stephen Richter, and
Lawrence D. Davis

8. Performing Organization Report No.

10. Work Unit No.

9. Performing Organization Name and Address

Harris Corporation
Government Aerospace Systems Division
Mail Stop 4849
Melbourne, FL 32902

11. Contract or Grant No.

NAS8 - 38575

13. Type of Report and Period Covered

Contractor Report-Final

12. Sponsoring Agency Name and Address

National Aeronautics and Space Administration
Marshall Space Flight Center
Huntsville, AL 35812

14. Sponsoring Agency Code

15. Supplementary Notes

Technical Monitor - Angelia P. Bukley, Marshall Space Flight Center

16. Abstract

The linear-quadratic-gaussian (LQG) compensator has been developed to facilitate the design of control laws for multi-input, multi-output (MIMO) systems. The compensator is computed by solving two algebraic equations for which standard closed-loop solutions exist. Unfortunately, the minimal dimension of an LQG compensator is almost always equal to the dimension of the plant and can thus often violate practical implementation constraints on controller order. This deficiency is especially highlighted when considering control-design for high-order systems such as flexible space structures. This deficiency has motivated the development of techniques that enable the design of optimal controllers whose dimension is less than that of the design plant. A homotopy approach based on the optimal projection equations that characterize the necessary conditions for optimal reduced-order control. Homotopy algorithms have global convergence properties and hence do not require that the initializing reduced-order controller be close to the optimal reduced-order controller to guarantee convergence. However, the homotopy algorithm previously developed for solving the optimal projection equations has sublinear convergence properties and the convergence slows at higher authority levels and may fail. This report describes a new homotopy algorithm for synthesizing optimial reduced-order controllers for discrete-time systems. Unlike the previous homotopy approach, the new algorithm is a gradient-based, parameter optimization formulation and has been implemented in MATLAB. The results reported here may offer the foundation for a reliable approach to optimal, reduced-order controller design.

17. Key Words (Suggested by Author(s))

discrete-time, control design,
numerical methods, homotopy

18. Distribution Statement

unclassified - unlimited

19. Security Classif. (of this report)

unclassified

20. Security Classif. (of this page)

unclassified

21. No. of pages

22. Price