Edmond A. Jonckheere (S'72–S'76–M'78–SM'83) was born in Belgium in 1950. He received the M.S. degree in electrical engineering from the University of Louvain, Belgium, in 1973, the Dr. Eng. degree in aerospace engineering from the University of Toulouse, France, in 1975, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1978.

From 1973 to 1975, he was a Research Fellow of the European Space Agency. From 1975 to 1978, he was a Research Assistant, and subsequently a Research Associate, in the Department of Electrical Engineering, University of Southern California. In 1979, he was with the Philips Research Laboratory, Brussels, Belgium. Since 1980, he has been with the Department of Electrical Engineering-Systems of the University of Southern California, where he is currently an Associate Professor. His research interests are in control, estimation, model reduction, and flexible structures.

J. William Helton (M'73), for a photograph and biography, see this issue, p. 1170.

The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton, and Moore

DAVID C. HYLAND AND DENNIS S. BERNSTEIN, MEMBER, IEEE

Abstract—**First-order necessary conditions for quadratically optimal reduced-order modeling of linear time-invariant systems are derived in the form of a pair of modified Lyapunov equations coupled by an oblique projection which determines the optimal reduced-order model. This form of the necessary conditions considerably simplifies previous results of Wilson [1] and clearly demonstrates the quadratic extremality and nonoptimality of the balancing method of Moore [2]. The possible existence of multiple solutions of the optimal projection equations is demonstrated and a relaxation-type algorithm is proposed for computing these local extrema. A component-cost analysis of the model-error criterion similar to the approach of Skelton [3] is utilized at each iteration to direct the algorithm to the global minimum.**

I. Introduction

The problem of approximating a high-order linear dynamical system with a relatively simpler system, i.e., the model-reduction problem, has received considerable attention in recent years. Among the myriad papers devoted to this problem are the notable contributions of Wilson [1], Moore [2], and Skelton [3] with which the present paper is concerned. In his 1970 paper, Wilson proposed an optimality-based approach to model reduction which involves minimizing the steady-state, quadratically weighted[1] output error when the original system and reduced-order model are subjected to white-noise inputs. For the resulting parameter-optimization problem, he obtained first-order necessary conditions which have the form of an aggregation (as, e.g., [4]) and which involve the solution of two Lyapunov equations each of order $n + n_m$, where n and n_m are the orders of the original and reduced-order models, respectively [5], [6].

Some time later, Moore proposed a quite different approach to model reduction based upon system-theoretic arguments as opposed to optimality criteria. Using the eigenvalues of the product of the controllability and observability gramians (which satisfy $n \times n$ Lyapunov equations), his method identifies subsystems which contribute little to the impulse response of the overall system. Such "weak" subsystems are thus eliminated to obtain a reduced-order model. This technique, known as balancing, has been vigorously developed in the recent literature [7]–[11]. Since this approach is completely independent of optimality considerations, there is, of course, no expectation that such reduced-order models are in any sense optimal.

A third approach to model reduction, proposed by Skelton [3], [12], also utilizes a quadratic optimality criterion as in [1]. However, rather than proceeding from necessary conditions as does Wilson, Skelton determines for a given basis the contribution (cost) of each state in a decomposition of the error criterion and truncates those with the least value. Although this approach is guided by optimality considerations, no rigorous guarantee of optimality is possible because of dependence on the choice of state space basis.

The present paper has five main objectives, the first of which is to show how the complex optimality conditions of Wilson can be transformed without loss of generality into much simpler and more tractable forms. The transformation is facilitated by exploiting the presence of an oblique (i.e., nonorthogonal) projection which was not recognized in [1][2] and which arises as a direct consequence of optimality. The resulting "optimal projection equations" constitute a coupled system of two $n \times n$

Manuscript received October 2, 1984; revised February 1, 1985. This paper is based on a prior submission of March 14, 1984. Paper recommended by Past Associate Editor, B. R. Barmish. This work was supported in part by the Air Force at Lincoln Laboratory/M.I.T., Lexington, MA.

The authors are with Harris Corporation, Government Aerospace Systems Division, Controls Analysis and Synthesis Group, Melbourne, FL 32901.

[1] The quadratic error criterion has been chosen for consideration in the present paper because of its relation to the standard engineering practice of stating specifications in terms of rms deviation.

[2] The projection was, however, pointed out in [28, p. 29].

0018-9286/85/1200-1201$01.00

modified Lyapunov equations [see (2.13), (2.14) or (2.21), (2.22)] whose solutions are given by a pair of rank-n_m controllability and observability *pseudogramians*. The highly structured form of these equations gives crucial insight into the set of local extrema satisfying the first-order necessary conditions.

The second objective of the paper is to show how the optimal projection equations provide a rigorous extremality context for Moore's balancing method and to clearly demonstrate its quadratic nonoptimality. Although for some problems the "weak subsystem" hypothesis leads to a nearly optimal reduced-order model, we construct examples for which the reduced-order model obtained from the balancing method is much worse with respect to the least-squares criterion than the quadratically optimal reduced-order model. In general, all that can be said is that the presence of a weak subsystem indicates that the reduced-order model obtained by truncation in the balanced basis may be in the proximity of an extremal of the quadratically optimal model-reduction problem; however, this extremal may very well be a global maximum. It should be noted that in a recent paper [13] Kabamba has used bounds on the model error to demonstrate the quadratic nonoptimality of the balancing method.

The third objective of the paper is to demonstrate via an example the mechanism responsible for the existence of multiple extrema of the optimal model-reduction problem. By characterizing the optimal projection as a sum of rank-1 eigenprojections of the product of the rank-deficient pseudogramians, it is immediately clear that the first-order necessary conditions of the problem are ambiguous in the sense that they fail to specify *which* n_m eigenprojections comprise the optimal projection corresponding to a solution (i.e., global minimum) of the optimal model-reduction problem. Specifically, since the pseudogramians can be rank deficient in $\binom{n}{n_m} = n!/n_m!(n - n_m)!$ ways, there may be precisely this many extremal projections corresponding to an identical number of local extrema.

The fourth objective of the paper is to propose a numerical algorithm for solving the optimal projection equations by exploiting their structure and taking advantage of the available insights. By expressing the modified Lyapunov equations in the form of "standard" Lyapunov equations, an iterative relaxation-type algorithm is developed. The crucial aspect of the proposed algorithm involves extracting an oblique projection at each step from the product of the solutions of the Lyapunov equations. Since $\binom{n}{n_m}$ rank-n_m projections can be extracted from the product of two $n \times n$ positive-definite matrices, it is quickly evident that the criterion by which the n_m eigenprojections are chosen determines which of the numerous local extrema will be reached. If, for example, the projection is chosen in accordance with the n_m largest eigenvalues of the product of the solutions of the Lyapunov equations, then it should not be surprising in view of the previous discussion that a global maximum may very well be reached. In this case, the first iteration of this algorithm involves Lyapunov equations whose solutions are the controllability and observability gramians and the eigenvalues in question are precisely the squares of the second-order modes [2, p. 24]. Thus, the first iteration coincides with the (nonoptimal) balancing approach of [2].

Since the optimal projection equations are a consequence of differential (local) properties, it should not be expected that they alone would possess the inherent ability to identify the global minimum. Moreover, because of the number of local extrema, second-order necessary conditions appear to be useless. Instead, we investigate an approach which chooses the eigenprojections according to a component-cost analysis of the model-error criterion. This technique can lead to a global minimum by effectively eliminating the local extrema which have considerably greater cost than the global minimum. This approach is philosophically identical to the component cost analysis of Skelton [3], [12]. Essentially, then, component cost analysis is utilized at each iteration to direct the algorithm to the global minimum. Although our application of this technique is admittedly heuristic, it should be noted that it is essentially proposed as a device for efficiently "sorting out" the local extrema which satisfy the otherwise mathematically rigorous necessary conditions. Hence, we propose component cost analysis as a crucial step in bridging the gap between local extremality and global optimality.

It should be pointed out that neither the numerical algorithm proposed in this paper nor the iterative algorithm developed in [4] and [5] has been proven to be convergent. The *principal contribution* of the present paper, however, is not a particular proposed algorithm but rather the *revelations concerning the structure of the first-order necessary conditions.* The proposed numerical algorithm should be considered but a prelude to a full investigation into numerical algorithms for the optimal projection equations. It should also be noted that the presence of the optimal projection was not exploited in developing the iterative algorithms in [4] and [5] (in fact, it did not even appear in [1]) and hence crucial insight into local extrema was lacking.

The fifth and last objective of the paper is to point out the connection between the optimal projection equations for model reduction obtained herein and the first-order necessary conditions obtained recently for two closely related problems, namely, reduced-order state estimation and fixed-order dynamic compensation.

The plan of the paper is as follows. Section II begins with general notation and definitions followed by the model-reduction problem statement and the main theorem which presents the optimal projection equations for model reduction. A series of remarks considers various aspects of the main theorem and sets the stage for discussing connections with [1] and [2]. Section III contains a detailed discussion of the sense in which the optimal projection equations simplify the necessary conditions given in [1], and Section IV shows how the approach of [2] is approximately extremal. Section V presents a simple example which clearly displays the possible existence of multiple extrema satisfying the optimal projection equations. This example shows that the balancing method of [2] may lead to a nonoptimal reduced-order model and suggests a heuristic procedure for selecting the eigenprojections comprising the projection corresponding to the global minimum, i.e., the *optimal* projection. In Section VI, a numerical algorithm for solving the optimal projection equations is proposed and applied to an illustrative example considered previously in [1] and [2] as well as to some interesting examples considered recently by Kabamba in [13]. Related results on reduced-order dynamic compensation and state estimation are briefly reviewed in Section VII and suggestions for further research are given in Section VIII. The proof of the main theorem appears in the Appendix.

II. Problem Statement and Main Result

The following notation and definitions will be used throughout the paper:

I_r	$r \times r$ identity matrix
Z^T	transpose of vector or matrix Z
Z^{-T}	$(Z^T)^{-1}$ or $(Z^{-1})^T$
$\rho(Z)$	rank of matrix Z
tr Z	trace of square matrix Z
$\|Z\|$	$[\text{tr } ZZ^T]^{1/2}$
Z_{ij}	(i, j)-element of matrix Z
diag $(\alpha_1, \cdots, \alpha_r)$	$r \times r$ diagonal matrix with listed diagonal elements
E_i	matrix with unity in the (i, i) position and zeros elsewhere
$\mathbb{E}$	expected value
$\mathbb{R}$, $\mathbb{R}^{r \times s}$	real numbers, $r \times s$ real matrices
stable matrix	matrix with eigenvalues in open left half plane
nonnegative-definite matrix	symmetric matrix with nonnegative eigenvalues
positive-definite matrix	symmetric matrix with positive eigenvalues

semisimple matrix	matrix similar to a diagonal matrix [14, p. 10]
nonnegative semisimple matrix	matrix similar to a nonnegative-definite matrix
positive-semisimple matrix	matrix similar to a positive-definite matrix
positive-diagonal matrix	diagonal matrix with positive diagonal elements
n, m, ℓ, n_m	positive integers, $1 \leq n_m \leq n$
x, u, y, x_m, y_m	n, m, ℓ, n_m, ℓ-dimensional vectors
A, B, C	$n \times n, n \times m, \ell \times n$ matrices
A_m, B_m, C_m	$n_m \times n_m, n_m \times m, \ell \times n_m$ matrices
R, V	$\ell \times \ell, m \times m$ positive-definite matrices

We consider the following problem.

Optimal Model-Reduction Problem: Given the controllable and observable system

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

$$y=Cx \tag{2.2}$$

find a reduced-order model

$$\dot{x}_m=A_mx_m+B_mu, \tag{2.3}$$

$$y_m=C_mx_m \tag{2.4}$$

which minimizes the quadratic model-reduction criterion[3]

$$J(A_m, B_m, C_m) \triangleq \lim_{t\to\infty} \mathbb{E}[(y-y_m)^TR(y-y_m)],$$

where the input $u(t)$ is white noise with positive-definite intensity V. To guarantee that J is finite, it is assumed that A is stable and we restrict our attention to the set of admissible reduced-order models

$$\mathcal{A} \triangleq \{(A_m, B_m, C_m): \quad A_m \text{ is stable}\}.$$

Since the value of J is independent of the internal realization of the transfer function corresponding to (2.3) and (2.4), we further restrict our attention to the set

$$\mathcal{A}_+ \triangleq \{(A_m, B_m, C_m)\in\mathcal{A}: \cdot (A_m, B_m) \text{ is controllable and } (A_m, C_m) \text{ is observable}\}.$$

The following lemma is needed for the statement of the main result.

Lemma 2.1: Suppose $\hat{Q}, \hat{P} \in \mathbb{R}^{n\times n}$ are nonnegative definite. Then $\hat{Q}\hat{P}$ is nonnegative semisimple. Furthermore, if $\rho(\hat{Q}\hat{P}) = n_m$ then there exist $G, \Gamma \in \mathbb{R}^{n_m\times n}$ and positive-semisimple $M \in \mathbb{R}^{n_m\times n_m}$ such that

$$\hat{Q}\hat{P}=G^TM\Gamma, \tag{2.5}$$

$$\Gamma G^T=I_{n_m}. \tag{2.6}$$

Proof: By [14, Theorem 6.2.5, p. 123], there exists $n \times n$ invertible $\hat{\Phi}$ such that the nonnegative-definite matrices $D_{\hat{Q}} \triangleq \hat{\Phi}\hat{Q}\hat{\Phi}^T$ and $D_{\hat{P}} \triangleq \hat{\Phi}^{-T}\hat{P}\hat{\Phi}^{-1}$ are both diagonal. Hence, $D_{\hat{Q}}D_{\hat{P}}$ is nonnegative definite and $\hat{Q}\hat{P} = \hat{\Phi}^{-1}D_{\hat{Q}}D_{\hat{P}}\hat{\Phi}$ is nonnegative semisimple. Next introduce $n \times n$ orthogonal U to effect a rearrangement of basis if necessary so that

$$\hat{Q}\hat{P}=\Phi\begin{bmatrix}\Lambda & 0\\ 0 & 0\end{bmatrix}\Phi^{-1},$$

where $\Phi \triangleq \hat{\Phi}U$ and $n_m \times n_m$ Λ is positive diagonal. Hence, for all $n_m \times n_m$ invertible S,

$$\hat{Q}\hat{P}=\Phi\begin{bmatrix}S\\ 0\end{bmatrix}(S^{-1}\Lambda S)[S^{-1} \quad 0]\Phi^{-1}$$

and thus, (2.5) and (2.6) hold with $G = [S^T \ 0]\Phi^T$, $M = S^{-1}\Lambda S$ and $\Gamma = [S^{-1} \ 0]\Phi^{-1}$. ■

For convenience in stating the main theorem, we shall refer to $G, \Gamma \in \mathbb{R}^{n_m\times n}$ and positive-semisimple $M \in \mathbb{R}^{n_m\times n_m}$ satisfying (2.5) and (2.6) as a *(G, M, Γ)-factorization of* $\hat{Q}\hat{P}$. Also, define the positive-definite controllability and observability gramians

$$W_c \triangleq \int_0^\infty e^{At}BVB^Te^{A^Tt}\, dt,$$

$$W_o \triangleq \int_0^\infty e^{A^Tt}C^TRCe^{At}\, dt,$$

which satisfy the dual Lyapunov equations

$$0=AW_c+W_cA^T+BVB^T, \tag{2.7}$$

$$0=A^TW_o+W_oA+C^TRC. \tag{2.8}$$

Main Theorem: Suppose $(A_m, B_m, C_m) \in \mathcal{A}_+$ solves the optimal model-reduction problem. Then there exist nonnegative-definite matrices $\hat{Q}, \hat{P} \in \mathbb{R}^{n\times n}$ such that, for some (G, M, Γ)-factorization of $\hat{Q}\hat{P}$, A_m, B_m, and C_m are given by

$$A_m=\Gamma AG^T, \tag{2.9}$$

$$B_m=\Gamma B, \tag{2.10}$$

$$C_m=CG^T, \tag{2.11}$$

and such that, with $\tau \triangleq G^T\Gamma$, the following conditions are satisfied:

$$\rho(\hat{Q})=\rho(\hat{P})=\rho(\hat{Q}\hat{P})=n_m, \tag{2.12}$$

$$0=\tau[A\hat{Q}+\hat{Q}A^T+BVB^T], \tag{2.13}$$

$$0=[A^T\hat{P}+\hat{P}A+C^TRC]\tau. \tag{2.14}$$

Several comments are in order. First, note that the main theorem consists of necessary conditions in the form of two modified Lyapunov equations (2.13) and (2.14) plus rank conditions (2.12) which must possess nonnegative-definite solutions $\hat{Q}$, $\hat{P}$ when an optimal reduced-order model exists. We shall call $\hat{Q}$ and $\hat{P}$ the *controllability* and *observability pseudogramians*, respectively, since they are analogous to W_c and W_o and yet have rank deficiency. The modified Lyapunov equations are coupled by the $n \times n$ matrix τ which is a projection (idempotent matrix) since

$$\tau^2=G^T\Gamma G^T\Gamma=G^TI_{n_m}\Gamma=\tau.$$

Note that, in general, τ is an *oblique* projection and not necessarily an orthogonal projection since it may not be symmetric. We shall refer to a projection τ corresponding to a solution (i.e., global minimum) of the optimal model-reduction problem as an "optimal projection." It should be stressed that the form of the optimal reduced-order model (2.7)–(2.9) is a direct consequence of optimality and not the result of an *a priori* assumption on the structure of the reduced-order model.

Since the optimal projection equations are first-order necessary conditions for optimality, they may possess multiple solutions corresponding to various local extrema such as local maxima, local minima, saddle points, etc. The following definition will prove useful.

Definition 2.1: Nonnegative-definite $\hat{Q}, \hat{P} \in \mathbb{R}^{n\times n}$ are *extremal* if (2.12)–(2.14) are satisfied. $(A_m, B_m, C_m) \in \mathcal{A}_+$ is

[3] J will occasionally be referred to as the "model-reduction error" or, simply, as the "cost."

extremal if there exist extremal $\hat{Q}$, $\hat{P}$ such that (A_m, B_m, C_m) is given by (2.9)–(2.11) for some (G, M, Γ)-factorization of $\hat{Q}\hat{P}$. The corresponding projection τ is an *extremal projection.*

Proposition 2.1: Suppose (A_m, B_m, C_m) is extremal. Then the model-reduction error is given by [4]

$$J(A_m, B_m, C_m) = 2\text{tr}\,[(\hat{Q}\hat{P} - W_c W_o)A]. \tag{2.15}$$

Proof: The proof is given at the end of Appendix A. ■

Remark 2.1: Noting the identities

$$-2\text{tr}\,[W_c W_o A] = \text{tr}\,[C^T R C W_c] = \text{tr}\,[BVB^T W_o], \tag{2.16}$$

which follow from (2.7) and (2.8), (2.15) can be written for extremal (A_m, B_m, C_m) as

$$J(A_m, B_m, C_m) = 2\text{tr}\,[\hat{Q}\hat{P}A] + \text{tr}\,[C^T R C W_c] = 2\text{tr}\,[\hat{Q}\hat{P}A] + \text{tr}\,[BVB^T W_o]. \tag{2.17}$$

For convenience in the following discussion, let $\hat{Q}$, $\hat{P}$, G, M, Γ, and τ correspond to some extremal (A_m, B_m, C_m). Now observe that if x_m is replaced by Sx_m, where S is an arbitrary nonsingular matrix, then an "equivalent" reduced-order model is obtained with (A_m, B_m, C_m) replaced by $(SA_mS^{-1}, SB_m, C_mS^{-1})$. Since $J(A_m, B_m, C_m) = J(SA_mS^{-1}, SB_m, C_mS^{-1})$, one would expect the main theorem to apply also to $(SA_mS^{-1}, SB_m, C_mS^{-1})$. Indeed, the following result shows that this transformation corresponds to the alternative factorization $\hat{Q}\hat{P} = (S^{-T}G)^T(SMS^{-1})(S\Gamma)$ and, moreover, that all (G, M, Γ)-factorizations of $\hat{Q}\hat{P}$ are related by an invertible transformation.

Proposition 2.2: If $S \in \mathbb{R}^{n_m \times n_m}$ is invertible, then $\bar{G} = S^{-T}G$, $\bar{\Gamma} = S\Gamma$ and $\bar{M} = SMS^{-1}$ satisfy

$$\hat{Q}\hat{P} = \bar{G}^T\bar{M}\bar{\Gamma}, \tag{2.5$'$}$$

$$\bar{\Gamma}\bar{G}^T = I_{n_m}. \tag{2.6$'$}$$

Conversely, if $\bar{G}, \bar{\Gamma} \in \mathbb{R}^{n_m \times n}$ and invertible $\bar{M} \in \mathbb{R}^{n_m \times n_m}$ satisfy (2.5)′ and (2.6)′, then there exists invertible $S \in \mathbb{R}^{n_m \times n_m}$ such that $\bar{G} = S^{-T}G$, $\bar{\Gamma} = S\Gamma$ and $\bar{M} = SMS^{-1}$.

Proof: The first part is immediate. The second part follows by taking $S \triangleq \bar{M}^{-1}\bar{\Gamma}G^TM$, noting $S^{-1} = M\Gamma\bar{G}^T\bar{M}^{-1}$ and using the identities $\bar{\Gamma}G^TM\Gamma\bar{G}^T = \bar{M}$ and $M\Gamma\bar{G}^T = \bar{\Gamma}G^T\bar{M}$. ■

The next result shows that there exists a similarity transformation which simultaneously diagonalizes $\hat{Q}\hat{P}$ and τ.

Proposition 2.3: There exists invertible $\Phi \in \mathbb{R}^{n \times n}$ such that

$$\hat{Q} = \Phi^{-1}\begin{bmatrix} \Lambda_{\hat{Q}} & 0 \\ 0 & 0 \end{bmatrix}\Phi^{-T}, \qquad \hat{P} = \Phi^T\begin{bmatrix} \Lambda_{\hat{P}} & 0 \\ 0 & 0 \end{bmatrix}\Phi, \tag{2.18}$$

$$\hat{Q}\hat{P} = \Phi^{-1}\begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}\Phi, \qquad \tau = \Phi^{-1}\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}\Phi, \tag{2.19a, b}$$

where $\Lambda_{\hat{Q}}, \Lambda_{\hat{P}} \in \mathbb{R}^{n_m \times n_m}$ are positive diagonal, $\Lambda \triangleq \Lambda_{\hat{Q}}\Lambda_{\hat{P}}$ and the diagonal elements of Λ are the eigenvalues of M. Consequently,

$$\hat{Q} = \tau\hat{Q}, \quad \hat{P} = \hat{P}\tau. \tag{2.20}$$

Proof: By [14, Theorem 6.2.5, p. 123], and by (2.12), there exists $n \times n$ invertible Φ such that (2.18) holds and thus (2.19a) also holds. Define

$$\bar{G} = [I_{n_m} \quad 0]\Phi^{-T}, \quad \bar{M} = \Lambda \quad \text{and} \quad \bar{\Gamma} = [I_{n_m} \quad 0]\Phi$$

so that (2.5)′ and (2.6)′ are satisfied. By the second part of Proposition 2.2 there exists invertible $S \in \mathbb{R}^{n_m \times n_m}$ such that $G = S^T\bar{G}$, $M = S^{-1}\bar{M}S$ and $\Gamma = S^{-1}\bar{\Gamma}$. Now (2.19b) follows from

$$\tau = G^T\Gamma = \bar{G}^T\bar{\Gamma} = \Phi^{-1}\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}\Phi. \qquad ■$$

It is useful to present an alternative form of the optimal model-reduction equations (2.13) and (2.14). For convenience, define the notation

$$\tau_\perp \triangleq I_n - \tau.$$

Proposition 2.4: Equations (2.13) and (2.14) are equivalent, respectively, to

$$0 = A\hat{Q} + \hat{Q}A^T + BVB^T - \tau_\perp BVB^T\tau_\perp^T, \tag{2.21}$$

$$0 = A^T\hat{P} + \hat{P}A + C^TRC - \tau_\perp^T C^TRC\tau_\perp. \tag{2.22}$$

Proof: By (2.20), (2.21) = (2.13) + (2.13)^T + (2.13)τ and (2.13) = τ(2.21). Similarly, (2.14) and (2.22) are equivalent. ■

Remark 2.2: Noting the identities

$$-2\ \text{tr}\,[\hat{Q}\hat{P}A] = \text{tr}\,[C^TRC\hat{Q}] = \text{tr}\,[BVB^T\hat{P}], \tag{2.23}$$

which follow from (2.20)–(2.22), (2.17) can be written for extremal (A_m, B_m, C_m) as

$$J(A_m, B_m, C_m) = \text{tr}\,[C^TRC(W_c - \hat{Q})] = \text{tr}\,[BVB^T(W_o - \hat{P})]. \tag{2.24}$$

To facilitate the discussion in the following sections, we consider the change of basis $\hat{x} \triangleq \Phi x$, where Φ is given by Proposition 2.3. Writing (2.1) and (2.2) as

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u, \tag{2.25}$$

$$y = \hat{C}\hat{x}, \tag{2.26}$$

where

$$\hat{A} \triangleq \Phi A\Phi^{-1}, \qquad \hat{B} \triangleq \Phi B, \qquad \hat{C} \triangleq C\Phi^{-1},$$

(2.9)–(2.11) become

$$A_m = \hat{\Gamma}\hat{A}\hat{G}^T, \tag{2.27}$$

$$B_m = \hat{\Gamma}\hat{B}, \tag{2.28}$$

$$C_m = \hat{C}\hat{G}^T, \tag{2.29}$$

where

$$\hat{\Gamma} \triangleq \Gamma\Phi^{-1}, \qquad \hat{G} \triangleq G\Phi^T$$

satisfy

$$\hat{G}^T\hat{\Gamma} = \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}, \qquad \hat{\Gamma}\hat{G}^T = I_{n_m}. \tag{2.30}$$

Note that (2.30) implies

$$\hat{\Gamma} = [S \quad 0], \qquad \hat{G} = [S^{-T} \quad 0], \tag{2.31}$$

for some $n_m \times n_m$ invertible S. Partitioning

$$\hat{x} = \begin{bmatrix} \hat{x}_m \\ \hat{x}_2 \end{bmatrix}, \qquad \hat{A} = \begin{bmatrix} \hat{A}_m & \hat{A}_{m2} \\ \hat{A}_{2m} & \hat{A}_{22} \end{bmatrix},$$

$$\hat{B} = \begin{bmatrix} \hat{B}_m \\ \hat{B}_2 \end{bmatrix}, \qquad \hat{C} = [\hat{C}_m \quad \hat{C}_2],$$

where $\hat{x}_m \in \mathbb{R}^{n_m}$ and $\hat{A}_m$, $\hat{B}_m$ and $\hat{C}_m$ are $n_m \times n_m$, $n_m \times m$ and

[4] The expressions (2.15)–(2.17) and (2.23)–(2.24) will be used in Sections V and VI.

$\ell \times n_m$, respectively, (2.27)–(2.29) and (2.31) yield

$$A_m = S\hat{A}_m S^{-1}, \qquad B_m = S\hat{B}_m, \qquad C_m = \hat{C}_m S^{-1}.$$

This shows that the optimal reduced-order model (modulo a state transformation) can be obtained by truncating the last $n - n_m$ states of the original system when it is expressed in the basis with respect to which $\hat{Q}$ and $\hat{P}$ have the diagonal forms

$$\begin{bmatrix} \Lambda_{\hat{Q}} & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} \Lambda_{\hat{P}} & 0 \\ 0 & 0 \end{bmatrix}.$$

Since the optimal projection τ has the simple form

$$\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix}$$

in this basis, we shall refer to (2.25) and (2.26) as an *optimal projection realization* of (2.1) and (2.2). Note that when (2.21) and (2.22) are expanded in an optimal projection basis (i.e., a basis corresponding to an optimal projection realization) they assume the form

$$0 = \hat{A}_m \Lambda_{\hat{Q}} + \Lambda_{\hat{Q}} \hat{A}_m^T + \hat{B}_m V \hat{B}_m^T, \tag{2.32}$$

$$0 = \hat{A}_{2m} \Lambda_{\hat{Q}} + \hat{B}_2 V \hat{B}_m^T, \tag{2.33}$$

$$0 = \hat{A}_m^T \Lambda_{\hat{P}} + \Lambda_{\hat{P}} \hat{A}_m + \hat{C}_m^T R \hat{C}_m, \tag{2.34}$$

$$0 = \Lambda_{\hat{P}} \hat{A}_{m2} + \hat{C}_m^T R \hat{C}_2. \tag{2.35}$$

If Φ in Proposition 2.23 is replaced by

$$\begin{bmatrix} (\Lambda_{\hat{Q}}^{-1} \Lambda_{\hat{P}})^{1/4} & 0 \\ 0 & I_{n-n_m} \end{bmatrix} \Phi,$$

which corresponds to a change of basis for the reduced-order model obtained by truncation, then $\Lambda_{\hat{Q}}$ and $\Lambda_{\hat{P}}$ are both replaced by $(\Lambda_{\hat{Q}}\Lambda_{\hat{P}})^{1/2}$ and hence this can be called a *balanced optimal projection basis*, utilizing the terminology of [2]. Thus, in a balanced optimal projection realization, $\Lambda_{\hat{Q}}$ and $\Lambda_{\hat{P}}$ appearing in (2.32)–(2.35) are equal.

The next result provides an interesting closed-form characterization of an extremal projection in terms of the Drazin generalized inverse of $\hat{Q}\hat{P}$. Since $(\hat{Q}\hat{P})^2 = G^T M^2 \Gamma$, and hence $\rho(\hat{Q}\hat{P})^2 = \rho(\hat{Q}\hat{P})$, the "index" of $\hat{Q}\hat{P}$ (see [15, p. 121]) is 1. In this case, the Drazin inverse is traditionally called the group inverse and is denoted by $(\hat{Q}\hat{P})^{\#}$ [15, p. 124]. Since, as is easily verified, $(\hat{Q}\hat{P})^{\#} = G^T M^{-1} \Gamma$, (2.6) leads to the following result.

Proposition 2.5: An extremal projection τ is given by

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#}. \tag{2.36}$$

An alternative representation for an extremal projection will prove useful for developing a numerical algorithm for solving (2.21) and (2.22). If $Q, P \in \mathbb{R}^{r \times r}$ are nonnegative definite then by Lemma 2.1 QP is nonnegative semisimple and thus there exists invertible $\Psi \in \mathbb{R}^{r \times r}$ such that

$$QP = \Psi^{-1} \Omega \Psi,$$

where $\Omega = \operatorname{diag}(\omega_1, \cdots, \omega_r)$ and $\omega_i \geq 0$ are the eigenvalues of QP. Now define the ith eigenprojection [16, p. 41]

$$\Pi_i[QP] \triangleq \Psi^{-1} E_i \Psi,$$

which is a rank-1 oblique projection. Note that QP has the decomposition

$$QP = \sum_{i=1}^{r} \omega_i \Pi_i [QP].$$

Proposition 2.6: An extremal projection τ is given by

$$\tau = \sum_{i=1}^{n_m} \Pi_i [\hat{Q}\hat{P}], \tag{2.37}$$

where the ith eigenprojection $\Pi_i[\hat{Q}\hat{P}[$ corresponds to the ith nonzero eigenvalue λ_i of $\hat{Q}\hat{P}$.

III. Relationship to Wilson's Form of the Necessary Conditions

The optimal model-reduction problem considered in the previous section is identical to the problem considered by Wilson in [1] with the minor exception that he sets $R = I_\ell$. In [1] G and Γ are denoted by θ_2^T and θ_1, (2.6) appears as (15), and (2.9)–(2.11) are given by (14a, b). Note that in [1], θ_1 and θ_2 depend upon the solutions of a pair of $(n + n_m) \times (n + n_m)$ Lyapunov equations [see (7), (9) of [1] or (A.2), (A.3) of the present paper] whose coefficients and nonhomogeneous terms depend in turn on A_m, B_m, and C_m [see (A.10)–(A.15)]. The advantage of the $n \times n$ optimal projection equations (2.21) and (2.22) over the form of the necessary conditions given in [1] [see (A.10)–(A.15)] is that the optimal projection equations are *independent of* A_m, B_m, and C_m. Hence, this permits the development of numerical algorithms which avoid the need to choose starting values for A_m, B_m, and C_m. To see this, note that although the unknowns A_m, B_m, and C_m appear explicity in (A.10)–(A.15), all data in the optimal projection equations (2.13) and (2.14) are known except for the solutions $\hat{Q}$ and $\hat{P}$. Moreover, the optimal projection τ, which was not recognized in [1], can be seen to play a fundamental role by coupling the modified Lyapunov equations (2.21) and (2.22) and determining (since $\tau = G^T\Gamma$) A_m, B_m, and C_m in (2.7)–(2.9).

IV. Relationship to Moore's Balancing Method

In contrast to Wilson's method for model reduction which is based on optimality principles, the approach due to Moore [2] relies on system-theoretic ideas. The main thrust of this approach "is to eliminate any weak subsystem which contributes little to the impulse response matrix" [2, p. 26]. The concept of a "weak subsystem" is defined by means of a dominance relation [2, p. 28] involving similarity invariants called second-order modes. Moore evaluates reduced-order models obtained in this way by computing the relative error in the impulse response given for MIMO systems by [2, p. 29]

$$\epsilon(A_m, B_m, C_m) \triangleq \left[\int_0^\infty \|H_e(t)\|^2 \, dt \Big/ \int_0^\infty \|H(t)\|^2 \, dt \right]^{1/2},$$

where $H_e(t) \triangleq H(t) - H_m(t)$, $H(t) \triangleq R^{1/2} C e^{At} B V^{1/2}$ and $H_m(t) \triangleq R^{1/2} C_m e^{A_m t} B_m V^{1/2}$. To discuss this approach in the context of the optimal model-reduction problem, we assume that $V = I_m$ and $R = I_\ell$.

Proposition 4.1: Suppose $(A_m, B_m, C_m) \in \mathfrak{A}$. Then

$$\begin{aligned} \epsilon(A_m, B_m, C_m) &= [-\tfrac{1}{2} J(A_m, B_m, C_m)/\operatorname{tr}(W_c W_0 A)]^{1/2} \\ &= [J(A_m, B_m, C_m)/\operatorname{tr}(C^T R C W_c)]^{1/2} \\ &= [J(A_m, B_m, C_m)/\operatorname{tr}(B V B^T W_0)]^{1/2}. \end{aligned} \tag{4.1}$$

Proof: The result follows from (A.1), (A.8), and (A.9) which hold without regard to either optimality or extremality. ■

Note that Proposition 4.1 shows that the relative error in the impulse response is minimized precisely when $J(A_m, B_m, C_m)$ is minimized. Actually, this result is to be expected since, as shown in [1], J can be obtained alternatively by taking $u(t)$ to be an impulse at $t = 0$.

To draw interesting comparisons with the results of [2], choose

$n \times n$ invertible Ψ such that $\Psi W_c \Psi^T$ and $\Psi^{-T} W_o \Psi^{-1}$ are both diagonal and hence

$$W_c W_o = \Psi^{-1} \Sigma^2 \Psi, \tag{4.2}$$

where $\Sigma \triangleq \text{diag}(\sigma_1, \cdots, \sigma_n)$ and the second-order modes σ_i (i.e., the positive square roots of the eigenvalues of $W_c W_o$) satisfy $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$. This transformation corresponds to replacing (2.1), (2.2) by

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}u, \tag{4.3}$$

$$y = \bar{C}\bar{x}, \tag{4.4}$$

where

$$\bar{x} \triangleq \Psi x, \quad \bar{A} \triangleq \Psi A \Psi^{-1}, \quad B \triangleq \Psi B, \quad \bar{C} \triangleq C\Psi^{-1}. \tag{4.5}$$

The transformed system (4.3), (4.4), called a principal axis realization [17], can further be chosen so that

$$\Psi W_c \Psi^T = \Psi^{-T} W_o \Psi^{-1} = \Sigma, \tag{4.6}$$

i.e., the balanced realization. Using (4.5), (2.7) and (2.8) become

$$0 = \bar{A}\Sigma + \Sigma \bar{A}^T + \bar{B} V \bar{B}^T, \tag{4.7}$$

$$0 = \bar{A}^T \Sigma + \Sigma \bar{A} + \bar{C}^T R \bar{C}. \tag{4.8}$$

The model-reduction procedure suggested in [2] involves partitioning

$$\bar{x} = \begin{bmatrix} \bar{x}_m \\ \bar{x}_2 \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} \bar{A}_m & \bar{A}_{m2} \\ \bar{A}_{2m} & \bar{A}_{22} \end{bmatrix},$$

$$\bar{B} = \begin{bmatrix} \bar{B}_m \\ \bar{B}_2 \end{bmatrix}, \quad \bar{C} = [\bar{C}_m \quad \bar{C}_2],$$

where $\bar{x}_m \in \mathbb{R}^{n_m}$ and $\bar{A}_m$, $\bar{B}_m$, and $\bar{C}_m$ have corresponding dimension, and extracting the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$. Hence, the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$ is extracted from (4.3), (4.4) in essentially the same way the optimal reduced-order model (A_m, B_m, C_m) is extracted from (2.25), (2.26). To see how the optimal-projection realization compares to a principal-axis realization, first note that (2.13) and (2.14) are satisfied by $\hat{Q} = W_c$ and $\hat{P} = W_o$ when the rank conditions (2.10) are ignored. Indeed, since W_c and W_o are positive definite, the rank conditions (2.12) do *not* hold. If, however, the system (2.1), (2.2) is expressed in the balanced coordinate system (4.3), (4.4) (so that $W_c = W_o = \Sigma$), then the assumption $\sigma_{n_m} \gg \sigma_{n_m - 1}$ implies that $\rho(W_c)$, $\rho(W_o)$ and $\rho(W_c W_o)$ are "approximately" equal to n_m and thus, in this sense, condition (2.10) *is* satisfied. This observation leads to the suggestion that when $\sigma_{n_m} \gg \sigma_{n_m+1}$, W_c and W_o are approximations to solutions $\hat{Q}$ and $\hat{P}$ of the optimal projection equations and the reduced-order model $(\bar{A}_m, \bar{B}_m, \bar{C}_m)$ of Moore is an approximation to some extremal (A_m, B_m, C_m). There is no guarantee, of course, that any particular extremum corresponds to the global minimum, or even to a local minimum.

V. Existence of Multiple Extrema and Component-Cost Ranking

In this section, we show by means of a simple example that the optimal projection equations may possess nonunique solutions corresponding to multiple extrema, e.g., local minima or maxima. We also show how decomposing the cost can identify the global minimum from among the numerous extrema. To begin, let $m = \ell = n$, $R = V = I_n$,

$$A \triangleq \text{diag}(-\alpha_1, \cdots, -\alpha_n),$$

where $\alpha_i > 0$, $i = 1, \cdots, n$, and suppose B and C are such that

$$BB^T = \text{diag}(\beta_1, \cdots, \beta_n), \qquad C^T C = \text{diag}(\gamma_1, \cdots, \gamma_n),$$

where $\beta_i > 0$, $\gamma_i > 0$, $i = 1, \cdots, n$. Hypothesizing diagonal solutions $\hat{Q}$ and $\hat{P}$ of (2.21) and (2.22) leads to

$$\hat{Q}_{ii} = \frac{\beta_i}{2\alpha_i}\delta_i, \qquad \hat{P}_{ii} = \frac{\gamma_i}{2\alpha_i}\delta_i,$$

where each δ_i, $i = 1, \cdots, n$ is either zero or one and exactly n_m of the δ_i's are equal to one. Hence $\tau = \text{diag}(\delta_1, \cdots, \delta_n)$. Note that there are $\binom{n}{n_m}$ such solutions of the optimal projection equations corresponding to $\binom{n}{n_m}$ local extrema.

Since

$$W_c = -\frac{1}{2}A^{-1}BB^T, \quad W_o = -\frac{1}{2}A^{-1}C^T C, \quad \hat{Q} = \tau W_c, \quad \hat{P} = \tau W_o$$

and A, W_c, and W_o commute, (2.15) becomes

$$J(A_m, B_m, C_m) = -\frac{1}{2}\text{ tr } \tau_\perp A^{-1} BB^T C^T C.$$

Hence,

$$J(A_m, B_m, C_m) = \sum_{i=1}^{n} \zeta_i (1 - \delta_i), \tag{5.1}$$

where

$$\zeta_i \triangleq \beta_i \gamma_i / 2\alpha_i.$$

To minimize J, it is clear that δ_i should be chosen to be unity for the largest n_m elements of the set $\{\zeta_i\}_{i=1}^n$ and zero otherwise. Although this choice is not necessarily unique, it does yield a global minimum. Note that choosing $\delta_i = 1$ is equivalent to selecting a particular eigenprojection $\Pi_i[W_c W_o]$ corresponding to the eigenvalue $\beta_i \gamma_i / 4\alpha_i^2$.

Remark 5.1: The expression in (5.1) can be regarded as a decomposition of the cost in terms of the state variables. The idea of deleting states based on their "component costs" is precisely the "component cost analysis" approach of Skelton [3], [12].

Using the example, it is easy to see that the balancing method of [2], which selects eigenprojections based upon the magnitude of the eigenvalues of $W_c W_o$, i.e., the (squares of the) second-order modes, may yield a grossly suboptimal reduced-order model. To this end, let

$$\alpha_1 = 1, \quad \alpha_2 = 10^6, \quad \beta_1 = 1, \quad \beta_2 = 10^6, \quad \gamma_1 = 1, \; \gamma_2 = 10^3$$

so that

$$\zeta_1 = 0.5, \qquad \zeta_2 = 500.$$

Clearly, J is minimized ($J = \zeta_1$) by choosing $\delta_1 = 0$, $\delta_2 = 1$, which corresponds to truncating the first state variable. If, however, the method of [2] is utilized, then judging by the second-order modes

$$\sigma_1 = 0.5, \qquad \sigma_2 = (2.5)^{1/2} \cdot 10^{-2} \approx 0.012,$$

the second state variable should be deleted. This, however, corresponds to choosing $\delta_1 = 1$, $\delta_2 = 0$ with the higher cost $J = \zeta_2$. The fact that the balancing approach of [2] fails to determine a solution of the optimal model-reduction problem should not be surprising in view of the fact that the error criterion plays no role in the balancing technique.

Although the above solution exploited the simple structure of this example, it is clear that choosing the global minimum from among the local extrema involves an eigenprojection decomposition of the cost J. To extend this idea to more general systems, we invoke the following heuristic approximation.

Approximation 5.1: Let Ψ define the balanced basis as in (4.6). Then Ψ also approximately defines a balanced optimal projection basis, i.e.,

$$\Psi\hat{Q}\Psi^T \approx \Psi^{-T}\hat{P}\Psi^{-1} \approx \bar{\tau}\Sigma^2, \tag{5.2}$$

where extremal

$$\bar{\tau} \triangleq \Psi\tau\Psi^{-1} = \text{diag}\ (\delta_1, \cdots, \delta_n) \tag{5.3}$$

and

$$\delta_i \in \{0,\ 1\}, \qquad \sum_{i=1}^{n} \delta_i = n_m.$$

Proposition 5.1: If Approximation 5.1 holds for extremal (A_m, B_m, C_m) then, with $\bar{\tau}_\perp \triangleq I_n - \bar{\tau}$,

$$J(A_m, B_m, C_m) \approx -2\text{tr}\ [\bar{\tau}_\perp \Sigma^2 \bar{A}] = 2\sum_{i=1}^{n} -\sigma_i^2 \bar{A}_{ii}(1-\delta_i). \tag{5.4}$$

Remark 5.2: From (4.7) and (4.8), it follows that (5.4) can be written either as

$$J(A_m, B_m, C_m) \approx \text{tr}\ [\bar{\tau}_\perp \Sigma \bar{B} V \bar{B}^T] = \sum_{i=1}^{n} \sigma_i (\bar{B} V \bar{B}^T)_{ii}(1-\delta_i) \tag{5.5}$$

or

$$J(A_m, B_m, C_m) \approx \text{tr}\ [\bar{\tau}_\perp \Sigma \bar{C}^T R \bar{C}] = \sum_{i=1}^{n} \sigma_i (\bar{C}^T R \bar{C})_{ii}(1-\delta_i). \tag{5.6}$$

Hence, Approximation 5.1 leads to the following component-cost ranking (again, in the sense of Skelton [3], [12]) of the $\binom{n}{n_m}$ extrema satisfying the optimal projection equations.

Component-Cost Ranking: Assume Approximation 5.1 is valid and choose the eigenprojections comprising extremal $\bar{\tau}$ such that

$\delta_i = 1$, if $-\sigma_i$ is among the n_m largest elements of the set $\{-\sigma_r^2 \bar{A}_{rr}\}_{r=1}^n$;

$\delta_i = 0$, otherwise.

For comparison purposes, we shall also consider the following ranking of the eigenprojections based upon the eigenvalues of $W_c W_o$ (i.e., second-order modes).

Eigenvalue Ranking: Choose the eigenprojections comprising extremal $\bar{\tau}$ such that

$\delta_i = 1$, if $-\sigma_i^2 \bar{A}_{ii}$ is among the n_m largest elements of the set $\{-\sigma_r\}_{r=1}^n$;

$\delta_i = 0$, otherwise.

Remark 5.3: The observation that the second-order modes alone may be a poor guide to determining an optimal reduced-order model has recently been made in [13] where bounds on the model-error criterion were given involving both the second-order modes and suitable weights called *balanced gains*. It can be seen from Proposition 5.1 that the role of balanced gains in our approach is played by the elements $-\sigma_i \bar{A}_{ii}$ when Approximation 5.1 holds. It can also be seen that the balanced gains of Kabamba yield bounds on the component costs of Skelton.

VI. Numerical Solution of the Optimal Projection Equations

Insofar as the ultimate aim of any model-reduction technique is to permit the development of numerical procedures for reducing high-order models, the optimal projection equations, comprising a coupled system of modified Lyapunov equations, appear promising in this regard. Therefore, we present an iterative computational algorithm that exploits the structure of these equations and the available insights. The reader is strongly reminded that the proposed algorithm is but a first attempt at solving these new equations and alternative algorithms may yet be devised. The basis of this algorithm is the ability to write the modified Lyapunov equations (2.21), (2.22) in the form of "standard" Lyapunov equations (6.1), (6.2) such that the pseudogramians $\hat{Q}$ and $\hat{P}$ are extracted at the final step (6.6). It follows from (2.32)–(2.35) that (2.21), (2.22) are indeed equivalent to (6.1), (6.2) (with $k = \infty$) and (6.6).

Algorithm:

Step 1) Initialize $\tau^{(0)} = I_n$.

Step 2) Solve for $\hat{\hat{Q}}^{(k)}$, $\hat{\hat{P}}^{(k)}$

$$0 = (A - \tau^{(k)} A \tau_\perp^{(k)})\hat{\hat{Q}}^{(k)} + \hat{\hat{Q}}^{(k)}(A - \tau^{(k)} A \tau_\perp^{(k)})^T + BVB^T, \tag{6.1}$$

$$0 = (A - \tau_\perp^{(k)} A \tau^{(k)})^T \hat{\hat{P}}^{(k)} + \hat{\hat{P}}^{(k)}(A - \tau_\perp^{(k)} A \tau^{(k)}) + C^T R C. \tag{6.2}$$

Step 3) Balance

$$\Phi^{(k)}\hat{\hat{Q}}^{(k)}(\Phi^{(k)})^T = (\Phi^{(k)})^{-T}\hat{\hat{P}}^{(k)}(\Phi^{(k)})^{-1} = \Sigma^{(k)}, \tag{6.3}$$

$$\Sigma^{(k)} = \text{diag}\ (\sigma_1^{(k)}, \cdots, \sigma_n^{(k)}), \qquad \sigma_1^{(k)} \geq \sigma_2^{(k)} \geq \cdots \geq \sigma_n^{(k)} \geq 0.$$

Step 4) If $k > 1$ check for convergence

$$e_k \triangleq \left[\frac{\text{tr}\ (C^T R C W_c) - \text{tr}\ (C^T R C \tau^{(k)} \hat{\hat{Q}}^{(k)} (\tau^{(k)})^T)}{\text{tr}\ (C^T R C W_c)}\right]^{1/2}. \tag{6.4}$$

If $|e_k - e_{k-1}| <$ tolerance then go to step 8); else continue;

Step 5) Select n_m eigenprojections

$$\Pi_{i_1}[\hat{\hat{Q}}^{(k)}\hat{\hat{P}}^{(k)}], \cdots, \Pi_{i_{n_m}}[\hat{\hat{Q}}^{(k)}\hat{\hat{P}}^{(k)}],$$

$$\Pi_i[\hat{\hat{Q}}^{(k)}\hat{\hat{P}}^{(k)}] \triangleq \Phi^{(k)} E_i (\Phi^{(k)})^{-1}.$$

Step 6) Update

$$\tau^{(k+1)} = \sum_{r=1}^{n_m} \Pi_{i_r}[\hat{\hat{Q}}^{(k)}\hat{\hat{P}}^{(k)}]. \tag{6.5}$$

Step 7) Check for convergence; if not, increment k and return to Step 2).

Step 8) Set

$$\hat{Q} = \tau^{(\infty)}\hat{\hat{Q}}(\tau^{(\infty)})^T, \qquad \hat{P} = (\tau^{(\infty)})^T \hat{\hat{P}} \tau^{(\infty)}. \tag{6.6}$$

For convenience, we shall adopt the notation $(A_m^{(k)}, B_m^{(k)}, C_m^{(k)})$, where $k > 0$, to denote the reduced-order model obtained as a result of applying the projection $\tau^{(k)}$, and we define (see Section IV)

$$\epsilon_k \triangleq \epsilon(A_m^{(k)}, B_m^{(k)}, C_m^{(k)}),$$

i.e., the relative error associated with $(A_m^{(k)}, B_m^{(k)}, C_m^{(k)})$. Note that, in general, $\epsilon_k \neq e_k$ since e_k denotes the relative error only for an extremum, i.e., when convergence has been reached.

It should be clear from the discussion in the previous section that the crucial step of the algorithm is Step 5)—the choice of the eigenprojections. For the examples which follow, we shall invoke consistently at Step 5) either the component-cost ranking based upon Approximation 5.1 or the eigenvalue ranking.

Remark 6.1: Note that in the special case $R = I_m$ and $V = I_\ell$, the first iteration of the algorithm yields $\hat{Q}^{(0)} = W_c$, $\hat{P}^{(0)} = W_o$. If, at Step 5), we choose $i_r = r$, $r = 1, \cdots, n_m$, i.e., the eigenprojections are selected according to the eigenvalue ranking, then $(A_m^{(1)}, B_m^{(1)}, C_m^{(1)})$ is precisely the reduced-order model obtained from balancing.

We shall first consider the following example which was treated by both Wilson and Moore. In this example, and those that follow, assume $R = I_m$, $V = I_\ell$.

Example 6.1:

$$A=\begin{bmatrix} 0 & 0 & 0 & -150 \\ 1 & 0 & 0 & -245 \\ 0 & 1 & 0 & -113 \\ 0 & 0 & 1 & -19 \end{bmatrix}, \quad B=\begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad C=[0\ 0\ 0\ 1].$$

Table I summarizes the results obtained for the three cases $n_m = 3, 2, 1$ utilizing the eigenvalue ranking. In each case, the proposed algorithm converged linearly in less than eight iterations and, in each case, improvement is evident over previously published results. As pointed out in [2], Wilson's result seems to imply a lack of final convergence. For this example, the balancing approach yields a reduced-order model close to the global minimum.

We now turn to a pair of interesting examples considered in [13].

Example 6.2:

$$A=\begin{bmatrix} -0.005 & -0.99 \\ -0.99 & -5000 \end{bmatrix}, \quad B=\begin{bmatrix} 1 \\ 100 \end{bmatrix}, \quad C=B^T.$$

Table II summarizes the results obtained using the eigenvalue ranking and Table III gives the results when the component-cost ranking is used. It is clear that the former method directs the algorithm to the global maximum whereas the latter approach yields the global minimum.

Example 6.3:

$$A=\begin{bmatrix} -0.25 & -0.4 \\ -0.4 & -0.72 \end{bmatrix}, \quad B=\begin{bmatrix} 1 \\ 1.2 \end{bmatrix}, \quad C=B^T.$$

Table IV reports the results obtained using either the component-cost ranking or the eigenvalue ranking which agree for this example. If the *alternative* eigenprojection is selected then, as expected, the algorithm converges to a global maximum (see Table V). The interesting aspect of this example, as discussed in [13], is that the error $\epsilon_1 = 0.5245$ (see [13]) for the reduced-order model obtained by either eigenprojection ranking is actually *greater* than $\epsilon_1 = 0.3849$ obtained by choosing the alternative reduced-order model. This situation seems to indicate that proper eigenprojection selection based upon a cost decomposition is able to direct the algorithm to the global minimum in cases for which the starting values are not nearby.

VII. The Optimal Projection Equations for Fixed-Order Dynamic Compensation and Reduced-Order State Estimation

We briefly discuss the relationship between the optimal projection equations for model reduction and analogous results for reduced-order control and estimation problems.

Fixed-Order Dynamic-Compensation Problem: Given the controlled system

$$\dot{x}=Ax+Bu+w_1, \tag{7.1}$$

$$y=Cx+w_2, \tag{7.2}$$

TABLE I
RELATIVE ERROR $e_\infty = \epsilon_\infty$

Order n_m	Wilson [1]	Moore [2]	Optimal Projection Equations
3	—	0.001311	0.001306
2	0.04097	0.03938	0.03929
1	—	0.4321	0.4268

TABLE II
EXAMPLE 6.2 WITH EIGENVALUE RANKING

k	e_k
1	0.9950371897
2	0.9950371691
3	0.9950371690

TABLE III
EXAMPLE 6.2 WITH COMPONENT-COST RANKING

k	e_k
1	0.0995037
2	0.0995449
3	0.0995924
4	0.0996520
5	0.0997346
6	0.0998648
7	0.1001125
8	0.1007724
9	0.1054569
10	0.0982006
11	0.0975409
12	0.0975342
13	0.0975330
14	0.0975329

TABLE IV
EXAMPLE 6.3 USING EITHER RANKING

k	e_k
1	0.646996
2	0.418341
3	0.220994
4	0.177276
5	0.176576

TABLE V
EXAMPLE 6.3 WITH THE OPPOSITE RANKING

k	e_k
1	0.7624928516
2	0.9999999961
3	0.9999999975
.	.
.	.
.	.
29	0.9999999999

design a fixed-order dynamic compensator

$$\dot{x}_c = A_c x_c + B_c y, \tag{7.3}$$

$$u = C_c x_c \tag{7.4}$$

which minimizes the performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t\to\infty} \mathbb{E}[x^T R_1 x + u^T R_2 u], \tag{7.5}$$

where $u \in \mathbb{R}^m$, $x_c \in \mathbb{R}^{n_c}$, $n_c \leq n$, w_1 is white disturbance noise, w_2 is nonsingular white observation noise, R_1 is nonnegative definite, and R_2 is positive definite.

Necessary conditions characterizing optimal (A_c, B_c, C_c) have been developed in [18]–[22] along the same lines as the main theorem. These conditions, called the optimal projection equations for fixed-order dynamic compensation, consist of four matrix equations (two modified Riccati equations and two modified Lyapunov equations) coupled by a projection. The modified Riccati equations, not surprisingly, are similar in form to the covariance and cost Riccati equations of LQG theory and the modified Lyapunov equations are similar to the optimal model-reduction equations (2.13) and (2.14). Hence, while the modified Riccati equations govern optimal estimation and optimal control, the additional modified Lyapunov equations characterize "optimal reduction." The important fact that all four equations are coupled supports the view that optimal fixed-order dynamic compensators *cannot*, in general, be designed by means of a stepwise procedure, e.g., by either open-loop model reduction followed by LQG or LQG followed by closed-loop model reduction.

Midway between the model-reduction and fixed-order dynamic-compensation problems lies the following problem.

Reduced-Order State-Estimation Problem: Given the observed system

$$\dot{x} = Ax + w_1, \tag{7.6}$$

$$y = Cx + w_2, \tag{7.7}$$

design a reduced-order state estimator

$$\dot{x}_e = A_e x_e + B_e y, \tag{7.8}$$

$$y_e = C_e x_e, \tag{7.9}$$

which minimizes the estimation criterion

$$J(A_e, B_e, C_e) \triangleq \lim_{t\to\infty} \mathbb{E}[(Lx - y_e)^T R (Lx - y_e)],$$

where $x_e \in \mathbb{R}^{n_e}$, $L \in \mathbb{R}^{p \times n_e}$ and L identifies the states, or linear combinations of states, whose estimates are desired. The order n_e of the estimator state x_e is determined by implementation constraints, i.e., by the computing capability available for realizing (7.8) and (7.9) in real time.

In view of the results already given, it should not be surprising (see [23]) that the optimal projection equations for reduced-order state estimation form a system of *three* matrix equations (a pair of modified Lyapunov equations along with a single modified Riccati equation) coupled by a projection which determines the gains of the optimal reduced-order estimator. This intrinsic coupling between the "operations" of optimal estimation (the modified Riccati equation) and optimal model reduction (the pair of modified Lyapunov equations) stresses the fact that reduced-order estimators designed by means of either model reduction followed by "full-order" state estimation or full-order estimation followed by estimator reduction will generally not be optimal for the given order.

VIII. Directions for Further Research

The most important area of research involves the further development of algorithms for solving the optimal projection equations. Although proving local convergence of the proposed algorithm appears possible, the more important problem is achieving global optimality via the component cost approach. Although the global minimum was attained for all examples attempted by the authors, it remains to treat considerably more complex systems.

An interesting extension of the main theorem involves the case in which the original system (2.1), (2.2) is a distributed parameter system, e.g., a partial differential equation or a functional differential equation. This generalization, which has been referred to as the "ultimate reduced-order problem" [24], may lead to the efficient generation of high-order discretizations for such systems. All of the mathematical machinery required to generalize the main theorem to this case has already been applied to fixed-order dynamic compensation in [25].

IX. Conclusion

First-order necessary conditions for quadratically optimal reduced-order modeling of a linear time-invariant plant are expressed in the form of a pair of $n \times n$ modified Lyapunov equations coupled by an oblique projection. This form of the necessary conditions considerably simplifies the original form given by Wilson in [1] and clearly reveals the possible presence of numerous extrema. The balancing method of Moore given in [2] is shown to yield a reduced-order model that is "close" to an extremal given by the necessary conditions. A numerical example shows, however, that this extremal may very well be the global maximum rather than the desired global minimum. An algorithm is proposed which exploits the presence of the optimal projection and computes the various local extrema by the choice of eigenprojections comprising the projection. A component-cost ranking of the eigenprojections, which is very much in the spirit of Skelton's method in [3] and [12], is used to direct the algorithm to the global optimum.

It should be pointed out that Moore's balancing appears to have strong ties with the L_∞ reduction problem via the Hankel norm [29]. Alternative settings for the Hankel operator, however, seem to indicate connections to the quadratic problem [30]. Finally, the robustness problem for reduced-order modeling, estimation, and control in a quadratic setting is discussed in [31].

Appendix

Proof of the Main Theorem

Introducing the augmented system

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u,$$

$$\tilde{y} = \tilde{C}\tilde{x},$$

where

$$\tilde{x} \triangleq \begin{bmatrix} x \\ x_m \end{bmatrix}, \quad \tilde{y} \triangleq y - y_m,$$

$$\tilde{A} \triangleq \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, \quad \tilde{B} \triangleq \begin{bmatrix} B \\ B_m \end{bmatrix}, \quad \tilde{C} \triangleq [C \quad -C_m],$$

leads to the expression

$$J(A_m, B_m, C_m) = \text{tr}\ \tilde{Q}\tilde{R}, \tag{A.1}$$

where $\tilde{R} \triangleq \tilde{C}^T R \tilde{C}$ and the nonnegative-definite steady-state covariance $\tilde{Q}$ of $\tilde{x}$ is given by the (unique) solution of

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}, \tag{A.2}$$

with $\tilde{V} \triangleq \tilde{B} V \tilde{B}^T$. To minimize (A.1) subject to the constraint (A.2), form the Lagrangian

$$L(A_m, B_m, C_m, \tilde{Q}) \triangleq \text{tr}\ [\lambda \tilde{Q}\tilde{R} + (\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V})\tilde{P}]$$

with multipliers $\lambda \geq 0$ and $\tilde{P} \in \mathbb{R}^{(n+n_m)\times(n+n_m)}$. Since $\mathfrak{A}_-$ is an open set, the standard Lagrange multiplier rule can be applied.

Using formulas for computing partial derivatives [26], it follows that

$$0=L_{\tilde{Q}}=\tilde{A}^T\tilde{P}+\tilde{P}\tilde{A}+\lambda\tilde{R}.$$

Since $\lambda = 0$ implies $\tilde{P} = 0$ (recall $\tilde{A}$ is stable), we can take $\lambda = 1$ without loss of generality. Hence, $\tilde{P}$ is the (unique nonnegative-definite) solution of

$$0=\tilde{A}^T\tilde{P}+\tilde{P}\tilde{A}+\tilde{R}. \tag{A.3}$$

Again using formulas from [26] and performing some manipulation, it follows that

$$0=L_{A_m}=Q_{12}^TP_{12}+Q_2P_2, \tag{A.4}$$

$$0=L_{B_m}=2(P_{12}^TB+P_2B_m)V, \tag{A.5}$$

$$0=L_{C_m}=2R(C_mQ_2-CQ_{12}), \tag{A.6}$$

where $\tilde{Q}$ and $\tilde{P}$ have been partitioned as

$$\tilde{Q}=\begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, \qquad \tilde{P}=\begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}. \tag{A.7}$$

Since (as will be seen shortly) Q_2 and P_2 are positive definite, define

$$G \triangleq Q_2^{-1}Q_{12}^T, \qquad \Gamma \triangleq -P_2^{-1}P_{12}^T, \tag{A.8}$$

so that (A.4)–(A.6) become (2.6), (2.10) and (2.11), respectively.

Next, define the nonnegative-definite matrices

$$\hat{Q} \triangleq Q_{12}Q_2^{-1}Q_{12}^T, \qquad \hat{P} \triangleq P_{12}P^{-1}P_{12}^T \tag{A.9}$$

and note that (A.4) implies that (2.5) holds with $M \triangleq Q_2P_2$. Since $Q_2P_2 = P_2^{-1/2}(P_2^{1/2}Q_2P_2^{1/2})P_2^{1/2}$, M is positive semisimple. The rank conditions (2.12) follow from Sylvester's inequality. Expanding (A.2) and (A.3) yields

$$0=AQ_1+Q_1A^T+BVB^T, \tag{A.10}$$

$$0=AQ_{12}+Q_{12}A_m^T+BVB_m^T, \tag{A.11}$$

$$0=A_mQ_2+Q_2A_m^T+B_mVB_m^T, \tag{A.12}$$

$$0=A^TP_1+P_1A+C^TRC, \tag{A.13}$$

$$0=A^TP_{12}+P_{12}A_m-C^TRC_m, \tag{A.14}$$

$$0=A_m^TP_2+P_2A_m+C_m^TRC_m. \tag{A.15}$$

Since A_m is stable and (A_m, B_m) is controllable, standard results (e.g., [27, p. 277]) imply that Q_2 is positive definite. Similarly, P_2 is positive definite.

It is easy to see at this point that A_m, B_m, and C_m are independent of Q_1 and P_1 and thus (A.10) and (A.13) can be ignored. Now, substituting (2.10), (2.11) and the identities

$$Q_{12}=\hat{Q}\Gamma^T, \qquad P_{12}=-\hat{P}G^T, \tag{A.16}$$

$$Q_2=\Gamma\hat{Q}\Gamma^T, \qquad P_2=G\hat{P}G^T, \tag{A.17}$$

into (A.11), (A.12), (A.14), and (A.15) yields

$$0=A\hat{Q}\Gamma^T+\hat{Q}\Gamma^TA_m^T+BVB^T\Gamma^T, \tag{A.18}$$

$$0=A_m\Gamma\hat{Q}\Gamma^T+\Gamma\hat{Q}\Gamma^TA_m^T+\Gamma BVB^T\Gamma^T, \tag{A.19}$$

$$0=A^T\hat{P}G^T+\hat{P}G^TA_m+C^TRCG^T, \tag{A.20}$$

$$0=A_m^TG\hat{P}G^T+G\hat{P}G^TA_m+GC^TRCG^T. \tag{A.21}$$

Computing (A.19)–Γ(A.18) implies

$$A_m=\Gamma A\hat{Q}\Gamma^T(\Gamma\hat{Q}\Gamma^T)^{-1}$$

which, since $\Gamma\hat{Q}\Gamma^T = Q_2$, yields (2.9). Alternatively, (2.9) can be obtained from (A.21)–G(A.20).

If we now substitute (2.9) into (A.18)–(A.21) and use the easily verified relations (2.20), it follows that (A.19) = Γ(A.18) and (A.22) = G(A.21), and thus (A.19) and (A.21), are redundant. Finally, G^T(A.18)^T and (A.20)Γ yield (2.13) and (2.14), respectively. Note that these last multiplications entail no loss of generality since $\rho(G) = \rho(\Gamma) = n_m$.

To show that the optimal projection equations entail no loss of generality over (A.2)–(A.6), let $\hat{Q}$, $\hat{P}$ be extremal and define Q_{12}, Q_2, P_{12}, P_2 by (A.16) and (A.17) for some (G, M, Γ)-factorization of $\hat{Q}\hat{P}$, and let Q_1, P_1 satisfy (A.10) and (A.13). Then it is straightforward to reverse the steps taken in the proof to arrive at (A.2)–(A.6). ■

Proof of Proposition 2.1: Extremal $\hat{Q}$, $\hat{P}$ leads to $\tilde{Q}$, $\tilde{P}$ as in (A.7) satisfying (A.2)–(A.6). Computing

$$J(A_m, B_m, C_m)=\text{tr}\ (Q_1C^TRC-2Q_{12}C_m^TRC)+\text{tr}\ (Q_2C_m^TRC_m)$$
$$=\text{tr}\ [C^TRC(W_c-\hat{Q})],$$

noting that (2.13), (2.14) are equivalent to (2.21), (2.22) because of (2.20) and using (2.23), leads to (2.15). ■

References

[1] D. A. Wilson, "Optimum solution of model-reduction problem," *Proc. IEE*, vol. 117, pp. 1161-1165, 1970.

[2] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 17-32, 1981.

[3] R. E. Skelton, "Cost Decomposition of linear systems with application to model reduction," *Int. J. Contr.*, vol. 32, pp. 1031-1055, 1980.

[4] M. Aoki, "Control of large-scale dynamic systems by aggregation," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 246-253, 1968.

[5] D. A. Wilson, "Model reduction for multivariable systems," *Int. J. Contr.*, vol. 20, pp. 57-64, 1974.

[6] R. N. Mishra and D. A. Wilson, "A new algorithm for optimal reduction of multivariable systems," *Int. J. Contr.*, vol. 31, pp. 443-466, 1980.

[7] L. Pernebo and L. M. Silverman, "Model reduction via balanced state space representations," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 382-387, 1982.

[8] K. V. Fernando and H. Nicholson, "On the structure of balanced and other principal representations of SISO systems," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 228-231, 1983.

[9] S. Shokoohi, L. M. Silverman, and P. M. Van Dooren, "Linear time-variable systems: Balancing and model reduction," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 810-822, 1983.

[10] E. I. Verriest and T. Kailath, "On generalized balanced realizations," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 833-844, 1983.

[11] E. A. Jonckheere and L. M. Silverman, "A new set of invariants for linear systems—Application to reduced-order compensator design," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 953-964, 1983.

[12] R. E. Skelton and A. Yousuff, "Component cost analysis of large scale systems," *Int. J. Contr.*, vol. 37, pp. 285-304, 1983.

[13] P. T. Kabamba, "Balanced gains and their significance for balanced model reduction," in *Proc. Conf. Inform. Sci. Syst.*, Princeton Univ., Princeton, NJ, 1984; also in *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 690-693, 1985.

[14] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.

[15] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. London, England: Pitman, 1979.

[16] T. Kato, *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 1966.

[17] C. T. Mullis and R. A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circ. Syst.*, vol. CAS-23, pp. 551-562, 1976.

[18] D. C. Hyland, "Optimality conditions for fixed-order dynamic compensation of flexible spacecraft with uncertain parameters," AIAA 20th Aerospace Sci. Meet., Orlando, FL, paper 82-0312, Jan. 1982.

[19] ——, "The optimal projection approach to fixed-order compensation: Numerical methods and illustrative results," AIAA 21st Aerospace Sci. Meet., Reno, NV, paper 83-0303, Jan. 1983.

[20] D. C. Hyland and D. S. Bernstein, "Explicit optimality conditions for fixed-order dynamic compensation," in *Proc. IEEE Conf. Dec. Contr.*, San Antonio, TX, Dec. 1983, pp. 161-165.

[21] ——, "The optimal projection equations for fixed-order dynamic

compensation," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1034–1037, 1984.

[22] D. C. Hyland, "Comparison of various controller-reduction methods: Suboptimal versus optimal projection," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 381-389.

[23] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 583–585, 1985.

[24] C. S. Sims, "Reduced-order modeling and filtering," *Contr. Dynamic Syst.*, vol. 18, pp. 55–103, 1982.

[25] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for fixed-order dynamic compensation of infinite-dimensional systems," in *Proc. AIAA Dynamics Specialists Conf.*, Palm Springs, CA, May 1984, pp. 396–400.

[26] M. Athans, "The matrix minimum principle," *Inform. Contr.*, vol. 11, pp. 592-606, 1968.

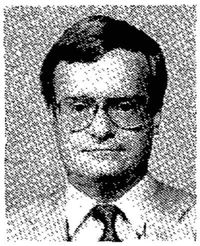
[27] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag, 1974.

[28] D. A. Wilson, "Reduction of the order of linear systems," Ph.D. dissertation, Univ. Cambridge, Cambridge, England, 1970.

[29] K. Glover, "All optimal Hankel-Norm approximations of linear multivariable systems and their L^∞—Error bounds," *Int. J. Contr.*, vol. 39, pp. 1115–1193, 1984.

[30] D. A. Wilson, "The Hankel operator and its induced norms," *Int. J. Contr.*, to be published.

[31] D. S. Bernstein and D. C. Hyland, "Optimal projection/maximum entropy stochastic modelling and reduced-order design synthesis," in *Proc. IFAC Workshop Model Errors, Compensation*, Boston, MA, June 1985.



David C. Hyland received the B.S., M.S., and Sc.D. degrees in aeronautics from the Massachusetts Institute of Technology (M.I.T.), Cambridge, in 1969, 1971, 1973, respectively.

After serving as a vibration specialist in a Cambridge-based Acoustics consulting firm, he joined the M.I.T. Lincoln Laboratory Staff in 1974. His work at Lincoln Laboratory included reentry vehicle dynamics, multibody spacecraft dynamics simulation, and spacecraft attitude control. He joined Harris Corporation, Government Aerospace Systems Division in 1983, where he leads the control systems analysis and synthesis group. His current research interests include stochastic modeling of parametric uncertainty in mechanical systems and optimization theory for fixed-order dynamic compensation.

Dennis S. Bernstein (M'82), for a photograph and biography, see p. 376 of the April 1985 issue of this TRANSACTIONS.