JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 17, No. 3, May–June 1994

Identification of Model Parameters and Associated Uncertainties for Robust Control Design

Valeri I. Karlov,* David W. Miller,† Wallace E. Vander Velde,‡ and Edward F. Crawley§
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The integration of system identification and robust control is considered. The identification algorithm is an extended Kalman filter, and the robust control algorithm is based on Petersen-Hollot's bounds (modified for random correlated parameters). The identification and control problems are coupled because the inputs for the identification experiment are selected to optimize the robust control performance. The optimization problem, interpreted as a form of Riccati equation control, is solved by exploiting the analytical properties of the Riccati equation in a nontraditional manner. The result appears an equivalent quadratic-linear boundary-value-problem, which allows a convergent numerical solution. An effective numerical algorithm is also offered for solving the extended Kalman filter equations in high-dimensional modal test problems. The algorithm is based on block decomposition of the modal state-space model. The developed approach is applied to the Middeck Active Control Experiment (MACE) testbed. MACE is an MIT STS flight experiment scheduled for launch in 1994.

I. Introduction

A. Problem Description

SYSTEM identification is an important part of any control design procedure for flexible, precision spacecraft. The real purpose of identification is to provide a control design model that improves the robust performance of a closed-loop system. This ideal can be formalized in the following scheme (Fig. 1). The essential element of this scheme is an ability to evaluate the residual uncertainties in the parameters and determine their cost in terms of robust control performance.

In present identification theory and practice (see Refs. 1–5) considerable effort is directed toward the estimation of uncertain parameters rather than evaluation of their accuracy characteristics (bounds). This is true, to a varying degree, of identification techniques such as the Empirical Transfer Function Estimation (ETFE), nonlinear Least-Squares and Maximum Likelihood Methods (LSM, MLM), Prediction Error Methods (PEM),[1] Eigensystem Realization Algorithm (ERA),[2,3] Q-Markov cover,[4] and others. Generally, these techniques allow the construction of models from measurement data without the assistance of a framework, such as the Finite Element Method (FEM), describing the model structure and its parameters. As a rule, these methods involve nonlinear operations, which eventually impede analytical error analysis. At the same time estimating identification accuracy is still possible,[1,5] especially for LSM and MLM techniques. However, in most cases it can be done only by sacrificing some or all of the dynamic nature of the problem. For example, nonlinear transformations of parameters can produce linear input-output models (ARMAX models), but in this case the original sensor and actuator noises become combined in one common error. This combination causes undesirable correlations and provides the most significant source of biased estimates.

An important question in identification is whether "standard" inputs, for example white noise, sufficiently excite the important dynamics. The answer to this question is not obvious because different modes have different costs, which, in turn, depend on the input itself. In other words, each mode requires its share of excitation. If it does not receive sufficient excitation in the limited time of the identification experiment (limited because of operational and computational considerations), it will not be estimated well, and in the case of "black-box" identification procedures, it may not appear in the model at all. It would be advantageous if the identification procedure, in conjunction with the analytical model (FEM), could signal the fact that some modes have not been identified well.

Recognizing the need for identification accuracy assessment, the objective of this paper is to derive identification techniques that take advantage of the physics of structural dynamics and can provide realistic bounds for all potential parameter uncertainties—parameters that are initially predicted by the FEM. Moreover, the developed approach includes input optimization that distributes excitation energy in such a way that the influence of residual uncertainties on robust control performance will be reduced. From a practical point of view, the input optimization allows one to achieve desirable identification accuracy (in terms of robust control performance) with much less structure excitation and with a shorter length of time-domain data—which is extremely important for tests in constrained operational environments such as space.

B. Approach to Structural Identification

The most natural way to meet the stated objective is to use a state-space model and apply the Extended Kalman Filter

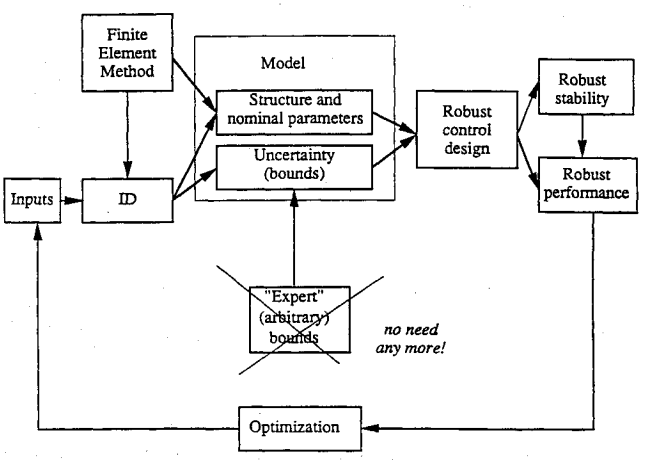

Fig. 1 Integration of identification and robust control.

Received Oct. 16, 1992; revision received July 23, 1993; accepted for publication July 25, 1993.

*Research Associate, Space Engineering Research Center.
†Principal Research Scientist, Space Engineering Research Center. Member AIAA.
‡Professor of Aeronautics and Astronautics, Space Engineering Research Center. Fellow AIAA.
§Professor of Aeronautics and Astronautics, MacVicar Faculty Fellow, Space Engineering Research Center. Associate Fellow AIAA.

(EKF) (see Ref. 6), which allows joint state and parameter estimation.

First, a modal state-space model has a clear physical interpretation, its structural matrices A, B, C, and D consist of natural modal parameters (frequencies, damping ratios, mode shapes), and an FEM provides a physical understanding of how the parameters are arranged in these matrices.

Second, the EKF itself is an augmented dynamic model that provides a complete description of the uncertainties in the original state-space model. Because of this, we are able to treat all noises properly (e.g., separate noises in sensors and actuators, introduction of colored noise models, etc.), model unstructured dynamics, use different nonlinear and adaptive modifications if needed, and so on.

Unfortunately, the application of this identification approach is restricted by having to know the structure of the model and a good initial guess for the parameters (<20–30%). Otherwise, we would face a very complicated problem of nonlinear filtering. So, we can state that the intent of the developed identification approach is measurement based, high-precision updating of a moderate-precision finite element (FE) model from noisy experimental data.

An observation can be made at this point: FEM and input-output identification techniques are two edges of a big area in which identification algorithms (EKF or other) construct models from measurement data using, to a greater or lesser degree, physical insight to the structural dynamics.

C. Main Results of this Paper

In terms of identification and control integration, three independent results are presented in this paper. First is the solution to the problem of input optimization. This problem is interpreted as a problem of controlling the Riccati equation for the augmented covariance matrix that describes the evolution of identification accuracy. The class of Riccati equation control problems dates to the 1960s and 1970s.[7,8] In Ref. 9, a new general approach to the solution of these problems is proposed. It allows transformation of the initially nonlinear control problems to an equivalent one that is linear with respect to phase variables, at the expense of a nontraditional usage of the Riccati equation's analytical properties. As a result of application of this approach to input optimization, there appears a quadratic-linear boundary-value problem that allows a convergent numerical solution. It is significant that some well-known results[10,11] on input optimization in linear systems, based on the parameter sensitivity function criterion, can be shown to be particular cases of this solution.

The second result of this paper is a dramatic improvement in the numerical capability of the EKF for modal models. The question of computational cost has always been a stumbling block for applying EKF to high-dimensional systems. But it seems to be obvious that one can take advantage of the special modal form of the matrices A, B, C, and D for controlled, flexible structures and obtain a block decomposition of the modal state-space model. As a result, the "modal" EKF can be derived.

The third result consists of modifying Petersen-Hollot bounds[12] for the case when parameter uncertainties are random and correlated. It should be noticed that some other robust control techniques (sensitizing methods[13–15] and optimal projection method[16]) are compatible with post-ID models of uncertainty in a more straightforward way and do not need such a modification. Petersen-Hollot's robust control formulation is chosen because it has excellent physical interpretation[17] in terms of uncertain potential energy, it provides robust stability under real parameter variations, and, therefore, it gives more reason to fight for robust performance by providing an optimal post-ID model of the remaining uncertainty.

The necessity of developing this integrated approach was motivated by work on the Middeck Active Control Experiment (MACE)[18,19] developed at the Space Engineering Research Center (SERC), MIT, and sponsored by the NASA In-Step and Control/ Structure Interaction (CSI) Offices. The objective of this shuttle manifested experiment is to investigate the behavior of controlled, flexible structures whose dynamics change between the 1-g and 0-g environments. In this paper, some preliminary 1-g results are given.

II. Statement of Active Identification Problem

A. Mathematical Model

The problem addressed concerns a linear stochastic system written in standard state-space notation

$$\dot{x} = A(\alpha)x + B(\beta)u + \xi, \qquad t \in (0, T) \tag{1}$$

$$y = C(\alpha, \beta)x + D(\beta)u + \eta \tag{2}$$

where $x(t)$ is the state vector ($n \times 1$); $u(t)$ is the input (control) ($m \times 1$); $\xi(t)$ is an external disturbance ($n \times 1$) interpreted as white noise with intensity matrix Ξ_ξ, i.e., $\xi(t) \in W(0, \Xi_\xi)$; $y(t)$ is the measurement vector ($l \times 1$); and $\eta(t)$ is the measurement error (white noise) ($l \times 1$), i.e., $\eta(t) \in W(0, \Xi_\eta)$. The white noises $\xi(t)$ and $\eta(t)$ are allowed to be correlated at each t with cross intensity $\Xi_{\xi\eta}$. Matrices A, B, C, and D have dimensions $n \times n$, $n \times m$, $l \times n$, and $l \times m$, respectively. The initial conditions for Eq. (1) are assumed to be random: $x(0) \in N(\hat{x}_0, \hat{P}_0)$. T is the duration of the identification experiment.

We will characterize uncertainties in the system (1), (2) by two vectors of parameters

$$\alpha = (\alpha_1 \cdots \alpha_N)^T, \qquad \beta = (\beta_1 \cdots \beta_M)^T \tag{3}$$

and assume that they can be described by the following shaping filters:

$$\dot{\alpha} = A_\alpha \alpha + d_\alpha + \theta_\alpha, \qquad \dot{\beta} = A_\beta \beta + d_\beta + \theta_\beta \tag{4}$$

where A_α and A_β are given (or assumed) matrices $N \times N$ and $M \times M$ (often diagonal); $d_\alpha(t)$ and $d_\beta(t)$ are deterministic vectors $N \times 1$ and $M \times 1$; and $\theta_\alpha(t)$ and $\theta_\beta(t)$ are white noises with intensities Ξ_α and Ξ_β, i.e., $\theta_\alpha(t) \in W(0, \Xi_\alpha)$ and $\theta_\beta(t) \in W(0, \Xi_\beta)$. The initial conditions in (4) are $\alpha(0) \in N(\hat{\alpha}, \hat{P}_\alpha)$, $\beta(0) \in N(\hat{\beta}, \hat{P}_\beta)$.

It should be emphasized that the introduced model of uncertainty is quite general for flexible structures, the dynamics of which are described in terms of physical modal parameters. So, the vector α is usually associated with natural frequencies and damping ratios while the vector β is associated with structural mode shapes. In system (1), (2) some general physical features have been taken into account. For example, if sensors and actuators are colocated, the matrices B and C include the same mode shapes. If sensors measure accelerations, the matrix C depends on frequencies and damping ratios and the matrices C and D consist of the same mode shapes. So, the approach under consideration provides physical insight into the identification process, which cannot be said about identification techniques that estimate matrices A, B, C, and D independently.

B. Criterion and Constraints

The goal of identification is to estimate the vector parameters α and β having input $u(t)$ and output $y(t)$ data plus physical information provided by a FEM. The optimization part of the problem consists of choosing the optimal input $u(t)$ over the time interval $(0, T)$ to enhance the identification accuracy at the terminal instant T.

There can be different criteria for optimization of $u(t)$. Most of them can be represented in the general form

$$J_a = \varphi\{W^T P_{\alpha_\Sigma} W\} \to \min_u \tag{5}$$

where P_{α_Σ} is the a posteriori covariance matrix of the combined vector $[(N+M)\times 1]$ of uncertain parameters $\alpha_\Sigma^T = (\alpha^T \vdots \beta^T)$; $W = (W_\alpha \vdots W_\beta)$ is a weighting matrix $[(N+M) \times (N+M)]$; and $\varphi\{\cdot\}$ is a given function.

If $\varphi\{\cdot\} = \mathrm{tr}\{\cdot\}$, the trace of the covariance matrix of the vector $w = W\alpha_\Sigma$ (average variance of w) is minimized. However, this criterion is not very specific because it reflects only a wish to increase the accuracy of estimating the given vector w, which is some arbitrary weighting (W) of the parameters. What is really of interest is increasing the accuracy of estimation in so far as robust control performance depends on this accuracy. Therefore, it is better to treat the criterion (5) as an indication of robust control performance. In this case, there is no vector w any more. Instead of this, we have to determine the required nonlinear function $\varphi\{\cdot\}$ and the matrix W by solving the problem of robust control (see Sec. V).

When optimizing the criterion (5), the following constraint on structural excitation must be considered:

$$J_e = \frac{1}{T} E\left[\int_0^T (x^T K x + u^T R u)\, dt\right] \le \bar{J}_e \tag{6}$$

where $E[\cdot]$ is the mathematical expectation operation and $K \ge 0$ and $R \ge 0$ are given matrices ($n \times n$) and ($m \times m$). The performance J_e has the same form as the one in the LQG control problem. In some cases, when $u(t)$ is a program (not a closed-loop input), the constraint (6) can be simplified and transformed into the deterministic one

$$J_e = \frac{1}{T}\int_0^T (x_0^T K x_0 + u^T R u)\, dt \le \bar{J}_e - J_d \tag{7}$$

where $x_0(t)$ is a nominal trajectory

$$J_d = \frac{1}{T} E\left[\int_0^T (x - x_0)^T K (x - x_0)\, dt\right]$$

In accordance with the Lagrange methodology, the criterion (5) and the constraint (7) can be combined in a generalized criterion

$$J_\Sigma = J_a + \mu J_e \tag{8}$$

where the scalar Lagrange multiplier μ shows the price of enhancing the robust control performance J_a by increasing the excitation J_e. Introduction of the generalized criterion (8) is convenient for solving the optimization problem. After the problem is solved, μ can be found from the condition $J_e(\mu) = \bar{J}_e - J_d$.

C. Interpretation of Problem as a Form of Riccati Equation Control

The dependence of criterion (5) on the input $u(t)$ can be expressed through the Riccati equation that describes the accuracy of the EKF estimation. To derive this equation, consider first some general points.

It is a well-known fact (see Ref. 6) that, if we introduce an augmented vector

$$x_\Sigma = \begin{pmatrix} x \\ \cdots \\ \alpha \\ \cdots \\ \beta \end{pmatrix} \quad [(n + N + M) \times 1] \tag{9}$$

and consider the combined system (1-4), the original identification problem will be transformed into a nonlinear filtering problem. This problem is very difficult to solve. In particular, the evolution of accuracy appears to be coupled with the estimation process itself. But, since we deal with a priori uncertainties about 20-30% (or lower) and nonlinearities in the filtering problem of simple form (see Sec. IV), we can take advantage of different linear and quasilinear techniques for solving the nonlinear filtering problem. These techniques will be considered in Sec. IV from the point of view of calculating estimates. As to the estimation of accuracy and optimization of the inputs to enhance it, we will limit ourselves to linear techniques in this section.

The linearization is performed with respect to the nominal parameters provided, for example, by a FEM. As a result, we have a linear filtering problem in the augmented system

$$\begin{cases} \dot{x}_0 = A_0 x_0 + B_0 u, & x_0(0) = \hat{x}_0 \\ \Delta\dot{x}_\Sigma = A_\Sigma(x_0, u)\Delta x_\Sigma + d_\Sigma + \xi_\Sigma, & \Delta x_\Sigma(0) \in N(0, \hat{P}_{\Sigma_0}) \end{cases} \tag{10}$$

with measurements

$$\Delta y = C_\Sigma(x_0, u)\Delta x_\Sigma + \eta \tag{11}$$

In (10) and (11), $x_0(t)$ is a nominal motion and $\Delta x_\Sigma(t)$ is a perturbation motion. The linearization is represented by the following notations:

$$\Delta x_\Sigma = \begin{pmatrix} \Delta x = x - x_0 \\ \cdots\cdots \\ \Delta\alpha = \alpha - \alpha_0 \\ \cdots\cdots \\ \Delta\beta = \beta - \beta_0 \end{pmatrix}, \quad A_\Sigma(x_0, u) = \left(\begin{array}{c|c|c} A_0 & H_A(x_0) & H_B(u) \\ \hline 0 & A_\alpha & 0 \\ \hline 0 & 0 & A_\beta \end{array}\right)$$

$$d_\Sigma = \begin{pmatrix} 0 \\ \cdots \\ d_\alpha \\ \cdots \\ d_\beta \end{pmatrix}, \quad \xi_\Sigma = \begin{pmatrix} \xi \\ \cdots \\ \theta_\alpha \\ \cdots \\ \theta_\beta \end{pmatrix}, \quad \hat{P}_{\Sigma_0} = \left(\begin{array}{c|c|c} \hat{P}_0 & 0 & 0 \\ \hline 0 & \hat{P}_\alpha & 0 \\ \hline 0 & 0 & \hat{P}_\beta \end{array}\right)$$

$$\Delta y = y - C_0 x_0 - D_0 u$$

$$C_\Sigma(x_0, u) = [C_0 \,\vdots\, H_C^\alpha(x_0) \,\vdots\, H_C^\beta(x_0) + H_D(u)]$$

$$H_A(x_0) = (A_1 x_0 \,\vdots\, \cdots \,\vdots\, A_N x_0), \quad H_B(u) = (B_1 u \,\vdots\, \cdots \,\vdots\, B_M u)$$

$$H_C^\alpha(x_0) = (C_1^\alpha x_0 \,\vdots\, \cdots \,\vdots\, C_N^\alpha x_0), \quad H_C^\beta(x_0) = (C_1^\beta x_0 \,\vdots\, \cdots \,\vdots\, C_M^\beta x_0)$$

$$H_D(u) = (D_1 u \,\vdots\, \cdots \,\vdots\, D_M u)$$

Here matrices $A_0 = A_0(\alpha_0)$, $B_0 = B_0(\beta_0)$, $C_0 = C_0(\alpha_0, \beta_0)$, and $D_0 = D_0(\beta_0)$ correspond to the nominal parameters α_0 and β_0. So do matrices

$$A_i = A_i(\alpha_0) = \partial A(\alpha)/\partial\alpha_i\,|_{\alpha=\alpha_0} \qquad (i = 1, N)$$

$$C_i^\alpha = C_i^\alpha(\alpha_0, \beta_0) = \partial C(\alpha, \beta)/\partial\alpha_i\Big|_{\substack{\alpha=\alpha_0 \\ \beta=\beta_0}} \qquad (i = 1, N)$$

$$C_i^\beta = C_i^\beta(\alpha_0) = \partial C(\alpha, \beta)/\partial\beta_i\,|_{\alpha=\alpha_0} \qquad (i = 1, M)$$

$$D_i = D_i(\beta_0) = \partial D(\beta)/\partial\beta_i\,|_{\beta=\beta_0} \qquad (i = 1, M)$$

Matrices $B_i = \partial B(\beta)/\partial\beta_i$ do not depend on the nominal parameters because β enters B in a linear way (see Sec. IV).

Now we can write the Riccati equation for the filtering problem (10), (11) and, as a result, arrive at the problem of Riccati equation control

$$\begin{cases} \dot{x}_0 = A_0 x_0 + B_0 u, & x_0(0) = \hat{x}_0 \\ \dot{P}_\Sigma = A_\Sigma(\cdot)P_\Sigma + P_\Sigma A_\Sigma^T(\cdot) - [P_\Sigma C_\Sigma^T(\cdot) + \Xi_{\xi\eta}]\Xi_\eta^{-1}[C_\Sigma(\cdot)P_\Sigma + \Xi_{\xi\eta}^T] + \Xi_{\xi_\Sigma}, & P_\Sigma(0) = \hat{P}_{\Sigma_0} \end{cases} \tag{12}$$

where $P_\Sigma(t)$ is the a posteriori covariance matrix $[(n+N+M) \times (n+N+M)]$ of the augmented vector $x_\Sigma(t)$.

To characterize identification accuracy, we can take only one block of this matrix, block

$$P_{\alpha_\Sigma}(t) = \begin{pmatrix} P_\alpha(t) & \vdots & P_{\alpha\beta}(t) \\ \cdots & \vdots & \cdots \\ P_{\beta\alpha}(t) & \vdots & P_\beta(t) \end{pmatrix}$$

which corresponds to the vector $\alpha_\Sigma^T = (\alpha^T \vdots \beta^T)$. So, the criterion for the control problem is written in the generalized form (8) based on the accuracy criterion (5) and the structural excitation constraint (7).

III. Solution of the Active Identification Problem

A. Linear Equivalent Problem

The optimization problem (12) and (10) is very difficult to solve due to the nonlinearity of the Riccati equation and its matrix structure. That is why the approach of Ref. 9 is implemented here. As was mentioned in the Introduction, it allows formulation of a linear equivalent problem by means of a nontraditional usage of the so-called analytical properties of the Riccati equation. Why nontraditional? If a linear Hamiltonian system [6] for the matrix variables $S_\Sigma(t)$ and $Q_\Sigma(t)$ [with dimension $(n+N+M)\times(n+N+M)$] is introduced, the covariance matrix $P_\Sigma(T)$ can be expressed in the form[6]

$$P_\Sigma(T) = Q_\Sigma(T) S_\Sigma^{-1}(T) \tag{13}$$

Although $S_\Sigma(t)$ and $Q_\Sigma(t)$ satisfy linear differential equations, nonlinearity in the problem would not be avoided—it would appear in the criterion—and the dimension would be doubled. At the same time, it is not necessary to care about the whole matrix P_Σ. Just a scalar convolution (5) of the block $P_{\alpha_\Sigma}(T)$ needs to be considered. It turns out that there is a sufficient number of Hamiltonian variables through which the criterion can be expressed and retain its linearity. [We speak about retaining the linearity of the "internal" criterion $\{\cdot\}$ in (5).]

The linear equivalent problem can be formulated as follows. It includes an equation for the reference motion $x_0(t)$ plus a projection of the Hamiltonian system for the rectangular matrices $S(t)$ and $Q(t)$ [with dimension $(n+N+M)\times(N+M)$]

$$\begin{cases} \dot{x}_0 = A_0 x_0 + B_0 u, & x_0(0) = \hat{x}_0 \\ \dot{Q} = \tilde{A}_\Sigma(\cdot) Q + \tilde{\Xi}_{\xi_\Sigma} S \\ \dot{S} = -\tilde{A}_\Sigma^T(\cdot) S + C_\Sigma^T(\cdot)\, \Xi_\eta^{-1} C_\Sigma(\cdot) Q \end{cases} \tag{14}$$

with linear constraints

$$\hat{P}_{\Sigma_0} S(0) - Q(0) = 0 \tag{15}$$

on the left end of the interval $(0, T)$

$$S(T) = W_\Sigma = \begin{pmatrix} 0_{n\times(N+M)} \\ \cdots\cdots\cdots \\ W \end{pmatrix} \tag{16}$$

on the right end, as well as the linear-quadratic criterion

$$J_\Sigma = \varphi\{W_\Sigma^T Q(T)\} + \mu J_e \to \min_\mu \tag{17}$$

The nonlinear operation $\varphi\{\cdot\}$ is "external" and represents, in a general case, the robust control performance. In (14), $\tilde{A}_\Sigma(\cdot)$ is matrix $A_\Sigma(\cdot)$ with $\tilde{A}_0 = A_0 - \Xi_{\xi\eta}\Xi_\eta^{-1} C_0$ instead of A_0; and $\Xi_{\xi_\Sigma} - \tilde{\Xi}_{\xi_\Sigma}$ when $\Xi_\xi \to \tilde{\Xi}_\xi = \Xi_\xi - \Xi_{\xi\eta}\Xi_\eta^{-1}\Xi_{\xi\eta}^T$.

The equivalence of criterion (17) to (8) is proved in Ref. 9.

B. Quadratic-Linear Boundary-Value Problem

The equivalent problem (14–17) can be solved on the basis of the maximum principle. Since the system (14) is linear with respect to S and Q as well as the constraints (15) and (16), and the criterion (17) is linear-quadratic, the solution will be much simpler than the solution to the original problem in terms of P_Σ.

The application of the maximum principle is quite straightforward, and all intermediate arithmetic is omitted. For the same reason, the final boundary-value problem (BVP) is given only for the case of uncertainties in the matrix A (vector α) and with $\Xi_{\xi\eta} = 0$. In this way the essence of the numerical algorithm can be shown and, at the same time, even more complicated block expressions can be avoided.

So, the BVP has the following quadratic-linear form:

$$\begin{aligned} \dot{x}_0 &= A_0 x_0 + (\tfrac{1}{2}\mu) B_0 R^{-1} B_0^T \lambda_0 \\ \dot{\lambda}_0 &= -A_0^T \lambda_0 + 2\mu K x_0 + 2\sum_{i=1}^{N} A_i S_x Q_\alpha^T [\varphi\dot{P}] e_i \\ \dot{Q}_x &= A_0 Q_x + H_A(x_0) Q_\alpha + \Xi_\xi S_x \\ \dot{Q}_\alpha &= A_\alpha Q_\alpha + \Xi_\alpha S_\alpha \\ \dot{S}_x &= -A_0^T S_x + C_0^T \Xi_\eta^{-1} C_0 Q_x \\ \dot{S}_\alpha &= -A_\alpha^T S_\alpha - H_A(x_0) S_x \end{aligned} \tag{18}$$

with constraints

$$\begin{gathered} t = 0\colon x_0(0) = \hat{x}_0, \qquad \hat{P}_0 S_x(0) - Q_x(0) = 0 \\ \hat{P}_\alpha S_\alpha(0) - Q_\alpha(0) = 0 \end{gathered} \tag{19}$$

$$t = T\colon \lambda_0(T) = 0, \qquad S_x(T) = 0, \qquad S_\alpha(T) = W_\alpha \tag{20}$$

In (18–20) Q_x, S_x and Q_α, S_α are blocks $(n\times N)$ and $(N\times N)$, which are partitions of Q and S

$$Q = \begin{pmatrix} Q_x \\ \cdots \\ Q_\alpha \end{pmatrix}, \qquad S = \begin{pmatrix} S_x \\ \cdots \\ S_\alpha \end{pmatrix}$$

The matrix $[\varphi\dot{P}]$ $(N\times N)$ is caused by the nonlinearity of the criterion (5)

$$[\varphi\dot{P}] = \frac{\partial}{\partial P_\alpha(T)} \varphi[P_\alpha(T)] \Big|_{P_\alpha(T) = W^T Q_\alpha(T)}$$

The following $N\times 1$ vectors $e_1 = (1\ 0 \ldots 0)^T, \ldots, e_N = (0\ 0 \ldots 1)^T$ are also involved.

After the BVP is solved, the optimal input $u(t)$ can be obtained in the form

$$\tilde{u}(t) = \frac{1}{2\mu} R^{-1} B_0^T \lambda_0 \tag{21}$$

It is interesting to notice that the received result generalizes a linear BVP of Ref. 10 in which input optimization was carried out with respect to the parameter sensitivity function criterion. So, to obtain the BVP of Ref. 10 from the BVP under consideration here, one has to assume $Q_\alpha = I_N$ (I_N is unity matrix) $N\times N$, $[\varphi\dot{P}] = I_N$, $\Xi_\xi = 0$, $A_\alpha = 0$, and $\Xi_\alpha = 0$. In this case, we turn off the extra variables S_α and Q_α, which form the criterion (5), which takes into account dynamic and noise characteristics.

It is obvious that the complexity of the quadratic-linear BVP in comparison with the linear BVP is a price for optimizing a direct accuracy criterion instead of an indirect sensitivity function criterion.

C. Numerical Algorithm for Optimization

The algorithm consists of the following sequence of operations (superscript index is the number of the iteration):

1) Set any initial input $u^0(t)$.

2) Solve the equation for $x_0(t)$ in the system (14).

3) Solve the conditional BVP in system (14) for $Q_x(t)$, $Q_\alpha(t)$, $S_x(t)$, and $S_\alpha(t)$ with $x_0(t) = x_0^0(t)$, $u(t) = u^0(t)$. This problem is linear and allows an effective solution—for example, by the transition matrix method or by the Riccati equation method.

4) Solve the equations for $\lambda_0(t)$, $Q_x(t)$, $Q_\alpha(t)$, $S_x(t)$, and $S_\alpha(t)$ in the system (18) backward in time with boundary conditions $\lambda_0(T) = 0$, $Q_x(T) = Q_x^0(T)$, $Q_\alpha(T) = Q_\alpha^0(T)$, $S_x(T) = 0$, and $S_\alpha(T) = W_\alpha$ as well as with $x_0(t) = x_0^0(t)$.

5) Construct the next approximation to the input

$$u^1(t) = \nu\tilde{u}(t) + (1-\nu) u^0(t)$$

where $\tilde{u}(t)$ is the input defined from the optimal structure (21) with $\lambda_0(t) = \lambda_0^0(t)$, and ν is a scalar $(0<\nu<1)$ chosen in a way to satisfy the condition $J_\Sigma^1(\nu)<J_\Sigma^0$. The latter condition can always be achieved in a few subiterations if we set $\nu=\nu^j=\frac{1}{2}^{j-1}$ $(j=1,\ldots)$. Of course, there can be different representations ν^j, but there is no need to make it complicated.

Then operations 2–5 are reiterated with a change of iteration number 0 by 1, 1→2, etc. The iteration process should be stopped when $|J_\Sigma^k - J_\Sigma^{k-1}|<\epsilon$, where $\epsilon>0$ is a given small value. The convergence of a numerical algorithm of this kind is proved in Ref. 9.

IV. Realization of EKF

A. Principal EKF Equations

After the active identification experiment has been planned and the optimal—based on the FE model a priori—input is found we can carry out the experiment, record measurement data $y(t)$, and then process it. Thereby, the on-line identification algorithm is an EKF. But, unlike the optimization problem, where we dealt with the characteristics of the joint state-parameter estimation accuracy (covariance matrix P_Σ), the filtering process is more sensitive to linearization errors. In other words, in the identification experiment design, it was not especially necessary to find an absolute exact optimal input—and we can use the simple linearized model for P_Σ. But, in the filtering problem, neglecting the linearization errors can cause divergence of the identification process.

Taking this into account, we will use the EKF in its quasi-linear version. The final on-line identification algorithm can be written in the form

$$\begin{cases}\dot{x}_\Sigma^* = \bar{A}_\Sigma(\alpha^*)x_\Sigma^* + K(x_\Sigma^*, u)[y - \bar{C}_\Sigma(\alpha^*, \beta^*)x_\Sigma^* - D(\beta^*)u] + \bar{B}_\Sigma(\beta^*)u + d_\Sigma \\ \dot{P}_\Sigma = A_\Sigma(x^*, \alpha^*, u)P_\Sigma + P_\Sigma A_\Sigma^T(x^*, \alpha^*, u) - K(x_\Sigma^*, u)\Xi_\eta K^T(x_\Sigma^*, u) + \Xi_{\xi_\Sigma}\end{cases} \tag{22}$$

where $K(\cdot) = [P_\Sigma C_\Sigma^T(\cdot) + \Xi_{\xi\eta}]\Xi_\eta^{-1}$ is the Kalman gain. Initial conditions in (22) are $x_\Sigma^*(0) = \hat{x}_{\Sigma_0}$, $P_\Sigma(0) = \hat{P}_{\Sigma_0}$. The matrices

$$\bar{A}_\Sigma(\alpha^*) = \left[\begin{array}{c|c|c} A(\alpha^*) & 0 & 0 \\ \hline 0 & A_\alpha & 0 \\ \hline 0 & 0 & A_\beta \end{array}\right], \qquad \bar{B}_\Sigma(\beta^*) = \left[\begin{array}{c} B(\beta^*) \\ \hline 0 \\ \hline 0 \end{array}\right]$$

$$\bar{C}_\Sigma(\alpha^*, \beta^*) = [C(\alpha^*, \beta^*) \vdots 0 \vdots 0]$$

reflect the fact that nonlinearities are taken into account directly when calculating the estimate x_Σ^*. At the same time, the matrices $A_\Sigma(\cdot)$ and $C_\Sigma(\cdot)$ are products of linearization when calculating the covariance matrix P_Σ. But this linearization is made in the vicinity of current estimates unlike the linearization in the model (10), (11).

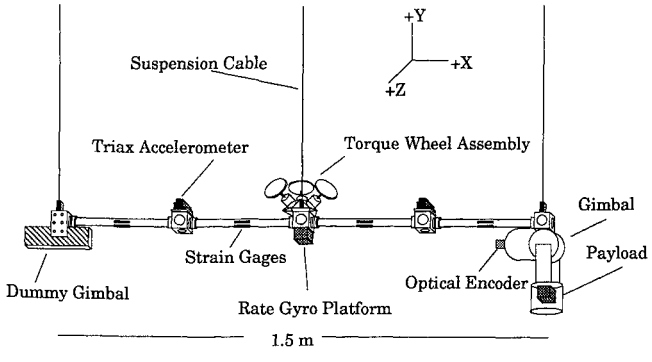

Fig. 2 Middeck Active Control Experiment—schematic.

B. EKF Modal Decomposition

The results obtained earlier are applicable to any system that can be described in terms of Eqs. (1) and (2). But, in high-dimensional systems (say, $n>50$, $N+M>500$), the possibility of applying the EKF is under question. Counting on future more powerful computers is not so optimistic as one could think. The reason is in the cubic dependence of computational expense on the system dimension.

But, if we limit ourselves to a certain class of systems—namely, to flexible structures dynamics that allow modal representation, we can significantly reduce the numerical complexity of the EKF. The main idea here is to take advantage of the obvious fact that the structural matrices A, B, C, and D are padded with zeros, and most nonzero elements are arranged as diagonals.

Let us introduce the further decomposition of the augmented vector (9)

$$x_\Sigma = \begin{bmatrix} x \\ \cdots \\ \alpha \\ \cdots \\ \beta \end{bmatrix} \Rightarrow x = \begin{pmatrix} z \\ \cdot\cdot \\ \nu \end{pmatrix}, \qquad \alpha = \begin{pmatrix} \omega \\ \cdot\cdot \\ \zeta \end{pmatrix}, \qquad \beta = \begin{bmatrix} \phi_d \\ \cdots \\ \phi_r \\ \cdots \\ \phi_a \\ \cdots \\ \phi_c \end{bmatrix} \tag{23}$$

where z is an $N_m \times 1$ vector of modal displacements; $\nu = \dot{z}$ is an $N_m \times 1$ vector of modal velocities; ω is an $N_m \times 1$ vector of natural frequencies; ζ is an $N_m \times 1$ vector of damping ratios; and ϕ_d, ϕ_r, ϕ_a, and ϕ_c are, correspondingly, $N_m \cdot N_d \times 1$, $N_m \cdot N_r \times 1$, $N_m \cdot N_a \times 1$, and $N_m \cdot N_c \times 1$ vectors of mode shapes at the locations where displacement (d), rate (r), and acceleration (a) sensors as well as actuators (c) are installed.

Moreover, N_m is the number of modes; N_d, N_r, N_a, and N_c are the numbers of sensors (d, r, a) and actuators (c).

Here is the well-known modal representation of the system (1), (2)

$$\underbrace{\begin{pmatrix} \dot{z} \\ \cdot\cdot \\ \dot{\nu} \end{pmatrix}}_{\dot{x}} = \underbrace{\left(\begin{array}{c|c} 0 & I \\ \hline -\mathrm{diag}(\omega_i^2) & -\mathrm{diag}(2\zeta_i\omega_i) \end{array}\right)}_{A(\alpha)} \cdot \underbrace{\begin{pmatrix} z \\ \cdot\cdot \\ \nu \end{pmatrix}}_{x} + \underbrace{\begin{pmatrix} 0 \\ \cdot\cdot \\ \Phi_c \end{pmatrix}}_{B(\beta)} u + \underbrace{\begin{pmatrix} 0 \\ \cdot\cdot \\ \xi_\nu \end{pmatrix}}_{\xi} \tag{24}$$

$$\underbrace{\left\{\begin{array}{c} y_d \\ \cdots \\ y_\nu \\ \cdots \\ y_a \end{array}\right\}}_{y} = \underbrace{\left\{\begin{array}{c|c} \Phi_d & 0 \\ \hline 0 & \Phi_r \\ \hline -\Phi_a\mathrm{diag}(\omega_i^2) & -\Phi_a\mathrm{diag}(2\zeta_i\omega_i) \end{array}\right\}}_{C(\alpha,\beta)} \cdot \underbrace{\begin{pmatrix} z \\ \cdot\cdot \\ \nu \end{pmatrix}}_{x} + \underbrace{\left\{\begin{array}{c} 0 \\ \cdots \\ 0 \\ \cdots \\ \Phi_a\Phi_c \end{array}\right\}}_{D(\beta)} u + \underbrace{\left\{\begin{array}{c} \eta_d \\ \cdots \\ \eta_r \\ \cdots \\ \eta_a \end{array}\right\}}_{\eta} \tag{25}$$

where Φ_d, Φ_r, Φ_a, and Φ_c are matrices $(N_d \times N_m$, $N_r \times N_m$, $N_a \times N_m$, and $N_c \times N_m)$ reshaped from vectors ϕ_d, ϕ_r, ϕ_a, and ϕ_c, correspondingly.

The corresponding blocks of the matrices A_i $(i=1, N)$, B_i $(i=1, M)$ C_i^α $(i=1, N)$, and C_i^β and D_i $(i=1, M)$ (see Sec. II) can be analytically formed by taking derivatives of the structural matrices $A(\alpha)$, $B(\beta)$, $C(\alpha, \beta)$, and $D(\beta)$ in (24) and (25) with respect to $\alpha^T = (\omega^T \vdots \zeta^T)$ and $\beta^T = (\beta_d^T \vdots \beta_r^T \vdots \beta_a^T \vdots \beta_c^T)$.

From the computational point of view, it is convenient to represent Eqs. (24) and (25) in descrete time. This representation can be done easily thanks to the analytical expression of

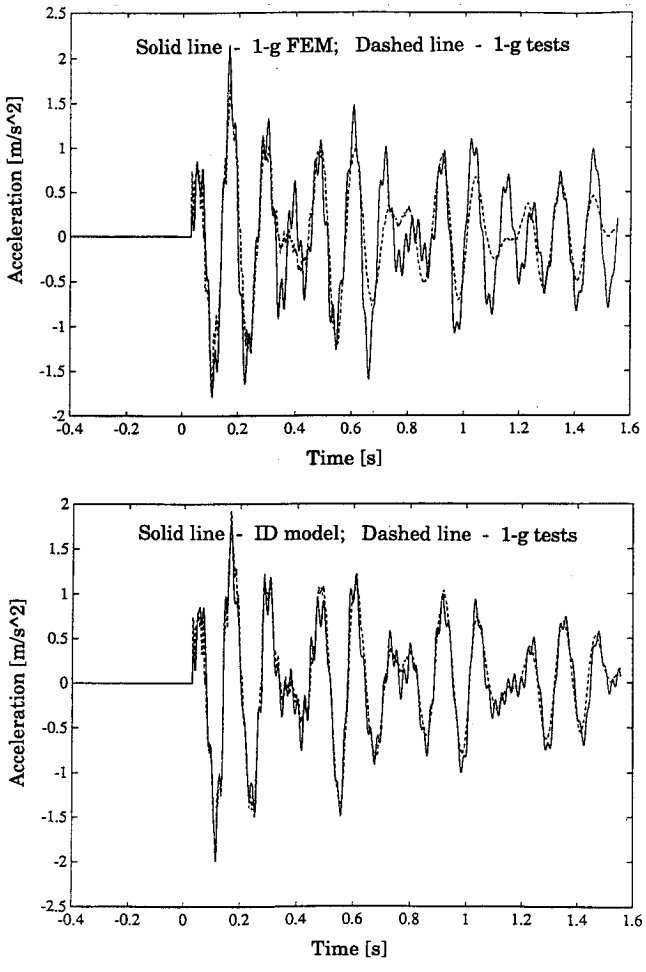

Fig. 3 Output/data set 1.

the transition matrix for modal systems in terms of $\sin(\cdot)$, $\cos(\cdot)$, and $\exp(\cdot)$.

Finally, the equations of the modal EKF should be written in block form according to the decomposition (23) of the augmented vector x_Σ. This operation is laborious but routine. For example, we have to deal, at least, with 8 blocks of the vector x_Σ and with 36 blocks of the symmetrical covariance matrix P_Σ. But, in this way, we are able to exclude all zero-blocks from the calculations and reduce block multiplications (if one of the blocks is diagonal) to array operations. Moreover, some additional transformations can be made to improve numerical capability. For example, one can successively process scalar measurements y_{d_i} $(i = 1, N_d)$, y_{r_i} $(i = 1, N_r)$, and y_{a_i} $(i = 1, N_a)$ (when the matrix Ξ_η is diagonal), use the analytical properties of the Riccati equation (see Sec. III), and so on.

The goal and space limitations of this paper do not allow us to cover all of these issues. But they are realized in developing the corresponding software. Here we limit ourselves to stating that the modal EKF has approximately linear (not cubic, as in a standard EKF) dependence of computational expense on the number of modes. This dependence was confirmed by solving real high-dimensional problems (see Sec. VI).

V. Design of Robust Control

A. Taking into Account Identification

The last element of the integrated approach under consideration is the robust control problem. As was mentioned in the Introduction, we will focus only on the Petersen-Hollot method.[12] Moreover, we will limit ourselves only to full-state feedback control with uncertainties only in the matrix A to show the integration of identification and robust control problems. At the same time, if more general dynamic and output feedback techniques are available, they can be relatively easily incorporated into this integrated approach.

The main area of our interest is the fact that active identification provides, by the beginning of control, a post-ID model of uncertainty in the form

$$\Delta\alpha = \alpha - \alpha^* \in N(0, P_\alpha) \tag{26}$$

while Petersen-Hollot's bounds require the following form:

$$|\Delta\alpha_i| = |\alpha_i - \alpha_i^*| < \Delta_i, \qquad (i = 1, N) \tag{27}$$

Of course, the statistical information (26) can be converted into the bounds (27). For example, one could choose $\Delta_i = 3\sigma_{\alpha_i}$ where σ_{α_i} is the standard deviation of α_i. But, in this case

$$P\{|\alpha_i - \alpha_i^*| < \Delta_i\} \le p = 0.9975^N \tag{28}$$

and, if N is large enough, p can be inadmissibly small (say, $p < 0.9$). The more delicate way would be to set $p(p \to 1)$ and solve (28) for $\Delta_i (i = 1, N)$. But this constraint does not define a unique solution, and there remains the problem of how to use the remaining $N - 1$ degrees of freedom.

Moreover, all of the mentioned transformations (26)→(27) have an overbounding character because they completely lose information about the correlations between parameter errors. The latter information is very important because errors in parameters can cancel each other or, on the contrary, amplify. The correlations can be taken into account by setting quadratic bounds

$$(\alpha - \alpha^*)^T P_\alpha^{-1} (\alpha - \alpha^*) \le \Delta^2(p), \qquad p \to 1 \tag{29}$$

instead of interval ones (27). But, again, if $N \to \infty$, this method will introduce additional conservatism to the robust control algorithm.

So, the best way to deal with the delicate post-ID information (26) is to take it into account directly. This method requires a modification to the Petersen-Hollot method.

B. Standard Petersen-Hollot Method

We review briefly the approach of Ref. 12. The statement of the robust control problem includes the system

$$\dot{x} = (A_0 + \sum_{i=1}^{N} \Delta\alpha_i A_i)x + B_0 u, \qquad x(0) = \hat{x}_0 \tag{30}$$

the model of uncertainty

$$|\Delta\alpha_i| \le \Delta_i, \qquad A_i = l_i m_i^T, \qquad l_i - n \times 1$$
$$m_i - n \times 1 \qquad (i = 1, N) \tag{31}$$

and the control performance

$$J = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau (x^T K x + u^T R u)\, dt \tag{32}$$

Then the Lyapunov function $V(x) = x^T \Lambda x$, $\Lambda(n \times n)$ is introduced and, in the function $\dot{V}(x) = \dot{V}_0(x) + \dot{V}_\alpha(x)$, the term $\dot{V}_\alpha(x)$, depending of $\Delta\alpha$, is bounded so that $\dot{V}(x) \le 0 \; \forall \, \Delta\alpha$. This guarantees robust stability. The bounding operation is based on the following sequence of inequalities:

$$\dot{V}_\alpha(x) = x^T \Lambda \left(\sum_{i=1}^{N} \Delta\alpha_i A_i \right) x + x^T \left(\sum_{i=1}^{N} \Delta\alpha_i A_i^T \right) \Lambda x$$

$$\le \left| 2x^T \left(\sum_{i=1}^{N} \Delta\alpha_i \Lambda l_i m_i^T \right) x \right| \le 2 \sum_{i=1}^{N} \Delta_i |(x^T \Lambda l_i)(m_i^T x)|$$

$$\le \sum_{i=1}^{N} \Delta_i^{l^2} (x^T \Lambda l_i)^2 + \sum_{i=1}^{N} \Delta_i^{m^2} (m_i^T x)^2$$

$$= x^T \Lambda L \, \Delta_L L^T \Lambda x + x^T M \Delta_M M^T x \tag{33}$$

where $\Delta_i = \Delta_i^l \Delta_i^m$, $\Delta_L = \text{diag}(\Delta_i^{l^2})$, $\Delta_M = \text{diag}(\Delta_i^{m^2})$, $L = (l_1 \ldots l_N) - n \times N$, and $M = (m_1 \ldots m_N) - n \times N$.

Finally, the robust controller is

$$u = -R^{-1} B_0^T \Lambda x \tag{34}$$

where Λ is the solution to the robust Riccati equation

$$\Lambda A_0 + A_0^T \Lambda + (K + M \Delta_M M^T) - \Lambda (B_0 R^{-1} B_0^T - L \Delta_L L^T) \Lambda = 0 \tag{35}$$

C. Modified Petersen-Hollot Bounds

The problem statement now includes the same system (30) and the performance (32). But the model of uncertainty is (26) with random correlated parameters.

Now, in the function $\dot{V}(x)$, the term $\dot{V}_\alpha(x)$ can be bounded so that

$$P\{\dot{V}(x) \le 0\} = p, \qquad p \to 1 \tag{36}$$

In other words, we have to find a bounding term $\dot{V}_\alpha^p(x)$ from the condition

$$P\{\dot{V}_\alpha(x) \le \dot{V}_\alpha^p(x)\} = p \tag{37}$$

The function $\dot{V}_\alpha(x)$ is linear with respect to the parameters $\Delta\alpha_i (i = 1, N)$

$$\dot{V}_\alpha(x) = 2 \sum_{i=1}^{N} \Delta\alpha_i c_i(x) = 2 \Delta\alpha^T c(x) \tag{38}$$

where

$$c(x) = \begin{bmatrix} x^T \Lambda l_1 m_1^T x \\ \vdots \\ x^T \Lambda l_N m_N^T x \end{bmatrix}$$

This means that having the statistical characteristics

$$m_V = E[\dot{V}_\alpha(x)] = 0, \quad \sigma_V^2 = E\{[\dot{V}_\alpha(x) - m_V]^2\} = 4 c^T(x) P_\alpha c(x)$$

we can solve the probabilistic problem (36) or (37) and find

$$\dot{V}_\alpha^p(x) = r \sigma_V = 2r \sqrt{c^T(x) P_\alpha c(x)} \tag{39}$$

It is important that r does not depend on N [unlike Δ in (29)]. For example, if $p = 0.9975$, $r = 3$ $\forall N$. This circumstance reduces the conservatism of the robust control because it prevents unrealistic (hardly probable) combinations of parameters.

Unfortunately, $\dot{V}_\alpha^p(x)$ does not have quadratic dependence on x. That is why an overbounding operation must be involved. But, unlike the overbounding operation (33), which deals with functions $|\cdot|$, the function (39) has better analytical properties. Using them one can show that there is an upper estimate

$$\dot{V}_\alpha(x) \le \dot{V}_\alpha^p(x) = \rho (x^T \Lambda L P_L L^T \Lambda x + x^T M P_M M^T x) \tag{40}$$

where $P_L * P_M = r^2 P_\alpha$ ($*$ is an array operation: $P_{L_{ij}} \cdot P_{M_{ij}} = r^2 P_{\alpha_{ij}}$).

All of the routine transformations are omitted here. We will mention only that, if $\rho = 1$, the proof of the inequality (40) is quite straightforward. At the same time, assuming that there exists an optimal ρ in the range $0 < \rho < 1$, one can find it as a result of the optimization problem:

$$J = \rho \to \min, \qquad F(x, \rho) > 0, \qquad 0 < \rho < 1 \tag{41}$$

where

$$F(x, \rho) = \rho^2 [x^T \Lambda(\rho) L P_L L^T \Lambda(\rho) x + x^T M P_M M^T x]^2 - c^T(x) P_\alpha c(x)$$

The function $F(x, \rho)$ has good analytical properties so we can expect an acceptable numerical solution for (41).

Finally, the robust control has the same form (34) with the modified robust Riccati equation

$$\Lambda A_0 + A_0^T \Lambda + (K + \rho M P_M M^T) - \Lambda (B_0 R^{-1} B_0^T - \rho L P_L L^T) \Lambda = 0 \tag{42}$$

D. Robust Control Performance

After the controller (34), (42) is constructed, we can use it for controlling the stochastic system (1). The corresponding performance can be estimated by the following expression[6]:

$$J = \text{tr}\{S \Xi_\xi\} \tag{43}$$

Taking into account the connection of robust control design with the results of identification, we will have $J = J(\alpha^*, P_\alpha)$. The estimate α^* that appears in matrices $A_0 = A_0(\alpha^*)$ and $A_i = A_i(\alpha^*)$ $(i = 1, N)$ is available only a posteriori. But the input for identification should be optimized a priori. In this situation, the way out is in optimizing the mean of J

$$J_m(P_\alpha) = E[J(\alpha^*, P_\alpha)] = \int J(\alpha^*, P_\alpha) p(\alpha^*)\, d\alpha^* = \varphi\{P_\alpha\} \tag{44}$$

where $p(\alpha^*) \in N(\hat{\alpha}, \hat{P}_\alpha)$ is the a priori model of uncertainty.

The problem (42-44) can be solved numerically (for example, by the Monte-Carlo method). Alternatively, one can take advantage of the fact that the post-ID uncertainty is quite small. In this case, the nonlinear function (44) can be linearized

$$J_m(P_\alpha) = J_m(0) + \text{tr}\{W_\alpha^T P_\alpha W_\alpha\} \tag{45}$$

where

$$W_\alpha^2 = \left. \frac{\partial J_m(P_\alpha)}{\partial P_\alpha} \right|_{P_\alpha = 0} \tag{46}$$

The derivative (46) can be calculated before the optimization process (for example, by the method of finite increments). The second term in (45) is to be optimized by choosing the input $u(t)$. So, we have the linear variant of the criterion (5) with $\varphi\{\cdot\} = \text{tr}\{\cdot\}$. But, in this case, the weighting matrix W_α reflects the cost of uncertain parameters α in terms of robust control performance.

It should be emphasized that, from the physical point of view, the optimal input $\tilde{u}(t)$ is a probing feedback signal (21) in the conjugate system for $\lambda_0(t)$ [see Eq. (18)] representing the inverted dynamics of the structure. The input $\tilde{u}(t)$ provides the best distribution of energy over all important directions that can contribute to robust control performance.

VI. Application to MACE

This EKF procedure was applied to test data from the MACE. MACE is a NASA In-Step and CSI Office funded Space Shuttle flight experiment tentatively scheduled for launch in the summer of 1994. The objective of MACE is to study the dependence on gravity of structural dynamics of a flexible spacecraft. Of particular interest is the extent to which the controlled performance of a precision controlled, multiple payload spacecraft can be predicted prior to launch and then refined once on orbit. To refine a finite element model of the spacecraft, as it behaves suspended in 1 g, the EKF procedure detailed in the prior sections was applied to experimental input/output data.

A drawing of the MACE test article[18,19] is shown in Fig. 2. It consists of a segmented straight tubular bus with a two-axis pointing/scanning payload attached to the right end and a dummy mass attached to the left end. The dummy mass will be replaced by a second two-axis gimbal in the near future. The test article is instrumented with two angle encoders on the gimbal axes, two three-axis rate gyro platforms, and other assorted sensors (accelerometers, strain gauges). One rate gyro platform is mounted in the payload and the other is mounted under the torque wheel assembly. The two-axis gimbal allows

rotation about the x and z axes via two DC torque motors. The torque wheel assembly is comprised of three orthogonally mounted DC servo motors with an inertia wheel mounted to each.

The characteristics of the FE model are as follows: 58 nodal points (348 = 58 · 6 DOF), 40 modes in the range 1–1000 Hz (including 11 constrained rigid modes), and gravity and suspension effects are taken into account.[20] In addition to the FEM model, internal dynamics of the rate gyros (second-order filters) are included in the MACE model. Since MACE can be specified as a "very flexible and multivariable" experiment with parametric uncertainty and somewhat noisy sensors, this application is appropriate for the EKF approach.

The EKF approach was applied to two types of data. First, simulated data was generated by adding broadband measurement noise to time response data calculated using the finite element model. Nominal and error bound values for various model parameters were estimated using the EKF procedure. Second, actual data, measured using noisy sensors, was used to estimate various finite element model parameters. This process presented additional challenges because the statistics of the noises were not well known, the finite element model now exhibited both structured and unstructured uncertainties, and the test article exhibited significant nonlinear behavior in some axes.

Simulated ID data was calculated for all of the MACE channels [5 inputs: gimballed payload (axes x and z) and torque wheel (axes x, y, and z); 22 outputs: rate gyro on payload (axes x, y, and z), rate gyro on torque wheel assembly (axes x, y, and z), 8 strain gauges (vertical and horizontal), gimbal encoder (axes x and z), and 2 accelerometers (axes x, y, and z)]. Impulse and white noise inputs were used. The results show that the estimates converged to the true parameters, thereby up to 200–300 parameters were estimated. The desired

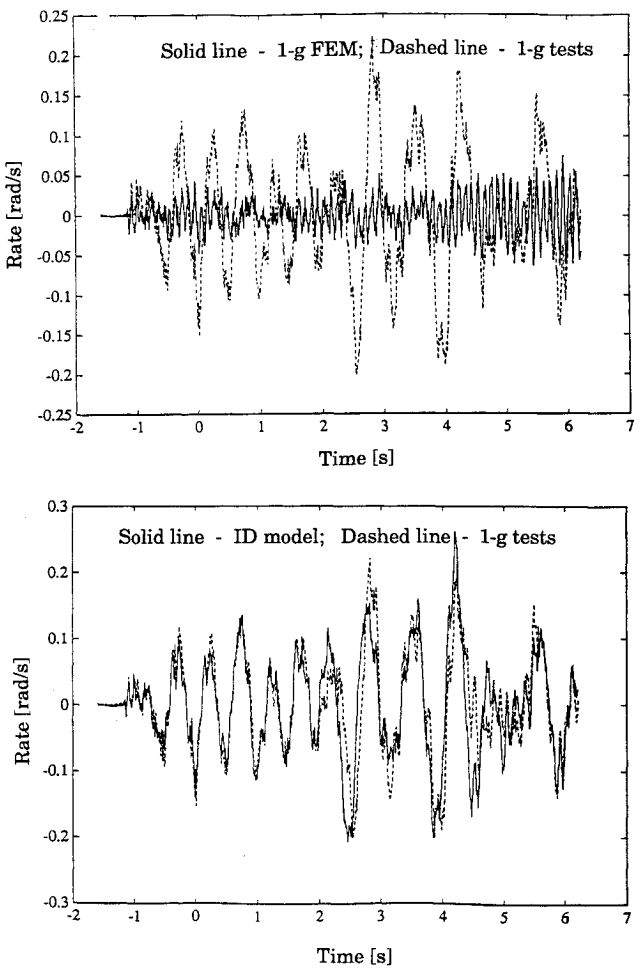

Fig. 4 Output/data set 2 (flexible contribution).

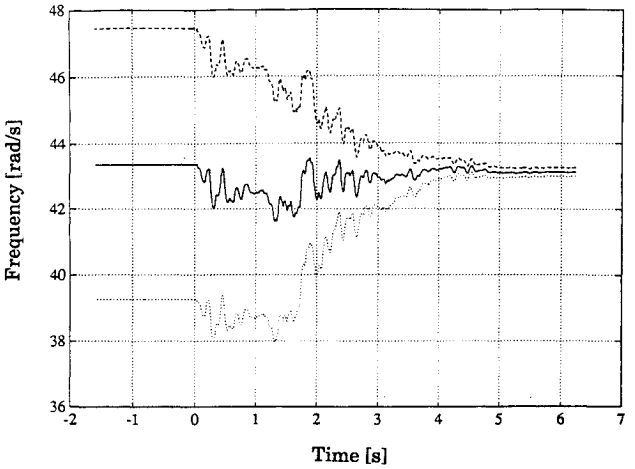

Fig. 5 Convergence of filtering process.

accuracy can be achieved with quite a short length N_d of time-domain data ($N_d \approx 1000$–2000), the computational cost is about 20–40 min (MATLAB, SUN Workstation) and the estimate of that cost for the standard EKF exceeds 24 hours.

Identification with real data was carried out for two data sets: 1) Gimballed payload (about z-axis)/2 accelerometers (along y-axis) and 2) Gimballed payload/Rate Gyro on payload (both about z-axis). Figure 3 presents the comparison of the 1-g FE model and the estimated model with respect to the measurement data (data set 1, impulse input). Figure 4 presents similar results for data set 2 with white noise input. In the latter case, only the contribution of the flexible modes is considered (i.e., the contribution of the first 11 nominal rigid, suspension, and payload articulation modes is subtracted). The improvement of the 1-g FE model, mainly in terms of the payload mode, is obvious; at the same time, the FEM provides quite an updatable guess.

It should be noticed that the residual discrepancy between the estimated and data responses is a result of experimental noises. But the confidence interval for all potential time responses, based on the parametric error bounds, contains the measured time response. That was confirmed by doing 20 independent experiments under the same conditions. Moreover, the accuracy characteristics analytically predicted by the EKF and based on the sensor-actuator noise specifications are well within the confidence intervals obtained by the Monte-Carlo method conducted by processing 20 EKF realizations.

It is important to understand that realistic accuracy characteristics (bounds) were obtained only after encountering and explaining a nonlinear effect in MACE. This effect consists of transferring energy from the higher modes to the lower ones through nonlinear damping (amplitude and velocity dependent Coulomb friction in the gimbal pivot). To deal with nonlinearities, a time-varying nature for the parameters was assumed and pertinent shaping filters (4) were selected. This approach allowed the estimation to stay in the framework of model interpretation "nominal parameters + bounds." At the same time there are preliminary results on identification of the nonlinear effect in terms of force-state mapping techniques.[21]

Figure 5 presents a typical time history of the bounds (in terms of 3σ) for one parameter (frequency of the third bending mode). Figure 6 provides a global look at how the a priori bounds (50% of uncertainty, for all natural frequencies) are improved. The bounds are introduced for two situations: realistic (when actual data was processed) and theoretical (when data was simulated in the linear MACE model). So, the scale of the nonlinear effect can be clearly seen. In Fig. 6, "not estimated" frequencies correspond to the modes that did not contribute to the response (for example, horizontal modes to vertical motion). The corresponding damping ratios were also updated with quite a high accuracy (up to 3–5%). The mode

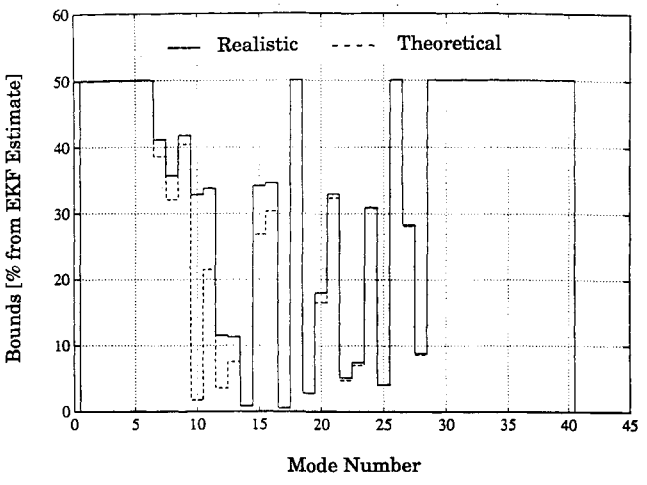

Fig. 6 3-sigma bounds for natural frequencies.

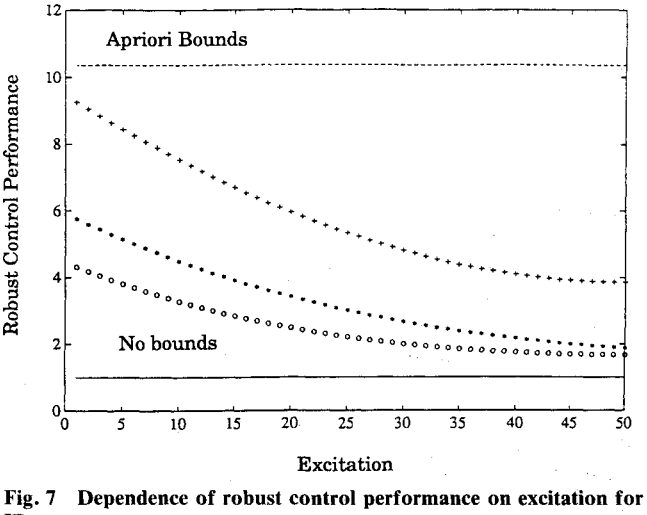

Fig. 7 Dependence of robust control performance on excitation for ID.

shapes in sensor and actuator locations were estimated reasonably only for mode No. 10 (defines eigen oscillation of the payload). However, the products $\phi_{r_i} \cdot \phi_{c_i}$, which represent input-output relations, were updated well for all contributing modes.

The series of experiments exhibited the dependence of identification accuracy on inputs. This makes the problem of input optimization actual. In this paper, we will limit ourselves to considering only the simplest model for the input optimization and robust control problems. This is a 1-DOF oscillator that corresponds to mode No. 10 ($\approx 90\%$ of contribution to the output) with nominal parameters $\omega_0 = 1.296$ Hz, $\zeta_0 = 0.1546$, $b_0 = \phi^0_{c_{10}} = 5.535$, and $c_0 = \phi^0_{r_{10}} = 5.596$. [The mode shapes correspond to data set 2.] The identification problem was solved on an interval with 1000 time points at a sample period of 0.0075 s. The robust control problem was solved on the next (infinite) interval. Thereby, the excitation for identification and robust control performance evaluation were introduced in a form similar to (7) and (32), with the same matrices K and R.

Figure 7 presents an investigation of how the level of excitation influences the robust control performance. All magnitudes here are presented in relevant units, where 1 is the control performance (32) for the nominal system (zero bounds). As one would expect, the white noise input is not so informative for identification. A much better result is provided by the sinusoidal input at the nominal natural frequency ω_0. Further improvement can be achieved by choosing the optimal input, which is a sinusoid with decaying (almost linearly) amplitude. Note that in the case of N-DOF and MIMO systems the results on input optimization will not be so obvious. There the optimal combination of directions and amplitudes can be found only as a result of numerical solution to the active identification problem.

VII. Conclusions

This paper has presented an approach to dynamic system identification taking into account the final goal—namely, to enhance robust control performance. The EKF was chosen, not as one more identification algorithm, but as the only stochastic counterpart to the original state-space model that provides the complete (in terms of two first moments) envelope of structural dynamics. The problems were how to solve the EKF equations in high-dimensional systems and to find optimal probing signals that reduce the effect of residual uncertainties on robust control performance.

Acknowledgments

This research is supported by the NASA In-Step and CSI Offices, with Gregory Stover as Technical Monitor. Special thanks to Erik Saarmaa for his excellent experimental work on MACE.

References

[1]Ljuing, L., *System Identification, Theory for the User,* Prentice-Hall, Englewood Cliffs, NJ, 1987.

[2]Juang, J. N., and Pappa, R. S., "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction," *Journal of Guidance, Control, and Dynamics,* Vol. 8, No. 5, 1985, pp. 620–627.

[3]Juang, J. N., Horta, L. G., and Phan, M., *System/Observer/Controller Identification Toolbox, User's Guide,* NASA Langley Research Center, Hampton, VA, Nov. 1991.

[4]King, A. M., Desai, U. B., and Skelton, R. E., "A Generalized Approach to Q-Markov Covariance Equivalent Realization of Discrete Systems," *Automatica,* Vol. 24, No. 4, 1988, pp. 507–515.

[5]Kosut, R. L., Lau, M. K., and Boyd, S. P., "Set-Membership Identification of Systems with Parametric and Nonparametric Uncertainty," *IEEE Transactions on Automatic Control,* Vol. 37, No. 7, 1992, pp. 929–941.

[6]Gelb, A., et al., *Applied Optimal Estimation,* MIT Press, Cambridge, MA, 1974, pp. 182–200.

[7]Athans, M., "On the Determination of Optimal cost Measurement Strategies for Linear Stochastic Systems," *Automatica,* Vol. 8, No. 4, 1972, pp. 397–412.

[8]Chernous'ko, F. L., "On Observation Process Optimization," *Prikladnaya Matematika i Mechanika,* Vol. 33, No. 2, 1969, pp. 101–111.

[9]Malyshev, V. V., Krasilshikov, M. N., and Karlov, V. I., *Optimization of Observation and Control Processes,* AIAA Education Series, AIAA, Washington, DC, 1992, pp. 48–81.

[10]Mehra, R. K., "Optimal Inputs for Linear System Identification," *IEEE Transactions on Automatic Control,* Vol. AC-19, No. 3, 1974, pp. 192–206.

[11]Kalaba, R. E., and Spingarn, K., "Optimal Inputs and Sensitivities for Parameter Estimation," *Journal of Optimization Theory and Applications,* Vol. 11, No. 1, 1973, p. 56.

[12]Petersen, I. R., and Hollot, C. V., "A Riccati Equation Approach to the Stabilization of Uncertain Linear Systems," *Automatica,* Vol. 22, No. 4, 1986, pp. 397–411.

[13]Ashkenazi, A., and Bryson, A. E., "Control Logic for Parameter Insensitivity and Disturbance Attenuation," *Journal of Guidance, Control, and Dynamics,* Vol. 5, No. 4, 1982, pp. 383–388.

[14]Okada, K., and Skelton, R. E., "Sensitivity Controller for Uncertain Systems," *Journal of Guidance, Control, and Dynamics,* Vol. 13, No. 2, 1990, pp. 321–329.

[15]Hagood, N. W., and Crawley, E. F., *Cost Averaging Techniques*

for Robust Control of Parametrically Uncertain Systems, Space Engineering Research Center, MIT, Rept. SERC #9-91, Cambridge, MA, 1991.

[16]Bernstein, D. S., and Hyland, 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.

[17]Douglas, J., and Athans, M., *Linear Quadratic Control for Systems with Structural Uncertainty,* Space Engineering Research Center, MIT, Rept. SERC #12-91, Cambridge, MA, 1991.

[18]Miller, D. W., de Luis, J., and Crawley, E. F., "Dynamics and Control of Multipayload Platforms: the Middeck Active Control Experiment (MACE)," IAF-90-292, 41st *Congress of the International Aeronautical Federation* (Dresden, Germany), International Astronautical Federation, 1990, pp. 1-11.

[19]Miller, D. W., Saarmaa, E., and Jacques, R. N., "Preliminary Structural Control Results from the Middeck Active Control Experiment (MACE)," AIAA Paper 92-2138, April 1992.

[20]Rey, D. A., "Modelling Gravity and Suspension Effects on Controlled Space Structures," M.S. Thesis, Massachusetts Inst. of Technology, Cambridge, MA, 1992.