AIR FORCE REPORT NO.
SAMSO-TR-72-131

AEROSPACE REPORT N
TR-0172(2441-02)-5

Synthesis of an Optimal Input to Enhance the Identification of a Linear Multistage System

Prepared by P. L. SMITH
Guidance and Control Systems Subdivision
Electronics Division

72 APR 04

Engineering Science Operations
THE AEROSPACE CORPORATION

Prepared for SPACE AND MISSILE SYSTEMS ORGANIZATION
AIR FORCE SYSTEMS COMMAND
LOS ANGELES AIR FORCE STATION
Los Angeles, California

APPROVED FOR PUBLIC RELEASE:
DISTRIBUTION UNLIMITED

34

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* The Aerospace Corporation El Segundo, California	2a REPORT SECURITY CLASSIFICATION Unclassified 2b GROUP

3. REPORT TITLE

SYNTHESIS OF AN OPTIMAL INPUT TO ENHANCE THE IDENTIFICATION OF A LINEAR MULTISTAGE SYSTEM

4 DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5 AUTHOR(S) *(First name, middle initial, last name)*

Patrick L. Smith

6 REPORT DATE 4 April 1972	7a TOTAL NO. OF PAGES 35	7b NO OF REFS 9

8a CONTRACT OR GRANT NO. F04701-71-C-0172	9a ORIGINATOR'S REPORT NUMBER(S) TR-0172(2441-02)-5
b PROJECT NO.	
c	9b OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* SAMSO-TR-72-131
d	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Space and Missile Systems Organization Air Force Systems Command Los Angeles, California

13. ABSTRACT

In fitting a transfer function model to input/output data, it is sometimes possible to select a particular input with the purpose of minimizing estimation error. In this report, some aspects of this optimal input synthesis problem are examined for a class of linear multistage systems. In particular, an input is selected from a class of equal-energy inputs which maximizes the trace of the Fisher information matrix with respect to the input sensitivity parameters, where it is assumed that the transient response and prediction accuracy of the unforced system is known. The optimal input is compared to a purely random input which has the same average energy, and numerical results for a second-order system are provided. Input synthesis with regard to the other model parameters which determine its unforced behavior is hampered by the fact that the Fisher information matrix is a function of these parameters.

Ia

DD FORM 1473 FACSIMILE

UNCLASSIFIED
Security Classification

14 KEY WORDS
Linear Multistage System Fisher Information Matrix Input Synthesis System Identification Kalman Filter Optimization Modeling

Distribution Statement (Continued)

Abstract (Continued)

I-b

Air Force Report No.
SAMSO-72-131

Aerospace Report No.
TR-0172(2441-02)-5

SYNTHESIS OF AN OPTIMAL INPUT TO ENHANCE THE IDENTIFICATION OF A LINEAR MULTISTAGE SYSTEM

Prepared by

P. L. Smith
Guidance and Control Systems Subdivision
Electronics Division

72 APR Ø4

Engineering Science Operations
THE AEROSPACE CORPORATION

Prepared for

SPACE AND MISSILE SYSTEMS ORGANIZATION
AIR FORCE SYSTEMS COMMAND
LOS ANGELES AIR FORCE STATION
Los Angeles, California

FOREWORD

This report is published by The Aerospace Corporation, El Segundo, California, under Air Force Contract No. F04701-71-C-0172.

This report, which documents research carried out from 1 July 1971 to 5 January 1972, was submitted to Major Eric W. Kurgas, SYX, on 20 April 1972 for review and approval.

The author acknowledges with gratitude the assistance of Ronald D. Sugar, Allan R. Stubberud, and Ronald G. Nishinaga.

Approved by

A. J. Schiewe, Director
Guidance and Control Systems Subdivision
Electronics Division
Engineering Science Operations

Publication of this report does not constitute Air Force approval of the report's findings or conclusions. It is published [illegible] for the exchange and stimulation of ideas.

Curtis D. Williams, Capt, USAF
Project Officer

ABSTRACT

In fitting a transfer function model to input/output data, it is sometimes possible to select a particular input with the purpose of minimizing estimation error. In this report, some aspects of this optimal input synthesis problem are examined for a class of linear multistage systems. In particular, an input is selected from a class of equal-energy inputs which maximizes the trace of the Fisher information matrix with respect to the input sensitivity parameters, where it is assumed that the transient response and prediction accuracy of the unforced system is known. The optimal input is compared to a purely random input which has the same average energy, and numerical results for a second-order system are provided. Input synthesis with regard to the other model parameters which determine its unforced behavior is hampered by the fact that the Fisher information matrix is a function of these parameters.

CONTENTS

FIGURES

TABLE

Preceding page blank

NOMENCLATURE

The following is a list of symbols used in the report. Underlined letters denote vectors.

Symbol	Description
A	Matrix of input coefficients
A_j	Partial of A with respect to γ_j
B	Matrix of noise input coefficients
b	Covariance of the steady-state measurement residuals
C	Sample covariance of the unconstrained likelihood function
F	N by N symmetric matrix which is a sum of similarity transforms of Σ
$\underline{h}$	Measurement matrix (a vector for scalar measurements)
I_n	Identity matrix of dimension n by n
J	Cost function
$\underline{k}$	Steady-state Kalman filter gain
$L(\underline{\theta})$	Unconstrained likelihood function
$L'(\underline{\theta})$	Constrained likelihood function
M	The Fisher information matrix
N	Total number of measurements
n	Dimension of $\underline{x}(i)$
$N_n(\underline{m}, Q)$	Multivariant normal distribution of dimension n with mean $\underline{m}$ and covariance Q

NOMENCLATURE (Continued)

Symbol	Description
Q	Covariance of the random input $\underline{r}(i)$
$\underline{r}(i)$	Random input at index i
$\underline{u}$	Composite vector of inputs
$u(i)$	Input at index i
$u^*(i)$	Optimal input at index i
$\tilde{u}(i)$	Random input at index i
$\underline{v}$	Composite vector of measurement residuals
$v(i)$	Measurement residual at index i
$\underline{w}(i)$	Observation vector at index i
$\underline{x}(i)$	State vector at index i
$\hat{\underline{x}}'(i)$	One-step-ahead predicted estimate of the state vector at index i
$\underline{z}$	Composite vector of measurements
$z(i)$	Output at index i
α	Total energy of input
$\underline{\gamma}$	Input sensitivity vector
ϵ	Efficiency of the random input compared to optimal input
$\underline{\theta}$	Unknown parameter vector
λ	Lagrange multiplier
λ_{ave}	Average of eigenvalues of F

NOMENCLATURE (Continued)

Symbol	Description
λ_{max}	Maximum eigenvalue of F
Σ	Covariance of unconstrained likelihood function
Φ	State transition matrix
$\underline{\phi}$	Parameters in state transition matrix

I. INTRODUCTION

Measurements of the input and output of an isolated physical system are made at uniformly spaced instants of time. The sets of N measurements of the input and output are denoted $\{u(i): i = 1, \cdots, N\}$ and $\{z(i): i = 1, \cdots, N\}$, respectively. The measurements are assumed to be error-free. Furthermore, $\{z(i): i = 1, \cdots, N\}$ are assumed to be the output of a multistage time-invariant linear system forced by the inputs $\{u(i): i = 1, \cdots, N\}$ and a postulated sequence of purely random inputs which are assumed to arise due to modeling errors. The multistage model contains unknown parameters which must be identified. The general problem of estimating the unknown parameters of this linear stochastic model from the input/output records is described in [1].

In some applications it is possible to apply a particular input sequence $\{u(i): i = 1, \cdots, N\}$ specifically for the purpose of enhancing the estimates of the unknown model parameters[1]. The basic conditions are:

1) The system to be identified is out of service; that is, one has complete control over the input during the identification phase.

2) The input resource (energy) is scarce.

If there is no penalty or constraint on the input, then optimal input synthesis is unnecessary. This is because any input which excites all the

[1] This problem is the dynamic analog to experimental design in statistics [2].

modes of the system, if applied long enough, is generally sufficient to allow convergence of the identification. There are several systems for which there is a time, magnitude, or energy constraint on the input. One example is a structure where the input is a thrust profile from a rocket engine which has a finite supply of fuel. Another example is process control where the amount of time spent in the parameter identification phase can be very costly.

The problem of selecting input signals to enhance the system identification process has been examined by several authors. Aoki and Staley [3] have obtained some general results for a wide class of discrete-time systems as well as justifying the use of the trace of the Fisher information matrix as a performance index. This paper retains the use of this convenient performance index, but formulates the problem differently by using the Kalman filter representation of th system. Levadi [4] considers identifying non-stationary, continuous systems with optimum energy-constrained and time-constrained input signals. The only source of noise considered, however, is output observation noise. Gagliardi [5] attacks the problem in a completely different manner by using multiple hypothesis testing.

In this paper a Kalman filter representation results in a considerable simplification of the problem. It is assumed that the input sequence is selected from a class of possible input sequences with equal total energy or equal total average energy if the input is random. This assumption attempts to bound the output of the system to prevent forcing it into a region where it is no longer linear. The optimal input is the sequence in this class which minimizes the trace of the Fisher information matrix. Some investigators have suggested the use of a purely random input to drive the system during the identification phase [6]. A comparative efficiency between a purely random input and the optimal input for the class of inputs studied is defined, and some numerical results for a second-order model are presented.

It should be emphasized at the outset that there is a fundamental limitation on the solution of the general optimal input synthesis problem: the Fisher information matrix is generally a function of the unknown parameters. Hence, one is faced with the impossible situation of requiring the true values of the unknown parameters in order to find the optimal input. In the general case, therefore, any physically realizable efficient input synthesis technique must be adaptive -- as better estimates of the parameters are obtained, the input is made to approach the appropriate optimal input. There is an exception in linear systems which is examined in this report: the Fisher information matrix is not a function of the input parameters.

II. THE MODEL

A linear multistage model which relates the input $\{u(i): i = 1, \ldots, N\}$ to the output $\{z(i): i = 1, \ldots, N\}$ is

$$\underline{x}(i + 1) = \Phi\underline{x}(i) + \underline{\gamma}\, u(i) + \underline{r}(i) \tag{1}$$

$$z(i) = \underline{h}^T\, \underline{x}(i) \tag{2}$$

where $\{\underline{r}(i): i = 1, \ldots, N\}$ is a set of independent vector-valued random variables each of which are distributed $N_n(\underline{0}, Q)$, where Q is a diagonal matrix with non-negative elements. The initial value $\underline{x}(i)$ is assumed to be $\underline{0}$. A canonical form of Eqs. (1) and (2) utilizes the matrix forms

$$\Phi = \left[\begin{array}{c|c} \underline{0} & I_{n-1} \\ \hline \multicolumn{2}{c}{-\underline{\phi}^T} \end{array}\right] \qquad \underline{h} = \left[\begin{array}{c} 1 \\ \hline \underline{0} \end{array}\right]$$

The unmeasured states $\{x_j(i): i = 1, \ldots, N; j = 2, \ldots, n\}$ can be estimated by the corresponding steady-state Kalman one-step-ahead predictor [7] given by

$$\hat{\underline{x}}'(i + 1) = \Phi\hat{\underline{x}}'(i) + \Phi\underline{k}v(i) + \underline{\gamma}u(i) \tag{3}$$

$$z(i) = \underline{h}^T\, \hat{\underline{x}}'(i) + v(i) \tag{4}$$

Preceding page blank

$\underline{k}$ is the steady-state Kalman filter gain. It is straightforward to estimate the unknown model parameters and $\underline{k}$ from the filter residuals [1]. If one assumes steady-state conditions, the measurement residuals $\{v(i): i = 1, \ldots, N\}$ are independent random variables each of which is distributed $N(0, b)$ as a result of the innovations property of the Kalman filter [8]. Equations (3) and (4) can also be used to represent the relationship between the input $\{u(i): i = 1, \ldots, N\}$ and the output $\{z(i): i = 1, \ldots, N\}$, since they are statistically equivalent [7].

III. DERIVATION OF AN UNCONSTRAINED LIKELIHOOD FUNCTION

The likelihood function which is most convenient for computing the maximum-likelihood estimates of the unknown model parameters is

$$L'(\underline{\theta}) = -\frac{N}{2}\log b - \frac{1}{2b}\sum_{i=1}^{N} v^2(i) \qquad (5)$$

where $\{v(i):\ i = 1, \ldots, N\}$ are constrained to satisfy Eq. (3), and where $\underline{\theta}$ is the vector of the unknown model parameters [1]. The constraint on the measurement residuals is cumbersome to work with analytically, so an equivalent likelihood function is derived in this section from the distribution of the measurements $\{z(i):\ i = 1, \ldots, N\}$, which do not have auxiliary constraints.

First, Eq. (3) is solved in closed form:

$$\hat{\underline{x}}'(n) = \sum_{j=1}^{n-1} \Phi^{n-1-j}\, \underline{\gamma}\, u(j) + \sum_{j=1}^{n-1} \Phi^{n-j}\, \underline{k}\, v(j) \qquad (6)$$

Since the measurements are assumed to be error-free, it follows that

$$z(n) = \underline{h}^T \left(\hat{\underline{x}}'(n) + \underline{k}\, v(n)\right) \qquad (7)$$

Substitute Eq. (6) into Eq. (7) to obtain a closed form expression for z(n):

$$z(n) = \sum_{j=1}^{n-1} \underline{h}^T \Phi^{n-1-j} \underline{\gamma}\, u(j) + \sum_{j=1}^{n} \underline{h}^T \Phi^{n-j} \underline{k}\, v(j) \tag{8}$$

Define the following composite N vectors:

$$\underline{z} \triangleq \begin{pmatrix} z(1) \\ \vdots \\ z(N) \end{pmatrix} \qquad \underline{u} \triangleq \begin{pmatrix} u(1) \\ \vdots \\ u(N) \end{pmatrix} \qquad \underline{v} \triangleq \begin{pmatrix} v(1) \\ \vdots \\ v(N) \end{pmatrix}$$

It follows from Eq. (8) that

$$\underline{z} = A\, \underline{u} + B\, \underline{v} \tag{9}$$

where A and B are lower triangular N by N matrices defined as

$$a_{\ell m} = \begin{cases} 0 & ; \quad \ell \le m \\ \underline{h}^T \Phi^{\ell-1-m} \underline{\gamma} & ; \quad \ell > m \end{cases} \tag{10}$$

$$b_{\ell m} = \begin{cases} 0 & ; \quad \ell < m \\ 1 & ; \quad \ell = m \\ \underline{h}^T \Phi^{\ell-m} \underline{k} & ; \quad \ell > m \end{cases} \tag{11}$$

Equation (9) shows that the measurements $\underline{z}$ are linear combinations of the inputs $\underline{u}$ and the measurement residuals $\underline{v}$ (which are assumed to result from modeling errors), assuming zero initial conditions.

To simplify the notation, define the n vector,

$$\underline{w}(n) \triangleq (\Phi^T)^n \underline{h} \tag{12}$$

and substitute it into Eqs. (10) and (11):

$$a_{\ell m} = \begin{cases} 0 & ; \quad \ell \le m \\ \underline{\gamma}^T \underline{w}(\ell - 1 - m) & ; \quad \ell > m \end{cases} \tag{13}$$

$$b_{\ell m} = \begin{cases} 0 & ; \quad \ell < m \\ 1 & ; \quad \ell = m \\ \underline{k}^T \underline{w}(\ell - m) & ; \quad \ell > m \end{cases} \tag{14}$$

From the innovations property $\underline{v}$ is distributed $N_N(\underline{0}, bI)$ and so by Eq. (9) $\underline{z}$ is distributed $N_N(A\underline{u}, bB\,B^T)$. The likelihood function $L(\underline{\theta})$ for the multivariate normal distribution is a standard result in statistics [2]:

$$L(\underline{\theta}) = -\frac{1}{2} \log |\Sigma| - \frac{1}{2} \mathrm{Tr}\, [\Sigma^{-1} C]$$

where

$$\Sigma \triangleq b\, B\, B^T$$

$$C \triangleq (\underline{z} - A\,\underline{u})\,(\underline{z} - A\,\underline{u})^T \tag{15}$$

IV. AN INPUT SYNTHESIS PROBLEM

In the most general modeling problems the values of $\underline{\phi}$, $\underline{k}$, $\underline{\gamma}$, and b are uncertain. For the remainder of this section, however, it is assumed that the transient response of the system ($\underline{\phi}$) and the one-step-ahead prediction accuracy of the unforced model ($\underline{k}$) are known or have been previously identified, but that the input sensitivity ($\underline{\gamma}$) is unknown. The true value of b is not required in this case. It is also assumed that the input $\underline{u}$ is to be selected from a class of inputs such that

$$\underline{u}^T \underline{u} = \alpha \tag{16}$$

where α is the total input energy. The question is: what form of the constrained input $\underline{u}$ will yield the most "information about $\underline{\gamma}$"? The measure of "information about $\underline{\gamma}$" is defined as the trace of the Fisher information matrix [3].

The Fisher information matrix is a measure of the dispersion of the gradient of the likelihood function with respect to the unknown parameters. At one extreme, if there is a one-to-one correspondence between the values of the unknown parameters and the measurements, then the information is a maximum. At the other extreme, if there is no correspondence, the information matrix is zero. See [3] for additional discussion.

The Fisher information matrix (denoted by M) for a multivariate normal distribution is a standard result in statistics [9].

Preceding page blank

The elements of M are

$$m_{ij} = E\left[\frac{\partial L(\underline{\gamma})}{\partial \gamma_i} \cdot \frac{\partial L(\underline{\gamma})}{\partial \gamma_j}\right] = \underline{u}^T A_i^T \Sigma^{-1} A_j \underline{u} \tag{17}$$

The elements of A_j are

$$\frac{\partial a_{\ell m}}{\partial \gamma_j} = \begin{cases} 0 & ; \quad \ell \le m \\ w_j(n-1-m) & ; \quad \ell > m \end{cases} \tag{18}$$

M is not a function of $\underline{\gamma}$, and so it is possible to determine M prior to estimating $\underline{\gamma}$. For the more general case where $\underline{\phi}$, $\underline{k}$, and b are also unknown, the Fisher information matrix is a function of ϕ, $\underline{k}$, and b, which reduces the value of the Fisher information matrix as a prior information measure. The expression for Tr[M] follows immediately from Eq. (17):

$$\mathrm{Tr}[M] = \frac{1}{b} \underline{u}^T F \underline{u} \tag{19}$$

where

$$F \triangleq \sum_{i=1}^{n} A_i^T (B B^T)^{-1} A_i \tag{20}$$

To proceed, define a scalar cost functional J which is proportional to Tr[M] and adjoin Eq. (17) with a Lagrange multiplier:

$$J = \underline{u}^T F \underline{u} + \lambda (\underline{u}^T \underline{u} - \alpha) \tag{21}$$

The gradient of J is

$$\nabla_{\underline{u}} J = F \underline{u} - \lambda \underline{u} \tag{22}$$

Setting the gradient to zero, the result is

$$F \underline{u} = \lambda \underline{u} \tag{23}$$

It follows from Eq. (23) that the input sequence which maximizes J is the eigenvector of F (denoted $\underline{u}^*$) which corresponds to the largest eigenvalue of F (denoted λ_{max}). It also follows from Eq. (23) that

$$\max \mathrm{Tr}[M] = \frac{\lambda_{max} \alpha}{b} \tag{24}$$

Note that neither J nor $\underline{u}^*$ is a function of b.

V. COMPARISON OF THE OPTIMAL INPUT WITH AN EQUIVALENT RANDOM INPUT

Some investigators have used random inputs to drive the system because a random input is most likely to excite all the modes of the system and is very easy to generate [6]. In this section, a purely random input is compared (in a method defined below) with the optimal input derived in the last section. First, let $\{\tilde{u}(i): i = 1, \ldots, N\}$ denote an input sequence which is a set of independent random numbers with mean zero, such that

$$E[\tilde{u}^2(i)] = \frac{\alpha}{N} \tag{25}$$

Therefore $\underline{\tilde{u}}$ has an average energy equal to α. Furthermore, suppose that $\{\tilde{u}(i): i = 1, \ldots, N\}$ are independent of the measurement residuals $\{v(i): i = 1, \ldots, N\}$. The <u>efficiency</u> of the random input is defined as the ratio

$$\epsilon = \frac{\overline{Tr[M]}}{\max\limits_{\underline{u}} Tr[M]} \tag{26}$$

where the overbar denotes the expected value of $Tr[M]$ using the random input. Since $\underline{\tilde{u}}$ is assumed to be independent of $\underline{v}$, it follows that

$$E\left[Tr[M]\right] = Tr\left[E[M]\right] = \frac{1}{b}E[\underline{\tilde{u}}^T F \underline{\tilde{u}}] = \frac{\alpha}{bN} Tr[F] = \frac{\lambda_{ave}\alpha}{b} \tag{27}$$

Preceding page blank

where λ_{ave} denotes the average of the eigenvalues of F. It follows from Eq. (24) that the efficiency of the random input compared to the optimal input is

$$\epsilon = \frac{\lambda_{ave}}{\lambda_{max}} \tag{28}$$

VI. EXAMPLE

The purpose of this example is two-fold: first, to illustrate the theory developed in this report and, second, to try to obtain some insight into the form of the optimal input for a certain class of identification problems. The following Kalman filter representation of a second-order system is examined:

$$\underline{\hat{x}}'(i+1) = \begin{bmatrix} 0 & 1 \\ -\phi_1 & -\phi_2 \end{bmatrix} \underline{\hat{x}}'(i) + \begin{bmatrix} 0 & 1 \\ -\phi_1 & -\phi_2 \end{bmatrix} \begin{bmatrix} 1 \\ k_2 \end{bmatrix} v(i) + \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} u(i) \qquad (29)$$

To repeat, the objective is to generate the input which most enhances the estimates of γ_1 and γ_2 from the input/output data.

The first step is to compute the elements of A_j and B, and then F.

A recursive equation for $\underline{w}(i)$ which greatly simplifies the calculation of the elements of A_j and B is derived next. By the Cayley-Hamilton theorem, Φ^T must satisfy its own characteristic equation, so it follows that

$$(\Phi^T)^2 = -\phi_2(\Phi^T) - \phi_1 I_2 \qquad (30)$$

Postmultiply by $\underline{h}$ and premultiply by $(\Phi^T)^{n-2}$ yields

$$(\Phi^T)^n \underline{h} = -\phi_2(\Phi^T)^{n-1}\underline{h} - \phi_1(\Phi^T)^{n-2}\underline{h} \qquad (31)$$

It follows from the definition of $\underline{w}(n)$ that

$$w(n) = -\phi_2 \underline{w}(n-1) - \phi_1 \underline{w}(n-2) \qquad (32)$$

for $n > 2$.

The second step is to compute the eigenvectors of F, the one corresponding to the largest eigenvalue of F being the optimal input.

For the numerical examples, the model parameters were chosen to simulate an underdamped and overdamped second-order system. The optimal input sequences were computed for the model parameters listed in Table I for N = 20 using a standard computer subroutine to find the eigenvalues and eigenvectors of F. The optimal inputs along with the system's unit response are graphed against the index i in Figs. 1 to 4.

TABLE I

PARAMETERS USED IN THE EXAMPLES

System Characterization	ϕ_1	ϕ_2	k_2	ϵ	Figure No.
Overdamped	-0.50	-0.25	0.25	0.05	1
	-0.50	-0.25	0.50	0.07	2
Underdamped	0.50	-0.25	0.25	0.37	3
	0.50	-0.25	0.50	0.21	4

The following discussion describes the physical significance of the optimal input. In all the cases where $k_2 = -\phi_2$ [no prediction error in $x_1(i)$], the majority of the input energy is spent in the first third of the interval (Figs. 1 and 3). This is apparently an effort to get the system moving as soon as possible. In cases where $k_2 > -\phi_2$ [prediction error in $x_1(i)$ as well as $x_2(i)$] the input energy is more dispersed throughout the interval (Figs. 2 and 4). This dispersion of the input energy appears to be a hedge

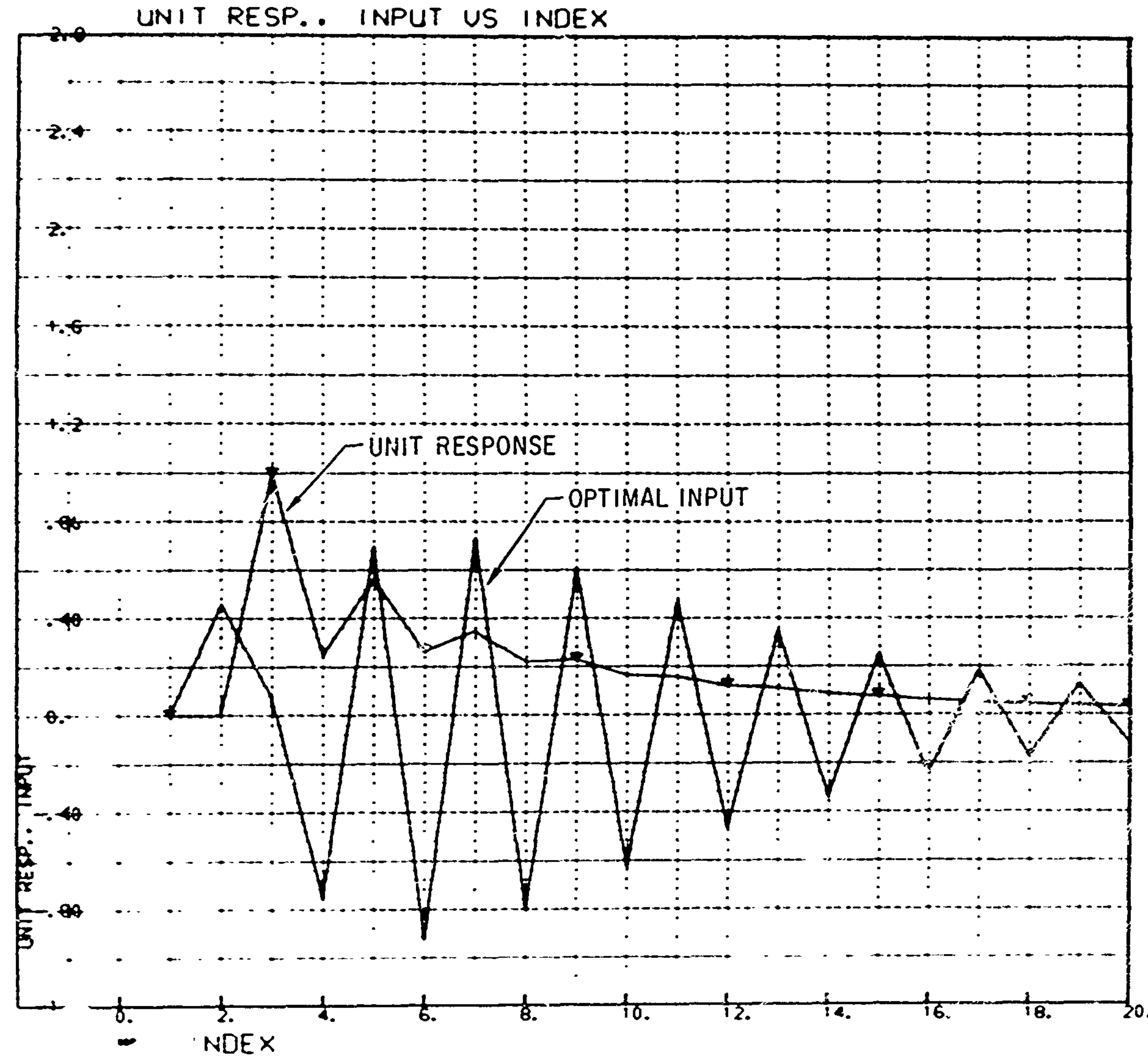

Fig. 1. Overdamped with k = 0.25.

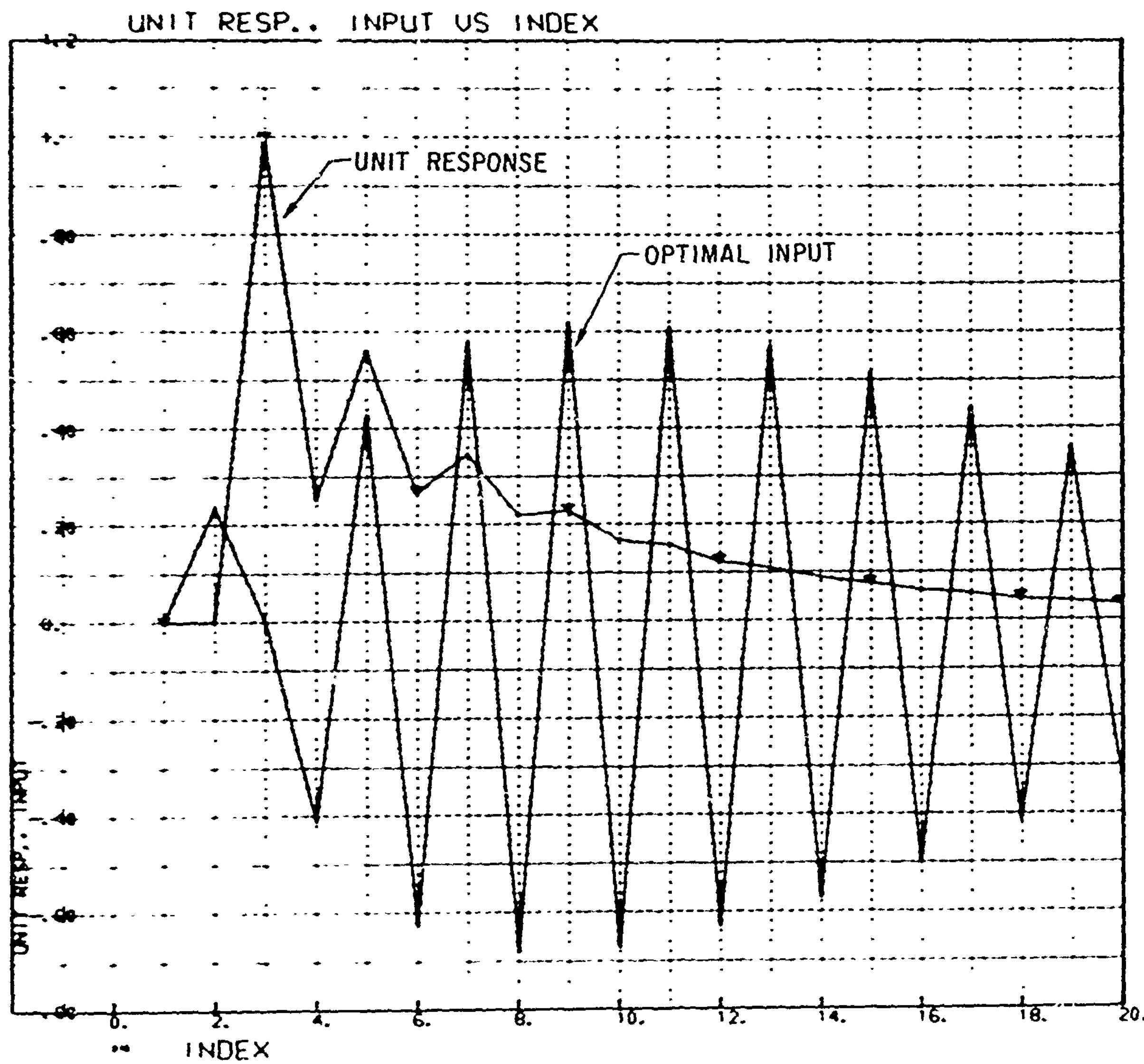

Fig. 2. Overdamped with k = 0.50.

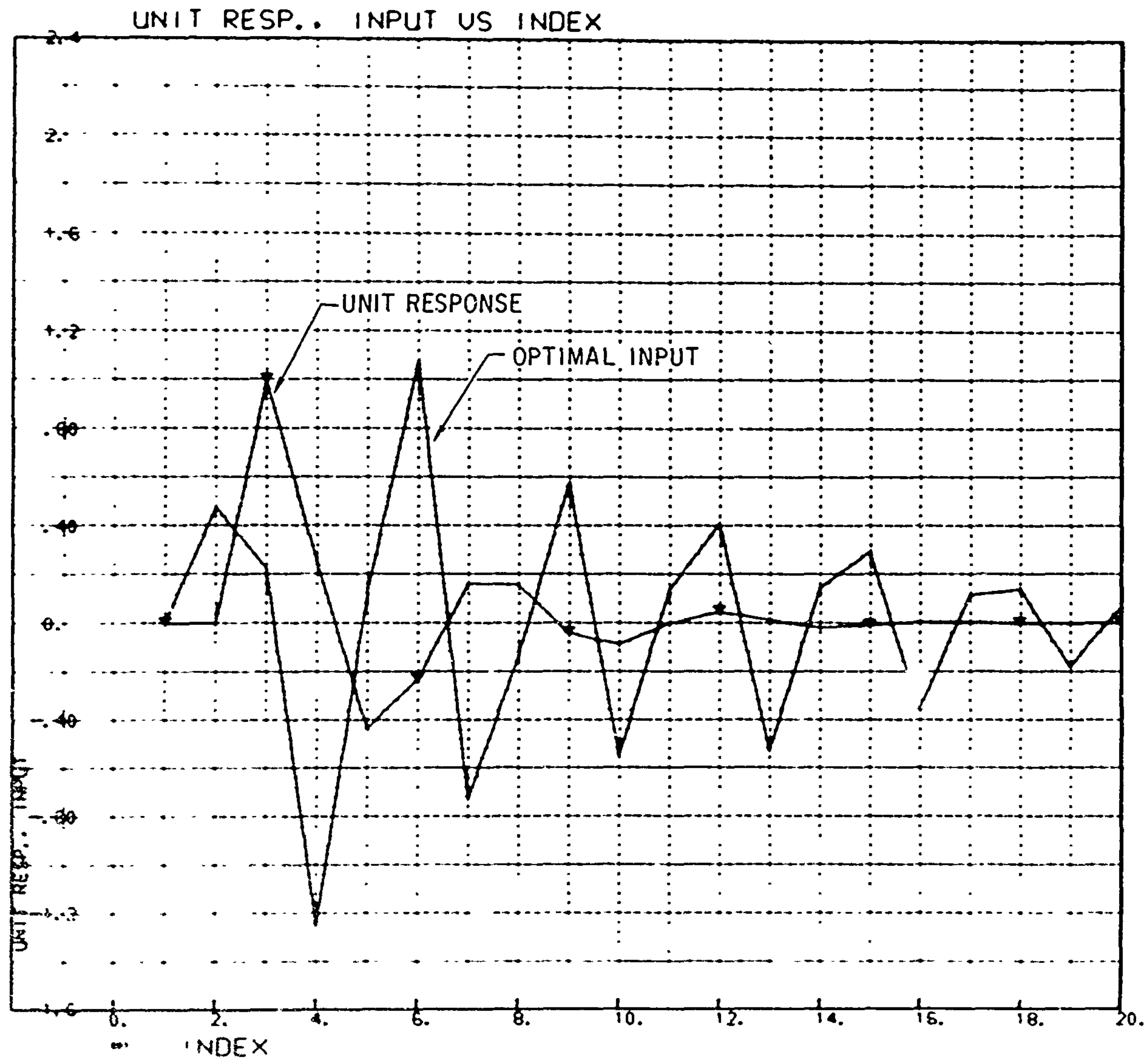

Fig. 3. Underdamped with k = 0.25.

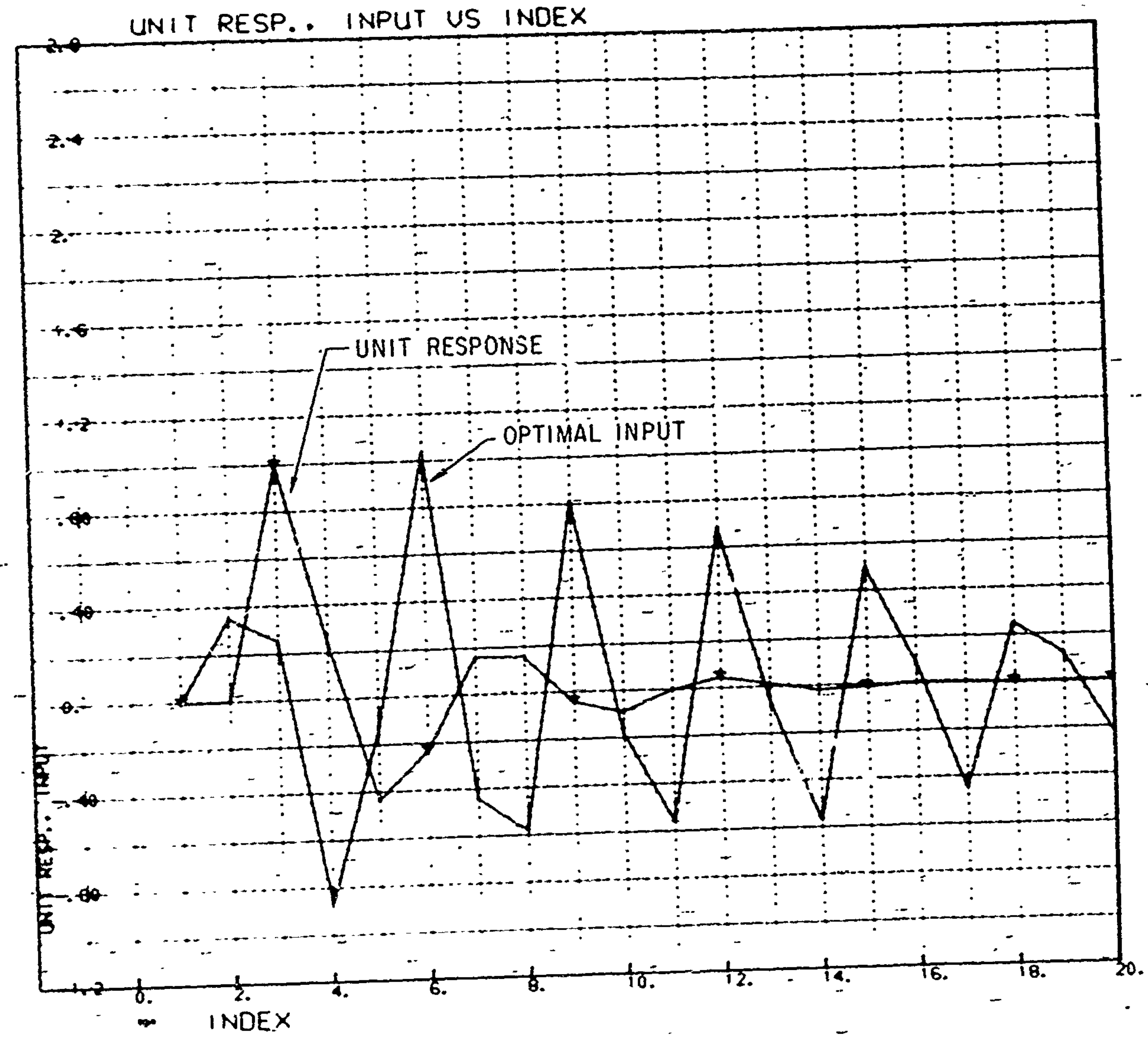

Fig. 4. Underdamped with k = 0.50.

against the poorer prediction accuracy of the plant (reflected in the larger value of k_2). The purely random input is better with faster systems than with slower ones, but the efficiency was never observed to be over .50 in any of the cases tried.

The majority of the energy of the input appears to be in frequencies higher than that contained in the unit response of the system, but not so high as to be greatly attenuated by the system. At least this appears to be one of the factors taken into account by the optimal input. It also explains why the random input is better for faster systems than for slower ones: more of the energy in the random input is attenuated by the slower system.

The analysis and the numerical results described in this example are useful for the situation where one has fairly accurate prior knowledge of the transient response and one-step-ahead prediction accuracy of the unforced systems, but no prior knowledge of the input sensitivity parameters. In this case, one can determine an input test signal which enhances the estimates of the input sensitivity parameters.

VII. INPUT SYNTHESIS FOR THE GENERAL CASE

In general, the parameters $\underline{\phi}$, $\underline{k}$, and $\underline{\gamma}$ must be identified. The residual covariance b can be estimated from the measurement residuals after the identification of the other parameters is complete [1]. Define $\underline{\theta}$ as the $3n \times 1$ vector of unknown model parameters:

$$\theta \triangleq \begin{pmatrix} \underline{\phi} \\ \underline{\gamma} \\ \underline{k} \end{pmatrix}$$

The elements of the Fisher information matrix in the general case are

$$m_{ij} = \frac{1}{2} \mathrm{Tr}\left[\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j}\right] + \underline{u}^T \frac{\partial A^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial A}{\partial \theta_j} \underline{u} \tag{33}$$

The expression for Tr[M] follows immediately:

$$\mathrm{Tr}[M] = \frac{1}{2} \sum_{i=1}^{3n} \mathrm{Tr}\left[\left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i}\right)^2\right] + \frac{1}{b} \underline{u} F \underline{u}^T \tag{34}$$

where

$$F \triangleq \sum_{i=1}^{3n} \frac{\partial A^T}{\partial \theta_i} (BB^T)^{-1} \frac{\partial A}{\partial \theta_i}$$

Preceding page blank

From Eq. (10) it follows that

$$\frac{\partial A}{\partial k_i} = 0$$

which shows that the input does not directly affect the identification of $\underline{k}$. Since B is a function of $\underline{k}$, the optimal input will depend on $\underline{k}$, but it is not possible to enhance the identification of $\underline{k}$ by manipulating the input. This makes sense because it is well known that the Kalman gain is independent of the input.

The problem one faces in maximizing Tr[M] with respect to $\underline{u}$ is that $\partial A/\partial\theta_i$ is a function of $\underline{\phi}$ and B is a function of $\underline{\phi}$ and $\underline{k}$ which are unknown. One way to synthesize an input is to use the a priori values of $\underline{\phi}$ and $\underline{k}$ in the optimization. If there is time during the identification process, the input can be reoptimized using the partially identified values of $\underline{\phi}$ and $\underline{k}$ to replace the a priori values.

VIII. CONCLUSION

An approach to optimal input synthesis for linear system identification has been presented. The significant result is: the optimal input is an eigenvector of a matrix related to the Fisher information matrix and in general it is a function of the unidentified system parameters.

REFERENCES

[1] P. L. Smith, "Fitting Linear Dynamic Models to Input/Output Data," The Aerospace Corporation Technical Report. To be published.

[2] M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, Vols. 2 and 3. New York: Hafner Publishing Co., 1961.

[3] M. Aoki and R. M. Staley, "On Input Signal Synthesis in Parameter Identification," Automatica, Vol. 6, pp. 431-440. Pergamon Press, 1970.

[4] V. S. Levadi, "Design of Input Signals for Parameter Estimation," IEEE Trans. Auto. Control AE-11, No. 2, April 1966.

[5] R. M. Gagliardi, "Input Selections for Parameter Identification in Discrete Systems," IEEE Trans. Auto. Control AC-12, No. 5, October 1967.

[6] G. E. P. Box and D. M. Jenkins, "Time Series Analysis, Forecasting and Control," Identification, Fitting, and Checking of Transfer Function Models, Chapter 11. San Francisco: Holden-Day, 1970.

[7] R. K. Mehra, "Identification of Linear Dynamic Systems." Paper presented at the IEEE Symposium on Adaptive Processes (8th), Pennsylvania, 17-19 November 1969.

[8] T. Kailath, "An Innovation Approach to Least Squares Estimation, Part I," IEEE Trans. Auto. Control 13:646-655, December 1968.

[9] P. L. Smith, "Estimation of Covariance Parameters in Time Discrete Linear Systems with Applications to Adaptive Filtering," The Aerospace Corporation Report No. TOR-0059(6311)-23, 1971.