Homotopy Approach to Optimal, Linear Quadratic, Fixed Architecture Compensation

Mathieu Mercadal*
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Optimal linear quadratic Gaussian compensators with constrained architecture are a sensible way to generate good multivariable feedback systems meeting strict implementation requirements. The optimality conditions obtained from the constrained linear quadratic Gaussian are a set of highly coupled matrix equations that cannot be solved algebraically except when the compensator is centralized and full order. An alternative to the use of general parameter optimization methods for solving the problem is to use homotopy. The benefit of the method is that it uses the solution to a simplified problem as a starting point and the final solution is then obtained by solving a simple differential equation. This paper investigates the convergence properties and the limitation of such an approach and sheds some light on the nature and the number of solutions to the constrained linear quadratic Gaussian problem. It also demonstrates the usefulness of homotopy on an example of an optimal decentralized compensator.

I. Introduction

THE linear quadratic Gaussian (LQG) methodology[1] offers a practical way of generating multi-input/multi-output (MIMO) dynamic output feedback systems with high nominal performance and guaranteed nominal stability for linear time invariant (LTI) systems. The methodology is an optimal control problem whose solution can be found algebraically by solving two uncoupled Riccati equations.[1] The existence of the solution depends on properties of the open-loop system, namely, observability and controllability. The solution is known to be the global minimum, which makes the tool particularly easy to use. The algebraic Riccati equations are solved most reliably using a Hamiltonian approach, which consists of finding and ordering the eigenvalues of a $2n$-dimensional matrix, where n is the dimension of the original system. The LQG methodology has been central in the development of multivariable design techniques.[1-3] However, one principal drawback of the methodology may be the complexity of the feedback laws that are generated. The order of an optimal LQG compensator is the sum of the order of the original plant and those of the various filters used to shape the process noise and/or the cost functional. The compensator is also centralized, each and every control command being based on the information coming from all of the sensors. Hence, for large-scale systems such as flexible structures, LQG compensators face implementability problems. As a result, simplified compensators have been sought. One preferred way has been to fix the order of the compensator: a compensator of order n_c is sought to control a system of order n, where n_c is defined by the designer and typically reflects the maximum number of states that can be implemented in the available digital filter. The order reduction can be performed directly[4-10] or indirectly.[11-14] Indirect methods consist of approximating the LQG compensator with one of specified order. These methods are numerically appealing since they rely on the full-order LQG solution and the order reduction is usually a straightforward, computationally nonintensive, noniterative, algebraic process. Direct methods consist of looking for the compensator of specified order that minimizes the quadratic cost. Such methods yield much better solutions[15]: unlike the indirect ones, they are guaranteed to be stable as long as there is one stabilizing compensator, and the direct minimization yields an optimal solution over the class of reduced-order compensator. Ad hoc reduction techniques, on the other hand, provide suboptimal solutions not guaranteed to stabilize the system, and the closed-loop performance may be greatly degraded.

Further simplification of the feedback law consists of fixing the information pattern.[16-18] A fixed architecture design is one where the control system is required to be made of p independent subcontrollers. Each subcontroller has a fixed order and is connected to selected sensors and actuators. Sensors, as well as actuators, may or may not be shared by different subcontrollers, resulting in very general control architectures[16,18] or stricter decentralized schemes.[17] The implementation requirements are taken directly into account by specifying a priori what can and cannot be achieved in terms of processing capability, networking, and data sampling. Given the architecture specifications, the best compensator is sought (i.e., the one minimizing the quadratic cost) in the class of controllers that has been defined. The approach marries the flexibility of the unconstrained LQG methodology with the necessity to obtain simplified feedback laws.

The main problem in using direct methods is that the optimization is tremendously complicated by the introduction of constraints in the compensator structure. Optimality conditions are easily obtainable, but they form a set of highly coupled equations. When the structure of the compensator is not specified, these equations separate and the problem splits into an optimal full state feedback problem and an optimal filter problem that are independent from one another. This property, known as the separation principle, does not hold when the order of the compensator is specified[8] and/or when the information pattern is constrained.[18] The order reduction implies that one cannot estimate all of the plant's states, and the selection of the information pattern requires the coordination of the various subcontrollers that form the feedback loop. Reducing the order leads to the optimal projection equations of Ref. 8. These equations are two modified Riccati equations coupled via a projection matrix by two Lyapunov equations. The modified Riccati equations cannot be solved using a Hamiltonian technique even when the projection is fixed, and the projection depends on the solution to the two Lyapunov equations. When the processing is decentralized, one can define again two modified Riccati equations, but the coupling

Received Aug. 16, 1990; revision received Jan. 21, 1991; accepted for publication Feb. 19, 1991.

*Postdoctoral Research Associate, Department of Aeronautics and Astronautics, Room 33-105, 77 Massachusetts Avenue; currently, Research Engineer, Advanced Research Department, Dassault Aviation, 78 quai Marcel Dassault, Saint Cloud, France.

does not have the form of a projection and occurs through four Lyapunov equations.[18] The problem appears, therefore, as a complicated parameter optimization, and very little global results have been found. Its solution can be approached using numerical optimization techniques.[5,6,9,10] These methods are based on local information (i.e., the gradient of the cost) and do not provide any global answers about the existence, number, and nature of solutions satisfying the optimality conditions.

Practically, general minimization algorithms necessitate good starting solutions in order to converge rapidly. The design of these initial solutions (i.e., fixed architecture compensators with the required structure) may be as difficult as the design of the optimal compensator. Homotopy algorithms have been proposed to solve difficult optimization and control design problems and offer a way to answer some theoretical questions as well.[19-23] A homotopy method consists of continuously varying the parameter of a problem while tracking its solutions, starting with a particular value of the parameter for which obvious solutions exist and changing this value to one corresponding to a more complex problem. These methods are also based on local information: the derivative of a solution with respect to the problem's parameter is computed and the solution for a new value of the parameter is obtained by integrating this derivative. Hence, the solution is obtained by changing the problem locally and finding the corresponding infinitesimal variation of its solution. The integration can be stabilized by using optimization steps to improve the accuracy. Such steps have high convergence rates since one is always close to the solution of the perturbed problem throughout the integration. The goal of this paper is to present a homotopy algorithm for solving the fixed architecture LQG and to investigate the number and the nature of its possible solutions. Section II states the fixed architecture LQG problem and derives the first- and second-order necessary conditions for optimality. The link between these conditions and the OPE is stated for the reduced-order LQG. Section III presents the homotopy algorithm, showing how to get problems for which solutions can be found easily and how to compute the derivative of the solution with respect to the homotopy parameter. Section IV investigates the convergence properties of such an algorithm and disproves the conjectures made in Refs. 21 and 22 using counter examples. This section shows that the introduction of constraints in the compensator structure has made the optimization considerably more complicated. The algorithm converges only for limited changes in the system and the choice of the simplified problem plays a very important role in the success of the method. Section V presents the design of a decentralized controller of a coupled beam system to illustrate the use of homotopy on a more realistic example. Section VI concludes the paper.

II. Constrained Linear Quadratic Gaussian Problem

A. General

Consider the LTI system:

$$\dot{x} = Ax + Bu + w \tag{1}$$

$$y = Cx + v \tag{2}$$

where $x \in R^n$, $u \in R^m$, $y \in R^l$, $w \in R^n$ is a white noise disturbance with intensity V, and $v \in R^l$ is a white measurement noise with intensity V_c. The pair (A,B) is assumed stabilizable and (C,A) detectable. The quadratic performance criterion to be minimized is

$$J = \tfrac{1}{2} \lim_{t\to\infty} E\{x(t)^T R x(t) + u(t)^T R_c u(t)\} \tag{3}$$

where R is a non-negative, and R_c a positive definite matrix. Furthermore, the pair $(\sqrt{R},A)$ is assumed detectable and $(A,\sqrt{V})$ stabilizable. Under such hypotheses, the optimal control law that minimizes J is an LTI dynamic compensator of order n and the closed-loop system is guaranteed to be stable.[1] The fixed architecture solution to the LQG problem is an LTI dynamic compensator also, but with a specified structure. The compensator has the generic state-space representation:

$$\dot{x}_c = A_c x_c + Ky \tag{4}$$

$$u = Gx_c \tag{5}$$

with the following restrictions:

$$x_c = [x_1^T, x_2^T, \ldots, x_p^T]^T \tag{6}$$

where $x_i \in R^{n_i}$ is the state vector of the ith subcontroller. The subcontrollers have independent dynamics; thus,

$$A_c = \text{blockdiag}\{A_1, A_2, \ldots, A_p\} \tag{7}$$

The matrix K is partitioned into l columns:

$$K = [K_1\, K_2 \cdots K_l] \tag{8}$$

and each column is, in turn, partitioned into p subcolumns according to the partitioning of the state vector x_c:

$$K_j^T = [K_{1j}^T\, K_{1j}^T \cdots K_{pj}^T], \qquad K_{ij} \in R^{n_i} \tag{9}$$

The subcolumn K_{ij}, which corresponds to the gains used to feed the jth measurement into the ith subcontroller, will be free if the sensor is connected to this subcontroller, and zero otherwise. Similarly, G is partitioned into m rows,

$$G^T = [G_1^T\, G_2^T \cdots G_m^T] \tag{10}$$

and each row is partitioned into p subrows,

$$G_i = [G_{i1}\, G_{i2} \cdots G_{ip}], \qquad K_{1j}^T \in R^{n_j} \tag{11}$$

The subrow G_{ij}, which corresponds to the gains on the states of the jth subcontroller used to generate the ith control, will be free if the actuator is connected to this subcontroller and zero otherwise. The architecture is completely defined by the following three sets that contain, respectively, the row and column indices of the matrices A_c, G, and K, which are not set to zero:

$$\mathcal{A} = \{(i,j),\ 1 \le i \le n_c,\ 1 \le j \le n_c,\ a_{ij} \text{ free in } A_c\} \tag{12a}$$

$$\mathcal{G} = \{(i,j),\ 1 \le i \le m,\ 1 \le j \le n_c,\ g_{ij} \text{ free in } G\} \tag{12b}$$

$$\mathcal{K} = \{(i,j),\ 1 \le i \le n_c,\ 1 \le j \le l,\ k_{ij} \text{ free in } K\} \tag{12c}$$

Defining E_{ij}^a, E_{ij}^g, and E_{ij}^k, respectively, as an $n_c \times n_c$, an $m \times n_c$, and an $n_c \times l$ matrix with all zero entries except the (i,j) equal to 1, the matrices A_c, G, and K can be written as

$$A_c = \sum_{(i,j)\in\mathcal{A}} E_{ij}^a a_{ij} \tag{13a}$$

$$G = \sum_{(i,j)\in\mathcal{G}} E_{ij}^g g_{ij} \tag{13b}$$

$$K = \sum_{(i,j)\in\mathcal{K}} E_{ij}^k k_{ij} \tag{13c}$$

The free variables of A_c, G, and K can be gathered into a vector ξ that contains all of the free variables of the problem. The optimization is to find ξ that minimizes $J(\xi)$ such that the corresponding closed-loop system is asymptotically stable. Assume that G, A_c, and K stabilize the system. The cost can then be written as

$$J = \tfrac{1}{2}\text{Tr}\{\tilde{Q}R_{cl}\} \tag{14}$$

where $\tilde{Q}$ is the closed-loop, steady-state covariance matrix given by

$$0 = A_{cl}\tilde{Q} + \tilde{Q}A_{cl}^T + V_{cl} \tag{15}$$

with

$$A_{cl} = \begin{bmatrix} A & BG \\ KC & A_c \end{bmatrix} \tag{16}$$

$$R_{cl} = \begin{bmatrix} R & 0 \\ 0 & G^T R_c G \end{bmatrix}, \qquad V_{cl} = \begin{bmatrix} V & 0 \\ 0 & KV_cK^T \end{bmatrix} \tag{17}$$

The set of stabilizing compensators is an open set, and assuming it is not empty, the solution must render the cost stationary. In order to compute the derivative of the cost with respect to the free parameters of the problem, one must define a Lagrangian to incorporate the dependency of $\tilde{Q}$ on A_c, G, and K through Eq. (15). Define L as

$$L = ½\mathrm{Tr}\{\tilde{Q}R_{cl} + \tilde{P}(A_{cl}\tilde{Q} + \tilde{Q}A_{cl}^T + V_{cl})\} \tag{18}$$

where $\tilde{P}$ is a matrix of Lagrange multipliers that represents the sensitivity of the cost to changes of intensity of the process and the measurement noises. Defining the following partition,

$$\tilde{P} = \begin{bmatrix} P_{00} & P_{0c} \\ P_{c0} & P_{cc} \end{bmatrix}, \qquad \tilde{Q} = \begin{bmatrix} Q_{00} & Q_{0c} \\ Q_{c0} & Q_{cc} \end{bmatrix} \tag{19}$$

the optimality conditions are obtained using the matrix minimum principle[24]:

$$0 = A_{cl}^T\tilde{P} + \tilde{P}A_{cl} + R_{cl} \tag{20}$$

$$0 = A_{cl}\tilde{Q} + \tilde{Q}A_{cl}^T + V_{cl} \tag{21}$$

$$0 = \mathrm{Tr}\{E_{ij}^{aT}(P_{c0}Q_{0c} + P_{cc}Q_{cc})\}, \qquad (i,j) \in \mathcal{A} \tag{22}$$

$$0 = \mathrm{Tr}\{E_{ij}^{gT}[R_cGQ_{cc} + B^T(P_{00}Q_{0c} + P_{0c}Q_{cc})]\}, \; (i,j) \in \mathcal{G} \tag{23}$$

$$0 = \mathrm{Tr}\{E_{ij}^{kT}[P_{cc}KV_c + (P_{c0}Q_{00} + P_{cc}Q_{c0})C^T]\}, \; (i,j) \in \mathcal{K} \tag{24}$$

Remark. When the compensator is centralized ($p = 1$), the optimality conditions, Eqs. (20–24), become simply Eqs. (3.1–5) of Ref. 8, which are used to derive the OPE. The optimal projection τ is, in that case,

$$\tau = -Q_{0c}Q_{cc}^{-1}P_{cc}^{-1}P_{c0} \tag{25}$$

and follows directly from Eq. (22), which is a complete matrix equation when the compensator is centralized and all of the elements in A_c are free.

Remark. When the two Lyapunov equations, Eqs. (20) and (21), are satisfied, $\tilde{P}$ and $\tilde{Q}$ are uniquely defined as functions of (G, A_c, K) (or ξ), except when A_{cl} has some purely imaginary eigenvalues. Since the closed loop is assumed to be asymptotically stable, Eqs. (20) and (21) have unique solutions. The derivatives of the Lagrangian found in Eqs. (22–24) are, in that case, the derivatives of the original cost J. Consider, for example, the derivative of J with respect to the (i,j) element of A_c, a_{ij}, and denote the derivative with respect to such an element J_a for simplicity. Then,

$$J_a = ½\mathrm{Tr}\{\tilde{Q}_aR_{cl}\} \tag{26}$$

since R_{cl} does not depend on A_c, $R_{cla} = 0$. $\tilde{Q}_a$ is found by differentiating Eq. (21), which yields

$$0 = A_{cl}\tilde{Q}_a + \tilde{Q}_aA_{cl}^T + A_{cla}\tilde{Q} + \tilde{Q}A_{cla}^T \tag{27}$$

where

$$A_{cla} = \begin{bmatrix} 0 & 0 \\ 0 & E_{ij}^a \end{bmatrix} \tag{28}$$

An expression for R_{cl} is obtained from Eq. (20) in terms of $\tilde{P}$ and A_{cl}. Replacing the expression in Eq. (26), one gets

$$\begin{aligned} J_a &= -½\mathrm{Tr}\{\tilde{Q}_a(A_{cl}^T\tilde{P} + \tilde{P}A_{cl})\} \\ &= -½\mathrm{Tr}\{\tilde{P}(A_{cl}\tilde{Q}_a + \tilde{Q}_aA_{cl}^T)\} \\ &= ½\mathrm{Tr}\{\tilde{P}(A_{cla}\tilde{Q} + \tilde{Q}A_{cla}^T)\}, \qquad \text{using Eq. (27)} \\ &= \mathrm{Tr}\{A_{cla}^T\tilde{P}\tilde{Q}\} \end{aligned} \tag{29}$$

which, using Eq. (28), is just the expression of Eq. (22), which defines L_a. The derivatives of J with respect to g_{ij} and k_{ij} are found using similar manipulations, differentiating Eqs. (20) and (21), leading to Eqs. (23) and (24).

B. Second-Order Optimality Conditions

A stationary point is a local minimum only if the Hessian $J_{\xi\xi}$, or matrix of second derivatives, is non-negative. The Hessian is also required to implement a continuous homotopy. $J_{\xi\xi}$ is obtained by differentiating Eqs. (22–24) using the two Lyapunov equations, Eqs. (20) and (21), to provide the derivatives of $\tilde{P}$ and $\tilde{Q}$. Consider the derivative of the vector J_ξ with respect to the jth element of ξ, ξ_j. The corresponding derivatives will be denoted with a prime. The derivatives of A_{cl}, R_{cl}, and V_{cl} are simply

$$A_{cl}' = \begin{bmatrix} 0 & BG' \\ K'C & A_c' \end{bmatrix} \tag{30}$$

$$R_{cl}' = \begin{bmatrix} 0 & 0 \\ 0 & G'^TR_cG + G^TR_cG' \end{bmatrix} \tag{31}$$

$$V_{cl}' = \begin{bmatrix} 0 & 0 \\ 0 & K'V_cK^T + KV_cK'^T \end{bmatrix} \tag{32}$$

where A_c', G', and K' are matrices filled with zeros except for the element corresponding to ξ_j, which is equal to 1. $\tilde{P}'$ and $\tilde{Q}'$ satisfy

$$0 = A_{cl}^T\tilde{P}' + \tilde{P}'A_{cl} + A_{cl}'^T\tilde{P} + \tilde{P}A_{cl}' + R_{cl}' \tag{33}$$

$$0 = A_{cl}\tilde{Q}' + \tilde{Q}'A_{cl}^T + A_{cl}'\tilde{Q} + \tilde{Q}A_{cl}'^T + V_{cl}' \tag{34}$$

Define $\tilde{M}' = \tilde{P}'\tilde{Q} + \tilde{P}\tilde{Q}'$ and block partition $\tilde{M}'$, $\tilde{P}'$, and $\tilde{Q}'$ as follows:

$$\tilde{M}' = \begin{bmatrix} M_{00}' & M_{0c}' \\ M_{c0}' & M_{cc}' \end{bmatrix} \tag{35}$$

$$\tilde{P}' = \begin{bmatrix} P_{00}' & P_{0c}' \\ P_{c0}' & P_{cc}' \end{bmatrix} \tag{36}$$

$$\tilde{Q}' = \begin{bmatrix} Q_{00}' & Q_{0c}' \\ Q_{c0}' & Q_{cc}' \end{bmatrix} \tag{37}$$

The column $J_{\xi\xi_j}$ is obtained element by element by differentiating Eqs. (22–24) using Eqs. (33) and (34). The computation yields

$$J_{A_{ij}\xi_j} = \mathrm{Tr}\{E_{ij}^{aT}M_{cc}'\} \tag{38}$$

$$J_{G_{ij}\xi_j} = \mathrm{Tr}\{E_{ij}^{gT}(R_cG'Q_{cc} + R_cGQ_{cc}' + B^TM_{0c}')\} \tag{39}$$

$$J_{K_{ij}\xi_j} = \mathrm{Tr}\{E_{ij}^{kT}(P_{cc}K'V_c + P'_{cc}KV_c + M'_{c0}C^T)\} \tag{40}$$

Equations (38–40) provide a closed-form expression for the entries of $J_{\xi\xi}$. The expression is too complicated, however, to provide any insight into the nature of the stationary point that might be obtained. In order for a stationary point to be a minimum, J must be non-negative. For the compensator to be stabilizing, $\tilde{P}$ and $\tilde{Q}$ must be non-negative as well.

III. Homotopy Algorithm

A. General

The following procedure was proposed originally in Refs. 21 and 22 to solve the reduced-order LQG and is adapted here for more general fixed architecture LQG problems. Define a family of LQG problems parameterized by a scalar $\alpha \in [0, 1]$ as follows:

$$A(\alpha) = A_0 + f_1(\alpha)(A_1 - A_0) \tag{41a}$$

$$B(\alpha) = B_0 + f_2(\alpha)(B_1 - B_0) \tag{41b}$$

$$C(\alpha) = C_0 + f_3(\alpha)(C_1 - C_0) \tag{41c}$$

$$R(\alpha) = R_0 + f_4(\alpha)(R_1 - R_0) \tag{41d}$$

$$V(\alpha) = V_0 + f_5(\alpha)(V_1 - V_0) \tag{41e}$$

$$R_c(\alpha) = R_{c0} + f_6(\alpha)(R_{c1} - R_{c0}) \tag{41f}$$

$$V_c(\alpha) = V_{c0} + f_7(\alpha)(V_{c1} - V_{c0}) \tag{41g}$$

The $f_i(\alpha)$ are right differentiable functions of such that $f_i(0) = 0$ and $f_i(1) = 1$. The problem of finding a solution for each α can be written as

$$0 = J_\xi(\xi, \alpha) \tag{42}$$

Because J_ξ is differentiable, one can find the derivative of the solution with respect to α, ξ_α by differentiating Eq. (42) and integrating Davidenko's differential equation:

$$0 = J_{\xi\xi}\xi_\alpha + J_{\xi\alpha} \tag{43}$$

Assuming a solution to the control problem is known for $\alpha = 0$, a solution to the problem for $\alpha = 1$ could be written as

$$\xi(1) = \xi(0) + \int_0^1 \xi_\alpha \, \mathrm{d}\alpha \tag{44}$$

provided that the integrand is well defined and that Eq. (43) has a solution. This procedure is the basis of the homotopy. Instead of solving a complicated system of nonlinear equations such as Eqs. (20–24), the problem has become a simple differential equation.

B. Diagonal Problem

Consider the LQG problem such that the matrices A_0, R_0, V_0 have a block-diagonal structure with $p + 1$ blocks:

$$A_0 = \mathrm{blockdiag}(A_1, A_2, \ldots, A_p, A_{p+1})$$

$$R_0 = \mathrm{blockdiag}(R_1, R_2, \ldots, R_p, R_{p+1})$$

$$V_0 = \mathrm{blockdiag}(V_1, V_2, \ldots, V_p, V_{p+1})$$

where, for $i \leq p$, A_i, R_i, and V_i are matrices of dimension n_i, the order of the ith subcontroller, with R_i and V_i non-negative. A_{p+1}, R_{p+1}, and V_{p+1} are square matrices of proper dimension such that A_0, R_0, and V_0 are n dimensional. It follows from the partitioning that $\Sigma n_i \leq n$. Assume that the matrices B_0 and C_0 have the following partition:

$$B_0 = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ 0 & B_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & B_{pp} \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$C_0 = \begin{bmatrix} C_{11} & 0 & \cdots & 0 & 0 \\ 0 & C_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & C_{pp} & 0 \end{bmatrix}$$

where B_{ii} is $n_i \times m_i$, and where C_{ii} is $l_i \times n_i$. Finally,

$$R_{c0} = \mathrm{diag}(r_{10}, r_{20}, \ldots, r_{m0}), \qquad r_{i0} \in R$$

$$V_{c0} = \mathrm{diag}(v_{10}, v_{20}, \ldots, v_{l0}), \qquad v_{i0} \in R$$

Assume that the control architecture specifies that the first m_1 actuators and l_1 sensors are used, but not exclusively, by subcontroller 1, the next m_2 actuators and l_2 sensors are used, but not exclusively, by subcontroller 2, etc. The hypothesis is not restrictive since one can always renumber the sensors and actuators. Further, assume that subcontroller 1 has order n_1, subcontroller 2, order n_2, etc. Because of the diagonal structure of all of the matrices in the problem, the unconstrained LQG solution to the problem does have the required architecture and the problem reduces to solving p independent LQGs of smaller dimension. Sensors and actuators can be shared by more than one loop without affecting the result since a given actuator (respectively, sensor) can affect (sense) one and one subsystem only. Thus, for any given architecture, and for $\Sigma n_i \leq n$ (i.e., the total order of the controller is less than that of the plant), it is always possible to find a set of parameters A_0, B_0, C_0, etc., of proper dimensions for which the unconstrained LQG solution has the desired control structure.

When one desires a controller whose total order is greater than that of the plant, a trick must be played. One can augment the size of the matrices A_1, B_1, C_1, etc., Eqs. (41), with stable uncontrollable and unobservable modes so that the final problem $\alpha = 1$ is unchanged and so that the constraint on the dimension is satisfied. If A denotes the dynamics of the system to be controlled, one can, for example, reproduce in the starting matrix A_0 a mode of A on which two distinct subcontrollers can have equal influence with the chosen control architecture. The matrices B_0 and C_0 are such that each subcontroller acts independently on one and only one of the duplicated modes for $\alpha = 0$; for $\alpha = 1$, the augmented matrices B_1 and C_1 permit the control and the observation of the actual mode, the original one in A, while A_1 contains a now useless stable uncontrollable and unobservable mode. One can also duplicate sensors and actuators as long as the added elements do not affect the final problem, but help find a diagonal starting point to the procedure. In every case, one can find a starting point to the homotopy that is easy to compute.

C. Computing $J_{\xi\alpha}$

The derivatives of $A(\alpha)$, $B(\alpha)$, $C(\alpha)$, $R(\alpha)$, $V(\alpha)$, $R_c(\alpha)$, and $V_c(\alpha)$ are obtained by differentiating Eqs. (41). A practical choice for the f_i is simply $f_i(\alpha) = \alpha$. $\tilde{P}_\alpha$ and $\tilde{Q}_\alpha$ satisfy

$$0 = A_{cl}^T\tilde{P}_\alpha + \tilde{P}_\alpha A_{cl} + A_{cl\alpha}^T\tilde{P} + \tilde{P}A_{cl\alpha} + R_{cl\alpha} \tag{45}$$

$$0 = A_{cl}\tilde{Q}_\alpha + \tilde{Q}_\alpha A_{cl}^T + A_{cl\alpha}\tilde{Q} + \tilde{Q}A_{cl\alpha}^T + V_{cl\alpha} \tag{46}$$

where

$$A_{cl\alpha} = \begin{bmatrix} A_\alpha & B_\alpha G \\ KC_\alpha & 0 \end{bmatrix} \tag{47}$$

$$R_{cl\alpha} = \begin{bmatrix} R_\alpha & 0 \\ 0 & G^T R_{c\alpha} G \end{bmatrix}, \qquad V_{cl\alpha} = \begin{bmatrix} V_\alpha & 0 \\ 0 & KV_{c\alpha}K^T \end{bmatrix} \tag{48}$$

Partitioning $\tilde{M} = \tilde{P}\tilde{Q}$, $\tilde{P}_\alpha$, $\tilde{Q}_\alpha$, and $\tilde{M}_\alpha = \tilde{P}_\alpha\tilde{Q} + \tilde{P}\tilde{Q}_\alpha$ similarly to Eqs. (36) and (37), one gets the following expressions for the elements of $J_{\xi\alpha}$:

$$J_{aij\alpha} = \mathrm{Tr}\{E_{ij}^{aT} M_{cc\alpha}\} \tag{49}$$

$$J_{gij\alpha} = \mathrm{Tr}\{E_{ij}^{gT}(R_{c\alpha}GQ_{cc} + R_cGQ_{cc\alpha} + B_\alpha^T M_{0c} + B^T M_{0c\alpha})\} \tag{50}$$

$$J_{kij\alpha} = \mathrm{Tr}\{E_{ij}^{kT}(P_{cc\alpha}KV_c + P_{cc}KV_{c\alpha} + M_{0c\alpha}C^T + M_{0c}C_\alpha^T)\} \tag{51}$$

Remark. The cost does not change when the state-space representation of the compensator is changed. Considering the continuous similarity transformation $T(\epsilon) = I + \epsilon N$, where N is block diagonal in order to preserve the architecture constraints, T is invertible for small ϵ and its derivative with respect to ϵ is equal to N. Vectors obtained by rearranging in a column form the elements of the matrices $G_\epsilon = GN$, $A_{c\epsilon} = -NA_c + A_cN$, and $K_\epsilon = -NK$ that correspond to the free entries of G, A_c, and K will be in the null space of the Hessian. Thus, the dimension of the null space is larger or equal to $\Sigma_{i=1}^p n_i^2$. One can show, however, that the vector $J_{\xi\alpha}$ is always perpendicular to the space spanned by these null vectors.[18] Hence, the system becomes singular only when the null space of $J_{\xi\xi}$ is not exclusively constituted by such vectors. Davidenko's differential equation, Eq. (43), can therefore be solved most of the time. Practically, one can use least squares, or insist on a canonical representation of the compensator. Times where it cannot be solved will correspond to bifurcation points.

D. Algorithm Structure

The homotopy algorithm is a numerical implementation of the integral formula, Eq. (44). One feature of homotopy is that one can check the accuracy of the solution by evaluating the norm of the gradient $\| J_\xi \|$ and improve it by performing an optimization at any given α. Because the integration step provides a near-optimal solution, the optimization will converge rapidly, and one can therefore rely on it strongly to obtain fast convergence and good accuracy, rather than utilizing fancier integration techniques and smaller steps. The algorithm is as follows:

Step 1. Find initial solution to diagonal problem.

Step 2. Compute $J_{\xi\xi}$, $J_{\xi\alpha}$, and ξ_α. If system is singular, stop.

Step 3. Find $\Delta\alpha_k$ such that

$$\alpha_{k+1} = \alpha_k + \Delta\alpha_k$$

$$\xi_{k+1}^0 = \xi_k + \Delta\alpha_k \xi_{\alpha_k}$$

$$0.5\epsilon_1 \le \| J_\xi(\xi_{k+1}^0, \alpha_{k+1}) \| \le 2.0\epsilon_1$$

Step 4. Minimize J at $\alpha_{k+1} = \alpha_k + \Delta\alpha_k$ until

$$\| J_\xi(\xi_{k+1}, \alpha_{k+1}) \| \le \epsilon_2$$

If minimization failed, halve ϵ_1. If $\epsilon_1 \le \epsilon$, stop. Else, repeat step 3 with new ϵ_1.

Step 5. Repeat steps 2–4 until $\alpha = 1$.

Remark. High accuracy is not required except for $\alpha = 1$. Hence, ϵ_2 can be coarse throughout the path and be much smaller at the end. Increasing ϵ_1 increases the integration step. One is limited, however, by the fact that the optimization of step 4 is less efficient with a poorer starting point. One can implement an adaptation on the size of ϵ_1 based on the number of steps required to minimize J.

Remark. A tridiagonal form is chosen for the dynamic matrices of the subcontrollers to reduce the singularity of $J_{\xi\xi}$ while handling problems of repeated eigenvalues, eigenvalue merging, etc. The Hessian is still singular, with a smaller null space, and modified Newton steps are used in the minimization. The search direction is obtained by solving in a least-squares sense the system

$$J_{\xi\xi_k}\Delta\xi_k = -J_{\alpha k} \tag{52}$$

where $\Delta\xi_k$ is the new direction and $J_{\alpha k}$ is the current value of the gradient.[25] Quadratic convergence is still obtained close to the minimum since the cost is invariant to all orders in modifications of the realization of the compensator transfer function.

Remark. Step 4 stabilizes the integration if and only if the solution remains a stabilizing compensator and a local minimum for the cost. This algorithm cannot follow solutions corresponding to saddle points or local maxima and will fail to track a solution path should a bifurcation arise.

IV. Homotopy Convergence

The homotopy approach presented in the previous section may fail for different reasons that must be studied here. Reference 21 shows the method to globally converge in the case of the unconstrained LQG, and a conjecture is proposed in Refs. 21 and 22 that all of the diagonal solutions will connect to solutions to the deformed problem in the case of the reduced-order LQG, proposing a bound on the possible number of solutions. This section proves otherwise and shows that the method does not have global convergence properties when architecture constraints are imposed, be they order reduction or processing decentralization constraints. The convergence will depend on the choice of the initial problem, the final problem, as well as the deformation that is undertaken. It is shown that 1) the nature of the solution is not invariant along one path, as the problem's parameters are changed, and 2) all of the solutions to the simple diagonal problem are not diagonal. It is shown, furthermore, that bifurcations occur when critical points are encountered and that the number of local minima depends not only on the architecture, but also on the parameters of the problem. Example 1 shows a continuous path of solutions where, for some values of α, the closed loop is stable, whereas for other values, the closed loop is unstable. Example 2 exhibits a diagonal problem having a nondiagonal solution that cannot be found by solving simplified LQG problems like the diagonal ones. Example 3 shows a bifurcation occurring as the solution on a path changes from being a local minimum to being a saddle point. Finally, example 4 shows two paths of solutions such that, depending on the value of α, the solutions correspond to two local minima or a minima and a saddle point.

A. Example 1: Closed-Loop Stability

Consider the second-order single-input single-output (SISO) system:

$$C = [3/4 - \alpha/2 \quad 1/4 + \alpha/2], \quad A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with $\alpha \in [0,1]$ and

$$R = V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad R_c = V_c = 1$$

The plant has two poles, at -1 and $+1$, and a zero at $-1/2 + \alpha$. Putting the compensator in controller canonical form ($k = 1$), the closed-loop characteristic polynomial becomes

$$\phi(s) = s^3 - a_c s^2 - 2gs + a_c + (2\alpha - 1)g/2 \tag{53}$$

The closed-loop system is stable only if $\alpha < ½$ or, in other words, if the zero of the system is minimum phase. Let a_{c0} and g_0 be an optimal solution for $\alpha_0 \leq ½$. Consider now closing the loop around the system where $\alpha = 1 - \alpha_0$ with a compensator such that $g = g_0$ and $a_c = -a_{c0}$. The closed-loop characteristic polynomial for the first system is

$$\phi(s) = s^3 - a_{c0} s^2 - 2g_0 s + a_{c0} + (2\alpha_0 - 1)g_0/2$$

and for the second system:

$$\phi(s) = s^3 + a_{c0} s^2 - 2g_0 s - a_{c0} - (2\alpha_0 - 1)g_0/2$$

Hence, the poles of the second system are the mirror images about the imaginary axis of those of the first system. If a_{c0}, g_0, and $k = 1$ make the cost stationary for $\alpha = \alpha_0$, then $g = g_0$ and $a_c = -a_{c0}$ are two values of the control parameters that, along with $k = 1$, satisfy the optimality conditions for $\alpha = 1 - \alpha_0$. The corresponding matrices $\tilde{P}$ and $\tilde{Q}$ for the new values or the controllers are, respectively,

$$\tilde{P}_1 = \begin{bmatrix} -p_{22} & -p_{12} & p_{23} \\ -p_{21} & -p_{11} & p_{13} \\ p_{32} & p_{31} & -p_{33} \end{bmatrix}, \quad \tilde{Q}_1 = \begin{bmatrix} -q_{22} & -q_{12} & q_{23} \\ -q_{21} & -q_{11} & q_{13} \\ q_{32} & q_{31} & -q_{33} \end{bmatrix}$$

if the following $\tilde{P}_0$ and $\tilde{Q}_0$

$$\tilde{P}_0 = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}, \quad \tilde{Q}_0 = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}$$

satisfy the optimality conditions for $\alpha = \alpha_0$, $a_c = a_{c0}$, and $g = g_0$. One can define a projection τ using the matrices $\tilde{P}_1$ and $\tilde{Q}_1$ in Eq. (25). (Such a solution is not considered the solution to the OPE in Ref. 8 since the closed loop is unstable, even though a projection can be found and the equations are satisfied.) The associated cost $J_1 = \mathrm{Tr}(\tilde{Q}_1 R_{cl})$ is equal to $-J_0$. As the zero crosses the imaginary axis, the cost becomes infinite. The negative values of the cost have no physical meaning. $\mathrm{Tr}(\tilde{Q}R_{cl})$ is still defined, however. Figure 1 shows the control parameters. The solution can be extended for $\alpha = ½$ by continuity, and the derivative of the solution can be defined as well. There exists, therefore, a continuous solution path such that the closed loop is either stable or unstable.

B. Example 2: Nondiagonal Solutions to Diagonal System

Consider the initial problem where

$$A = \mathrm{diag}(a_1, a_2, \ldots, a_n)$$

$$B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad B_1 = \mathrm{diag}(b_1, b_2, \ldots, b_m)$$

$$C = [C_1 \quad 0], \quad C_1 = \mathrm{diag}(c_1, c_2, \ldots, c_l)$$

where R, V, R_c, and V_c are diagonal as well. Such an LQG problem consists of controlling $\inf(l, m, n)$ completely decoupled first-order SISO systems with an n_c-dimensional compensator. One obvious solution is to select n_c out of $\inf(l, m, n)$ subsystems that are controllable and observable and control them independently using first-order controllers, solving n_c independent first-order LQG problems. The projections τ associated with such solutions are diagonal. The ith diagonal element is $\tau(i, i) = 1$ if one chooses to control the ith mode of the plant a_i and $\tau(i, i) = 0$ otherwise. Such solutions are designated diagonal solutions. Letting n_u be the number of unstable

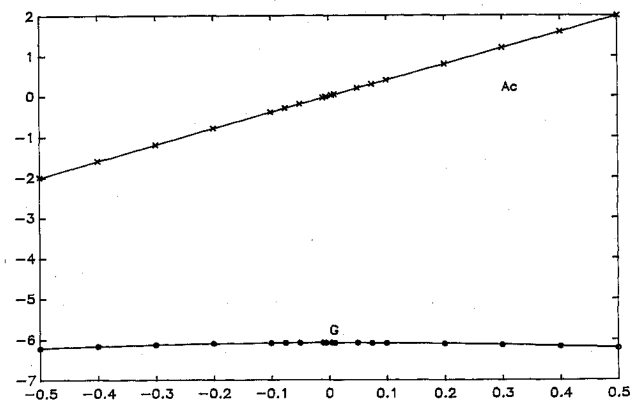

Fig. 1 Compensator parameters making the cost stationary.

poles, one must control each of these modes in order to have a stable closed-loop system; $n_c - n_u$ modes can still be controlled out of $\inf(n, l, m) - n_u$ remaining controllable and observable stable modes. The number of such possible combinations is

$$\begin{bmatrix} \inf(l, m, n) - n_u \\ n_c - n_u \end{bmatrix}$$

or 1 if the number is not defined. Consider now the second-order system:

$$A = \begin{bmatrix} -0.01 & 0 \\ 0 & -0.1 \end{bmatrix}, \quad B = C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$R = V = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad R_c = V_c = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

with a first-order compensator ($n_c = 1$). This example corresponds to the case where $n_c < \inf(l, m)$. The number of combinations is equal to 2 in this case, and there are two diagonal solutions to the problem. The first one is

$$a_c = -1.9901, \quad G = \begin{bmatrix} 0.9900 \\ 0 \end{bmatrix}$$

$$K = [-0.9900 \quad 0], \quad \tau = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

with a cost $J = 81.9605$, and the second one is

$$a_c = -3.9050, \quad G = \begin{bmatrix} 0 \\ 1.9025 \end{bmatrix}$$

$$K = [0 \quad -1.9025], \quad \tau = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

with a cost $J = 64.4961$. A nondiagonal solution can be found numerically by direct minimization. This solution is

$$a_c = -3.6802, \quad G = \begin{bmatrix} 0.7037 \\ -1.6535 \end{bmatrix}$$

$$K = [-0.7037 \quad 1.6535], \quad \tau = \begin{bmatrix} 0.1537 & -0.3606 \\ -0.3606 & 0.8463 \end{bmatrix}$$

The cost associated with this solution is $J = 56.9492$. The nondiagonal solution appears to be the minimum for the problem and yields a much smaller cost than both the diagonal ones.

If R and V are continuously changed from their current values to

$$R = V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Table 1 Eigenvalues of the Hessian

α	s_1	s_2	s_3	s_4	s_5
−0.0300	0.000	2.020	0.062	3.590	38.32
−0.0280	0.000	0.094	0.062	3.590	43.13
−0.0279	0.000	−0.018	0.062	3.590	43.39
−0.0270	0.000	−1.098	0.062	3.590	45.92
−0.0250	0.000	−4.080	0.062	3.590	52.48

the two diagonal solutions will subsist, the first one yielding the smallest cost, but the nondiagonal solution will quickly become a saddle point: the solution does not correspond to a local minimum any more. The algorithm based on minimization to stabilize the integration will leave the path of nondiagonal solution at the critical point and will converge to the first diagonal solution. The occurrence of critical points is illustrated in the next section.

C. Example 3: Critical Points and Bifurcations

A critical point occurs when $J_{\xi\xi}$ has a vector in its null space that does not correspond to a change in the realization of the compensator transfer function. Consider the following second-order system:

$$A = \begin{bmatrix} \alpha & 0 \\ 0 & 0.1 \end{bmatrix}$$

$$B = C = R = V = R_c = V_c = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

where α is the homotopy parameter and is varied in the interval $[-0.03, -0.01]$. A reduced-order compensator of order $n_c = 1$ is required. Because the second pole of the system is unstable and must be stabilized, the problem admits only one diagonal solution,

$$\tau = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

This particular solution is independent of α, but not its Hessian. The system has two inputs and two outputs, and since $n_c = 1$, ξ has five elements. Table 1 shows the eigenvalues of $J_{\xi\xi}$, $s_1, s_2, \ldots, s_5$, as a function of α for α between -0.03 and -0.025. Notice that s_1 is always zero. This corresponds to the predicted singularity of J, which results from the possibility to scale the compensator state variable without affecting the optimality of the solution. The occurrence of a critical point can be seen, however, on s_2. The eigenvalue is positive for $\alpha = -0.0280$ but becomes negative for $\alpha = -0.0279$. The solution starts out as a local minimum and then becomes a saddle point. The critical solution occurs for a value of $\alpha_c = -0.02791583$. For $\alpha \geq \alpha_c$, one can find numerically, by direct optimization, better nondiagonal solutions. A nondiagonal solution was found for $\alpha = -0.01$, and it was subsequently tracked by homotopy as a function of α, as α was moved back to its critical value. The result is that the path of nondiagonal solutions merges with the diagonal solution at α_c: a bifurcation has occurred at the critical point. Figure 2 shows the difference between the parameters of the diagonal solution that is originally followed and the parameters of the nondiagonal solution that appears at α_c; $dG(1)$ denotes the difference in the first entry of the gain matrix G, $dG(2)$, the difference in the second entry, and so forth. In this particular example, $K(1) = G(1)$ and $K(2) = -G(2)$, hence, dK is not plotted. The cost yielded by the diagonal solution increases rapidly as α reaches 0 and as the first mode becomes unstable. As α becomes positive, the diagonal solution becomes unstable, as in example 1.

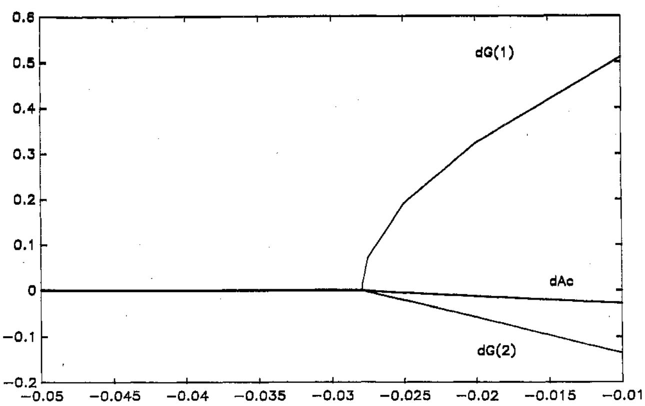

Fig. 2 Bifurcation along a path of solutions.

D. Example 4: Multiple Local Minima

Consider, finally, the fourth-order system:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & -0.1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & -0.1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1-\alpha \\ 0 \\ \alpha \end{bmatrix}$$

$$C = [1 \quad 0 \quad 1 \quad 0]$$

with the LQG parameters

$$R = V = I_4, \qquad R_c = V_c = 1$$

A second-order compensator is sought for the problem. A solution S_0 is found numerically for $\alpha = 0$, and a solution S_1 is found numerically for $\alpha = 1$. The solution S_0 is integrated from $\alpha = 0$ forward, and the solution S_1 is integrated from $\alpha = 1$ backward. At $\alpha = 0.055$, two solutions are found. From the forward integration, one gets

$$A_c = \begin{bmatrix} -0.6905 & 0.4230 \\ -1.3465 & -0.6744 \end{bmatrix}, \qquad K = \begin{bmatrix} 0.5652 \\ -0.1005 \end{bmatrix}$$

$$G = [0.0621 \quad -0.5307], \qquad J = 29.3008$$

The nonzero eigenvalues of the Hessian are, in that case,

$$\Lambda(J_{\xi\xi}) = \{46.6288, 13.6840, 1.7978, 0.0771\}$$

From the backward integration, one gets

$$A_c = \begin{bmatrix} -2.2471 & -0.9574 \\ -1.5384 & -2.1382 \end{bmatrix}, \qquad K = \begin{bmatrix} 1.2721 \\ 0.1405 \end{bmatrix}$$

$$G = [-0.1802 \quad -1.2304], \qquad J = 29.3030$$

The nonzero eigenvalues of the Hessian are, in that case,

$$\Lambda(J_{\xi\xi}) = \{47.7989, 5.8213, 0.3235, 0.0015\}$$

Both solutions are local minima for the cost. Notice that the fourth eigenvalue of the Hessian is very small in both cases. As α is increased, the path of solutions P_0 originating from S_0 encounters a critical point and the solutions are saddle points for larger values of α. Similarly, as α is decreased, the path of solutions P_1 originating from S_1 encounters a critical point and the solutions are saddle points for smaller α. The two compensators found for $\alpha = 0.055$ yield very similar values of the cost. Yet, as seen in Table 2, their pole-zero structures are very different: one compensator is minimum phase with two oscillatory poles, whereas the second has a fast and a slow real mode and is nonminimum phase.

Table 2 Compensator characteristics for $\alpha = 0.055$

	Forward solution		Backward solution	
Closed-loop poles	−0.4769	±j 0.5188	−3.4983	
			−0.4864	
	−0.2509	±j 1.0592	−0.2450	±j 1.0053
	−0.0546	±j 1.4135	−0.0554	±j 1.4128
Compensator poles	−0.6825	±j 0.7547	−0.9778	
			−3.4075	
Compensator zero	−5.2212		3.8647	
Optimal cost	29.3008		29.3030	

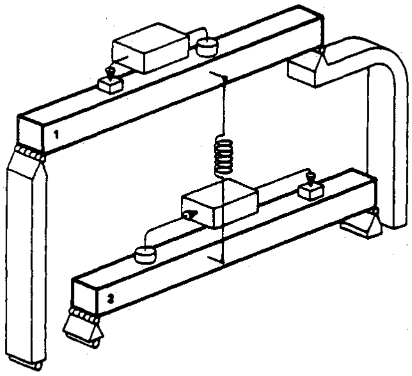

Fig. 3 Interconnected beams (from Ref. 17).

V. Decentralized Control Example

In this section, we consider a more complicated application, which is the decentralized control problem considered in Ref. 17. The system to control is made of a pair of simply supported Euler-Bernoulli beams connected by a spring (Fig. 3). Each beam has one rate sensor and one force actuator. Two vibrational modes are retained to describe each beam, and the state-space representation of the system is an eighth-order interconnected model. The expression for the A, B, and C matrices have been derived in Ref. 17:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}, \quad C = \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix}$$

where ω_{ij} and ζ_{ij} are, respectively, the jth modal frequency and damping ratio of the ith beam, k is the spring constant, c_i the position of the spring attachment, a_i the actuator location, and s_i the sensor location on the ith beam, all distances being nondimensionalized by the beam length. The parameters found in Ref. 17 are

$$\omega_{1i} = 1.00 \text{ rad/s} \qquad \omega_{2i} = 4.00 \text{ rad/s} \qquad \zeta_i = 0.005$$

$$a_1 = 0.3 \qquad s_1 = 0.65 \qquad c_1 = 0.6$$

$$a_2 = 0.8 \qquad s_2 = 0.2 \qquad c_2 = 0.4$$

The penalty and noise intensity matrices are

$$R = \text{diag}\{1,\ 1/\omega_{11},\ 1,\ 1/\omega_{21},\ 1,\ 1/\omega_{12},\ 1,\ 1/\omega_{22}\}$$

$$V = \text{diag}\{0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1\}$$

The controller consists of two decentralized compensators, each of which uses the sensor and actuator of only one beam. The optimal solution is sought for $k = 10$ using the homotopy algorithm and two different diagonal systems to get initial solutions.

A. Using Homotopy: Continuous Solution Path

Because of symmetry, the eigenstructure of the interconnected beam system can be separated into two distinct classes of modes: the group modes and the spring modes. The group modes are located at $\omega = 1.0$ and 4.0 rad/s and correspond to both beams bending along their first or second mode in such a way that the spring is never stretched. The spring modes are located at $\omega = 3.1077$ and 5.6870 rad/s and are dominated by the spring stiffness. This property can be exploited to find an initial diagonal problem. The system can be put in modal form and the modes ordered so that the first two are the group modes and the remaining ones are the spring modes. A simplified B_0 is found by making the spring modes uncontrollable from the first actuator and the group modes uncontrollable from the second actuator. Similarly a simplified C_0 is obtained by making the spring modes unobservable from the first sensor and the group modes unobservable from the second sensor. The matrices R and V, after transformation, happen to have the right block-diagonal structure in that case. The initial solution is found by solving two fourth-order unconstrained

where

$$A_{ii} = \begin{bmatrix} 0 & \omega_{1i} & 0 & 0 \\ -\omega_{1i} - k/\omega_{1i}\sin(\pi c_i)^2 & -2\zeta_i\omega_{1i} & -k/\omega_{2i}\sin(\pi c_i)\sin(2\pi c_i) & 0 \\ 0 & 0 & 0 & \omega_{2i} \\ -k/\omega_{1i}\sin(\pi c_i)\sin(2\pi c_i) & 0 & -\omega_{2i} - k/\omega_{2i}\sin(\pi c_i)^2 & -2\zeta_i\omega_{2i} \end{bmatrix}$$

$$A_{ij} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -k/\omega_{1j}\sin(\pi c_i)\sin(\pi c_j) & 0 & k/\omega_{2j}\sin(\pi c_i)\sin(2\pi c_j)0 & 0 \\ 0 & 0 & 0 & 0 \\ -k/\omega_{1j}\sin(\pi c_j)\sin(2\pi c_i) & 0 & k/\omega_{2j}\sin(2\pi c_i)\sin(2\pi c_j) & 0 \end{bmatrix}$$

$$B_{ii} = \begin{bmatrix} 0 \\ -\sin(\pi a_i) \\ 0 \\ -\sin(2\pi a_i) \end{bmatrix}, \qquad C_{ii} = [0 \quad \sin(\pi s_i) \quad 0 \quad \sin(2\pi s_i)]$$

LQG problems, one for the group and one for the spring modes. The solution to the complete problem is then obtained by changing B and C from the simplified diagonal form to their actual values, thus making the spring modes controllable by the first actuator and observable from the first sensor and making the group modes controllable by the second actuator and observable from the second sensor. Table 3 summarizes the steps of the integration.

The unconstrained LQG cost is $J = 11.0795$. Hence, the decentralized controller yields a relatively close performance with a simplified information structure. Note that the cost values both in the constrained and the unconstrained case are different in Ref. 17. A sequential design similar to Ref. 17 performed in Ref. 18 resulted in a solution similar to that produced by the homotopy, thus indicating a probable mismatched in Ref. 17 between the parameters used in the calculation and those reported.

B. Using Homotopy: Discontinuous Solution Path

The system formed by the two beams is naturally decoupled when the stiffness of the interconnecting spring is zero. The dynamic matrix A depends linearly on the spring constant k and can be written in the form:

$$A = A_0 + k\Delta A \tag{54}$$

where A_0 is block diagonal and represents the dynamics of the uncoupled beams. B, C, R, and V have the right block-diagonal form for $k = 0$ and need not be simplified. Hence, one can use the spring constant as the homotopy parameter ($\alpha = k/10$) and start from the decoupled beam system to find a solution to the coupled problem.

The initial compensators are the two fourth-order LQG solutions to the problem with $k = 0$. As the coupling becomes larger, the solution breaks down and a critical point occurs at $\alpha = 0.5724$ or $k = 5.724$. Increasing k slightly beyond the critical point, the minimization step (step 4) converges toward a local minimum not located on the path that was followed. Once a path of local minima has been found, the integration resumes until k reaches 10. The solution obtained is identical to the one previously found.

C. Using Homotopy: Multiple Solution Paths

The path of solutions initiating from the decoupled beam system $k = 0$ cannot be followed continuously passed $k = 5.724$. For values of $k \geq 5.724$, another path constituted of local minima was found, up to $k = 10$. In this section, we are interested in following this new path of solutions as k is decreased. Starting at $k = 10$ with a nondiagonal solution obtained in Sec. V.A, the path of solution is followed with

$$A(\alpha) = A_0 + 10(1 - \alpha)\Delta A$$

that is, $k = 10(1 - \alpha)$. The transformation of the parameters is the same as in Sec. V.B, except that it is followed now with decreasing k. The corresponding path of solutions is called the backward path, and the first one originating from the decoupled system with $k = 0$ is the forward path of solutions. As k keeps decreasing, the backward path encounters a critical point at $k = 1.795$. At that point, the minimization then converges to the solution that is on the forward path. For k between 1.795 and 5.724, one has two distinct continuous paths of solutions, each solution being a local minimum for the cost, as seen in Table 4. The path of solutions that started out considering a strong coupling between the beams breaks down as k becomes small and the corresponding solutions become saddle points for the cost. As reported previously, the forward path also breaks down, at $k = 5.724$, as the solutions become saddle points. For $k = 0$, the solution on the backward path will correspond to a saddle point for the cost and will not be diagonal. This solution could not be obtained numerically. Similarly, for $k = 10$, the solution on the forward path is likely to be a saddle point and could not be obtained numerically.

Table 3 Summary of a continuous homotopy procedure

α	$\lVert J_\xi \rVert$ Step 3	$\lVert J_\xi \rVert$ Step 4	J Step 3	J Step 4	Minimization step
0.1000	$9.1e-2$	$6.2e-5$	10.208	10.198	4
0.2875	$2.1e-1$	$5.9e-6$	10.693	10.637	7
0.3930	$2.2e-1$	$5.3e-5$	10.917	10.902	5
0.4867	$3.8e-1$	$6.6e-6$	11.153	11.111	6
0.5570	$3.9e-1$	$3.9e-7$	11.300	11.269	6
0.6098	$4.1e-1$	$7.5e-5$	11.408	11.376	6
0.7035	$4.0e-1$	$6.7e-5$	11.598	11.551	7
0.7973	$3.7e-1$	$1.9e-7$	11.725	11.707	6
0.9037	$4.0e-1$	$9.1e-8$	11.868	11.853	6
1.0000	$1.4e-1$	$1.8e-7$	11.953	11.945	4

Table 4 Optimal cost on two nonconnecting homotopic paths

k	J Forward path	J Backward path
0.000	7.8926	—
1.000	7.7483	—
1.800	8.0781	8.1654
2.000	8.1816	8.2587
3.000	8.7914	8.8221
4.000	9.4692	9.4779
5.000	10.0875	10.0787
5.700	10.4692	10.4459
5.724	10.4816	10.4581
6.000	—	10.5959
7.000	—	11.0844
8.000	—	11.5486
9.000	—	11.8875
10.000	—	11.9450

D. Discussion

The two-beam example has emphasized the fact that the convergence properties of the algorithm are only local. The solution to the constrained LQG problems is highly dependent on the actual parameters of the system and not on generic properties such as controllability and observability.

The homotopy algorithm developed here allows solution paths to be tracked as long as the nature of the solution remains the same. In order to provide a satisfactory solution, the following conditions must be met: 1) the starting diagonal problem and the final problem must have the same character; hence, the optimal compensator for the final problem is a continuous change of the optimal compensator for the simplified problem; and 2) the choice of the architecture is in accordance with the nature of the problem and does not allow for multiple minima. For strong coupling ($k \geq 5.724$), the problem admits only one minimum, and for weak coupling ($k \leq 1.795$), the problem also has apparently a single minimum. For intermediate values, however, multiple paths of minima exist, and the possibility to miss one path is greater. This could lead to missing a better solution, as illustrated in Table 4: for $k = 5$, the solution of the backward path is better than that of the forward path. The backward path, however, does not originate from the simplified diagonal solution chosen to start the forward path.

VI. Conclusions

This paper has developed a homotopy algorithm for solving fixed architecture LQG problems and has studied its convergence properties, yielding a better understanding of the solutions to the constrained LQG. Homotopy appears to be a practical tool since it is possible to define, for any architecture, a system for which the unconstrained LQG solution has the desired structure. By transforming the parameters of the initial system so that they become the parameters of the original one, the optimal solution can be tracked and the solution of a difficult coupled matrix equation problem can be found by solving a simple differential equation. It is found, however, that the nature of the solution may change along the solution path: a stabilizing compensator becoming a nonstabilizing compensator, a local minimum becoming a saddle point. The differential problem may not be defined everywhere over all possible changes in the system's parameter, and critical points and bifurcations can occur. Even if simplified problems can be defined, not all of their solutions can be found in a systematic manner, implying that some solution paths may be missed as well. Hence, homotopy is not a global tool for finding the global optimum to the fixed architecture LQG problem. Problems can have multiple minima and will be intrinsically difficult to solve. When the simplified systems as well as the architecture are carefully chosen, homotopy can be, however, a highly successful design tool for generating complex optimal feedback structures.

Further research is necessary on means to choose initial problems. A possible direction inspired by indirect compensator simplification methods[11-14] is 1) to find the diagonal problem closest, in some to-be-defined sense, to the original problem so that the distortion of the problem is minimal; or 2) to find the diagonal problem whose exact optimal solution is the approximate solution yielded by order reduction or compensator simplification techniques, so that the distortion is also minimal.

Further research is also necessary on means to select the control architecture since this has some importance on the convergence properties and is also a prime factor dictating the number of solutions.

Acknowledgment

This research was supported by NASA Langley Research Center under NASA Grant NAG-1-126.

References

[1]Kwakernaak, H., and Sivan, R., *Linear Optimal Control Systems*, Wiley, New York, 1972.

[2]Gupta, N. K., "Frequency-Shaped Cost Functionals: Extension of LQG Design Approach," *Journal of Guidance and Control*, Vol. 3, No. 6, 1980, pp. 529-535.

[3]Athans, M., "A Tutorial on the LQG /LTR Method," *Proceedings of the 1986 American Control Conference*, Seattle, WA, 1986.

[4]Johnson, T. L., and Athans, M., "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems," *IEEE Transactions on Automatic Control*, Vol. AC-15, 1970, pp. 658-660.

[5]Mendel, J. M., and Feather, J. J., "On the Design of Optimal Time Invariant Compensators for Linear Stochastic Time Invariant Systems," *IEEE Transactions on Automatic Control*, Vol. AC-20, No. 9, 1975, pp. 653-657.

[6]Ly, U. L., "A Design Algorithm for Robust Low Order Controller," Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, Stanford University, Stanford, CA, June 1982.

[7]Kabamba, P. T., "An Integrated Approach to Reduced Order Control Theory," *Optimal Control Applications and Methods*, Vol. 4, 1983, pp. 405-415.

[8]Hyland, D. S., and Bernstein, D. C., "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Transactions on Automatic Control*, Vol. AC-29, No. 11, 1984, pp. 1034-1037.

[9]Ly, U. L., Bryson, A. E., and Cannon, R. H., "Design of Low Order Compensators Using Parameter Optimization," *Automatica*, Vol. 21, No. 3, 1985, pp. 537-539.

[10]Moerder, D. D., and Calise, A. J., "Convergence of a Numerical Algorithm for Calculating Optimal Output Feedback Gains," *IEEE Transactions on Automatic Control*, Vol. AC-30, No. 9, 1985, pp. 900-903.

[11]Wilson, D. A., "Optimum Solution of Model Reduction Problems," *Proceedings of the Institution of Electrical Engineers*, Vol. 117, 1970, pp. 1161-1165.

[12]Moore, B. C., "Principal Component Analysis in Linear Systems: Controllability, Observability and System Reduction," *IEEE Transactions on Automatic Control*, Vol. AC-26, No. 1, 1981, pp. 17-32.

[13]Enns, D., "Model Reduction for Control System Design," Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, Stanford University, Stanford, CA, June 1984.

[14]Liu, Y., and Anderson, B. D. O., "Controller Reduction via Stable Factorization and Balancing," *International Journal of Control*, Vol. 44, No. 2, 1986, pp. 507-531.

[15]Hyland, D. C., and Richter, S., "On Direct Versus Indirect Methods for Reduced Order Controller Design," *IEEE Transactions on Automatic Control*, Vol. AC-35, No. 3, 1990, pp. 377-379.

[16]Wenk, C. J., and Knapp, C. H., "Parameter Optimization in Linear Systems with Arbitrarily Constrained Controller Structure," *IEEE Transactions on Automatic Control*, Vol. AC-25, No. 3, 1980, pp. 496-500.

[17]Bernstein, D. S., "Sequential Design of Decentralized Dynamic Compensators Using Optimal Projection Equations," *International Journal of Control*, Vol. 46, No. 5, 1987, pp. 1569-1577.

[18]Mercadal, M., "*H*2 Fixed Architecture Control Design for Large Scale Systems," Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology, Cambridge, MA, June 1990.

[19]Lefebvre, S., Richter, S., and DeCarlo, R., "A Continuation Algorithm for Eigenvalue Assignment by Decentralized Constant Output Feedback," *International Journal of Control*, Vol. 41, No. 5, 1985, pp. 1293-1299.

[20]Sebok, D. R., Richter, S., and DeCarlo, R., "Feedback Gain Optimization in Decentralized Eigenvalue Assignment," *Automatica*, Vol. 22, No. 4, 1986, pp. 433-447.

[21]Richter, S., "Reduced Order Control Design via the Optimal Projection Equations: a Homotopy Approach for Global Optimality," *Proceedings of the 1987 American Control Conference*, 1987, pp. 1527-1531.

[22]Richter, S., and Collins, E. G., "A Homotopy Algorithm for Reduced Order Compensator Design Using the Optimal Projection Equations," *Proceedings of the 28th Control and Decision Conference*, IEEE, Tampa, FL, 1989, pp. 506-511.

[23]Kabamba, P. T., Longman, R. W., and Jian-Guo, S., "A Homotopy Approach to the Feedback Stabilization of Linear Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 10, No. 5, 1987, pp. 422-432.

[24]Athans, M., "The Matrix Minimum Principle," *Information and Control*, Vol. 11, 1968, pp. 592-606.

[25]Scales, L. E., *Introduction to Nonlinear Optimization*, Springer-Verlag, New York, 1985.