NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

ADAPTIVE MODELING OF THE DYNAMICS

OF AUTONOMOUS LAND VEHICLES

by

James H. Kessler IV

June 1986

Thesis Advisor: Robert B. McGhee

Approved for public release; distribution is unlimited.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School	6b. OFFICE SYMBOL (If applicable) 62
6c. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5100	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)
8c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5100
	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)

ADAPTIVE MODELING OF THE DYNAMICS OF AUTONOMOUS LAND VEHICLES

12. PERSONAL AUTHOR(S)
Gessler, James, H

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Engineer's Thesis	FROM _____ TO _____	June 1986	119

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Parameter identification with unknown inputs, parameter estimation of land vehicles, comparison of identification techniques.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This study examines the feasibility of identifying the dynamic parameters of autonomous land vehicles. Initially a simulation study is done using a simplified vehicle model and three different identification schemes. The parameter estimator which has the best characteristics is then used in an improved model to determine if this approach can be used to obtain the parameters of muli-wheeled vehicles.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Prof. Robert B. McGhee	22b. TELEPHONE (Include Area Code) (408) 646- 2095
	22c. OFFICE SYMBOL 52Mz

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

Adaptive Modeling of the Dynamics of Autonomous Land Vehicles

by

James H. Kessler, IV
Captain. United States Marine Corps
B.S., Rensselaer Polytechnic Institute, 1977

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING
and
ELECTRICAL ENGINEER

from the

NAVAL POSTGRADUATE SCHOOL
June 1986

ABSTRACT

This study examines the feasibility of identifying the dynamic parameters of autonomous land vehicles. Initially a simulation study is done using a simplified vehicle model and three different identification schemes. The parameter estimator which has the best characteristics is then used on an improved model to determine if this approach can be used to obtain the parameters of multi-wheeled vehicles.

TABLE OF CONTENTS

I. INTRODUCTION 7

A. INTRODUCTION 7

B. PROBLEM STATEMENT 7

C. APPROACH 8

D. LIMITATIONS 8

E. STARTING POINT 9

F. ORGANIZATION 9

G. NOTATION 9

II. SURVEY OF PREVIOUS WORK 11

A. INTRODUCTION 11

B. LEAST-SQUARES APPROXIMATION 11

C. KALMAN FILTER IDENTIFIER 16

D. STOCHASTIC GRADIENT ESTIMATION 19

E. SUMMARY 24

III. LEAST-SQUARES IDENTIFIER 26

A. INTRODUCTION 26

B. SIMPLIFIED MODEL FOR SIMULATION STUDY 27

C. INPUT DISTURBANCE AND INITIAL CONDITIONS 31

D. STATE MEASUREMENT NOISE 38

E. STATE DERIVATIVE MEASUREMENT NOISE 38

F. COMBINED NOISE 40

G. COMPUTATIONAL AND STORAGE REQUIREMENTS 43

H. SUMMARY 44

IV. KALMAN FILTER IDENTIFIER 46

A. INTRODUCTION 46

B. INITIALIZATION OF THE KALMAN FILTER 47

C. SYSTEM EXCITATION 51

D. STATE MEASUREMENT NOISE 57

E. STATE DERIVATIVE MEASUREMENT NOISE 65

F. COMBINED NOISE 67

G. INSTRUMENTATION 70

H. COMPUTATIONAL AND STORAGE REQUIREMENTS 71

I. SUMMARY 72

V. STOCHASTIC GRADIENT IDENTIFIER 74

A. INTRODUCTION 74

B. BIAS ELIMINATION 74

C. CHOICE OF THE GAIN MATRIX R(k) 75

D. INITIALIZATION OF R(k) 77

E. SUMMARY 84

VI. IMPROVED MODEL IDENTIFICATION 87

A. INTRODUCTION 87

B. AN IMPROVED MODEL 87

C. IMPROVED VEHICLE PARAMETER IDENTIFICATION 92

1. P(0) Initialization 93

2. Convergence Rate of the Improved Model 95

3. Initial Conditions and Road Noise 97

4. System Identification with Pulse Inputs 99

5. Identification with State Measurement Noise 99

D. TWO INPUTS TO THE IMPROVED MODEL 102

E. COMPUTATIONAL AND STORAGE REQUIREMENTS 106

F. SUMMARY 109

VII. SUMMARY AND CONCLUSIONS 111

A. INTRODUCTION 111

B. RESULTS 111

C. FURTHER STUDY 114

LIST OF REFERENCES 116

I. INTRODUCTION

A. INTRODUCTION

In recent years the concept of autonomous vehicles has become a reality. Martin Marietta Corp. has built an autonomous wheeled vehicle [Ref. 1]. FMC Corporation is developing a tracked autonomous vehicle [Ref. 2]. Ohio State University is constructing a six legged vehicle [Ref. 3].

As autonomous vehicles progress it will become necessary for them to identify their own dynamic parameters in order to operate effectively. This will allow the vehicle to identify its mass and other characteristics and determine if it can operate within the envelope for which it was designed. This would also enable the vehicle to continuously monitor its parameters and identify maintenance problems or other potential problems and take the necessary steps to prevent damage to itself. This capability would give the vehicle what all natural autonomous systems have, the ability to "feel" their own dynamic parameters.

B. PROBLEM STATEMENT

The objective of this thesis is to determine the feasibility of using parameter identification techniques to identify the dynamic parameters of an autonomous vehicle, on line, and to monitor variations in those parameters during operation. The term on-line in this case means, while operating normally. The desire here is

to allow the vehicle to passively identify its parameters without the need to actively participate in developing inputs specifically selected for identification.

In this thesis, continuous time parameters will be identified. This will increase the hardware requirements, but will allow for certain other desirable features required for multi-wheeled vehicles. Thus, the problem of this thesis is to passively identify and track the continuous time dynamic parameters of a wheeled vehicle.

C. APPROACH

The number of different parameter identification techniques available is extensive, but many of the approaches are similar. Two major distinction are, the ability to identify nonlinear systems, and to allow stochastic inputs. The work of this thesis is directed toward problems solvable by stochastic linear identifiers. Here, three approaches will be considered in detail. First each technique will be presented mathematically and then used in a simulation study to analyze their respective characteristics. A simplified vehicle model will be used for this purpose. The scheme found to have the most desirable properties will then be used to identify the parameters of a more realistic model. This will determine if the approach is valid for the identification of wheeled vehicles. From these results, the feasibility of identifying and tracking the vehicle parameters will be considered.

D. LIMITATIONS

In the development of a simulation study some assumptions must necessarily be made. In this thesis, it will be assumed that the input from the road is white noise. In addition, it will be assumed that position, velocity, and acceleration are available for the vehicle's body and all of its wheels.

E. STARTING POINT

The field of parameter identification is well developed. The standard approach to parameter estimation is with respect to scalar discrete time systems. In this analysis a multi-input multi-output continuous-time parameter identifier is used. This will require certain changes to the standard algorithms used in the literature.

F. ORGANIZATION

The mathematical development of the three identification schemes will be presented in the next chapter. In the following three chapters, each approach will be analyzed using a simulation study on a simplified model that is presented in Chapter III. In Chapter VI, the technique which demonstrated the best characteristic during the simulation study will be used on an improved model developed in that chapter. Chapter VII will present the concluding remarks concerning the feasibility of this from of parameter identification when used on land vehicles.

G. NOTATION

The notational convention used in this thesis conforms to the standard used in the literature. Capital letter denote matrices, as in P(k), where P is a discrete-time varying matrix. Lower case letter denote scalars or vectors. In most cases, this will be clear from the context and the dimensionality of the equation. Thus, v(t) is a continuous time variable, and in this case would be described in the text as a vector of dimension n. The transpose of a vector or matrix is denoted by a superscript T. An element of either a vector or matrix is shown with a subscripted number. For example, the element of the third row, second column of the A matrix is shown by a_{32}. If a matrix is made by concatenating a sequence of matrices, then it is bold faced. This is also true for vectors, and therefore the vector $\mathbf{v}(t)$ is made by concatenating several v(t) vectors together. This notation sets the stage for the next chapter where each technique is mathematically presented.

II. SURVEY OF PREVIOUS WORK

A. INTRODUCTION

Parameter identification has received much attention, and is frequently used in modern aerospace vehicle control[Ref. 4:p. 32]. There are numerous approaches to this identification problem, only a few of which will be mentioned here. Many of these techniques have several names and can be applied in different situations with only slight changes to the algorithm. This discussion will be confined to three well-defined approaches that have been used on several different types of problems with good results. These techniques are: least-squares approximation, Kalman filter identifiers or minimum-variance estimation, and stochastic gradient parameter estimation or stochastic approximation. In this chapter, each technique for parameter identification will be presented. This will establish the mathematical notation to be used in the rest of this thesis and will make the limitations of each approach more transparent. In the following chapters each scheme will then be presented in the context of a simplified vehicle parameter identification simulation, which will allow comparison of the various techniques.

B. LEAST-SQUARES APPROXIMATION

The most straightforward parameter identification technique is the least-squares approximation approach. Thus, its mathematical development provides a

good starting point for the discussion of this chapter. Mendel [Ref. 4:pp. 5-20] has established a notation, particularly suited to the analysis of different parameter identifiers and his notation shall, as much as possible, be used here. Following this notation, let the linear system with unknown constant parameters be defined by

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2.1}$$

where the elements of the A and B matrices are unknown, but the state vector x(t), of dimension n, is available either by direct measurement or signal estimation. The input vector u(t) will for the present be 0. This continuous time system is then measured at discrete time intervals t_k, k=0,1,2.. . . . The discrete time measurements produce a state vector $x(t_k)$ and its derivative $\dot{x}(t_k)$ at time t_k. These vectors will be denoted as $x(k)$ and $\dot{x}(k)$ for notational convenience.

A word of caution is in order here, for the techniques presented in this thesis do not conform to the standard usage of discrete time systems. Here, the use of the state vector and its derivative will produce an asynchronous identifier. The asynchronous identifier is not subject to the same timing constraints as its synchronous counterpart. In the asynchronous case, the continuous time parameters can be obtained with no restriction on the time between samples. This arbitrary selection of the sample interval will be important in the elimination of input-output crosscorrelation in Chapter VI. The derivations presented here are inspired by, but not identical to the work of Mendel, who describes the single-input single-output discrete time identifiers discussed here. As a specific example,

consider the identification of a system with a bandwidth of 100 rad/sec. In order to preserve the output signal and prevent aliasing, such a signal would normally be sampled at perhaps 3 times its highest frequency, or 48 samples per second. This would require the processing of one sample every 0.02 seconds. For some identification schemes this may not be a limitation, but for many recursive algorithms this would be a constraint. In the asynchronous identifier proposed here there are no timing constraints, except those imposed on data acquisition. The requirement on data acquisition is simply that the measurement of $x(k)$ and $\dot{x}(k)$ must be taken at the same instant. This requirement is quite easily handled with a few data buffers and analog to digital converters.

If the parameter vector

$$\theta = (\, a_{11},\, a_{12},\, \ldots,\, a_{nn-1},\, a_{nn})^{T} \tag{2.2}$$

is formed from the matrix A, and if the matrix H(k) is defined as

$$H(k)=\begin{pmatrix} x_1(k)x_2(k) & \cdots & x_n(k) & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & x_1(k)x_2(k) & \cdots & x_n(k) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & & & & & & & \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & x_1(k) & \cdots & x_{n-1}(k)x_n(k) \end{pmatrix} \tag{2.3}$$

then the system can be described using the parameter vector θ as

$$\dot{x}(k) = H(k)\theta \tag{2.4}$$

Now if x and $\dot{x}$ can be measured perfectly, then in n measurements of the state vector and its derivative, θ can be determined from

$$\begin{pmatrix} \dot{x}(n-1) \\ \dot{x}(n-2) \\ \vdots \\ \dot{x}(0) \end{pmatrix} = \begin{pmatrix} H(n-1) \\ H(n-2) \\ \vdots \\ H(0) \end{pmatrix} \theta \tag{2.5}$$

Specifically, if

$$\dot{\mathbf{x}} = (\dot{x}^T(n-1), \dot{x}^T(n-2), \ldots, \dot{x}^T(0))^T \tag{2.6}$$

and

$$\mathbf{H} = (H^T(n-1), H^T(n-2), \ldots, H^T(0))^T \tag{2.7}$$

then

$$\theta = \mathbf{H}^{-1}\dot{\mathbf{x}} \tag{2.8}$$

This case corresponds to the minimum number of measurements needed to obtain a unique solution for θ.

However, if $\dot{x}$ is contaminated with noise then the measured quantity will be

$$z = \dot{x} + v \tag{2.9}$$

or

$$\mathbf{z} = \mathbf{H}\theta + \mathbf{v} \tag{2.10}$$

The vectors $\mathbf{v}$ and $\mathbf{z}$ are of dimension l and $\mathbf{v}$ is assumed to be zero mean gaussian white noise, with variance R, or [N(0,R)]. Here l is some multiple of n, giving redundant measurements. The requirement for redundant measurements is a result of the uncertainty of the derivative of the state vector $\dot{x}$, due to measurement noise v.

To minimize the distance between θ, the parameter vector and $\hat{\theta}$ the estimated parameter vector, a cost function is defined as [Ref. 5:p. 24]

$$J = (\mathbf{z} - H\hat{\theta})^T(\mathbf{z} - H\hat{\theta}) \tag{2.11}$$

If J is minimized, then

$$\frac{\partial J}{\partial \hat{\theta}} = \underline{0} \tag{2.12}$$

thus

$$\mathbf{H}^T\mathbf{H}\hat{\theta} = \mathbf{H}^T\mathbf{z} \tag{2.13}$$

or

$$\hat{\theta} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{z} \tag{2.14}$$

The estimate $\hat{\theta}$ will be unbiased if v and H are statistically independent. This is easily seen if one takes the expectation of the previous equation. That is,

$$E[\hat{\theta}] = E[(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T(\mathbf{H}\theta + \mathbf{v})] \tag{2.15}$$

or

$$E[\hat{\theta}] = E[\theta] + (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T E[\mathbf{v}] = \theta \tag{2.16}$$

where the second term in the above equation vanishes because the vector v was assumed to be zero mean.

Mathematically, the result of Equation 2.16 is quite nice; however in practice the x vector will also be contaminated with noise. Thus, it is appropriate to define a new vector r as.

$$r(k) = x(k) + w(k), \qquad w(k) \tilde{} N(0, Q) \tag{2.17}$$

so that

$$\begin{pmatrix} z(l-1) \\ z(l-2) \\ \vdots \\ z(0) \end{pmatrix} = \begin{pmatrix} H_r(l-1) \\ H_r(l-2) \\ \vdots \\ H_r(0) \end{pmatrix} \theta + \mathbf{v} \tag{2.18}$$

where

$$H_r(k) = \begin{pmatrix} r_1(k)r_2(k) \ldots r_n(k) & 0 & \ldots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & r_1(k)r_2(k) \ldots r_n(k) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & & & & & \\ 0 & 0 & \ldots & 0 & 0 & 0 & 0 & r_1(k) \cdots r_{n-1}(k)r_n(k) \end{pmatrix} \tag{2.19}$$

Clearly, in this case, H_r and $\mathbf{v}$ are not statistically independent since,

$$v_1 = v_1 - \theta_1 w_1 - \theta_2 w_2 - \ldots - \theta_n w_n \tag{2.20}$$

$$v_2 = v_2 - \theta_{n+1} w_1 - \theta_{n+2} w_2 - \ldots - \theta_{2n} w_n$$

This correlation between H and v will almost certainly produce a biased estimation of θ [Ref. 4:pp. 132-33].

At this point several characteristics of this identification technique are apparent. This is a batch processing approach, which means that all of the data, that is **z** and **H**, must be stored in memory before the procedure can begin. With the current availability of low-cost large-volume memory this may not be a problem. However, it also requires the multiplication of two $n^2 \times l$ matrices, which is fairly computationally intensive, because l is very large. Again, technology may have reached the level where this is not a severe difficulty. Finally, the bias on the parameter estimates due to the state measurement noise must be determined, and its impact on the control of the autonomous system considered.

C. KALMAN FILTER IDENTIFIER

The Kalman filter is used to obtain the minimum-variance estimates of signals from noisy measurements recursively; i.e. in real time. It is assumed that the reader has a general knowledge of Kalman filter theory. which has received widespread attention in the literature. [Ref. 5]

For a system described as

$$x(k+1) = \Phi(k+1.k)x(k) + Lu(k) \tag{2.21}$$

$$z(k+1) = H(k+1)x(k+1) + v(k+1) \tag{2.22}$$

where u and v are gaussian random variables with zero mean and variances of Q and R, respectively, it can be shown [Ref. 5:pp. 102-11] that the optimal estimate of $x(k)$, denoted $\hat{x}(k)$ is recursively obtained by

$$\hat{x}(k+1|\,k) = \Phi(k+1,k)\hat{x}(k|\,k) + Lu(k) \tag{2.23}$$

$$\hat{x}(k+1|\,k+1) = \hat{x}(k+1|\,k) + G(k+1)[z(k+1) - H(k+1)\hat{x}(k+1|\,k) \tag{2.24}$$

The $G(k+1)$ matrix is the Kalman gain matrix and is defined as

$$G(k+1) = P(k+1|\,k)H^T(k+1)[H(k+1)P(k+1)H^T(k+1) + R(k+1)]^{-1} \tag{2.25}$$

where $P(k+1|\,k)$ is the estimation error covariance matrix of $(x-\hat{x})$ defined by

$$P(k+1|\,k) = \Phi(k+1,k)P(k|\,k)\Phi^T(k+1,k) + LQL^T \tag{2.26}$$

and

$$P(k+1|\,k+1) = [I - G(k+1)H(k+1)]P(k+1|\,k) \tag{2.27}$$

If the output error vector given by

$$e(k+1) = z(k+1) - H(k+1)\hat{x}(k+1|\,k) \tag{2.28}$$

which is the quantity multiplied by $G(k+1)$ in Equation 2.24, is compared with

Equation 2.22 rewritten as.

$$v(k+1) = z(k-1) - H(k+1)\hat{x}(k+1) \tag{2.29}$$

then as $\hat{x} \to x$, $e \to v$.

In order to transform the state estimation Kalman filter into a parameter identifier, the system defined in Section 2.B as

$$\dot{x}(k) = Ax(k) + Bu(k) \tag{2.30}$$

where u is again set to zero, is changed to

$$\theta(k+1) = \theta(k) \tag{2.31}$$

and from Equations 2.4 and 2.9

$$z(k) = \dot{x}(k) + v(k) = H(k)\theta + v(k) \tag{2.32}$$

Here $v \sim [N(0,R)]$ is the state derivative measurement noise. The forward estimation of the parameter vector in Equation 2.23 is now changed to an identity equation, since intuitively we wish

$$\hat{\theta}(k|\,k-1) = \hat{\theta}(k-1|\,k-1) \tag{2.33}$$

for a time invariant system. Thus $\Phi(k+1,k) = I$, the identity matrix, and the present estimate becomes

$$\hat{\theta}(k|\,k) = \hat{\theta}(k|\,k-1) + G(k)[z(k) - H(k)\hat{\theta}(k|\,k-1)] \tag{2.34}$$

Examining again the error vector e, now given as

$$e(k) = z(k) - H(k)\hat{\theta}(k|\,k-1) \tag{2.38}$$

and Equations 2.28 and 2.29, here rewritten as

$$v(k) = z(k) - H(k)\theta \tag{2.39}$$

shows that as $\hat{\theta} \to \theta$, $e \to v$. Since E[v] = 0, then $\hat{\theta}$ will be an unbiased estimate of θ.

From the above analysis and definitions, the equations for the Kalman gain matrix of the parameter identifier are

$$G(k) = P(k|k-1)H^T(k)[H(k)P(k|k-1)H^T(k) + R(k)]^{-1} \tag{2.40}$$

where $P(k|k-1)$ is now defined as

$$P(k|k-1) = P(k-1|k-1) \tag{2.41}$$

and

$$P(k|k) = [I - G(k)H(k)]P(k|k-1) \tag{2.42}$$

and R is the covariance matrix of v. This identifier is recursive, but suffers from the same limitation that the least-squares approach exhibited. Namely, the Kalman filter identifier does not eliminate the bias produced by state measurement noise. For this reason the stochastic gradient approach of the next section is considered, since theoretically it will eliminate the bias caused by state measurement noise.

D. STOCHASTIC GRADIENT ESTIMATION

The stochastic gradient approach is a recursive identifier and therefore has a common structure with other recursive parameter identification schemes. The structure is

$$\textit{new estimation} = \textit{old estimation} + \textit{gain matrix} \times \textit{error} \tag{2.43}$$

To begin, consider again the system described using the parameter vector θ given as

$$\dot{x}(k) = H(k)\theta + v_d(k) \tag{2.44}$$

where the measurement of $\dot{x}(k)$ is further contaminated with noise as in

$$z(k) = \dot{x}(k) - v_m(k) \tag{2.45}$$

Thus there is an element of both disturbance and measurement noise corrupting $\dot{x}(k)$. It is also assumed that the $x(k)$ state vector cannot be measured perfectly so that its measured value is

$$r(k) = x(k) + w(k) \tag{2.46}$$

making the contaminated $H(k)$ matrix appear as

$$H_r(k) = H(k) + N(k) \tag{2.47}$$

where

$$N(k) = \begin{pmatrix} w_1(k)w_2(k) \cdots w_n(k) & 0 & \cdots 0 & 0 & \cdots 0 & 0 \\ 0 \quad 0 \quad \cdots 0 & w_1(k)w_2(k) \cdots w_n(k) & 0 & \cdots 0 & 0 \\ \vdots \quad \vdots \quad \vdots & & & & & \\ 0 \quad 0 \quad \cdots 0 & 0 \quad 0 \quad 0 \quad 0 & w_1(k) \cdots w_{n-1}(k)w_n(k) \end{pmatrix} \tag{2.48}$$

The random variable $w(k)$ is zero mean with variance Σ_w and the variance of N(k) is then

$$E[N(k)^T N(k)] = \Sigma_N \tag{2.49}$$

The measurable value of $\dot{x}(k)$ is $z(k)$ which is obtained from

$$z(k) = H(k)\theta + v(k) \tag{2.50}$$

where

$$v(k) = v_d(k) + v_m(k) \tag{2.51}$$

However, the only available approximation of z(k), denoted $\hat{z}(k)$ is

$$\hat{z}(k) = H_r(k)\hat{\theta}(k) \tag{2.52}$$

where in this case $\hat{\theta}(k)$ is the k*th* estimation of θ.

The error in the estimate of $z(k)$ is

$$\tilde{z}(k) = z(k) - \hat{z}(k) \tag{2.53}$$

and if $\tilde{z}(k)$ is to be minimized, then a cost function is established as

$$J[\hat{\theta}(k)] = \frac{1}{2}\tilde{z}^T(k)\tilde{z}(k) \tag{2.54}$$

It can easily be shown that the gradient of $J[\hat{\theta}(k)]$ with respect to $\hat{\theta}(k)$ is

$$\nabla J_{\hat{\theta}} = - H_r^T(k)\tilde{z}(k) \tag{2.55}$$

This leads to the stochastic gradient algorithm

$$\hat{\theta}(k+l) = \hat{\theta}(k) + R(k)H_r^T(k)\tilde{z}(k) \tag{2.56}$$

where R(k) is the gain matrix to be determined and l is the spacing parameter to be chosen such that $H(k)$ and $H(k-l)$ are independent. In a continuous time system, it will be assumed that 4 time constants of the system are enough to make $H(k)$ and $H(k-l)$ independent. The reason for this requirement will be discussed shortly.

Certainly an unbiased estimate of θ is desirable. The nature of sequential estimation is such that only asymptotic unbiasness is possible [Ref. 4:p. 233]. Looking closer at Equation 2.56 to determine the bias imposed by the estimation process, the expectation of both sides is taken, giving

$$\begin{aligned} E[\hat{\theta}(k+l)] &= E[\hat{\theta}(k)] + R(k)E[H_r^T(k)\{H(k)\theta + v(k) - H_r(k)\hat{\theta}(k)\}] \\ &= E[\hat{\theta}(k)] - R(k)\{E[E[H_r^T(k)H_r(k)|\,\hat{\theta}(k)]\hat{\theta}(k)] \\ &\qquad - E[E[H_r^T(k)v(k)|\,\hat{\theta}(k)]] \\ &\qquad - E[E[H_r(k)H(k)\theta|\,\hat{\theta}(k)]]\} \end{aligned} \tag{2.57}$$

where it has been implicitly assumed that $R(k)$ is independent of $H_r(k)$ and $\hat{\theta}(k)$.

Examining each term of Equation 2.57 to determine its effect on the estimate of θ will demonstrate the source of the bias. The first term multiplied by $R(k)$ is dissected, taking the inner expectation giving [Ref. 4:p. 233-42]

$$E[H_r^T(k)H_r(k)|\hat{\theta}(k)] = E[\{H^T(k)H(k) + H^T(k)N(k) + N(k)H(k) + N^T(k)N(k)\}|\hat{\theta}(k)] \quad (2.58)$$

$$= \Omega_H + \Sigma_N \quad (2.59)$$

or

$$E[E[H_r^T(k)H_r(k)|\hat{\theta}(k)]\hat{\theta}(k)] = \{\Omega_H + \Sigma_N\}E[\hat{\theta}(k)] \quad (2.60)$$

Here the independence of $H(k)$ and $H(k-l)$ has been used to assure that

$$E[H^T(k)H(k)|\hat{\theta}(k)] = E[H^T(k)H(k)] \equiv \Omega_H \quad (2.61)$$

In a like manner it can be shown that

$$E[E[H_r^T(k)v(k)|\hat{\theta}(k)]] = E[H^T(k)v(k)] \quad (2.62)$$

and

$$E[E[H_r^T(k)H(k)\theta|\hat{\theta}(k)]] = \Omega_H\theta \quad (2.63)$$

Then rewriting Equation 2.57 gives

$$E[\hat{\theta}(k+l)] = E[\hat{\theta}(k)] - R(k)\{(\Omega_H + \Sigma_N)E[\hat{\theta}(k)] + E[H^T(k)v(k)] + \Omega_H\theta\} \quad (2.64)$$

Clearly, the steady state solution should be

$$\lim_{k\to\infty} E[\hat{\theta}(k+l)] = E[\hat{\theta}(k)] \quad (2.65)$$

thus

$$(\Omega_H + \Sigma_N)\lim_{k\to\infty} E[\hat{\theta}(k)] = E[H^T(k)v(k)] + \Omega_H\theta \quad (2.66)$$

or

$$\lim_{k\to\infty} E[\hat{\theta}(k)] = (\Omega_H + \Sigma_N)^{-1}\{E[H^T(k)v(k)] + \Omega_H\theta\} \quad (2.67)$$

Notice that if

$$\Sigma_N = E[H^T(k)v(k)] = 0 \tag{2.68}$$

then

$$\lim_{k\to\infty} E[\hat{\theta}(k)] = \Omega_H^{-1}\Omega_H\theta = \theta \tag{2.69}$$

Unfortunately, in the final system analyzed here both Σ_η and $E[H^T(k)v(k)]$ will be non-zero.

To obtain an unbiased estimate in the system of interest it will be necessary to remove the biasing terms from Equation 2.64 . More specifically, the new algorithm will be

$$\hat{\theta}(k+1) = [I+R(k)\Sigma_\eta]\hat{\theta}(k) + R(k)H_r^T(k)\tilde{z}(k) - R(k)E[H^T(k)v(k)] \tag{2.70}$$

The term $E[H^T(k)v(k)]$ is not available, but assume that it is a linear function of θ as in

$$E[H^T(k)v(k)] = \lambda + M\theta \tag{2.71}$$

However, since θ is unknown, this equation can only be estimated as

$$\hat{\Lambda} = \lambda + M\hat{\theta} \tag{2.72}$$

where λ and M are application dependent and must be calculated prior to implementation.

The choice of $R(k)$ remains to be determined. Extensive analysis of the selection of R(k) has been presented in the literature.[Ref. 4:pp.193-210][Ref. 6:pp. 42-51] In this section two choices for the gain matrix R(k) will be stated as presented in [Ref. 4:pp. 193-203, 242-251].

Using Lyapunov's Main Stability Theorem, it can be shown that the Lyapunov- optimum weighting matrix is

$$R_j(k)=\frac{diag[h_1(k),\ h_2(k),\ \ldots,\ h_{n^2}(k)]}{\Sigma_{i=1}^{n} h_i(k) r_i^2(k)} \qquad \text{for } j=1,2,\ \ldots,n \tag{2.73}$$

Note here that $R(k)$ is now correlated with H(k).

Using Venter's theorem and its conditions for convergence, the R(k) matrix becomes

$$R(k) = \frac{1}{k^p} diag[h_1(k),\ h_2(k),\ \ldots,\ h_n(k)] \quad \text{for } \frac{1}{2}<p\leq 1 \tag{2.74}$$

where in both cases $h_i(k)$ must conform to

$$0< h_L \leq h_i(k) \leq h_U < \infty \tag{2.75}$$

for i=1, 2, . . ., n^2, and all $k\geq 0$.

Although Equation 2.74 is somewhat simpler to compute. this advantage is more than offset by its slower convergence. It has been suggested that because of the faster convergence of Equation 2.73 then for $k\leq k'$ use [Ref. 4:p 268]

$$R(k)=\frac{diag[h_1(k),\ h_2(k),\ \ldots,\ h_{n^2}(k)]}{\Sigma_{i=1}^{n} h_i(k) r_i^2(k)} \tag{2.76}$$

and, for $k > k'$. use

$$R(k) = \frac{1}{k^p} diag[h_1(k),\ h_2(k),\ \ldots,\ h_n(k)] \tag{2.77}$$

where k' can be chosen quite arbitrarily.

E. SUMMARY

The parameter identification techniques described here are but three of the most well known and in no way span the variety of identification schemes. In fact the first two parameter estimator described here are very closely related. While

the Kalman filter and least squares approach look quite different. the Kalman filter is in fact the recursive form of the weighted least squares algorithm [Ref 6:p. 20].

The Kalman filter and the stochastic gradient approaches are related in that they are both recursive. This gives them the same structure as described in Equation 2.43 , from this perspective, the stochastic gradient approach described here uses the simplest possible choice for the gain matrix while the Kalman filter uses a much more sophisticated choice.

III. LEAST-SQUARES IDENTIFIER

A. INTRODUCTION

In this chapter a simulation study will be done using the least-squares identification approach. The purpose of the study is to determine the capabilities and limitations of the least-squares approach, with regard to various forms of input noise.

Recall from Chapter II that the state derivative measurement was given by

$$\mathbf{z} = \mathbf{H}\theta + \mathbf{v} \tag{3.1}$$

and that $\hat{\theta}$ was then determined by

$$\hat{\theta} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{z} \tag{3.2}$$

In this chapter the possibility of input disturbance noise, u(t), will be included as well as the state measurement noise, w(t). These additional sources of noise will produce the new relations

$$\mathbf{z} = \mathbf{H}\theta + \mathbf{v} + \mathbf{B}\mathbf{u} \tag{3.3}$$

and

$$\hat{\theta} = (\mathbf{H}_r^T\mathbf{H}_r)^{-1}\mathbf{H}_r^T\mathbf{z} \tag{3.4}$$

where $\mathbf{B}$ is defined as

$$\mathbf{B} = \begin{pmatrix} B & 0 & 0 & \cdots & 0 \\ 0 & B & 0 & \cdots & 0 \\ 0 & 0 & B & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & B \end{pmatrix} \tag{3.5}$$

In this equation, B is as defined in Equation 2.1, and 0 is a zero matrix of appropriate dimension. Recalling that the kth noise corrupted state measurement is denoted

$$r(k) = x(k) + w(k) \tag{3.6}$$

while the corresponding noisy state derivative measurement is

$$z(k) = \dot{x}(k) + v(k) \tag{3.7}$$

it follows that Equations 3.3 and 3.4 reduce to Equations 3.1 and 3.2 when $u(t)$ and $w(t)$ are not present.

This new set of equations will allow a realistic assessment of the accuracy of the least-squares identification approach with various amounts of noise present in both the measurements of the system and input from the road. These noise sources will each be considered in detail and then in combination to examine their overall effect on the estimation of the system parameters.

B. SIMPLIFIED MODEL FOR SIMULATION STUDY

In preparation for a simulation study, it is necessary to define a simplified model of a suspension system. A fourth order, two-degree-of-freedom system was selected. This choice will provide a rich enough model to allow analysis of the results obtained from the three identification techniques discussed in Chapter II.

Consider the system depicted in Figure 3-1. The lower mass represents the wheels and the upper mass is the body. The equations of motion are

$$\Sigma_{F_{x_1}} = -\ddot{x}_1 m - (k_1 + k_2)x_1 + k_2 x_2 - b\dot{x}_1 + b\dot{x}_2 + k_1 u(t) = 0 \tag{3.8}$$

$$\Sigma_{F_{x_2}} = -\ddot{x}_2 M - k_2 x_1 + k_2 x_2 - b\dot{x}_1 + b\dot{x}_2 \tag{3.9}$$

which can be written in the state space form

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{3.10}$$

as

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-(k_1+k_2)}{m} & \frac{-b}{m} & \frac{k_2}{m} & \frac{b}{m} \\ 0 & 0 & 0 & 1 \\ \frac{k_2}{M} & \frac{b}{M} & \frac{-k_2}{M} & \frac{-b}{M} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \frac{k_1}{m} \\ 0 \\ 0 \end{bmatrix} u(t) \tag{3.11}$$

where

$$x(t) = \begin{bmatrix} x_1(t) \\ \dot{x}_1(t) \\ x_2(t) \\ \dot{x}_2(t) \end{bmatrix} \tag{3.12}$$

For the purpose of simulation, $u(t)$, the input vector, will be gaussian noise of zero mean. In order to obtain the five constants (M, m, k_1, k_2, b) which make up the parameters of the system, and arrive at a unique value for each constant, it will

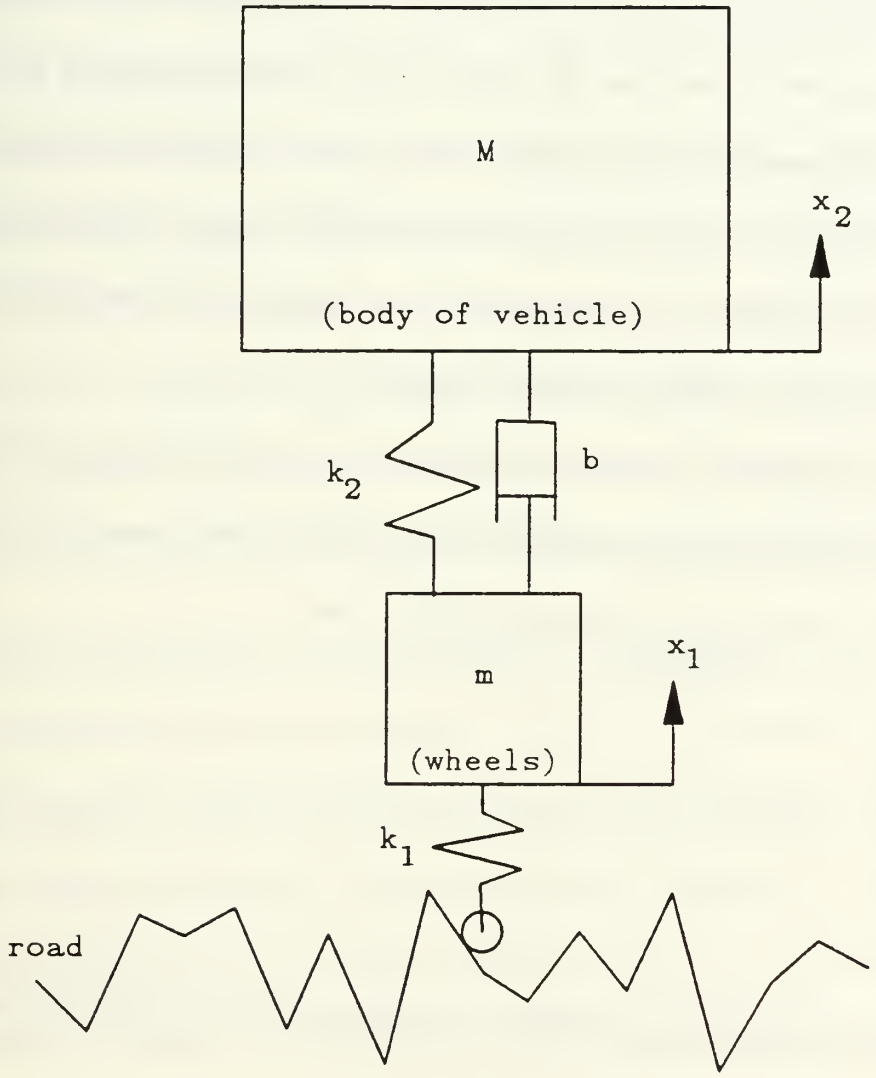

Figure 3-1 Simplified vehicle

be necessary to know one of the masses (M , m). *a priori* since one of the masses appears in all of the parameter values.

To obtain the B matrix, from observation of the vehicle response only, it is evidently necessary to be able to measure the input $u(t)$. This will not be possible for the system analyzed here, thus the value of the B matrix must be known *a priori*. This requirement, though inconvenient, is not a serious complication since the spring constant and mass of the wheels should be easily estimated from experimentation on a relatively small element of the vehicle. This also fulfills the requirement for having a value for one of the masses.

Starting with Equation 3.11, values were then chosen for the five constants in the system to be used during the simulation study. The values, chosen to achieve a reasonable time constant and damping coefficient, were :

$$\begin{bmatrix} M \\ m \\ k_1 \\ k_2 \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \\ 90 \\ 50 \\ 40 \end{bmatrix} \tag{3.13}$$

This produces a dominant time constant for the system of 1.4 sec, with poles at

$$s = \begin{bmatrix} -.727 \pm j2.30 \\ -1.90 \\ -40.6 \end{bmatrix} \tag{3.14}$$

The units of the system constants are left arbitrary and must simply be consistent.

Finally, it will be assumed that measurements of both the state $x(k)$ and its derivative $\dot{x}(k)$ are available, but are possibly corrupted by noise. Thus, there are three sources of noise in the identification process: the input disturbance $u(t)$, the state measurement noise $w(t)$, and the noise associated with the measurement of the state derivative $v(t)$.

In the remainder of this chapter the effects of these noise sources on the least squares identifier shall be examined. The criterion for acceptable identification will be when all parameters are within 10% of the actual value. Since the parameters in the second and fourth rows of Equation 3.11 are the unknown values, only their estimation error shall be considered, and the parameters of the first and third rows will be assumed known, because of the structure of the plant.

C. INPUT DISTURBANCE AND INITIAL CONDITIONS

The effect of the input disturbance $u(t)$ will be observed with and without initial conditions. Here the initial conditions will be considered another form of input to excite the system. The system shall be analyzed under three different forms of excitation: first by initial conditions alone, then with white noise from the road, and finally from a pulse input. The resulting parameter estimates will be used to determine how best to excite the system to obtain acceptable identification.

If initial conditions are the only form of excitation or noise in the system then, as shown in Chapter II, in n measurements of the system the parameters

can be found exactly. On the other hand, with no initial conditions and with white noise input from the road, in the absence of other noise, the parameters of the upper mass can be measured prefectly, but the lower mass parameter estimates are significantly corrupted. Referring to Figure 3-1, this comes about because the net force acting on the lower mass is a function of both the present state and the unknown suspension system displacement from the road. The force on the upper mass, however, is a function of only the present state which is known exactly. Thus, with no measurement noise, the fourth line of Equation 3.11 is known exactly, while the second line is corrupted by unknown road noise and the parameter can not be found precisely in a finite number of measurements. However, combining Equations 2.14 and 3.3 gives

$$\hat{\theta} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T[\mathbf{H}\theta + \mathbf{v} + \mathbf{B}\mathbf{u}] \tag{3.15}$$

or taking the expected value

$$E[\hat{\theta}] = E[\theta] + (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T E[\mathbf{v} + \mathbf{B}\mathbf{u}] = \theta \tag{3.16}$$

since both v and u are zero mean random processes. Therefore it is possible to identify the parameters of the system with only road noise as excitation, but quite impractical since, as experimental results in Chapter IV will show, matrices with dimensions larger than 8 by 10,000 would be required.

Figure 3-2 shows a measure of the normalized mean square error of the parameters, given by

$$\xi_k = \sum_{i=1}^{8} \left(\frac{\theta_i - \hat{\theta}_i(k)}{\theta_i} \right)^2 \qquad (3.17)$$

For this curve, and throughout the remainder of this thesis, the response of the continuous time system is obtained by numerical integration of the state equations using a fourth order Runge-Kutta integration formula. The time interval between samples of the resulting response is in every case 0.05 seconds.

On Figure 3-2, ξ_k is shown for four different identification simulation runs, each with the same standard deviations of input noise, for the case of white road noise and no measurement error in $x(k)$ or $\dot{x}(k)$. Thus, only the actual disturbance ensembles are different. Note the large variance in the resulting identification. This variance is quite undesirable, because of its inconsistent approximation results. From the linearity of the state equation, it is expected that the size of σ_u has no effect on the accuracy of the parameter estimation process. This has been verified by changing the scale factor on $u(t)$ and observing from the simulation that the parameters are uneffected. Also, since ξ_k is defined as in Equation 3.17, then a necessary condition to obtain an acceptable estimate is

$$\xi_{final} \leq 0.08 \qquad (3.18)$$

That is, if the average parameter estimation errors do not exceed 10 percent, then this condition must be satisfied. Examination of the four curves of Figure 3-2 shows that attainment of this accuracy is likely to require, on the average, many times more than 500 iterations of the least squares procedure.

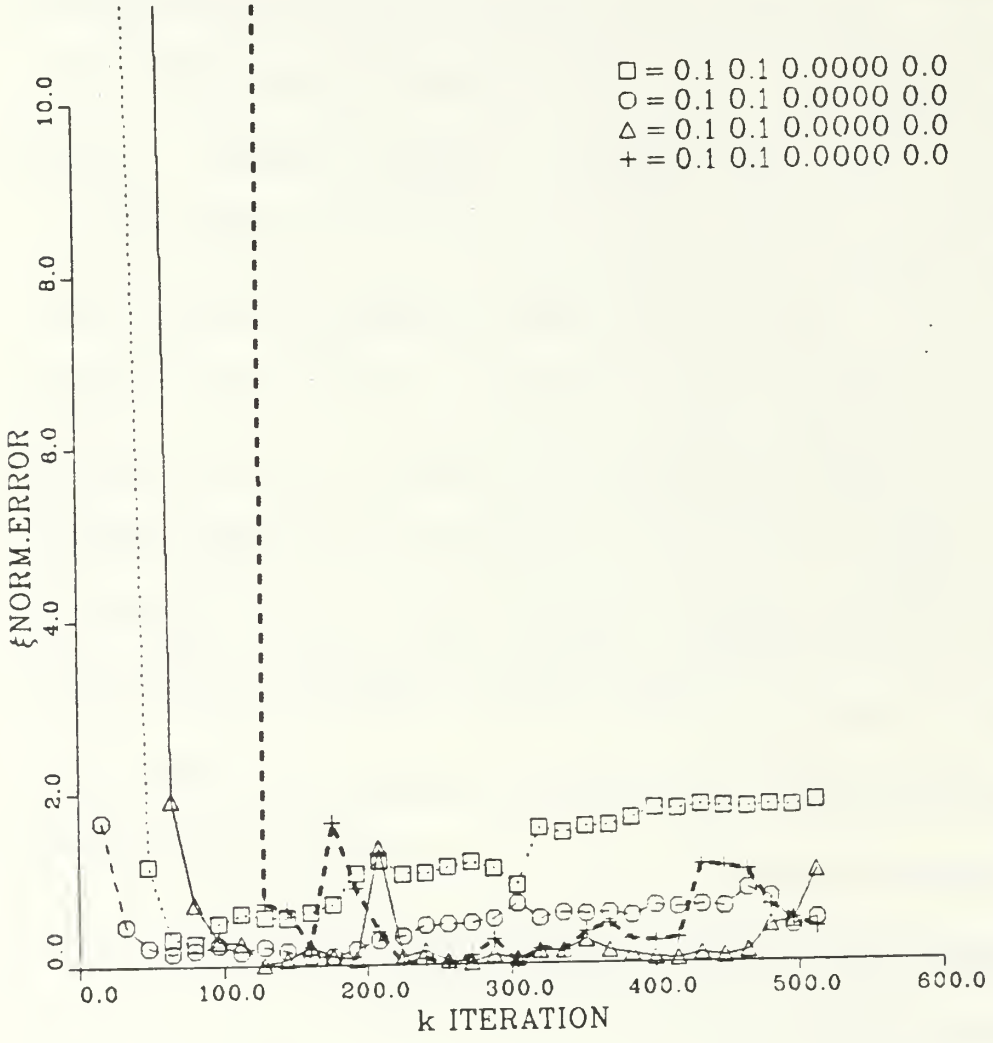

Figure 3-2 Parameter estimation error vs. number of iterations for four different ensembles of road noise.
Initial conditions. $\sigma = 0.1$
Road noise, $\sigma = 0.1$
State measurement noise, $\sigma = 0.0$
State derivative measurement noise, $\sigma = 0.0$

If both random initial conditions and white noise are applied to the vehicle, the result is essentially that of the dominant form of excitation. Figure 3-3 illustrates the relationship between ξ_k and the excitation of σ_{ic} and σ_u. From the curves, the effect of various amounts of road noise for a given amount of initial excitation can be seen. If the road noise is small, then the identifier can estimate the parameters in just a few measurements. This is exhibited in the curve marked by "+" symbols. As road noise increases, the identifier requires more measurements before acceptable estimation is reached. In the curve marked by squares, the identifier cannot reach a correct estimation within 500 measurements. If more memory was available to store $x(k)$ and $\dot{x}(k)$ then the estimation ought to improve.

If a pulse of various widths is used to excite the system, then the parameter estimates are less accurate than if initial conditions are used to excite the system. When the pulse is input during only one measurement cycle, then so little energy is imparted to the vehicle that the identifier is essentially trying to estimate the parameters through the road noise and will require more than 125 measurements to identify the parameters, as shown by the curve marked by squares. If pulses wider than one measurement cycle are used then the system receives more energy, but the input becomes correlated with the state, which as will be seen in Chapter IV, produces a biased estimate. The other three curves in Figure 3-4 show pulse widths of two, three and four measurement cycles. Thus, a pulse input does not seen to improve the identification process.

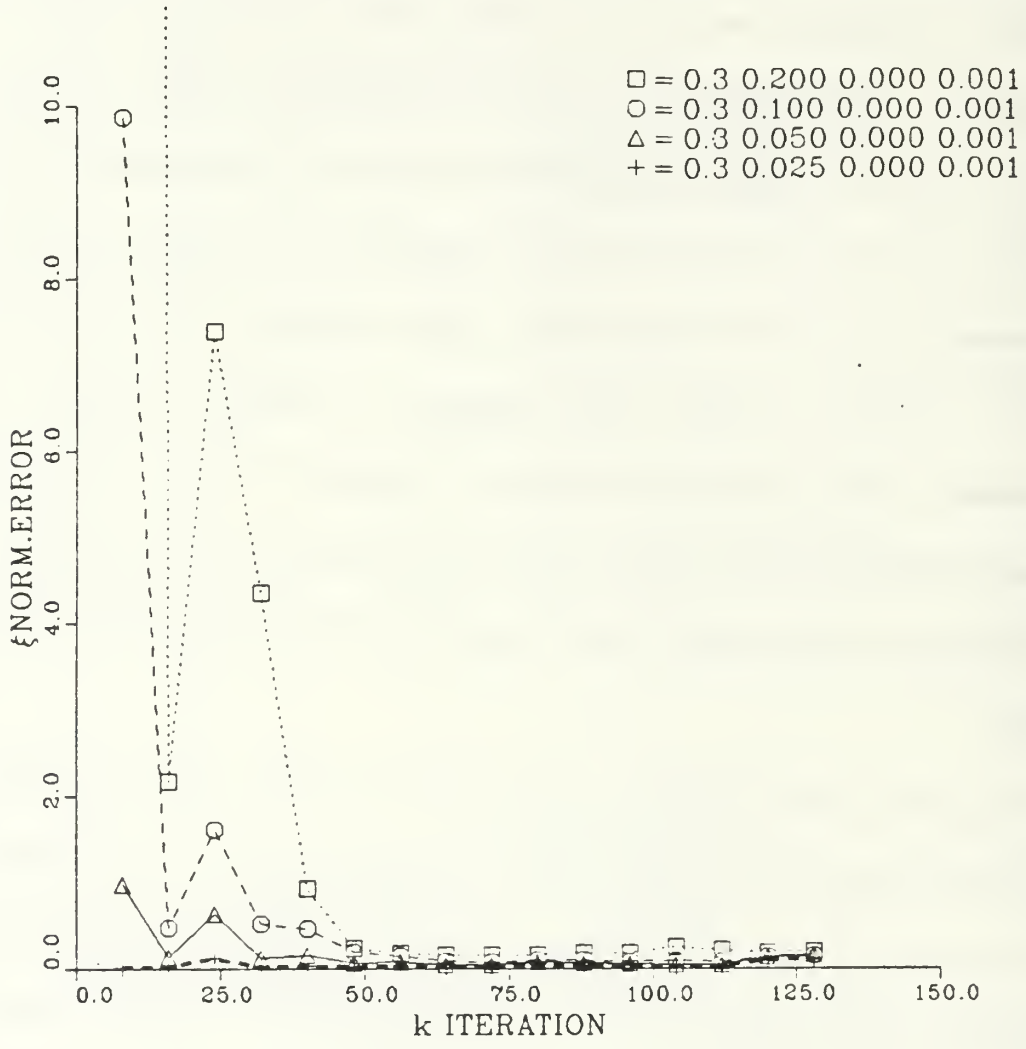

Figure 3-3 Parameter estimation error vs. number of iterations when four different levels of road noise and constant initial conditions excite the system.

Initial conditions, $\sigma = 0.3$

Road noise, $\sigma = 0.2, 0.1, 0.05, 0.025$

State measurement noise, $\sigma = 0.0$

State derivative measurement noise, $\sigma = 0.0$

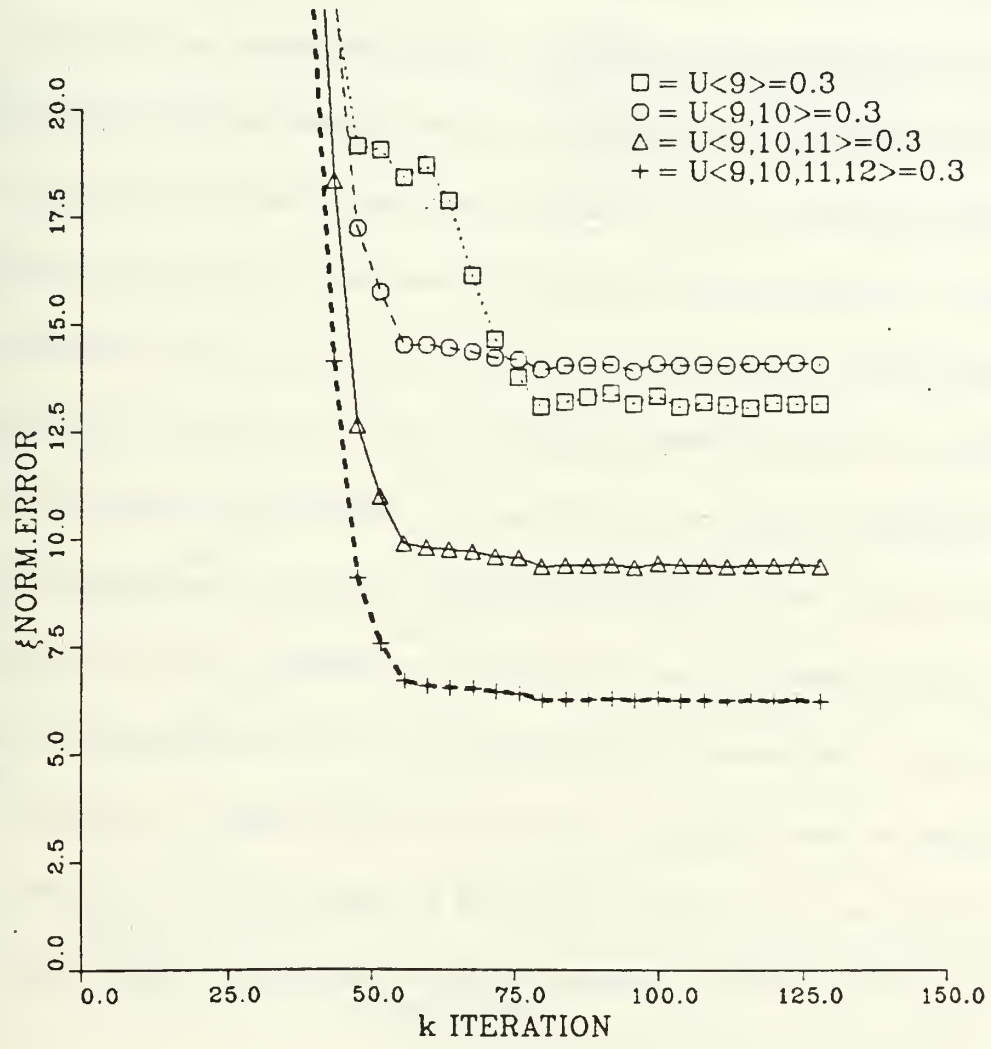

Figure 3-4 Parameter estimation error when pulses of four different widths are used to excite the system.
Initial conditions, $\sigma = 0.01$
Road noise, $\sigma = 0.001$
State measurement noise. $\sigma = 0.0$
State derivative measurement noise, $\sigma = 0.0$

D. STATE MEASUREMENT NOISE

The effect of state measurement noise on the least squares identifier is devastating. Using only initial conditions to excite the system, which would in the absence of noise produce exact results, the least squares identifier deteriorates rapidly with small values of σ_w, the standard deviation of the state measurement noise. This is as expected since the technique is based on an assumption of perfect state measurements. Figure 3-5 shows the effect of varying σ_w on the normalized mean square error of the estimated parameters. Since the excitation is the result of initial conditions, which are diminishing with time, the error induced by σ_w increases as the state vector magnitude decreases. Thus there is an optimum time to extract the parameter at about 32 measurement intervals. The curve marked by circles has a measurement error of 1 % of the initial conditions. Even this small measurement noise standard deviation is too large to allow the estimator to find the parameters. If σ_w is reduced to 0.5% of σ_{ic} then satisfactory results can be expected, if the parameters are taken at about 32 measurements. This is quite a constraint on the measurement equipment for $x(k)$.

E. STATE DERIVATIVE MEASUREMENT NOISE

Using initial conditions to excite the system, the analysis of the effects of σ_v, the standard deviation of the state derivative measurement noise, will be carried out by varying σ_v until ξ_k falls below the required value of .08. Figure 3-6 shows that if σ_v is 20% of σ_{ic} then acceptable results can be expected. The same noise

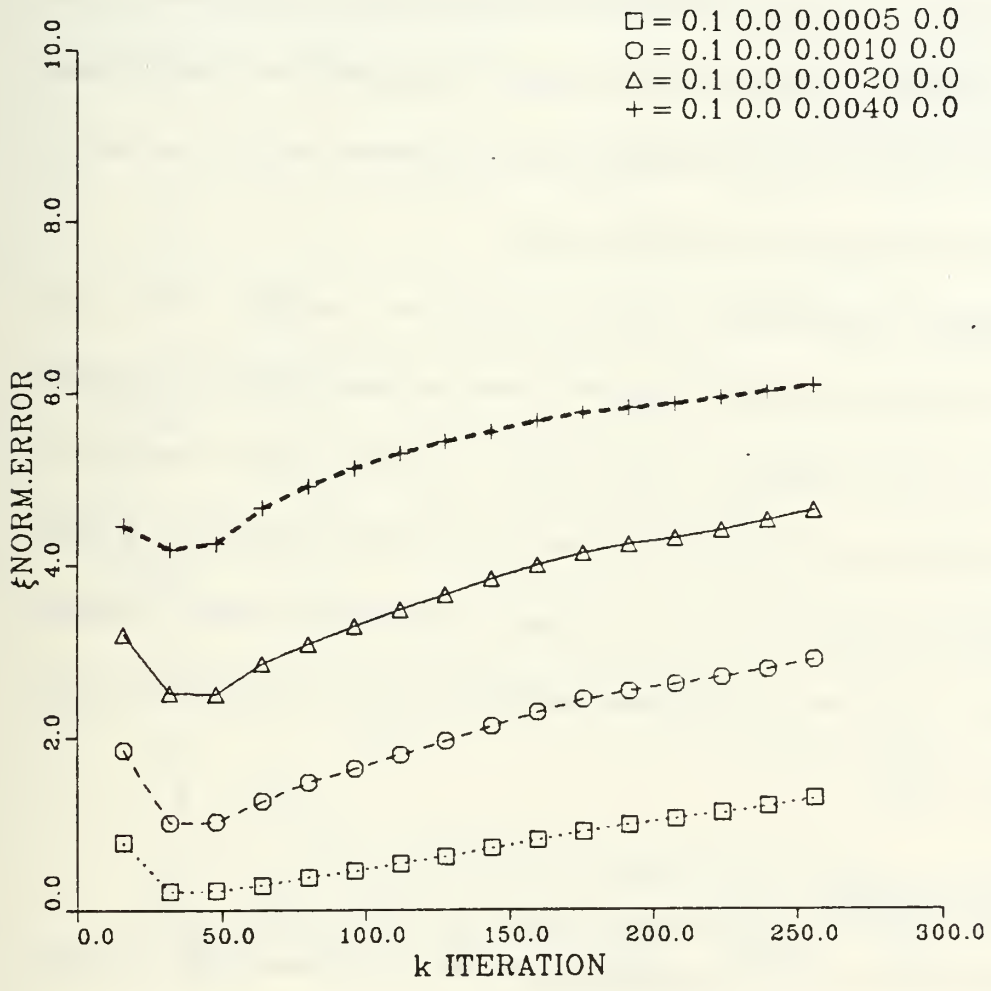

Figure 3-5 Parameter estimation error vs. number of iterations when four different levels of state measurement noise are applied.
Initial conditions, $\sigma = 0.1$
Road noise, $\sigma = 0.0$
State measurement noise, $\sigma = 0.0005, 0.0010, 0.0020, 0.0040$
State derivative measurement noise, $\sigma = 0.0$

sample was used for all four curves on this figure with only a change in scale factor to increase the effective noise level.

The error introduced by the state derivative noise source is much larger in the upper mass parameters than in the lower mass. This should be expected since intuitively, larger accelerations are produced from small state vector values on the lower mass, thus overshadowing the noise on $\ddot{x}_1(k)$.

State derivative measurement noise acts upon the identifier in the same manner that road noise does. This being true then, as observation time increases the estimates of the parameters will improve. However, unlike road noise measurement noise does not excite the system, and within 5 seconds (or k=100) the state vector will reduce to 3 % of its original magnitude, eventually becoming zero. Therefore, some form of continuous excitation must be present during the identification process.

F. COMBINED NOISE

The combination of all the noise sources results in essentially the same identification performance as would be expected from the worst individual noise source applied. This should not be surprising since the three noise sources are independent. Figure 3-7 illustrates this, for as each noise source becomes dominant, its influence on the parameter estimates becomes more apparent. This is particularly evident in the curve marked with "+" symbols, where the increasing parameter error is due to state measurement noise σ_w. Because of the

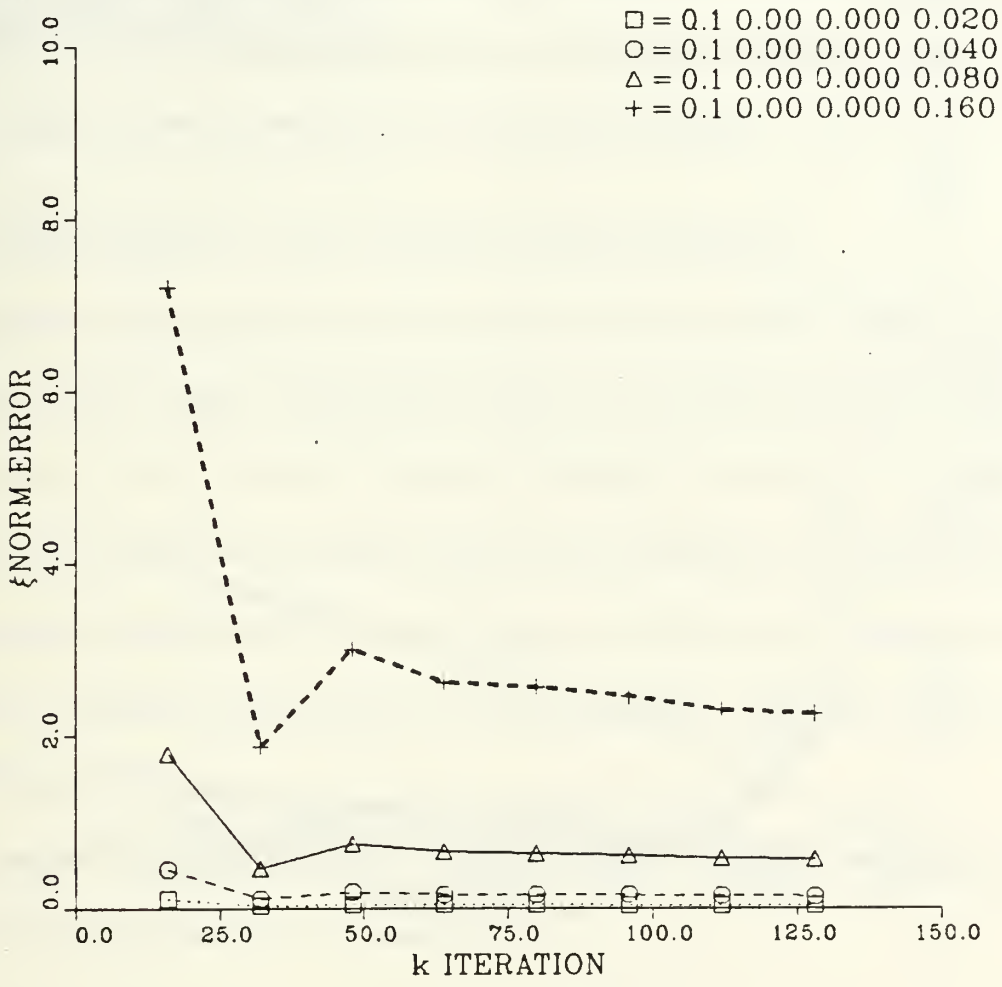

Figure 3-6 Parameter estimation error vs. number of iterations when four different levels of state derivative measurement noise are applied.
Initial conditions, $\sigma = 0.1$
Road noise, $\sigma = 0.0$
State measurement noise, $\sigma = 0.0$
State derivative measurement noise, σ =0.020, 0.040, 0.080, 0.160

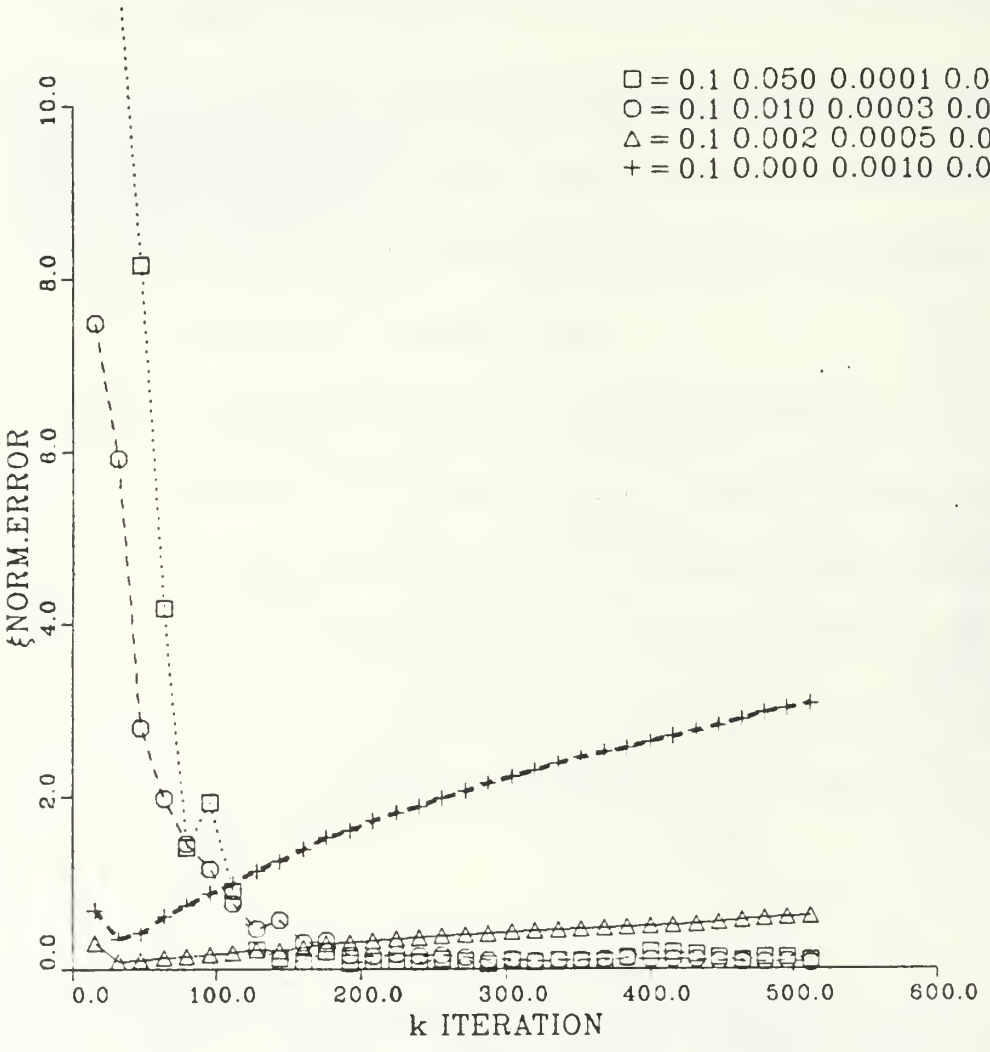

Figure 3-7 Parameter estimation error vs. number of iterations when four different noise sets are applied.
Initial conditions, $\sigma = 0.1$
Road noise, $\sigma = 0.050, 0.010, 0.002, 0.0004$
State measurement noise, $\sigma = 0.000125, 0.00025, 0.0005, 0.0010$
State derivative measurement noise, σ =0.040, 0.020, 0.010, 0.005

large initial conditions, the estimations become nearly acceptable quite rapidly, but the additive effect of state measurement noise quickly produces a bias in the parameter estimates. In the curve marked by triangles, the estimates do become acceptable, but the bias associated with even this small state measurement noise is too much to allow the estimates to remain correct for long.

G. COMPUTATIONAL AND STORAGE REQUIREMENTS

The least squares algorithm has one rather obvious limiting feature. It is not recursive and at each measurement interval, the state vector and the two derivatives, $\dot{x}_1$ and $\dot{x}_2$, must be stored. Therefore, it requires the storage of $(n+2)l+2\times 2n^2$ floating point numbers, before the identification procedure begins. The number of measurements, l, is quite large, and therefore the storage needed is excessive. In the final computation, two matrices with $(2n)^2$ elements will be used for $\mathbf{H}^T\mathbf{H}$ and its inverse.

From the above analysis, the total number of bytes of storage required for the simple system discussed in this chapter is

$$l\times[(4+2)\times 4] + 2\times 4\times 4^2\times 4 = l\times 24 + 512 \qquad (3.19)$$

where each floating point number is assumed to be 4 bytes or 32 bits. Thus, if the least squares algorithm program uses 1000 32 bit words of memory, then with 1 Megabyte of storage, l could be larger than 40,000 measurements.

Computationally, the requirements during the data collection process would be minimal, but during the matrix manipulation operation the process is

extremely intensive. The matrix multiplication of $\mathbf{H}^T\mathbf{H}$ would result in 320.000 multiplies, 320.000 adds, 160.000 index register increments, 320.000 loads, and 64 stores. Assuming 6 μsec/multiply, and 2 μsec for all other operations, the computation of the $\mathbf{H}^T\mathbf{H}$ is completed in 3.52 seconds. Assuming the inversion process, the $[\mathbf{H}^T\mathbf{H}]^{-1}\mathbf{H}^T$ computation, and the multiplication by the extended $\mathbf{z}$ vector all take a similar amount of time, the total identification computation is completed in 14 seconds.

Finally, if data measurements are taken every 0.05 seconds, the memory would be filled in 33.3 minutes, and in 33.6 minutes the identification process would be complete.

H. SUMMARY

The least squares identifier is elegantly simple mathematically, but very memory intensive. For the simple model employed in this chapter, its use is almost feasible. However, as the state size increases, so does the memory requirement. This limitation can easily be overcome with the recursive algorithm that is presented in the next chapter. The recursive scheme used in Chapter IV has all of the benefits of the least squares approach and eliminates the large memory requirement.

The second drawback of the least squares identifier is the need for very low noise state measurement equipment. In Section D it was seen that the standard deviation of the state measurement noise could not exceed 0.5 % of the standard

deviation of the initial conditions. and that the best parameter estimate was achieved in 32 measurements. The system time constant is 1.4 second. which is roughly equivalent to 32 measurements, meaning that the initial conditions have diminished to 37 % of their original value. At this point in time, the measurement signal to noise ratio is

$$\frac{S}{N} = \frac{\sigma_x}{\sigma_w} = \frac{0.037}{0.005} = 7.3 \tag{3.20}$$

or 17.3 dB. This signal to noise ratio is easy to achieve for a sensor of this type. In Chapter IV the signal to noise requirement will be examined in considerable detail in order to determine the requirements on each sensor.

In this chapter very little was said about the sensors. The actual sensor scheme was quite unrealistic, because it was assumed that $\dot{x}_1$ and $\dot{x}_2$ were each measured twice: once for the state vector, and once for the state derivative vector. In the following chapter these measurements will be taken once and used in both vectors. A closer examination of this measurement discrepancy reveals that it has no significant effect on the unknown parameter estimates and only effects the identification of those parameters already assumed known.

IV. KALMAN FILTER IDENTIFIER

A. INTRODUCTION

This chapter deals with the Kalman filter identifier developed in Chapter II. Using a simulation study as in Chapter III, the limitations of the Kalman filter approach will be explored. The recursive nature of the Kalman filter approach makes it much better suited for on-line parameter identification. This will alleviate the excessive storage needed for the least squares approach, but will require considerable computation between system measurements during the identification process.

Returning to the development in Chapter II, and using the more complete noise sources of Chapter III. Equations 2.31 and 2.32 are transformed to

$$\dot{x}(k) = H(k)\theta \tag{4.1}$$

$$z(k) = \dot{x}(k) + v(k) + Bu(k) \tag{4.2}$$

where u(k) and v(k) are random with covariance R_u and R_v. Rewriting the Kalman filter equations to illuminate the changes incorporated above, gives

$$\hat{\theta}(k|k-1)=\hat{\theta}(k-1|k-1) \tag{4.3}$$

$$\hat{\theta}(k|k) = \hat{\theta}(k|k-1) + G(k)[z(k) - H_r(k)\hat{\theta}(k|k-1)] \tag{4.4}$$

$$G(k) = P(k|k-1)H_r^T(k)\left[H_r(k)P(k|k-1)H_r^T(k) + R_v(k) + BR_uB^T\right]^{-1} \tag{4.5}$$

$$P(k|k-1) = P(k-1|k-1) \tag{4.6}$$

$$P(k,k) = [I - G(k)H_r(k)]\,P(k,k-1) \tag{4.7}$$

where $H_r(k)$ is the matrix defined by Equation 2.19 .

It will be assumed that the measurements of both the state and its derivative are available, but possibly corrupted by noise. Applying the three sources of noise in the identification process: u(t) the road noise, w(t) the state measurement noise, and v(t) the state derivative measurement noise, the effect of each source will be analyzed. In this chapter, as in the last, the error measurement ξ_k will be used to determine how effectively the parameters are being estimated, with the minimum acceptable estimates being within 10 percent of the actual values.

B. INITIALIZATION OF THE KALMAN FILTER

In Chapter II the specifics of initializing the Kalman filter were not addressed. These initial values play an important role in the operation of the filter and therefore should not be overlooked. The parameter estimation trajectory for the recursive least squares identifier has been given for the single input single output case as [Ref. 4:p.21]

$$\hat{\theta}(k) = \left[P^{-1}(0) + \sum_{m=1}^{k} \frac{1}{R} h(m)h^T(m)\right]^{-1}\left[P^{-1}(0)\hat{\theta}(0) + \sum_{m=1}^{k} \frac{1}{R} h(m)z(m)\right] \tag{4.8}$$

This equation can be rewritten for the case of multiple outputs as

$$\hat{\theta}(k) = \left[P^{-1}(0) + \sum_{m=1}^{k} H^T(m)R^{-1}H(m)\right]^{-1}\left[P^{-1}(0)\hat{\theta}(0) + \sum_{m=1}^{k} H^T(m)R^{-1}z(m)\right] \tag{4.9}$$

Substituting Equation 4.2 for $z(k)$ gives

$$\hat{\theta}(k) = \left[P^{-1}(0) - \sum_{m=1}^{k} H^T(m)R^{-1}H(m)\right]^{-1} \tag{4.10}$$

$$\left[P^{-1}(0)\hat{\theta}(0) + \sum_{m=1}^{k} H^T(m)R^{-1}H(m)\theta - \sum_{m=1}^{k} H^T(m)R^{-1}v(m)\right]$$

where $v(k)$ denotes both disturbance noise and measurement noise without loss of generality. If Equation 4.10 is rearranged, collecting like terms, then

$$\hat{\theta}(k) = [I + P(0)P^{-1}(k)]^{-1}\hat{\theta}(0) + [P(k)P^{-1}(0) + I]^{-1}\theta + \tag{4.11}$$

$$\left[P^{-1}(0) + P^{-1}(k)\right]^{-1} \sum_{m=1}^{k} H^T(m)R^{-1}v(m)$$

where the equality

$$P(k) = \left[\sum_{m=1}^{k} H^T(m)R^{-1}H(m)\right]^{-1} \tag{4.12}$$

has been used to simplify Equation 4.11 . The third term on the right side of Equation 4.11 is related to the state measurement noise and will be discussed in Section D of this chapter. The other two terms relate to initialization and represent the identification trajectory without state measurement noise.

In order for $\hat{\theta}(\infty)$ to approach θ, the first term on the right side of Equation 4.11 must approach 0 and the element multiplied with θ in the second term must become an identity matrix. The first requirement can be satisfied by setting $\hat{\theta}(0)$ equal to zero. The second requirement, restated as

$$\lim_{k\to\infty}[P(k)P^{-1}(0) + I]^{-1}\theta \to \theta \tag{4.13}$$

is fulfilled if

$$\lim_{k\to\infty} P(k) \to 0 \tag{4.14}$$

However, since the selection of $P^{-1}(0)$ very small would help reduce $P(k)P^{-1}(0)$ more

quickly. it is commonly suggested that [Ref. 6:p. 21]

$$P(0) = Ic \tag{4.15}$$

where c is some large constant.

In the case of the identification process, since the values of the parameters are unknown, the exact value of the P matrix cannot be known either. In fact the best that can be done is to initialize the P matrix diagonally with values that are approximately the square of the parameter values one anticipates. Thus if nothing is known about the system except that the general range of the parameters is 10 to 1000, then a diagonal matrix with values of 10^4 may be a best first approximation.

Figure 4-1 illustrates the effect of choosing the P matrix incorrectly. The first two curves marked by squares and circles, show the effect of P(0|-1) initialized too small. The identifier assumes it is closer to the actual value than it is and therefore allows only small changes in the parameters This produces very slow convergence to the correct values.

The curve marked with triangles is obtained when the diagonal of the P matrix is initialized with 1000. This will be the initialization scheme used throughout this chapter. since it appears to have both good initial response and noise suppression in the later stage of the identification process.

The curve marked with "+" symbols shows that increasing the diagonal elements of the P matrix to 10,000 produces still further improvement in the

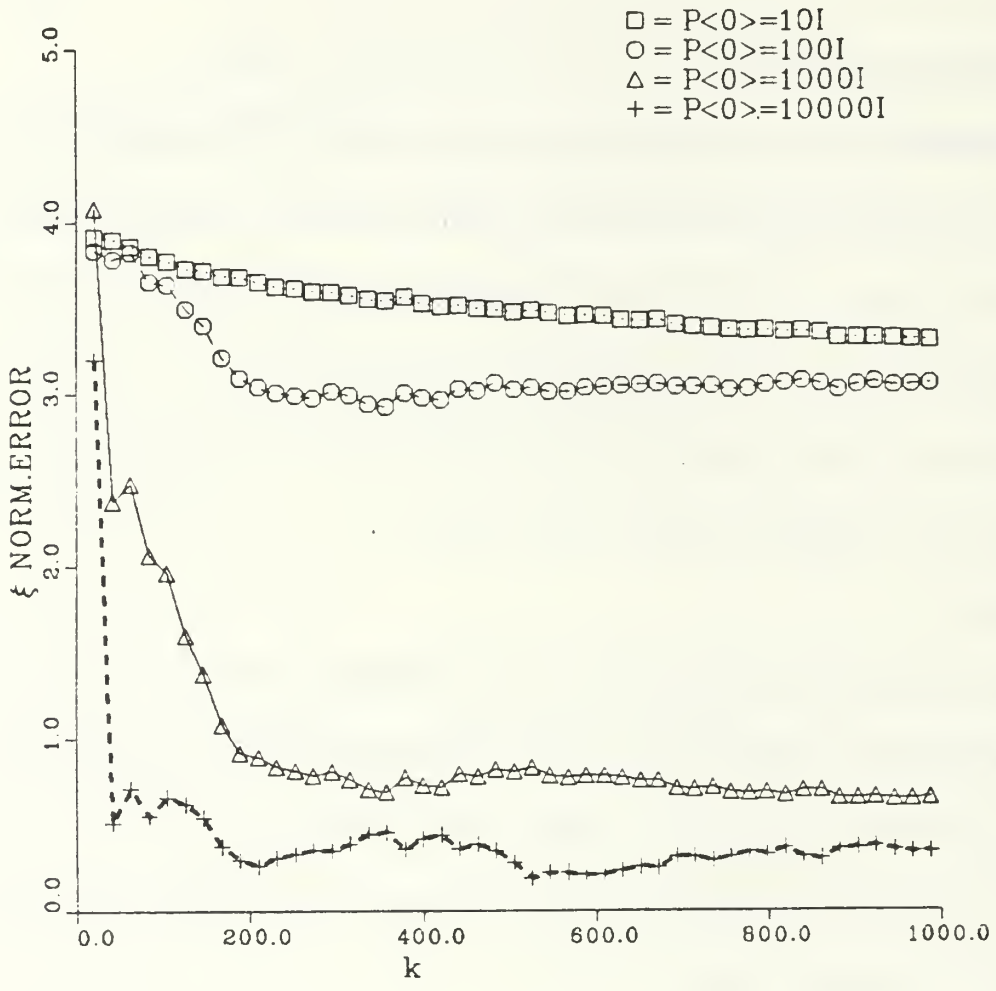

Figure 4-1 Parameter estimation vs. number of iterations for four different initializations of the error covariance matrix.
Initial conditions, σ =0.1
Road noise, σ =0.1
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.005.

convergence rate of $\hat{\theta}$. However, other experiments showed a less significant effect, and a value of 1000 was therefore used.

The R matrix in the gain matrix equation, Equation 4.5, is in this case the sum of two matrices: the state derivative noise covariance matrix R_e, and the road noise covariance matrix BR_uB^T. The R matrix controls the size of the gain matrix, reflecting the confidence the identifier has in the measurements it is receiving. If the confidence is high, the R matrix is small and the gain matrix is large, which in turn decreases the error covariance matrix P more rapidly. Therefore, the uncertainty due to the road noise only influences the parameters of the second row in Equation 3.11 .

C. SYSTEM EXCITATION

It has been stated that the Kalman filter identifier described here is simply a recursive weighted least-squares algorithm[Ref. 7:p. 252]. This being true, then the use of this recursive approach will alleviate some of the computational difficulties observed in Chapter III. Also, since the least squares identifier is able to correct for noise introduced into the state derivative vector, given the proper number of iterations, the error in the parameter estimates could be made arbitrarily small, with an arbitrarily large noise source.

The recursive identification process will make it possible to find the parameters with only road noise as system excitation. However, as can be seen from Figure 4-2, the identification requires more than 40,000 iterations to obtain

satisfactory values, for those parameters corrupted by the road noise; i.e., the lower mass parameter estimates. Figure 4-2 shows the identification error measure ξ_k as k increases. With no initial conditions the identification proceeds slowly but does converge to the proper values as k becomes large. The upper mass parameters, on the other hand, reach their final value within a much shorter time, usually less than 1000 iterations, depending upon the value of σ_r. Notice the flat section of the curve between 15,000 and 20,000 iterations. This plateau is 5000 iterations wide and could be mistaken for some form of parameter bias if only 20,000 iterations were done. There is an obvious difference between plateauing and bias in that a bias will actually increase the parameter estimation error consistently for many iterations, while the plateau will remain relatively flat with small variations up and down.

The slow convergence of parameter values observed with road noise as the only excitation cannot be accelerated with the inclusion of initial conditions alone. Along with initial conditions, the road noise must be small in order to increase the filter's confidence in z_2, the measured value for x_2, and allow the identification process to proceed more rapidly. Figure 4-3 demonstrates this quite clearly. The curve marked with squares has initial conditions slightly larger than the standard deviation of the road noise and consequently cannot converge to the proper parameter estimates (for the second row parameters) before the initial conditions have diminished and the identifier has only road road noise for excitation while identifying the second row parameters embedded in the road noise. The curve

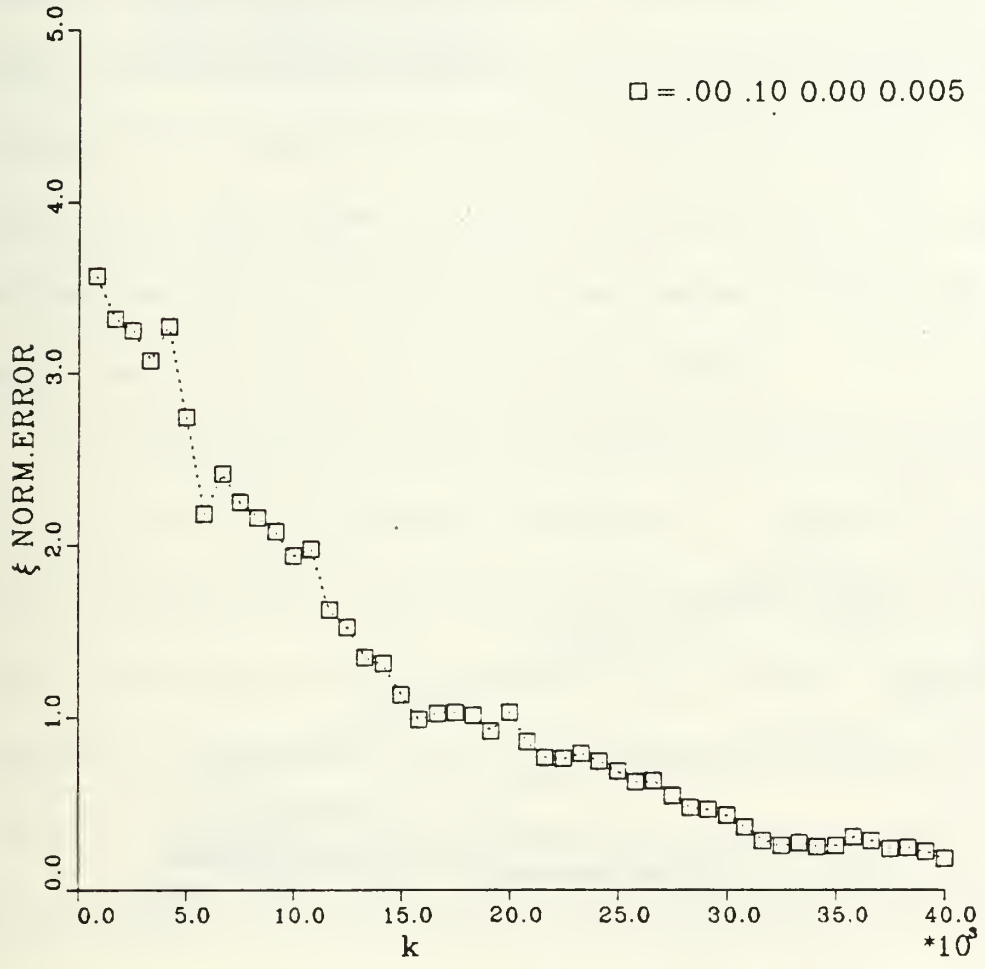

Figure 4-2 Parameter estimation vs. number of iterations with road noise as the only system excitation.
Initial conditions, σ =0.0
Road noise, σ =0.1
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.005.

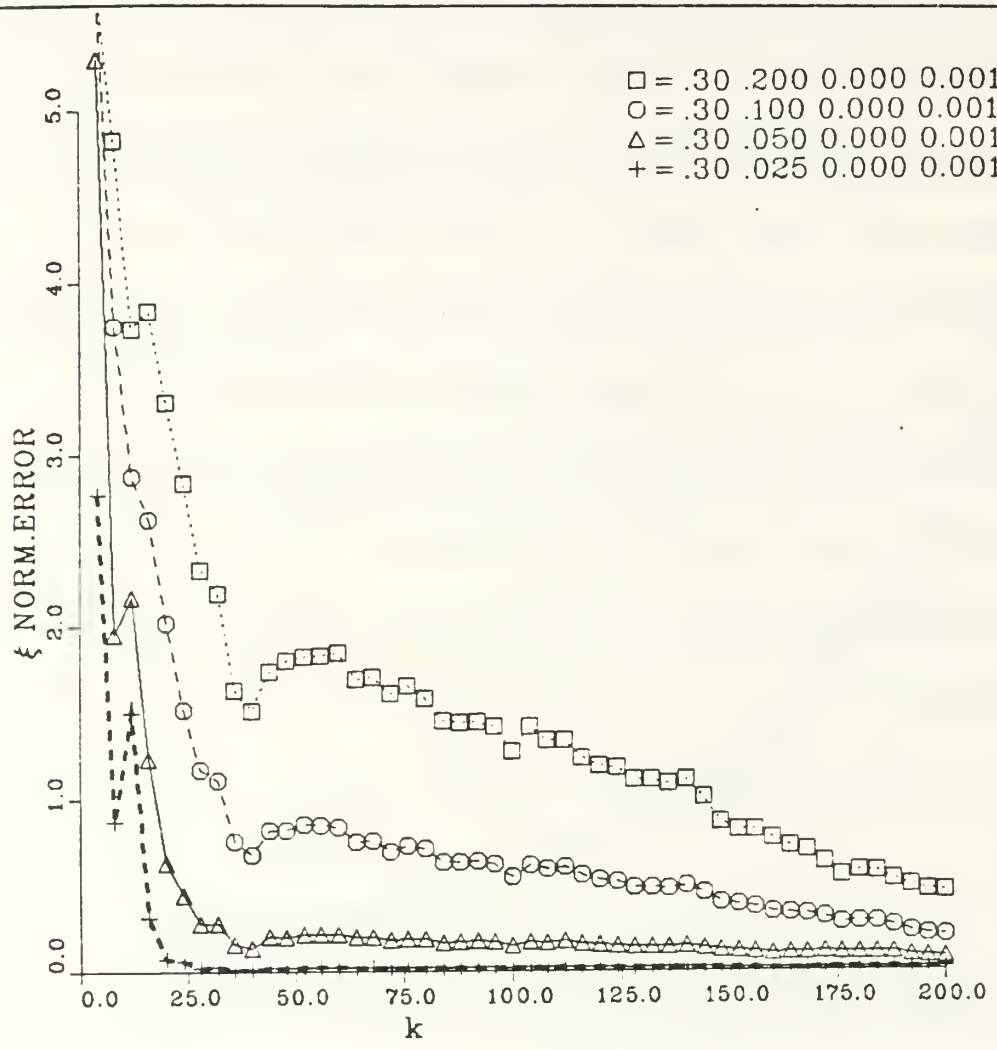

Figure 4-3 Parameter estimation vs. number of iterations with initial conditions and various levels of road noise.
Initial conditions, σ =0.3
Road noise, σ =0.20, 0.10, 0.05, 0.025
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.001.

labeled with circles shows the same leveling off after about 32 iterations. This is because the fourth row parameter estimates have converged to the correct values. while the second row which is covered by noise could not converge as quickly. In the bottom curve marked by "+" symbols both the second and the fourth row parameters converge before the initial conditions diminish. The second row estimates were able to converge in this case because the road noise was much smaller than the excitation due to the initial conditions.

Since 40.000 iterations to reach an acceptable result is unappealing, the use of an input pulse seems appropriate to decrease the required observation time. Several important factors affect the Kalman filter's performance with an input pulse as the source of excitation. The size and duration of the pulse are very influential in determining the effect of the pulse on the identifier. The energy absorbed by the system must be sufficiently large to overcome the corruption of the state derivative vector by the road noise. More importantly, the width of the pulse must be less than one measurement cycle in order to prevent correlation between the input and the output and thus produce a biased estimate. Figure 4-4 shows the identification error when a pulse of various durations is applied to the vehicle, and the other sources of noise are very small. When the pulse is present for only one measurement of the system then the identifier produces acceptable estimates within a few iterations. However, if the pulse is present during more than one measurement cycle then it produces an estimate which will require 20.000 iterations to reach acceptable results. In the case of the pulse width of 3

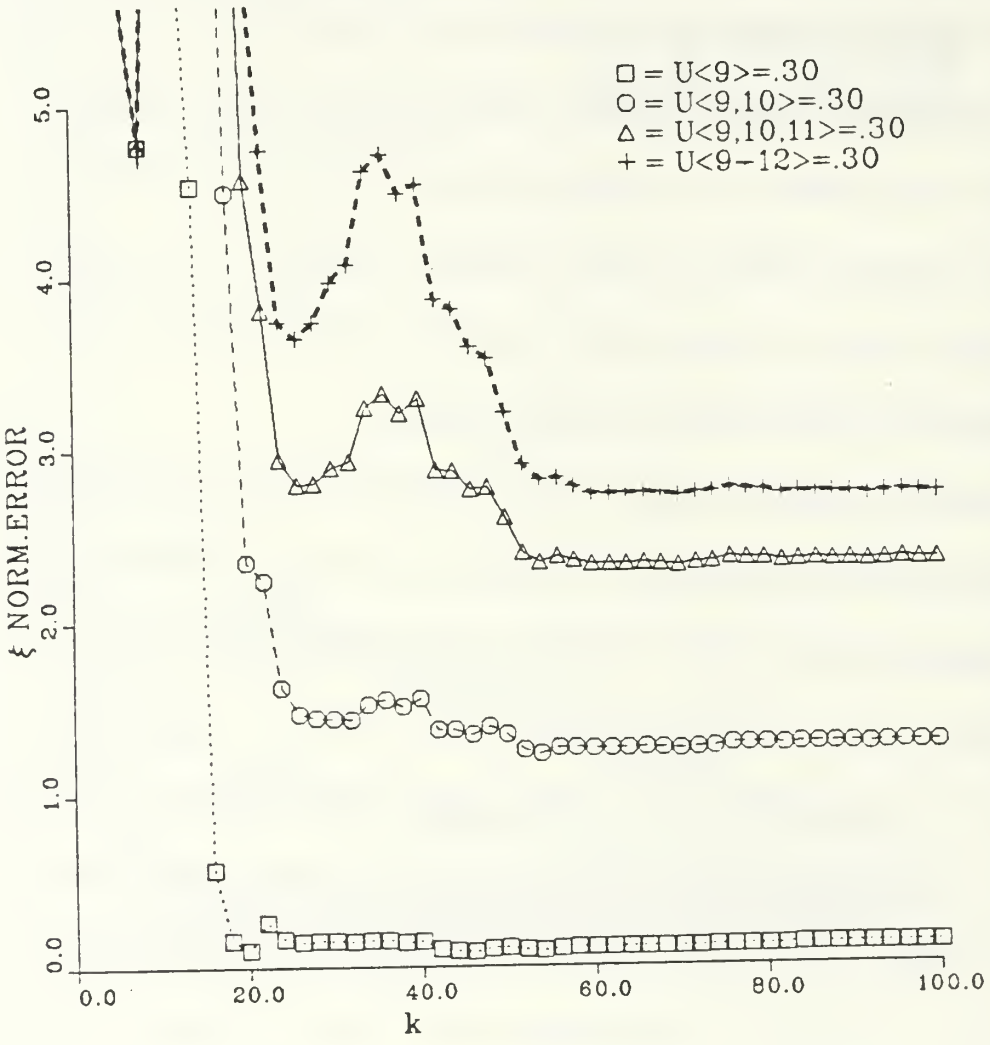

Figure 4-4 Parameter estimation error for various input pulse durations shown in the legend.
Initial conditions, σ =0.01
Road noise, σ =0.001
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.01.

measurement cycles, another 30.000 iterations will be needed to find the correct parameters. In the final curve, where the pulse was present during 4 measurement cycles almost nothing is gained, and identification will take another 35.000 iterations.

D. STATE MEASUREMENT NOISE

The least squares identifier, and more specifically its recursive form, the Kalman filter identifier, is derived with the assumption that the measurement of the state is perfect. Therefore, the bias resulting from this noise source is unavoidable. Returning to the parameter estimate trajectory in Equation 4.11, the last term on the right side of the equation is the state measurement noise bias. It was shown in Equations 2.19 and 2.20 that when the state measurements are noisy, $H_r(k)$ and $v(k)$ become correlated.

The third term of Equation 4.11 can be rewritten as

$$[P^{-1}(0) + P^{-1}(k)]^{-1} \sum_{m=1}^{k} H_r^T(m) R^{-1} v(m) = \left[\frac{1}{k} P^{-1}(0) + \frac{1}{k} P^{-1}(k)\right]^{-1} \frac{1}{k} \sum_{m=1}^{k} H_r^T(m) R^{-1} v(m) \quad (4.16)$$

where the left side is multiplied by $\frac{k}{k}$. Clearly, as $k \to \infty$ then

$$\lim_{k \to \infty} kP(k) E[H_r^T(k) R^{-1} v(k)] \approx cE[H_r^T(k) R^{-1} v(k)] \quad (4.17)$$

where c is a constant, since $P(k)$ is essentially a linearly decreasing function of k. This has several implications. If significant state measurement noise is present, then it is necessary to obtain the parameter estimates in as few iterations as possible, because the correlation bias is a cumulative process. Rapid identification

implies large initial excitation which would have the added quality of producing a rapidly decreasing $P(k)$ matrix causing c in Equation 4.17 to be small.

Figure 4-5 shows this cumulative bias quite clearly, for three different values of measurement noise. Notice the two lower curves marked by triangles and "+" symbols, which show the error measurement ξ_k for σ_w equal to 1 % of the standard deviation of the road noise, and no state measurement noise, respectively. These curves indicate that for small values of σ_w the identification process is only slightly biased, but even for this small value of measurement noise there is a cumulative bias. However, if σ_w is 2% of the standard deviation of the input road noise, the identification is seriously impeded, as shown by the curve marked with circles. The curve marked with squares is for σ_w equal to 4 % of σ_u, and it is clear that in 10000 iterations no improvement has been made in the normalized estimation error ξ_k. In fact for values of σ_w larger than 2 % of σ_u, the parameter estimation is divergent.

Figure 4-6 is an illustration of the bias accumulated over 40,000 iterations when σ_w is 2 % of σ_u. Figure 4-6 should be compared with Figure 4-2, where after 40,000 iterations the difference between the two error measures ξ_k is about 1.

Since the affect of σ_w is so profound on the estimation process, it seems appropriate to obtain a signal to noise ratio for the state vector which will achieve acceptable results. Recall that the standard deviation of the output of a linear system is given by [Ref. 8:p. 184]

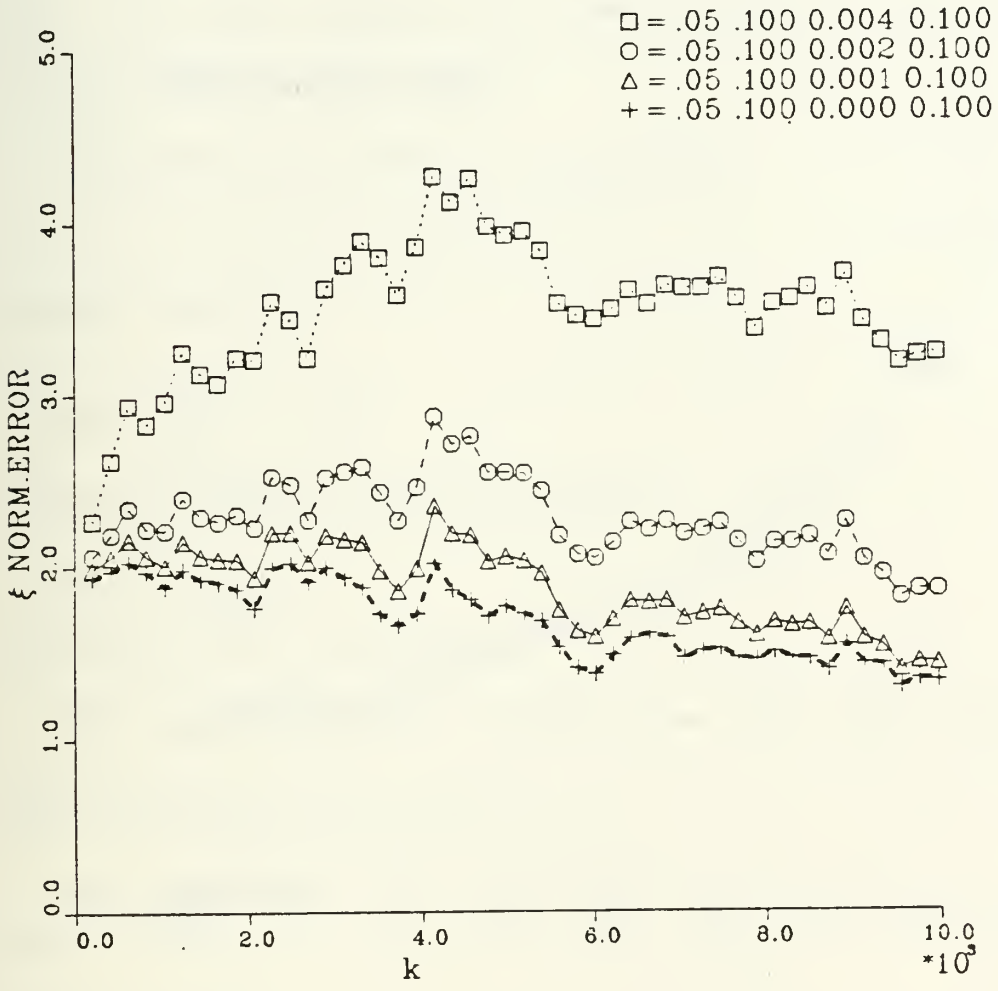

Figure 4-5 Parameter estimation vs. number of iterations with road noise as excitation and various level of state measurement noise.
Initial conditions, σ =0.05
Road noise, σ =0.1
State measurement noise, σ =0.004, 0.002, 0.001, 0.0005
State derivative measurement noise, σ =0.1.

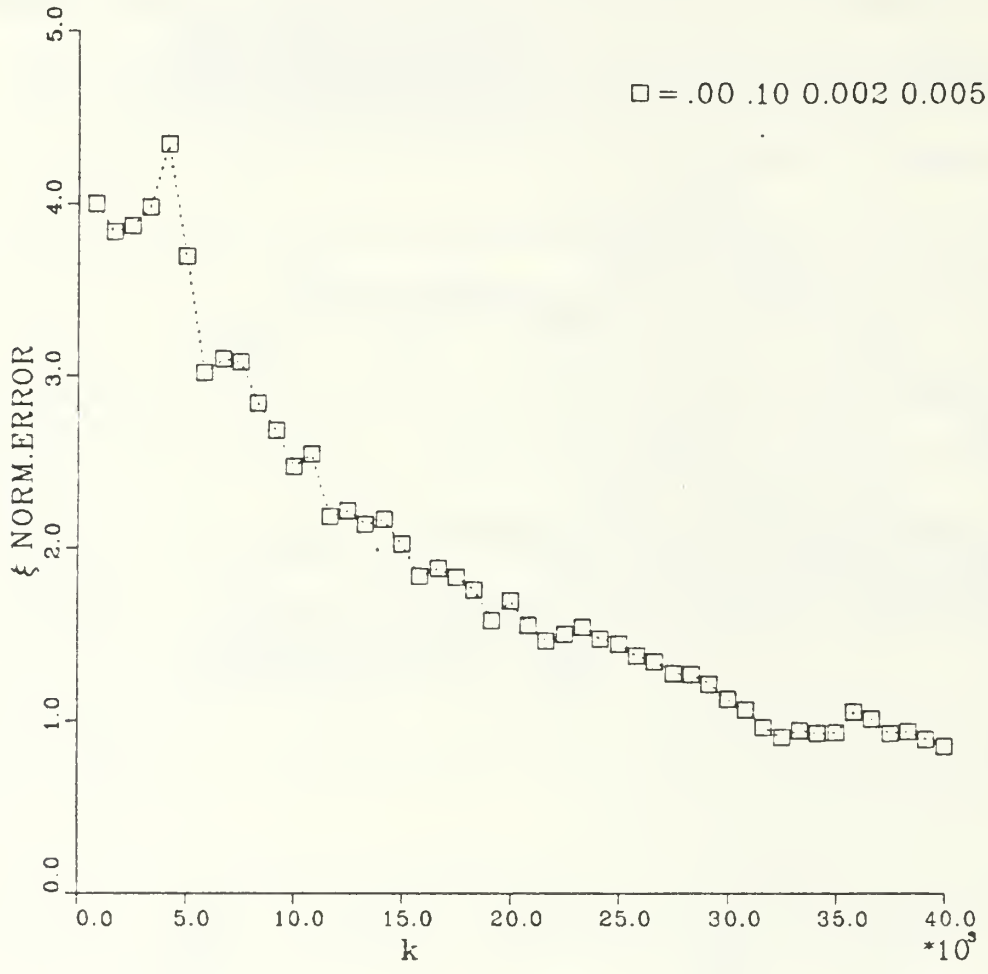

Figure 4-6 Parameter estimation vs. number of iterations with road noise as the only system excitation.
Initial conditions, σ =0.0
Road noise, σ =0.1
State measurement noise, σ =0.002
State derivative measurement noise, σ =0.005.

$$\sigma^2_{out} = \frac{1}{\pi}\int_0^{\infty} |H(\omega)|^2 S_u(\omega)\, d\omega \tag{4.18}$$

where $|H(\omega)|$ is the magnitude of the transfer function and $S_u(\omega)$ is the power spectral density of the input signal.

The transfer function H(s) can be obtained from

$$X(s) = (sI-A)^{-1}x(0) + (sI-A)^{-1}BU(s) \tag{4.19}$$

or if x(0) is zero. then

$$H(s) = \frac{X(s)}{U(s)} = (sI-A)^{-1}B \tag{4.20}$$

The magnitude squared of $H(s)$ is derived from

$$|H(s)|^2 = H(s)H(-s) \tag{4.21}$$

and replacing s by jw produces $|H(w)|^2$.

The system is excited by a discrete noise source whose autocorrelation function can easily be derived. Since

$$R_u(\tau) = E[u(t)u(t-\tau)] \tag{4.22}$$

then

$$R_u(0) = E[u^2] = \sigma_u^2 \tag{4.23}$$

and

$$R_u(\tau) = 0 \quad \text{for} \quad |\tau| > T \tag{4.24}$$

so that

$$R_u(\tau) = \begin{cases} \sigma_u^2\left[1 - \dfrac{|\tau|}{T}\right] & ,|\tau| \le T \\ 0 & ,|\tau| > T \end{cases} \tag{4.25}$$

where T is the pulse width of the discrete noise source [Ref. 8:p. 121]. Figure 4-7

depicts the autocorrelation function of the discrete road noise. The power spectral density of this process is. [Ref. 8:App. E]

$$S_u(\omega) = \sigma_u^2 \left[T \frac{\sin^2(\omega T/2)}{(\omega T/2)^2} \right] \tag{4.26}$$

Therefore, the standard deviation of the output is

$$\sigma_i^2 = \frac{T\sigma_u^2}{\pi} \int_0^\infty |H_i(\omega)|^2 \left[\frac{\sin(\omega T/2)}{(\omega T/2)} \right]^2 d\omega \tag{4.27}$$

where i denotes the associated state variable $i = 1,2,3,4$. Using numerical integration. the output standard deviations σ_z of the state vector are

$$\sigma_z = \begin{bmatrix} .386 \\ 1.84 \\ .402 \\ 1.03 \end{bmatrix} \sigma_u \tag{4.28}$$

The observed standard deviations of the state vector during simulation of the simplified model for four different ensembles of road noise and average over 10000 measurements are as follows:

$$\frac{\sigma_z}{\sigma_u} = \begin{bmatrix} .366 \\ 1.78 \\ .368 \\ .941 \end{bmatrix}, \begin{bmatrix} .390 \\ 1.83 \\ .403 \\ 1.05 \end{bmatrix}, \begin{bmatrix} .352 \\ 1.76 \\ .372 \\ .922 \end{bmatrix}, \begin{bmatrix} .372 \\ 1.80 \\ .388 \\ 1.03 \end{bmatrix} \tag{4.29}$$

These results are in excellent agreement with the analytic results of Equation 4.28 . Some interesting observations can be made from these results. Note that for a given level of road noise σ_u, the wheel velocities σ_{x_2} are almost twice as large as the input noise values. Written in equation form this becomes

$$\sigma_{z_2} = 1.84\sigma_u \tag{4.30}$$

The vehicle body velocities reflect the input noise almost exactly. Also, the

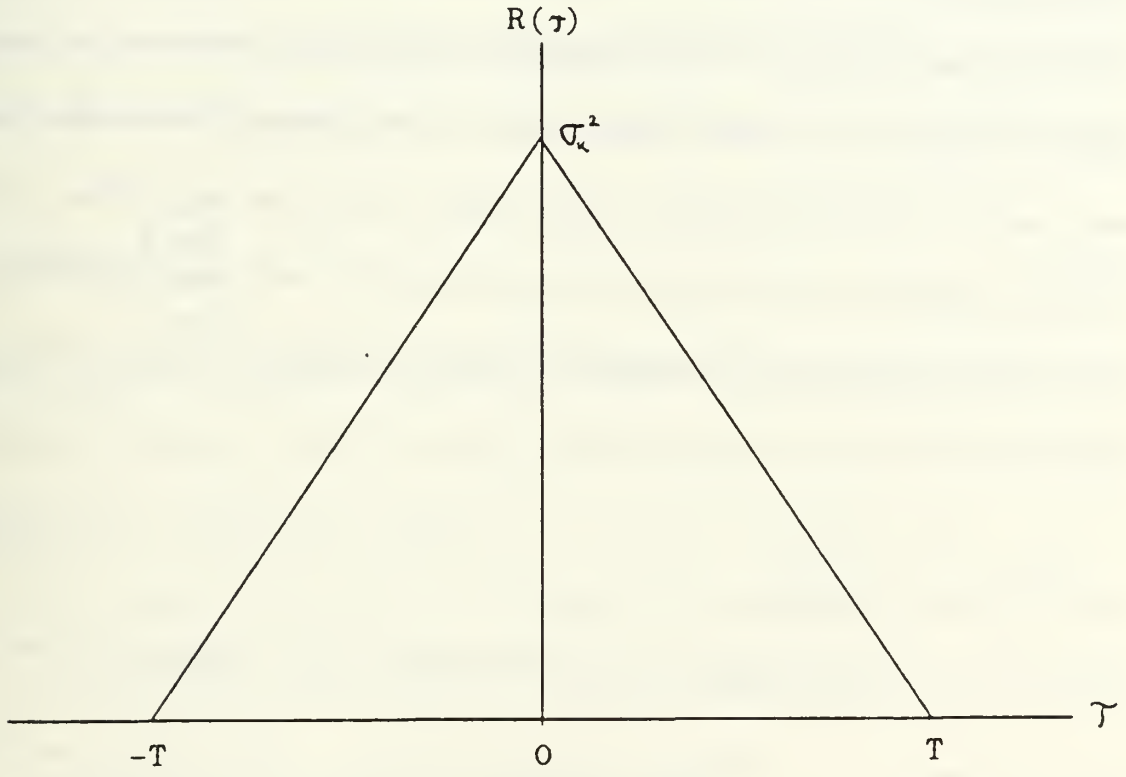

Figure 4-7 Autocorrelation function for discrete road noise with T =0.05

position change of the wheels and the body attenuates the input road noise significantly.

The results in Equation 4.28, give the standard deviation of the state vector with only road noise as input to the simplified vehicle. Evidently, if measurement noise is applied uniformly to the component measurements of the state, some of them will suffer more corruption than others. In order to dissect the identification error caused by the noise added to each element of the state vector when σ_w is uniformly applied to each element, the noise was added to only one element of the state vector at a time during simulation. The signal to noise ratio at each sensor is given by

$$\frac{S_i}{N} = \eta_i = \frac{\sigma_i}{\sigma_w} = \frac{\sigma_i}{\sigma_u}\frac{\sigma_u}{\sigma_w} \tag{4.31}$$

where the i denotes the element of the state vector and $\frac{\sigma_i}{\sigma_u}$ is given in Equation 4.28 . Therefore, if the road noise standard deviation σ_u is 25 times greater than the state measurement noise standard deviation σ_w, then the signal to noise ratio at each sensor is

$$\frac{S_i}{N} = \begin{bmatrix} .386 \\ 1.84 \\ .402 \\ 1.03 \end{bmatrix} 25 = \begin{bmatrix} 9.65 \\ 46.0 \\ 10.0 \\ 25.8 \end{bmatrix} = \begin{bmatrix} 19.7 \\ 33.3 \\ 20.0 \\ 28.2 \end{bmatrix} dB. \tag{4.32}$$

Define $e_i(k)$ as

$$e_i(k) = \xi_{w_i}(k) - \xi_0(k) \tag{4.33}$$

where $\xi_{w_i}(k)$ is the error measure with state measurement noise added to only the

ith element of the state vector and $\xi_0(k)$ is the error measure after k iterations without any state measurement noise applied.

By adding state measurement noise to only one element of the state vector at a time, the effect of the added noise on each element can be seen. Three simulations were run, each with a different ensemble of noise and the average error caused by adding measurement noise to each element of the state vector was

$$e_i(8000) = \begin{bmatrix} .605 \\ .001 \\ .705 \\ .032 \end{bmatrix} \tag{4.34}$$

Notice here that the sensor with a 33.3 dB signal to noise ratio produces only a very small bias error at 8000 iterations, while the sensor with a sensitivity of 28.2 dB. produces a bias 30 times as large.

More analysis could be done on this topic, but it is sufficient to conclude that the necessary signal to noise ratio for each sensor is about 33 dB.

E. STATE DERIVATIVE MEASUREMENT NOISE

The Kalman filter identifier can identify the parameters of the simplified vehicle through the state derivative noise introduced by the random vector $v(k)$. Since this algorithm is a weighted least squares method, and is recursive, the size of σ_v determines only the length of time required to identify the parameters. The accuracy of the estimates can be arbitrarily close to the actual values given the appropriate number of iterations.

However, should the standard deviation of the vector $v(t)$ be underestimated the result will be a biased estimate of the parameters. Because the $P(k)$ matrix diminishes too rapidly, the parameters cannot be estimated before the gain matrix becomes very small and no significant change in the parameter estimates can be expected. This bias leads to a very serious limitation of the Kalman filter identifier, which requires either extended observation time or acceptance of a bias produced by the excess road noise. The long observation time is required since the standard deviation of the road noise must be estimated large to allow for large road noise inputs, which in turn will cause very slow convergence.

The need to eliminate this excess excitation bias points to a reinitialization scheme, which will restart the gain matrix and allow identification to continue. This scheme is used whenever the filters perceived error in the error covariance matrix becomes exceedingly small, while the actual measured error given by

$$e(k) = z - H\hat{\theta} = [\, e_1, \ldots, e_n(k) \,]^T \tag{4.35}$$

remains relatively large. The reinitialization of the identifier will eliminate this form of bias, but will not improve the required estimation time directly. The advantage of choosing the R matrix smaller than the actual value of the noise covariance is obvious from Figure 4-8. What is less obvious is exactly how to determine when to reinitialize and how much to increase the diagonal of the P matrix. Figure 4-8 demonstrates the effectiveness of a reinitialization procedure that multiplies the diagonal elements of the P matrix by ten and doubles the elements of the R matrix when the sum of the diagonal elements of the P matrix

become smaller than ten times the sum of the error vector in Equation 4.35 . This scheme is given in equation form as

$$10\sum_{i=1}^{n} |e_i(k)| > \sum_{j=1}^{n^2} P_{jj}(k) \qquad (4.36)$$

Figure 4-8 illustrates the advantage of a reinitialization procedure. The curve marked with squares was produced using an identifier with the correct values in the noise covariance matrix R. The other three curves denote identifiers which initially had smaller values in the R matrix, as shown in the legend. The reinitialization scheme is visible at several point along the curve marked by "+" symbols. The most obvious reinitialization happens at about 5000 iterations, when both the curve marked by triangles and the one marked with "+" symbols are reinitialized. It seems evident that reinitialization can be applied whenever the state derivative noise is larger than anticipated, or very little is known about the noise source, as in the case of road noise.

F. COMBINED NOISE

The combination of all the noise sources results in essentially the same identification performance as would be expected from the worst individual source, just as it did for the least squares approach of Chapter III. Figure 4-9 shows that the effect of state derivative measurement noise is insignificant when compared with the effect of even very small values of state measurement noise. The correlation of the $H(k)$ matrix and the input noise $v(k)$ as described in Equation 2.20 is the cause of the bias due to state measurement noise. In the case where

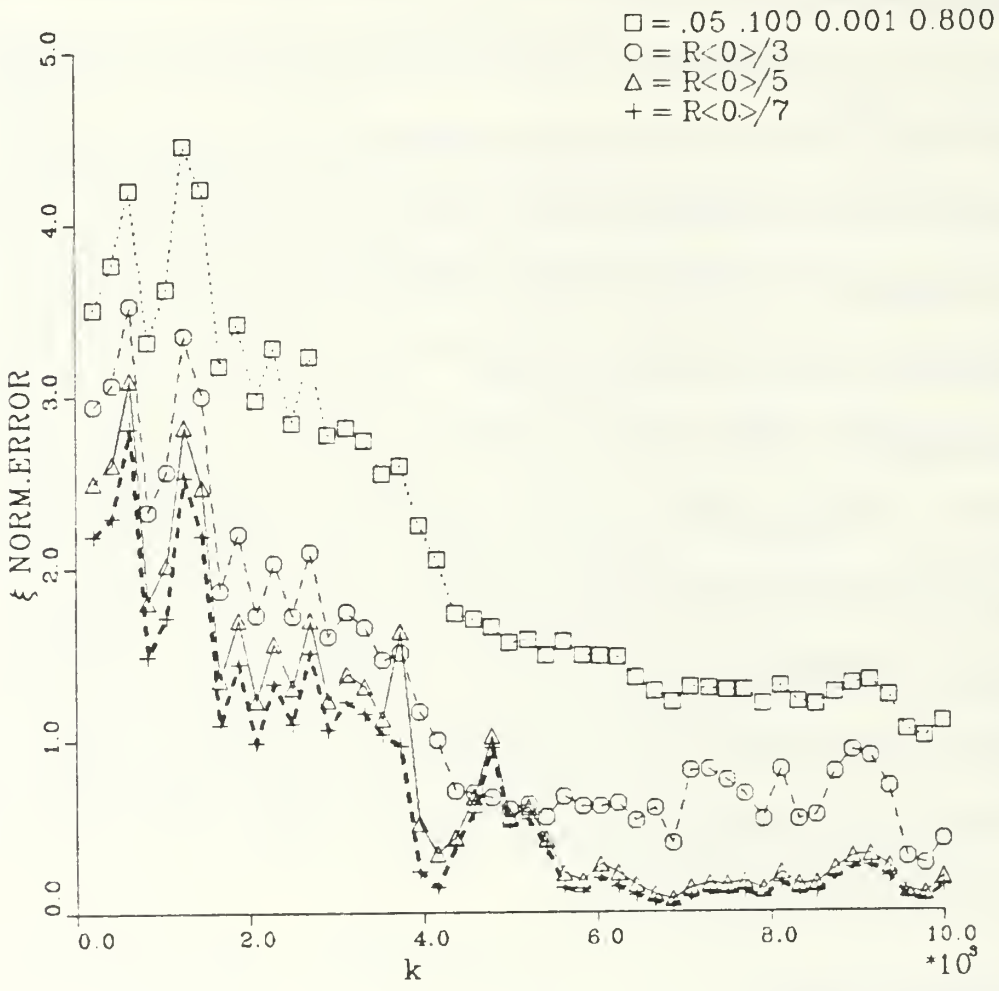

Figure 4-8 Parameter estimation vs. number of iterations with incorrect selections of the road noise covariance matrix using reinitialization to continue the identification process.
Initial conditions, σ =0.05
Road noise, σ =0.1
State measurement noise, σ =0.001
State derivative measurement noise, σ =0.8.

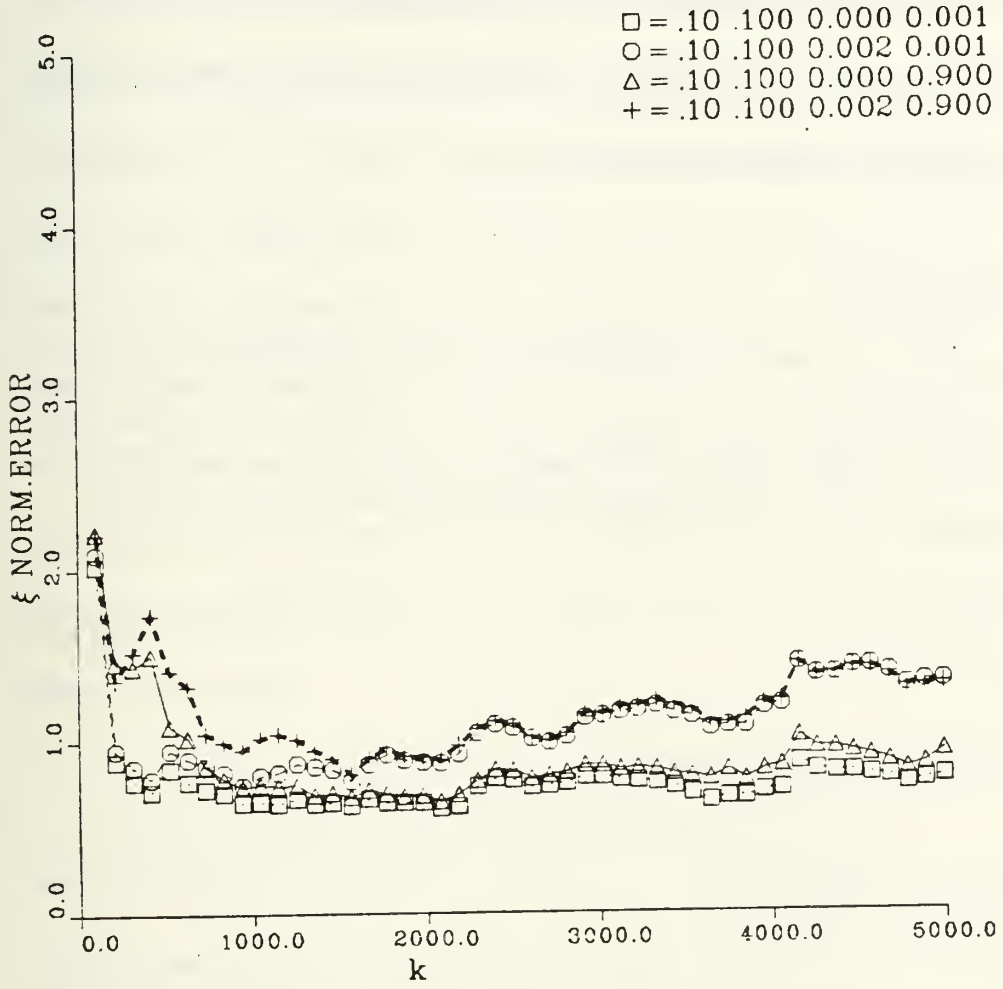

Figure 4-9 Parameter estimation vs. number of iterations for various levels of statemeasurement noise and state derivative measurement noise.
Initial conditions, σ =0.1
Road noise, σ =0.1
State measurement noise, σ =0.000, 0.002, 0.000, 0.002
State derivative measurement noise, σ =0.001, 0.001, 0.900, 0.900.

only σ_t was applied, as in the curve marked by triangles, it has very little effect on the identifier. When state measurement noise is present however, the error introduced is quite significant. When both noise sources are applied the outcome is equivalent to only state measurement noise.

G. INSTRUMENTATION

In Chapter III it was assumed that the state vector measurement and the state derivative vector measurement were obtained from separate instruments. In this chapter, a more realistic instrumentation philosophy was employed. Since

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} \tag{4.37}$$

then the noisy measurements are given by

$$\begin{pmatrix} z_1 \\ z_3 \end{pmatrix} = \begin{pmatrix} \dot{x}_1 + v_1 \\ \dot{x}_3 + v_3 \end{pmatrix} = \begin{pmatrix} x_2 + w_2 \\ x_4 + w_4 \end{pmatrix} = \begin{pmatrix} r_2 \\ r_4 \end{pmatrix} \tag{4.38}$$

Thus, when only one measurement is taken of each physically distinct component of the state vector, it follows that

$$\begin{pmatrix} v_1 \\ v_3 \end{pmatrix} = \begin{pmatrix} w_2 \\ w_4 \end{pmatrix} \tag{4.39}$$

This does not significantly affect the results of Chapter III, for the reasons given in the summary of that chapter, but it does reduce the number of measurements of the system from 8 to 6.

The measurements taken on each mass are the acceleration $\ddot{x}$, the velocity $\dot{x}$, and the position x. While the acceleration measure can be allowed to be noisy, the velocity and the position must have a signal to noise ratio of 33 dB. The acceleration will no doubt be obtained from an accelerometer, but the velocity and position measurements could be obtained in a variety of ways. Since integration is inherently a noise reducing process, it should be possible to integrate the acceleration to obtain the velocity and integrate the velocity to obtain the position. This could be done with analog or digital techniques and the choice might very well depend on the computational load on the microprocessor used in the identifier. Some drift correction would of course be required over long measurement times.

H. COMPUTATIONAL AND STORAGE REQUIREMENTS

The Kalman filter identifier places a large burden on the microcomputer during its operation. Table I presents a rough calculation of the typical cycle time for the 4 state, 6 measurement Kalman filter examined in this chapter. Adding up the times in the bottom row of this table gives a total cycle time of 3622 μ seconds. While the numbers used for the time required for each operation are somewhat arbitrary, they are typical, and thus it seems quite possible to process a measurement every 0.05 seconds, as was done in the simulations of this chapter. The storage requirement for this algorithm was 352 bytes of memory. This is a

Table I. Cycle time for the Kalman filter identifier

to compute	operation				
	mult	divid	add	inc/load	store
G=PH[HPH+R]	80	2	56	158/104	30
P=[I-GH]P	160	0	104	160/152	104
x=x+Ge	16	0	18	35/ 46	18
total ops	256	2	178	353/302	152
time/op (μsec)	6	8	2	2	2
total time	1536	16	356	1310	304

considerable improvement over the least squares identifier requirement of 1 Megabyte.

It is quite possible that this identifier could be implemented on a very small circuit board with only the microprocessor. a 1 kilobyte ROM chip. 512 bytes of static ram and 6 buffer chips to receive the discrete state and state derivative measurements.

I. SUMMARY

The Kalman filter identifier has several advantages over the least squares approach, with very few disadvantages. Because the Kalman filter is a recursive weighted least squares algorithm, it requires less memory. More importantly, since this approach is a Kalman filter, it provides an organized and meaningful way of choosing the weightings and the initial conditions. The identifier is capable of finding the parameters even when considerable road noise and state derivative measurement noise are present. This approach is not capable of producing an unbiased estimate of the parameters in the presents of state measurement noise. The signal to noise ratio of the sensors must be better than 33 dB in order to produce an unbiased estimate of the parameters. This signal to noise ratio is fairly close to the one found in Chapter III, using a very simple approach. The requirement for near perfect state measurement is a limitation of both the least squares and Kalman filter approachs. This limitation, though constraining, is not unacceptable, due to the recent advances in accelerometers and ring laser gyros. It will make the use of expensive hardware a necessity, however. In the next chapter, the stochastic gradient approach is evaluated. The stochastic gradient approach promises to remove the need for perfect state measurement and still produce an unbiased estimate.

V. STOCHASTIC GRADIENT IDENTIFIER

A. INTRODUCTION

The stochastic gradient identifier is perhaps the simplest algorithm that can be applied for parameter identification. This simplicity has the price of slower convergence. In this chapter the stochastic gradient approach will be analyzed using a simulation study and the same simple vehicle model as in the last two chapters. This algorithm promises to eliminate the need for perfect state measurements which is a very desirable feature.

Recall from Chapter II, that for the system defined by

$$z(k) = H_r(k)\theta + v(k) \tag{5.1}$$

where $v(k)$ represents both disturbance and measurement noise, and $H_r(k)$ is defined in Equation 2.19, that the stochastic gradient algorithm is

$$\hat{\theta}(k+1) = [I + R(k)\Sigma_N]\hat{\theta}(k) + R(k)H_r^T(k)\tilde{z}(k) - R(k)E[H^T(k)v(k)] \tag{5.2}$$

This equation was derived under the assumption that all three noise sources used in Chapters II and III would be present, and therefore the algorithm will produce an unbiased estimate of the parameters of the vehicle in the presence of: $u(t)$ the road noise, $w(t)$ the state measurement noise, and $v(t)$ the state derivative measurement noise.

B. BIAS ELIMINATION

If Equation 5.2 is broken up into components associated with either parameter identification or bias elimination the result is

$$\hat{\theta}(k+1) = \hat{\theta}(k) + R(k)H_r^T(k)\tilde{z}(k) \tag{5.3}$$

$$+ R(k)\Sigma_N\hat{\theta}(k) - R(k)E[H^T(k)v(k)]$$

The two elements on the second line of Equation 5.3 are used to reduce the bias: the first uses Σ_N, the covariance matrix of the state measurement noise N(k), and the second uses the input-output cross correlation $E[H^T(k)v(k)]$.

The value of the state measurement noise covariance matrix, Σ_N is easily obtained from

$$\Sigma_N = E[N^T(k)N(k)] \tag{5.4}$$

which is a block diagonal matrix, with n blocks of $n \times n$ elements equal to σ_w^2.

The value of the input-output crosscorrelation matrix is somewhat more difficult to obtain, but in this simplified model with a single input which is white noise the crosscorrelation is

$$E[H^T(k)v(k)] = E[H^T(k)]\ E[v(k)] = 0 \tag{5.5}$$

because $v(k)$ is assumed to be zero mean. This simplification leads to the algorithm

$$\hat{\theta}(k+1) = [I + R(k)\Sigma_N]\hat{\theta}(k) - R(k)H_r^T(k)\tilde{z}(k) \tag{5.6}$$

where the determination of $R(k)$ is the subject of the next section.

C. CHOICE OF THE GAIN MATRIX R(k)

There are many possible choices for the gain matrix R(k), only two of which were introduced in Chapter II. The R(k) matrices discussed in Chapter II were chosen because of their computational simplicity. However, another requirement is that R(k) must be independent of H(k) and $\hat{\theta}(k)$, which was a requirement for the derivation of Equation 5.2 . Two rather substantial problems become apparent immediately: first the need to make $H(k+l)$ and $H(k)$ statistically independent, and second the slow convergence of the algorithm due to the simple steepest descent gain matrix chosen as R(k). In order to make $H(k+l)$ and $H(k)$ independent in this simplified vehicle they must be separated in time by four time constants of the system or 5.6 seconds. This separation means that very extended observation times will be needed to reach an acceptable parameter approximation. The simple selection of the gain matrix R(k), which at first glance may seem to be a computational advantage, will require more iterations to obtain acceptable results than many of its more sophisticated counterparts.

One final observation can be made concerning the stability of Equations 5.2 and 5.6 , before proceeding into the simulation study. Consider again Equation 5.6 in the dynamic systems sense and notice the time varying state transition matrix given as

$$\Phi(k) = |I + R(k)\Sigma_N \tag{5.7}$$

For this discrete time system, the characteristic equation will certainly have poles on the unit circle, and if state measurement noise is present then the roots will all

be outside the unit circle, since both $R(k)$, the gain matrix and Σ_N, the state measurement noise covariance matrix are positive definite. It seem reasonable that the magnitude of $R(k)$ and its rate of decay will play a very important role in determining whether successful identification is made. However, there is very little information available concerning how to select an optimal $R(k)$ matrix. Thus the first part of the simulation study is devoted to the proper selection of $R(k)$.

D. INITIALIZATION OF R(k)

In Chapter II, two choices for the gain matrix $R(k)$ were presented. The first based on Venter's theorem was

$$R(k) = \frac{1}{k^p} diag[h_1(k), h_2(k), \ldots, h_{n^2}] \qquad \text{for } \frac{1}{2} < p \leq 1 \tag{5.8}$$

and the second based on Lyapunov's Main Stability theorem was

$$R(k) = \frac{diag[h_1(k), h_2(k), \ldots, h_{n^2}]}{\sum_{i=1}^{n^2} h_i(k) r_i^2(k)} \tag{5.9}$$

where $h_i(k)$ can vary for different anticipated values of θ. The second gain matrix, in Equation 5.9 is not independent of H(k) and although Mendel [Ref. 4:p. 268] suggests its use, it will probably result in a biased estimate.

Figure 5-1 shows the difference in convergence, when the two different choices of $R(k)$ are used in the absence of measurement noise while the system is excited by road noise. Again, the error measure ξ_k is used to determine how close the parameters estimates are to the actual parameters. The curve marked with

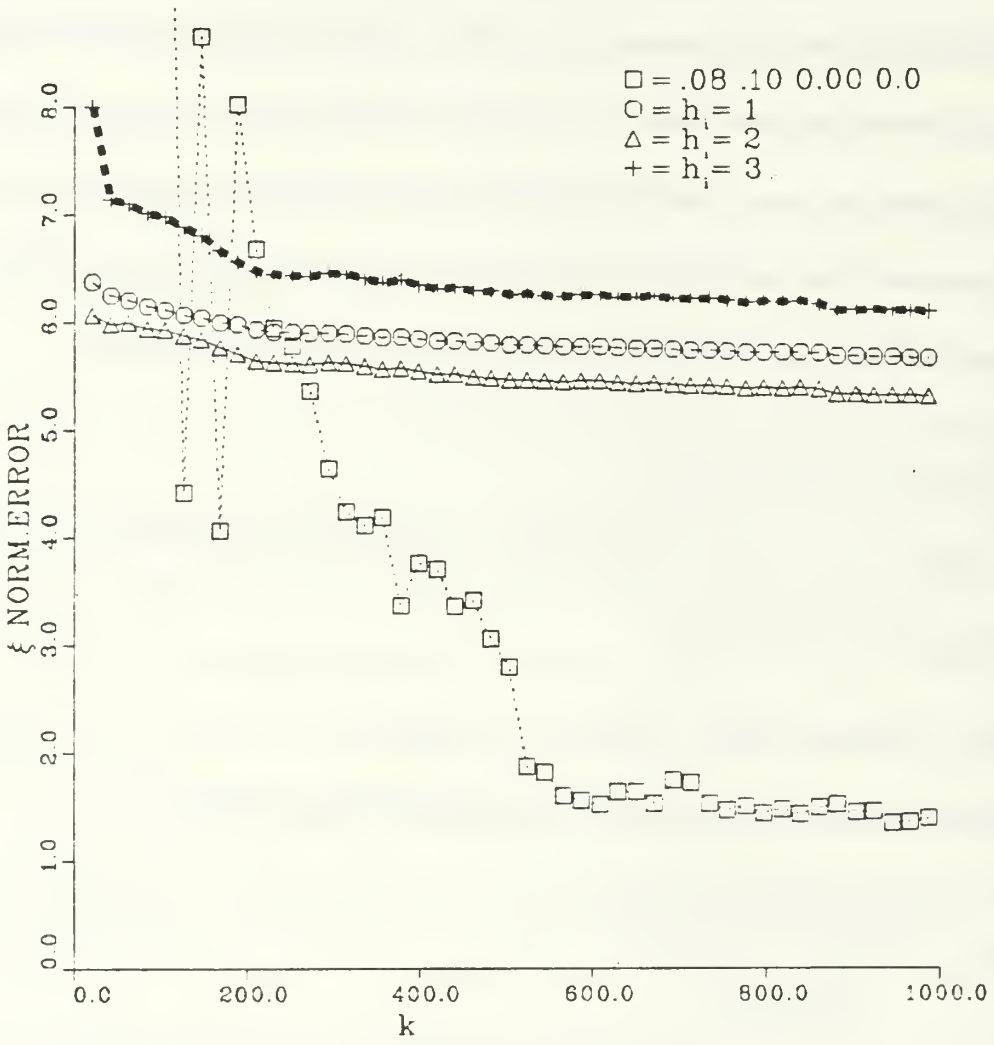

Figure 5-1 Parameter estimation error for Lyapunov gain matrix and three different values of diagonal elements for Venters gain matrix.
Initial conditions, $\sigma = 0.08$
Road noise, $\sigma = 0.1$
State measurement noise, $\sigma = 0.0$
State derivative measurement noise, $\sigma = 0.0$

squares is the error measure ξ_k when $R(k)$ is chosen as in Equation 5.9, while the other three curves are for various values of $h_i(k)$ using Equation 5.8 to determine $R(k)$, while holding p constant at 0.51. The curve derived from the use of Equation 5.9 initially reaches the smallest value and then stablizes as do the other three curves. This is due to the more responsive nature of Equation 5.9 and the larger variations in the bottom curve when compared with the upper three curves also demonstrates this. The error measure marked by circles has the smallest elements in the R(k) matrix and is therefore the least responsive as can be seen in the figure. The curve marked with "+" symbols has larger elements in the R(k) matrix and represents the largest elements that can be expected to produce acceptable convergence. The curve marked with triangles performs the best with smaller elements of R(k), and will produce more consistent results, particularly in the presents of noise. For this simulation, the separation time between measurements was 0.05 seconds, thus for the 1000 iterations represented here the total observation time is about 50 seconds. This time period has produced results which are worse than the Kalman filter approach normally produces in the first 5.0 seconds of it's operation, when starting with the same initial conditions.

This slow convergence in the absence of noise is quite troublesome. In fact if the identifiers convergence rate remained constant, which it certainly will not since convergence is asymptotic, it would still require approximately 18 hours to reach an acceptable approximation in the presence of measurement noise. There

are few identification scenarios in which 18 hours is considered appropriate for system identification. If the system is initially known and one was attempting to follow time varying parameters this approach might be more appealing.

The road noise in the vehicle model produced very large deviations from the actual second row parameters if the $h_i(k)$ values for that row became too large. The identifier produced the best results when

$$10h_i(k) = h_{i+4}(k) \qquad \text{for } i=1,2,3,4 \tag{5.10}$$

Figure 5-2 shows the error measurement curve for the same noise ensemble as used in Figure 5-1 but with state measurement noise added. Comparing the effect of the added noise on the Venter gain matrix $R_v(k)$ and the Lyapunov gain matrix $R_l(k)$, it can easily be seen that $R_l(k)$ is profoundly effected while the curve resulting from the use of $R_v(k)$ is nearly unchanged. This change in the error measurement curve due to the use of $R_l(k)$ is the result of the correlation of $R_l(k)$ and $H(k)$ which were assumed uncorrelated in Equation 2.57 . Therefore, the use of $R_l(k)$ will produce a bias and the use of $R_v(k)$ produces extended observation time. Since the stochastic gradient approach was chosen because it promised to produce an unbiased approximation of the parameters, the gain matrix $R_l(k)$ will not be explored further.

The slow convergence rate of the steepest descent stochastic gradient identifier must be improved in order to make the identification scheme acceptable. There are several factors which affect the convergence rate which can be explored in an attempt to improve the convergence. In Figures 5-1 and 5-2, the value of p

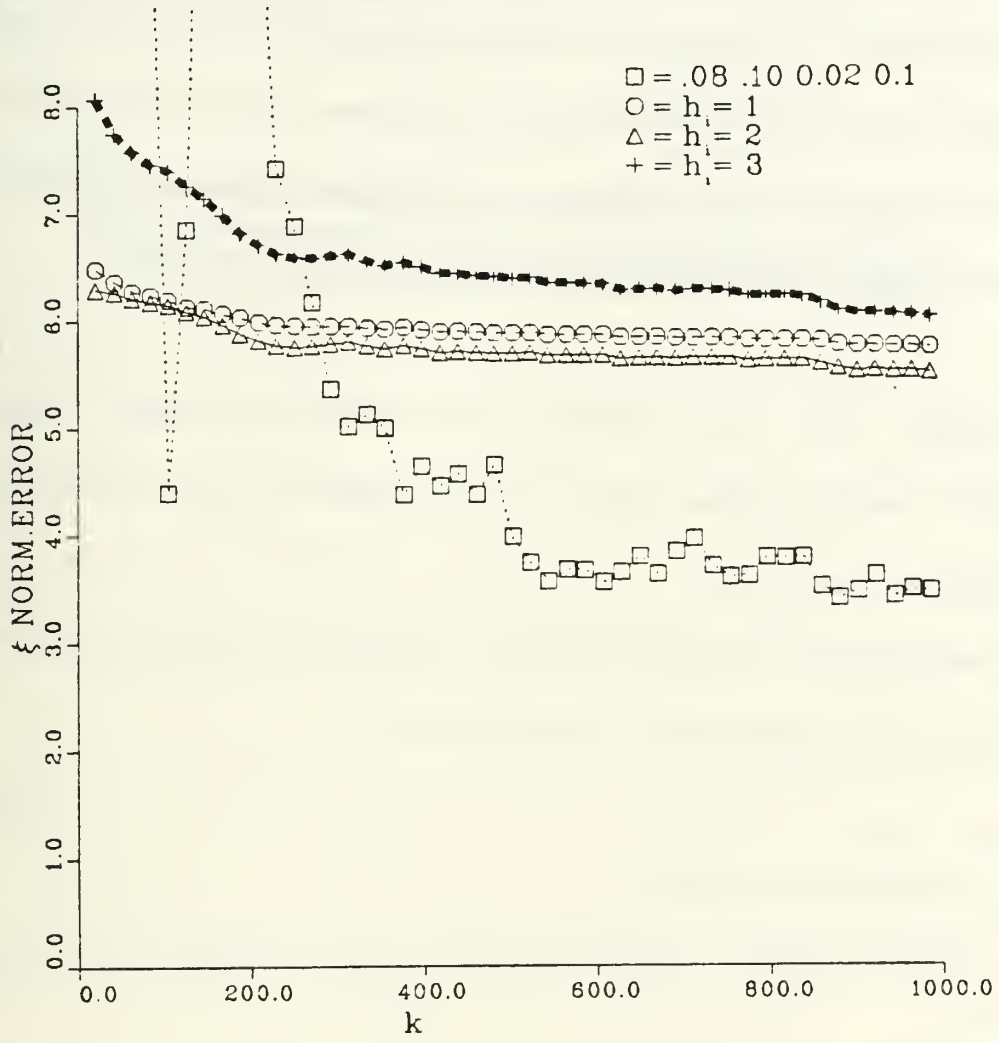

Figure 5-2 Parameter estimation error for Lyapunov gain matrix and three different values of diagonal elements for Venters gain matrix.
Initial conditions. $\sigma = 0.08$
Road noise. $\sigma = 0.1$
State measurement noise. $\sigma = 0.02$
State derivative measurement noise. $\sigma = 0.1$

the power to which k is raised in Equation 5.8, was 0.51. Figure 5-3 demonstrates the effect changing p to 0.95 has on the error measurement in the presence of measurement noise. Since p is now larger, which will force $R(k)$ to diminish more rapidly, one might have expected that the use of larger values of $h_1(k)$ would produce better results. This was not the case, and

$$h_1(k)=4 \tag{5.11}$$

was the largest value that converged in the presence of noise. Without measurement noise much larger values of $h_1(k)$ are possible, but the convergence rate is only slightly improved.

It is possible that the optimum value for p is between 0.51 and 0.95. However, Ljung [Ref. 6:p. 279] suggests that it is desirable to let

$$t\gamma(t)\rightarrow 1 \quad \text{as} \quad t\rightarrow\infty \tag{5.12}$$

where $\gamma(t)$ is defined in this thesis as

$$\gamma(k) = \frac{1}{k^p} \tag{5.13}$$

This implies that

$$\gamma(k) = \frac{1}{k} \tag{5.14}$$

It is also suggested that for small and intermediate values of k, $\gamma(k)$ should be chosen differently. Thus define

$$\gamma(k) = \frac{1}{1+\dfrac{\lambda(k)}{\gamma(k-1)}} \tag{5.15}$$

and

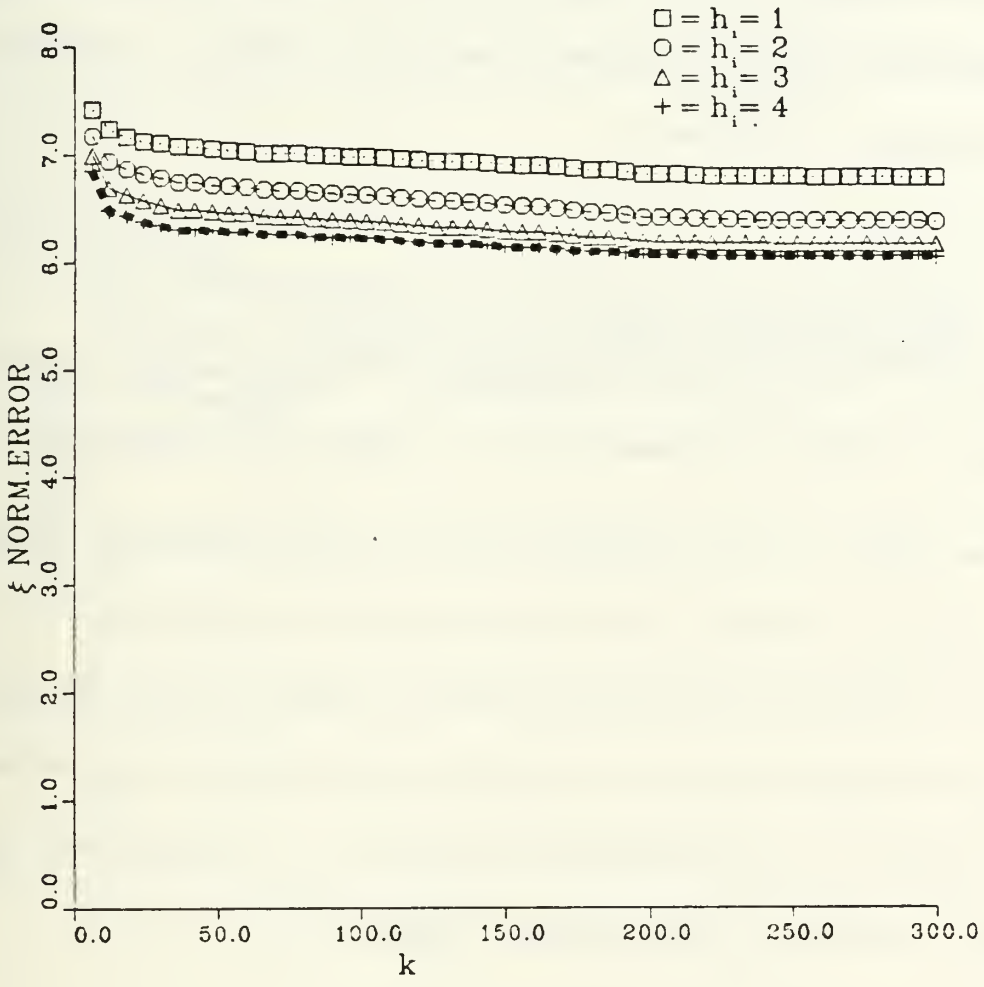

Figure 5-3 Parameter estimation error for four
different values of diagonal elements for Venters gain matrix.
and p=0.95.
Initial conditions. σ = 0.08
Road noise, σ = 0.1
State measurement noise, σ = 0.02
State derivative measurement noise, σ = 0.1

$$\lambda(k) = \lambda_0 \lambda(k-1) + (1-\lambda_0) \tag{5.16}$$

Here $\lambda(k)$ is considered to be a forgetting factor, reducing the identifiers responsiveness to early values of the error vector $z(k)-z(k)$. Figure 5-4 shows the improved convergence obtained with

$$\lambda_0 = 0.99 \tag{5.17}$$

and

$$\lambda(0) = 0.95 \tag{5.18}$$

when using $\gamma(k)$ as defined in Equation 5.15 to replace $\frac{1}{k^p}$ in Equation 5.8 . The smallest value of ξ_k has decreased below 5, but the improvement is not enough to produce an acceptable results in an acceptable observation time.

This slow convergence is predictable according to Ljung [Ref. 6:pp. 290-299]. It is suggested that the convergence rate can be improved with the Gauss-Newton algorithm. However, the Gauss-Newton algorithm has no advantage over the Kalman filter approach for the system being considered here.

E. SUMMARY

The stochastic gradient approach was chosen because it theoretically produces an unbiased estimate even in the presence of measurement noise. The simplicity of computation is a desirable feature. However, the computational simplicity of the steepest descent algorithm also means very slow convergence. The slow convergence of the algorithm could be improved with a more sophisticated choice of search direction. The improvement achieved by using an algorithm which is

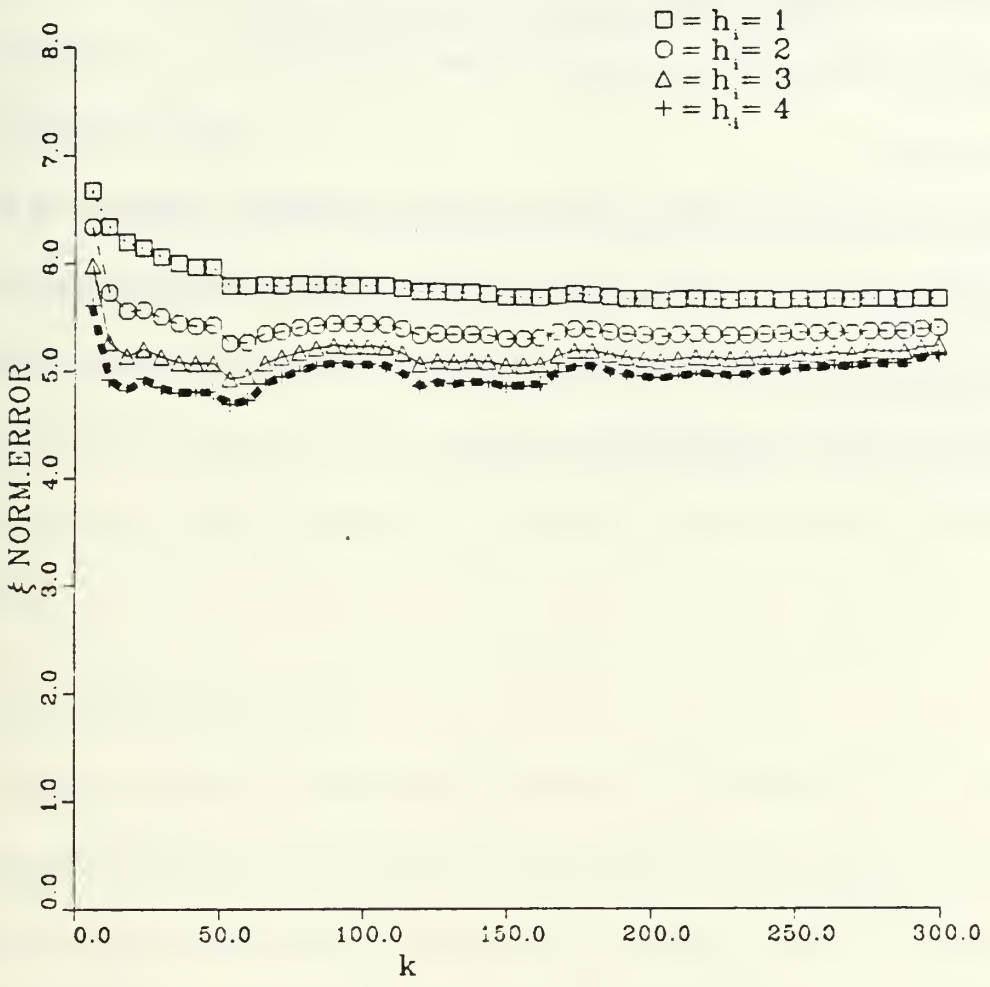

Figure 5-4 Parameter estimation error for four
different values of diagonal elements for Venters gain matrix.
and the use of a forgetting factor.
Initial conditions, $\sigma = 0.08$
Road noise, $\sigma = 0.1$
State measurement noise, $\sigma = 0.02$
State derivative measurement noise, $\sigma = 0.1$

correlated with $H(k)$ would be offset by the bias produced from this correlation. The use of the forgetting factor $\lambda(k)$ does improve the convergence, but this improvement is not significant enough to allow identification in an acceptable observation time.

Because of the slow convergence of the stochastic gradient approach, for the remainder of this thesis, it is abandoned in favor of the Kalman filter identifier. In the next chapter an improved model is introduced, which will introduce another facet of the land vehicle identification problem.

VI. IMPROVED MODEL IDENTIFICATION

A. INTRODUCTION

In the last three chapters a very simple vehicle model was used to test various identification algorithms. In this chapter a more realistic model is presented. The improved model is compared with the model developed in Chapter III, to determine how increased model size effects the identification process. Finally, the problem of two road noise inputs is discussed and the resulting identifier is analyzed.

B. AN IMPROVED MODEL

In Chapter III the simplified vehicle used for the simulation study of various identification schemes was developed. This simplified model has one very serious limitation that must be taken into account for an actual vehicle. The limitation is associated with the fact that it is a single input system, whereas an actual vehicle would have several inputs. Specifically, the inputs to the front and rear wheels will be correlated, and even if the white road noise approximation is considered valid, this correlation will produce a bias in an actual vehicle parameter identification. In order to analyze possible bias elimination schemes or in fact to determine if any approach is valid when there is correlation between the inputs, a

more realistic model is needed. In this section a eighth order, four-degree-of-freedom model is presented.

Consider the vehicle depicted in Figure 6-1. In this model the separation of the suspension into front and rear units, necessitates the introduction of rotational freedom about the lateral axis. This allows both rotational and vertical motion of the upper mass, the vehicle body.

The equations of motion for the improved vehicle are

$$\Sigma_{F_{x_1}} = -m_1\ddot{x}_1-[x_1-(x_3+\phi l_1)]k_{12} - [\dot{x}_1-(\dot{x}_3+\dot{\phi} l_1)]b_1 + (u_1(t)-x_1)k_{11} \tag{6.1}$$

$$\Sigma_{F_{x_2}} = -m_2\ddot{x}_2-[x_2-(x_3-\phi l_2)]k_{22} - [\dot{x}_2-(\dot{x}_3-\dot{\phi} l_2)]b_2 + (u_2(t)-x_2)k_{21} \tag{6.2}$$

$$\Sigma_{F_{x_3}} = -M\ddot{x}_3+[x_1-(x_3+\phi l_1)]k_{12}+[x_2-(x_3-\phi l_2)]k_{22} + [\dot{x}_1-\dot{x}_3+\dot{\phi} l_1)]b_1 + [\dot{x}_2-(\dot{x}_3-\dot{\phi} l_2)]b_2 \tag{6.3}$$

$$\Sigma_{M_\phi} = -J\ddot{\phi} + \{[x_1-(x_3+\phi l_1)]k_{12} + [\dot{x}_1-(\dot{x}_3+\phi l_1)]b_1\}l_1 - \{[x_2-(x_3-\phi l_2)]k_{22} + [\dot{x}_2-(\dot{x}_3-\dot{\phi} l_2)]b_2\}l_2 \tag{6.4}$$

where the equations have been linearized for small values of ϕ. The resulting state equations produce Equation 6.6, where the state vector is

$$x = \begin{bmatrix} x_1 \\ \dot{x}_1 \\ x_2 \\ \dot{x}_2 \\ x_3 \\ \dot{x}_3 \\ \phi \\ \dot{\phi} \end{bmatrix} \tag{6.5}$$

The upper mass M, is now supported by two wheels m_1 and m_2. The wheels

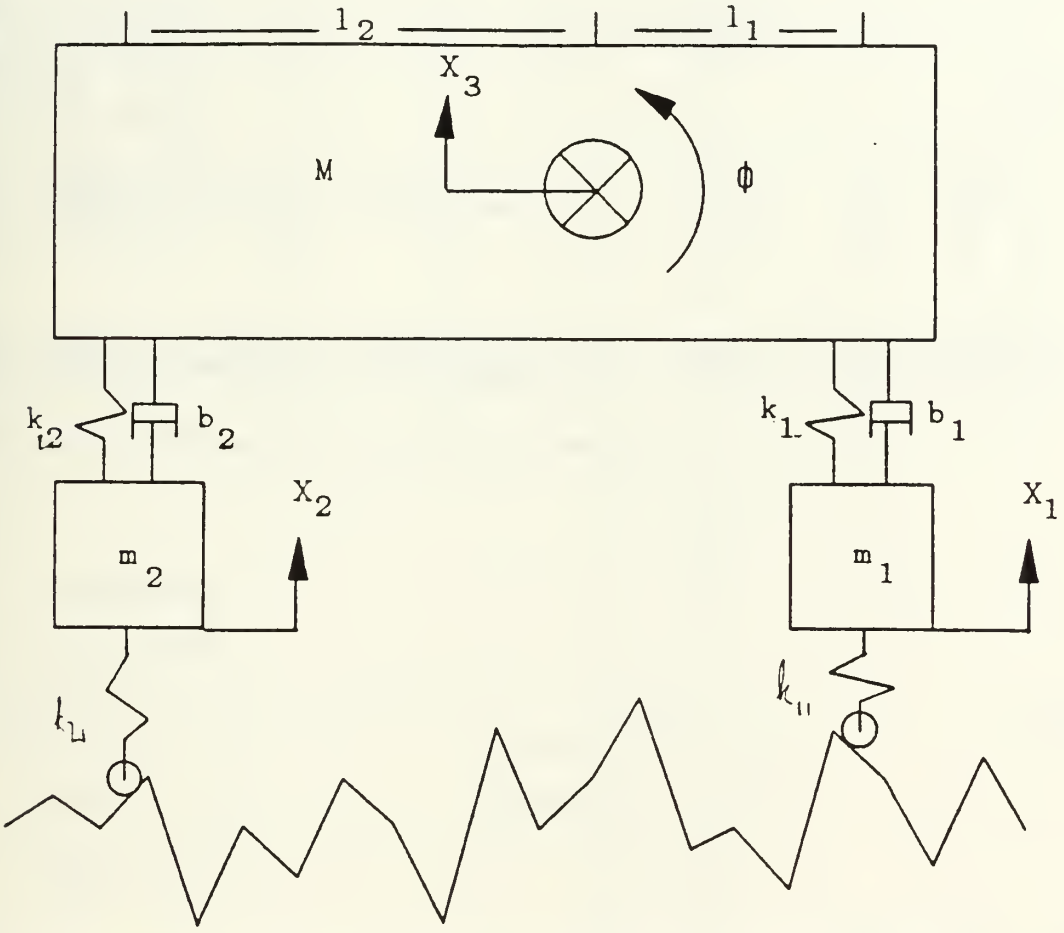

Figure 6-1 Improved Vehicle Model

$$\dot{x} = \tag{6.6}$$

$$\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{-(k_{11}+k_{12})}{m_1} & \frac{-b_1}{m_1} & 0 & 0 & \frac{k_{12}}{m_1} & \frac{b_1}{m_1} & \frac{k_{12}l_1}{m_1} & \frac{b_1 l_1}{m_1} \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{-(k_{21}+k_{22})}{m_2} & \frac{-b_2}{m_2} & \frac{k_{22}}{m_2} & \frac{b_2}{m_2} & \frac{-k_{12}l_2}{m_2} & \frac{-b_2 l_2}{m_2} \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\frac{k_{12}}{M} & \frac{b_1}{M} & \frac{k_{22}}{M} & \frac{b_2}{M} & \frac{-(k_{12}+k_{22})}{M} & \frac{-(b_1+b_2)}{M} & \frac{l_1k_{12}-l_2k_{22}}{M} & \frac{b_1l_1-b_2l_2}{M} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\frac{l_1k_{12}}{J} & \frac{l_1b_1}{J} & \frac{-l_2k_{22}}{J} & \frac{-l_2b_2}{J} & \frac{-l_2k_{22}+l_1k_{12}}{J} & \frac{-l_2b_2+l_1b_1}{J} & \frac{-l_1^2k_{12}-l_2^2k_{22}}{J} & \frac{-l_1^2b_1-l_2^2b_2}{J}
\end{pmatrix}$$

$$\times x + \begin{pmatrix}
0 & 0 \\
\frac{k_{11}}{m_1} & 0 \\
0 & 0 \\
0 & \frac{k_{21}}{m_2} \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0
\end{pmatrix} \begin{pmatrix} u(k) \\ u(k-\tau) \end{pmatrix}$$

are free to move vertically, while the body M, is free to move both vertically and to rotate about its center of mass, which need not be equidistant from the wheels. The angle of rotation ϕ, and the displacement x_3 describe the vehicle body, while x_1 pertains to the front wheel displacement and x_2 is the rear wheel displacement.

For the simulation study of this chapter the values selected for the dynamic components were

$$\begin{bmatrix} k_{11},\ k_{21} \\ k_{12},\ k_{22} \\ b_1,\ b_2 \\ m_1,\ m_2 \\ M \\ J \\ l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} 100 \\ 40 \\ 60 \\ 1 \\ 20 \\ 60 \\ 2 \\ 1 \end{bmatrix} \tag{6.7}$$

These values for the dynamic components of the system produce a vehicle model with roots of the characteristic equation located at

$$s = \begin{bmatrix} -.656 \pm j\ 3.11 \\ -.691 \pm j\ 2.08 \\ -.738 \\ -.837 \\ -61.5 \\ -65.3 \end{bmatrix} \tag{6.8}$$

As in the model developed in Chapter III, the input distribution matrix must be known, since the input vector is unknown, and the state vector x(k) and its derivative $\dot{x}(k)$ are measurable, but corrupted by noise.

In the improved model there are now two road noise inputs, which are correlated and shall be written as $u(k)$ and $u(k-\tau)$, where τ is the time delay between a road input to the front wheel and the same road input to the rear wheel. The value of τ is given in terms of the vehicle's velocity v, as

$$\tau = \frac{l}{v} \tag{6.9}$$

where l is the distance between the front and rear wheels. The correlation between $u(k)$ and $u(k-\tau)$ will be considered in detail in Section D. The other two noise sources remain the same, where $w(t)$ is the state measurement noise and $v(t)$ is the state derivative measurement noise.

In this chapter the parameter error measure ξ_k will be used again, with one minor change. The improved model parameters of Equation 6.5 have two zero values in each of the second and fourth rows. These zero values shall not be considered in the error measure ξ_k even though they appear in rows with unknown parameters. This assumption is a requirement of the definition of ξ_k given in Equation 3.17. This allows the use of the same error measure given by

$$\xi_k = \sum_{i=1}^{28} \left(\frac{\theta_i - \hat{\theta}_i(k)}{\theta_i} \right)^2 \tag{6.10}$$

The summation goes to 28 because the four zero values of rows 2 and 4 are not considered, and 32 of the 64 parameters are assumed known.

In the next section, the new model will be compared with the vehicle of Chapter IV to determine the effect increasing the state size has on system identification.

C. IMPROVED VEHICLE PARAMETER IDENTIFICATION

The opportunity to analyze the effect of changing state size on parameter identification should not be neglected, since only the sparsest reference is made to it in the literature. The change from a fourth order model to an eighth order model correspondingly changes the number of parameters from 16 to 64. As in the case of the fourth order system, half of the 64 parameter values are assumed known from the structure of the plant, and four more will not be considered in the parameter error measure ξ_k for the reasons discussed in the last section, leaving 28 parameters to be estimated. This rather substantial increase in parameters to be identified has a significant influence on the identification process. Making use of the results of Chapter IV, a comparison shall be made with the improved model under similar noise conditions. In this section the model will be excited using noise into the front suspension unit only. In the following section the improved model will be excited with road noise entering both suspension units and comparisons will be made.

1. P(0) Initialization

The initial value of the diagonal of P_0 plays an important role in the convergence rate of the Kalman filter identifier. Figure 6-2 shows how various values of the diagonal of P_0, the error covariance matrix, effect the identification process. Several observations can be made concerning the convergence rates as depicted by the parameter error measure ξ_k, especially when compared with the curves in Figure 4-1. The curve marked with "+" symbols produces the fastest

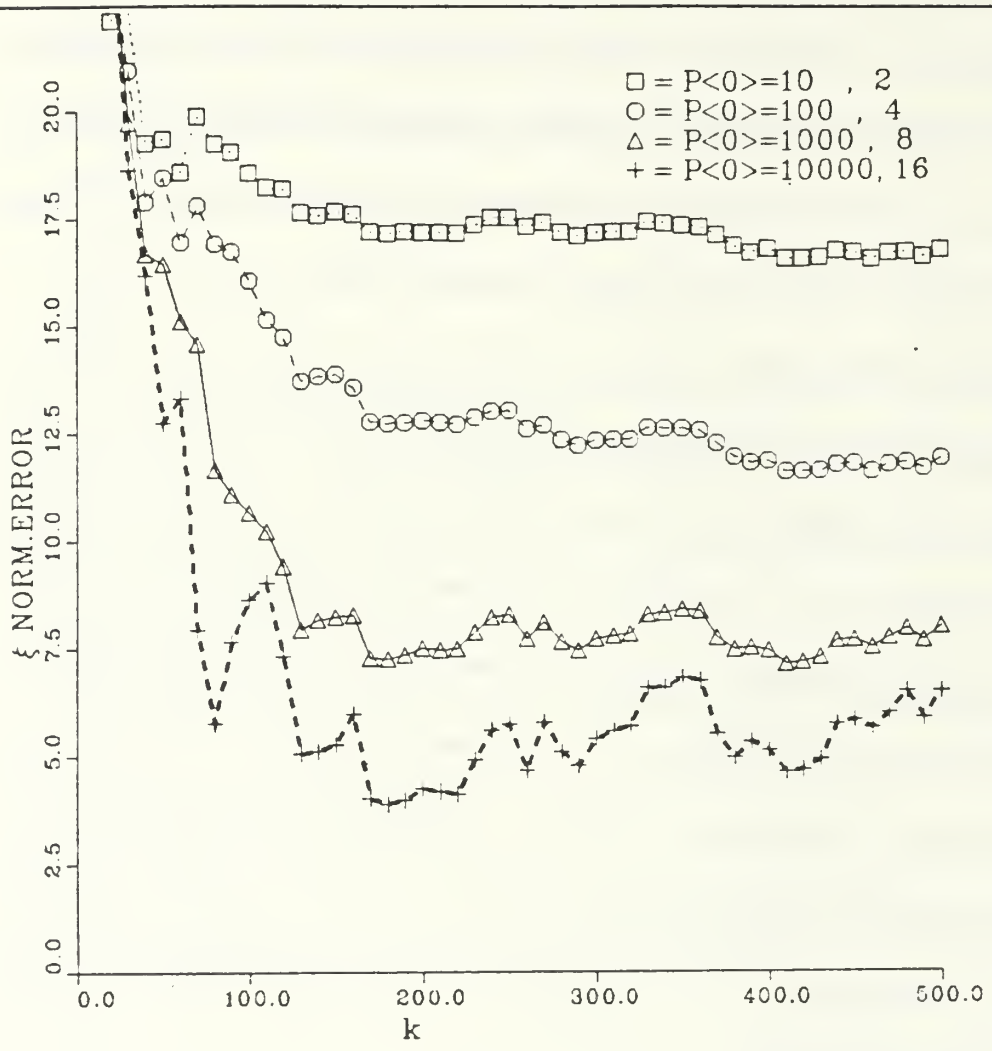

Figure 6-2 Parameter estimation error vs. number of iterations for four different initializations of the covariance matrix.
Initial conditions. σ =0.1
Road noise. σ =0.1
State measurement noise. σ =0.0
State derivative measurement noise. σ =0.005

initial convergence although the curve marked by triangles is approaching it and the triangled curve shows much better consistency during the early part of the identification. These remarks are true for both Figure 6-2 and 4-1. The higher order model must be broken into two diagonal blocks, because the size of the parameters in the sixth and eighth rows are so much smaller than the parameters of the second and fourth rows. Thus the second number in the legend is the value of the lower half diagonal of P(0). Because of its more consistent response. the value for the upper half diagonal of P_0 was chosen at 1000 and the lower half diagonal was selected at 8. The smaller value for the lower half diagonal of the P_0 matrix is due to the much smaller values of the elements of the sixth and eighth row parameters.

2. Convergence Rate of the Improved Model

The higher order of the improved model would lead to the assumption that convergence to acceptable parameter estimates will take longer than for the fourth order model. Figure 6-3 demonstrates the convergence rate for the eighth order model with only road noise excitation. Figure 4-2 shows the convergence rate for the fourth order model under the same noise conditions. It is difficult is compare convergence rates of systems of different order, but several observation can be made. In Figure 4-2, the plateauing effect is visible between 15,000 and 20,000 iteration. In higher order models the tendency for the parameter estimates to plateau is much greater and may require more iterations before improvement begins again. The plateauing effect in higher order models can be confused with a

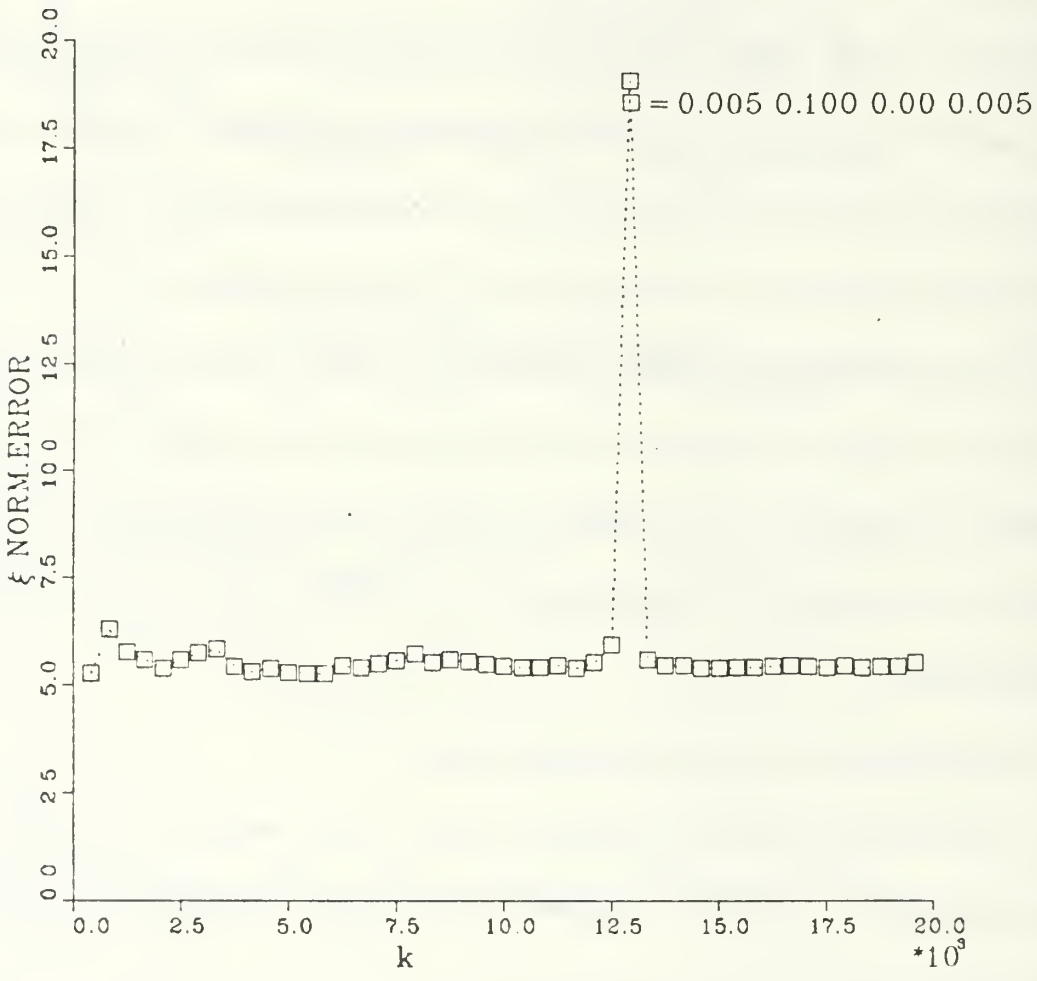

Figure 6-3 Parameter estimation error vs. number of iterations with road noise as the only system excitation.
Initial conditions, σ =0.0
Road noise, σ =0.1
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.005

bias if one does not look carefully. The difference is that in the case of a bias the estimations will tend to become worse and there will be an increase in the estimation error measure ξ_k, while plateauing will result in essentially a flat error measure. In fact, during the simulation study of this eighth order model a plateau of 20,000 iterations was encountered. How to detect these long intervals with very little improvement in parameter estimation and what to do when a plateau is encountered is not addressed here. The fourth order model in Chapter IV required 40,000 iterations to identify the system parameters, while the eighth order model shows no improvement after 20,000 iterations for the same noise scenario. The parameter error measurement spike that occurs at approximately 13,000 iterations is somewhat mysterious. It seems likely, however, that it is caused by some form of system reflection where the input is reflected back from the rear wheel, appearing as a second input to which the identifier respond incorrectly.

3. Initial Conditions and Road Noise

Figure 6-4 illustrates the effect that large initial conditions and various amounts of road noise have on the Kalman filter identifier. Since only 6 of the 8 parameters in the second row are considered in the parameter error measure ξ_K, and there is road noise on only the second row estimates, the identifier is able to extract all of the other parameters except the second row. As the second row parameters are covered with road noise, little progress is made toward their

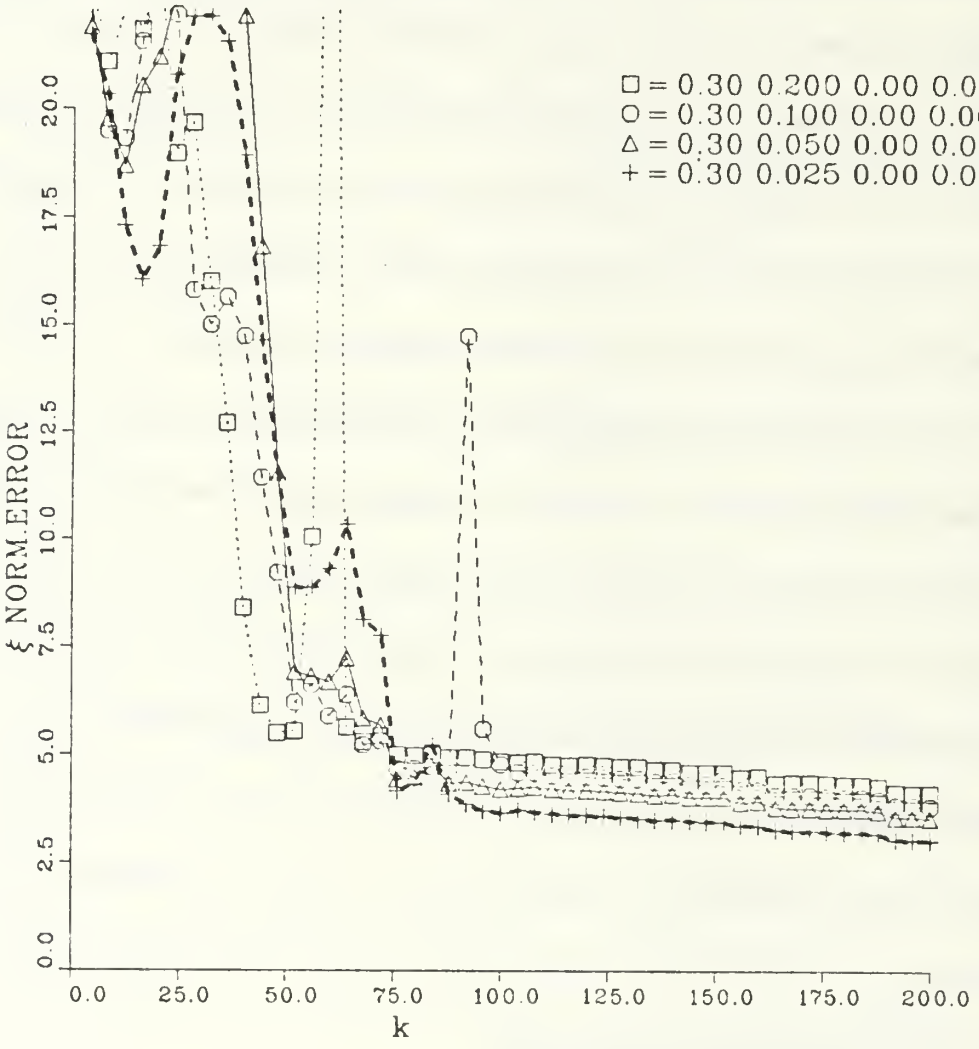

Figure 6-4 Parameter estimation error with various levels of road noise and constant initial conditions.
Initial conditions. σ =0.3
Road noise. σ =0.2. 0.1. 0.05. 0.025
State measurement noise. σ =0.0
State derivative measurement noise, σ =0.001

identification before the initial conditions diminish and the system's only excitation is from the road.

Comparing Figure 6-4 with Figure 4-3, some other observations can be made. Because of the larger search space of the eighth order system, it will usually require more iterations to reach acceptable identification. This can be seen in the slower convergence of Figure 6-4.

4. System Identification with Pulse Inputs

Figure 6-5 was produced using a P_0 matrix diagonally initialized with 1000. The increase in the estimators responsiveness is quite apparent. The curves in Figure 6-5 show the identification process when pulses of various widths are used. The curve marked by squares is the estimation error when the system is excited with an impulse, during the ninth measurement cycle. In the case of the fourth order system the energy absorbed by the system was enough to allow the identifier to estimate the parameters. In the eighth order model, because of the larger parameter inertia, the identifier is unable to reach acceptable identification before the impulse energy has dissipated.

5. Identification with State Measurement Noise

Figure 6-6 illustrates the tremendous effect state measurement noise has on the Kalman filter identifier when used in an eighth order system. The curve marked with "+" symbols has no state measurement noise, and the other three curves have varying levels of measurement noise. The system is excited by road noise. The curve marked with circles has measurement noise of only 1 % of the

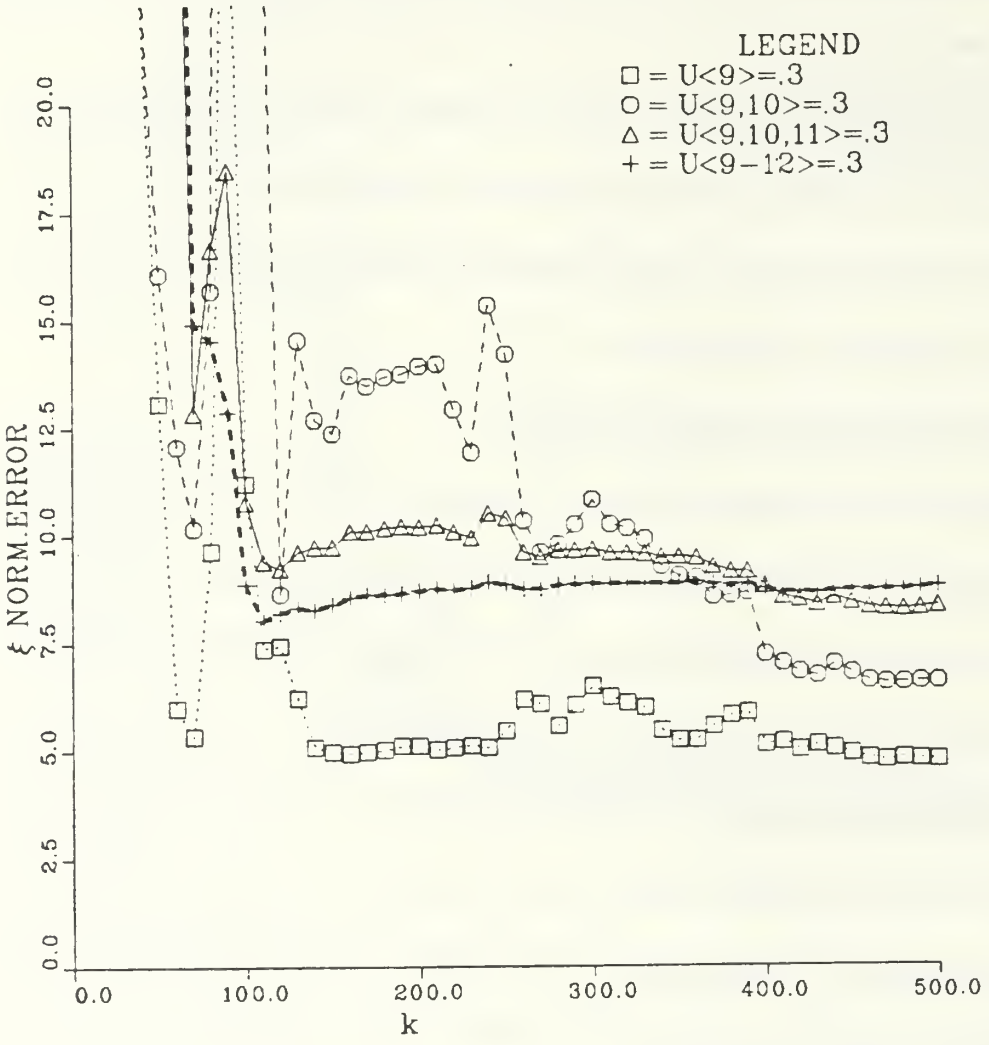

Figure 6-5 Parameter estimation error when system is excited with input pulses of various durations. Pulse duration shown in legend.
Initial conditions, σ =0.01
Road noise, σ =0.001
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.01

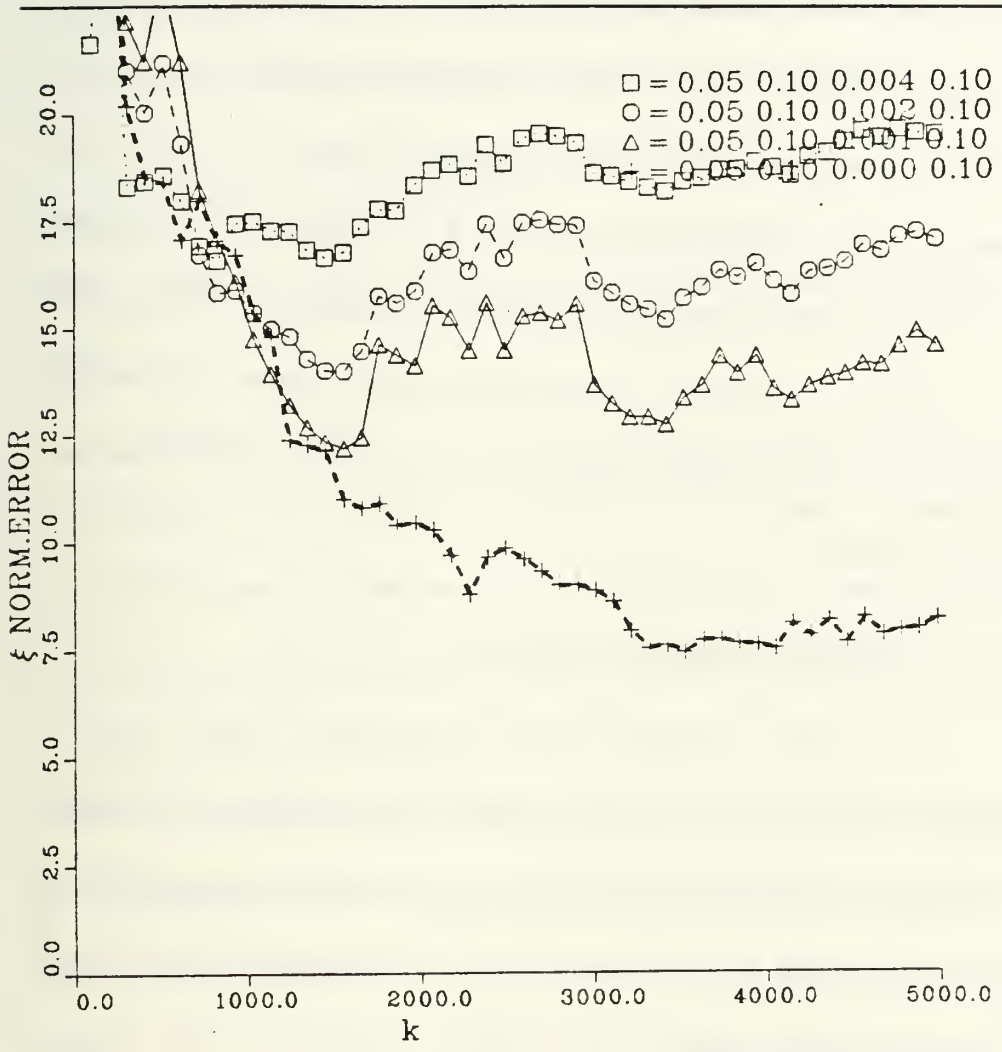

Figure 6-6 Parameter estimation error when various levels of measurement noise are added to the state vector.
Initial conditions, σ =0.05
Road noise, σ =0.1
State measurement noise, σ =0.002, 0.001, 0.0005, 0.0000
State derivative measurement noise, σ =0.1

road noise input to the system, and yet at 5000 iteration the parameter estimates are diverging. Figure 4-5 shows the identifiers estimation error for some of the same values of state measurement noise. In the simplified model, it can be seen that at 5000 iteration when measurement noise was 4 %, the identifier produced similar results. From this observation it is reasonable to assume that that eighth order model is 4 times as sensitive to state measurement noise and would therefore require sensors with a signal to noise ratio at least 6 dB greater than the fourth order model required.

D. TWO INPUTS TO THE IMPROVED MODEL

The use of a single road noise input into the improved vehicle model is somewhat artificial but does provide some insight into the effect of increased model size on the parameter identification process. The single input analysis of the last section also provides insight into results to be expected when two road noise inputs are applied to the model.

In this section the model will be excited with road noise through both wheels. The road noise into the front and rear suspension system of the vehicle will be correlated in time, since the front wheel will encounter an obstacle τ seconds before the rear wheel. Recall from Chapter IV Equation 4.11 that the bias on $\hat{\theta}(k)$ is the result of the input-output crosscorrelation as described in Equation 4.16. If the improved model is excited through the front wheel with an obstacle in the road, then the output, the state of the system will be correlated with the input to

the rear suspension when it encounters that same obstacle τ seconds later. Then from Equation 4.11, the bias produced by this correlation will be

$$[P^{-1}(0) + P^{-1}(k)]^{-1} \sum_{m=1}^{k} H^T(m) R^{-1} v(m) \qquad (6.11)$$

where v(m) is the total noise vector including $u(k-\tau)$, which is correlated with $H^T(m)$.

A possible solution to the input-output crosscorrelation bias that will undoubtedly result in slower convergence is as follows. Assume for a moment that the parameters of the second row of Equation 6.5 are known exactly. This being true, then the first element of the error vector is defined by

$$e_1(k) = z_2 - \mathrm{x}\theta_{1-8} = v_2(k) + b_{21}u(k) \qquad (6.12)$$

where θ_{1-8} represents the first 8 elements of the $\hat{\theta}$ vector, which are known exactly. The discrepancy in the subscripts resides in the fact the the first and third rows of the A matrix are assumed known and therefore are not part of the identification process. If it is further assumed that the measurement noise is negligible compared to the road noise, then

$$b_{21}u(k) \approx e_1(k) \qquad (6.13)$$

In order to eliminate the effect of the road noise from the identifier, the measurement samples must be taken when the rear wheel is exactly over the same position on the road that was samples τ iterations before when the front wheel was over that point. This can be accomplished because of the asynchronous nature of the identification scheme presented here. If this element of the error vector is

then subtracted from the second element of the error vector τ seconds later, this will eliminate the road noise from the error vector and the identifier will produce an unbiased estimate. This can be expressed mathematically as

$$e_2(k) = z_4 - x\hat{\theta}_{9-16} - e_1(k-\tau) = v_4(k) + g(x(k), \theta_{9-16}, \hat{\theta}_{9-16}) + b_{42}u(k) - b_{21}u(k-\tau) - v_2(k-\tau) \quad (6.14)$$

where the two input term on the right side of the equation cancel. The assumptions used earlier may now be lifted since $e_1(k-\tau)$ is white noise. This can be seen if Equation 6.12 is considered as a whitening filter. Then $e_1(k-\tau)$ is an innovation sequence and it can be shown that [Ref. 7:p. 250]

$$E[e_1(k-\tau) \mid z_2(k-\tau-1), \ldots, z_2(k_0)] = 0 \quad (6.15)$$

The convergence rate of this approach will certainly be affected, since the variance of the noise on the fourth row parameters is very large until the parameter estimates of the second row of Equation 6.5 become more accurate.

Before analyzing the bias elimination scheme proposed above, a look will be taken at how the identifier responds to two independent road noise inputs and then the bias itself will be shown. Figure 6-7 illustrates the parameter error when two different road noise ensembles are input to the front and rear suspension units of the improved model. As one might have anticipated, the added noise source results in less effective identification of the fourth row parameters which translates into higher values of the parameter error measure. However, the identification is somewhat better than one might expect. When two different road noise inputs are applied to the vehicle the resulting parameter error is not twice as bad as when a single input is applied. The plateau is again present for the last 4000 iterations.

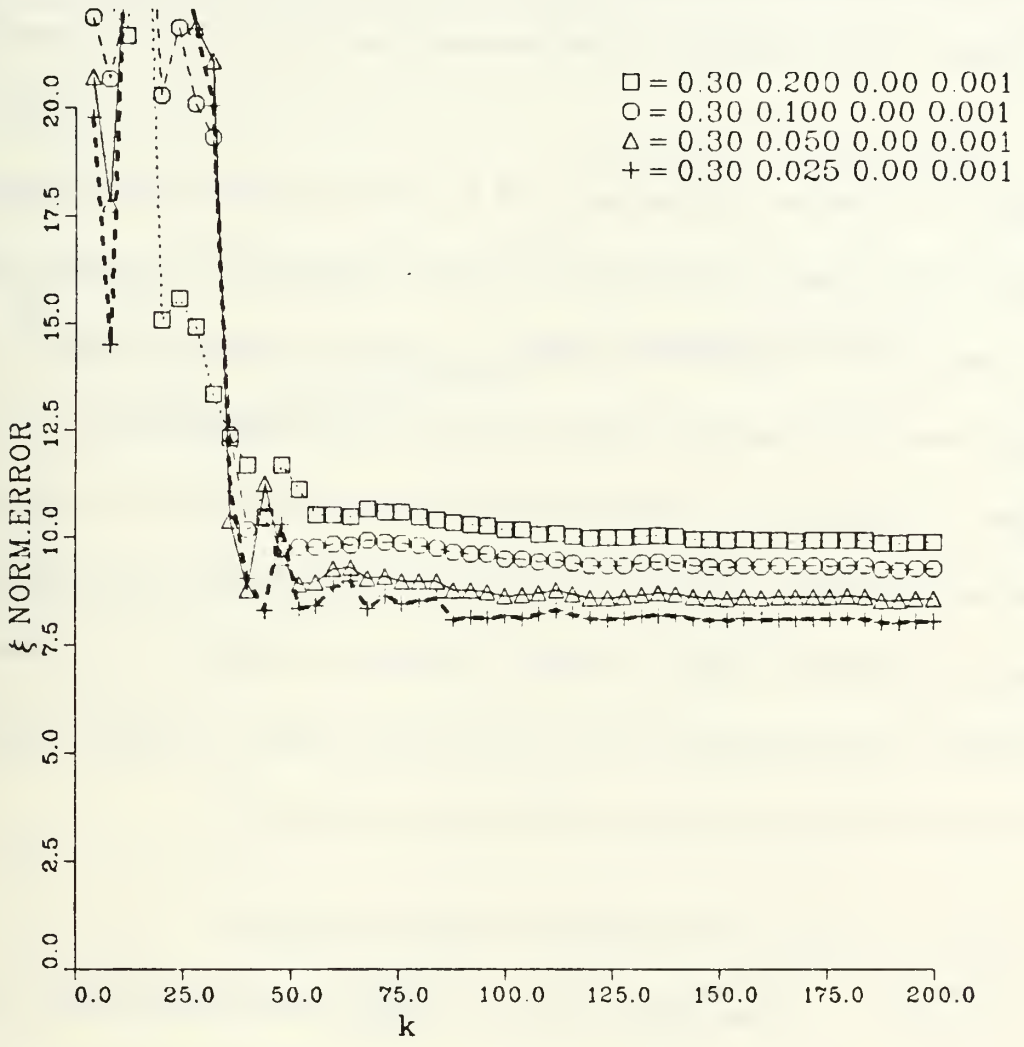

Figure 6-7 Parameter estimation error vs. number of iterations for four different values of road noise, holding initial conditions constant, for two independent noise inputs.
Initial conditions, σ =0.3
Road noise, σ =0.2, 0.1, 0.05, 0.025
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.001

One other observation concerning Figure 6-7 is that the plateaus appear to be quite distinct. There is a plateau at the parameter error value of 9 and another at 7.5.

When two correlated inputs are used to excite the vehicle the resulting parameter error measure is biased. This is shown in Figure 6-8. Notice that after 1000 iterations the identification process is producing estimates which are becoming worse. This bias is the result of the input-output crosscorrelation.

Figure 6-9 illustrates the improved identifier performance when the the bias elimination developed in Equations 6.12-6.14 are incorporated in the parameter estimation algorithm. Using the bias elimination scheme, the parameter measurement error approaches that of a single input shown in Figure 6-3. This curve was made by setting P_0 equal to 1000 along the diagonal.

E. COMPUTATIONAL AND STORAGE REQUIREMENTS

In Chapter IV it was found that the computational requirements of the algorithm were within the capability of a microprocessor. This is not the case for the eighth order system analyzed in this chapter. Also, since the parameters of the fourth, sixth and eighth rows are less corrupted by noise, the algorithm could stop searching for those parameters long before the parameters of the second row are found. This capability can be accomplished by using the structure of this identification approach. Recall that Equation 2.3, constructs the $H(k)$ matrix by

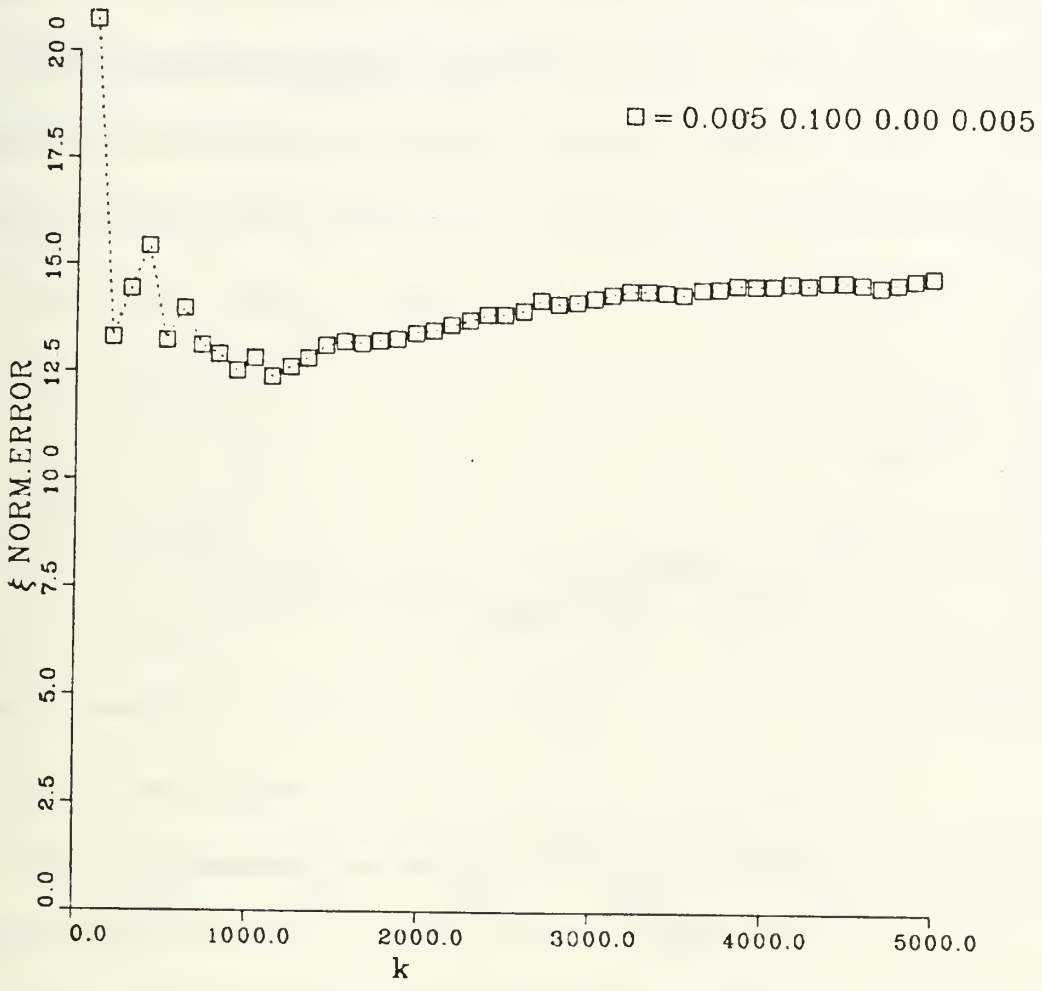

Figure 6-8 Parameter estimation error vs. number of iterations with correlated road noise as the only form of excitation.
Initial conditions, σ =0.0
Road noise, σ =0.1
State measurement noise, σ =0.0
State derivative measurement noise, σ =0.005

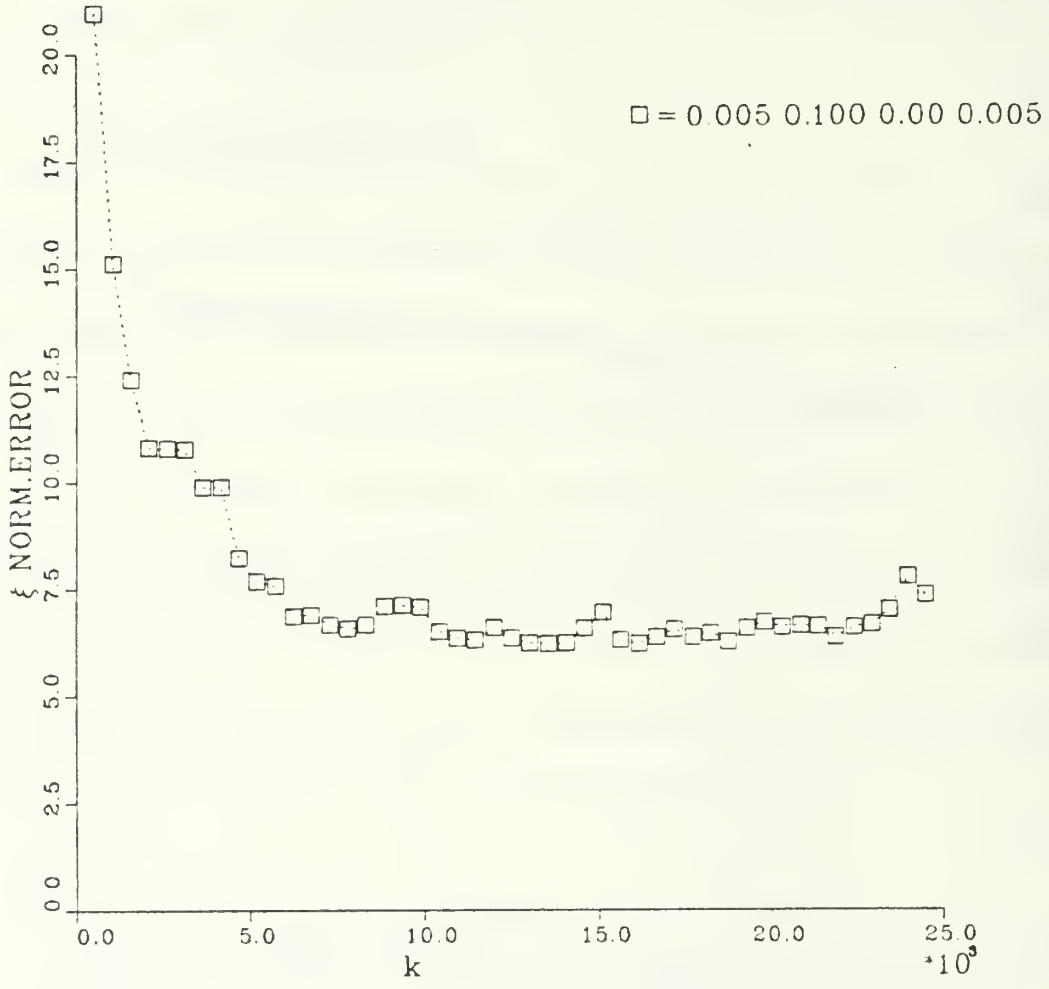

Figure 6-9 Parameter estimation error vs. number of iterations with correlated road noise as the only form of excitation and bias elimination is used.
Initial conditions. σ =0.0
Road noise. σ =0.1
State measurement noise. σ =0.0
State derivative measurement noise. σ =0.005

using disjoint state vectors. Thus, the identifier may be considered as four separate identifiers, which are associated by the state vector alone.

This allows the algorithm to be broken into four separate scalar identifiers, which eliminates the need to perform a matrix inversion. If each of the new, smaller parameter vectors is denoted by θ', then the new algorithm is given by

$$\theta'(k) = \theta'(k-1) + g'(k)[z_i(k) - x^T\theta'(k-1)] \tag{6.16}$$

$$g'(k) = \frac{P'(k-1)x}{x^T P'(k-1)x + r_i} \tag{6.17}$$

$$P'(k) = [\, I - g'(k)x^T]P'(k-1) \tag{6.18}$$

where $g'(k)$ is now the gain vector. The separation of the parameter vector will require this algorithm to be performed four times during each iteration, but as a row of parameters is identified it can be eliminated from the algorithm. This will reduce the computational requirements and thus the cycle time.

Each time through the algorithm the microprocessor will perform 270 multiplications, 8 divisions, 255 additions, 344 index register increments, 277 loads, and 96 stores. Using the same operation times as in Chapter IV, the total time required per shortened parameter vector is 3628 μseconds. Therefore, one iteration requires 14,512 μseconds or 0.0145 seconds. The asynchronous algorithm developed in this thesis does not depend on the time interval between measurements. This will allow the algorithm to reduce the interval between measurements as parameter rows are identified and dropped from the algorithm. The memory requirement is 2800 bytes for all four algorithms.

F. SUMMARY

Parameter identification in high order systems is fraught with peril. As the number of parameters to be identified increases, so does the dimension of the search space. This increase in the search volume results in plateauing. When the noise added to the state derivative is large as in the case of road noise, then these plateaus become very expansive, and it is not clear that identification is even possible. However, in most applications the parameters of the body are probably more important than those of the wheels.

The observant reader might have noticed that in none of the parameter estimation error curve of this chapter, did the error reduce below about 5. In fact, on several simulations not shown, the error did reach acceptable values. These low errors were obtained when the initial conditions were very large and the noise added to the measurements was very small.

VII. SUMMARY AND CONCLUSIONS

A. INTRODUCTION

In the previous six chapters, the theoretical results of parameter identification were presented, and applied to the problem of estimating the dynamic parameters of a land vehicle. The simulation studies performed in the last four chapters demonstrated the capabilities of each approach. In this chapter an overview of the result is given, and suggestions for further study presented.

B. RESULTS

The batch-process least squares procedure of Chapter III was judged to be impractical because it required the storage of all the data before the identification process could begin. The recursive form of this algorithm is the Kalman filter identifier without the theoretical insight of how to choose P(0).

The identification of the dynamic parameters of a land vehicle is possible with the Kalman filter identifier. The practicality of this approach is in question, however. There are several practical drawbacks to the identification scheme presented here that make it extremely costly to implement. Most autonomous vehicles have an inertial navigation system on board, thus the elements of the state vector dealing with the vehicle bodies position, velocity, acceleration, and their angular counterparts are available. These measurements will probably be

quite accurate and have reasonable signal to noise ratios. The wheel positions, velocities and accelerations are another matter. It is not likely that an accelerometer would be mounted on each wheel to measure its acceleration. Each wheel's acceleration could then be integrated to obtain its velocity, which in turn could be integrated to find the wheel's position. It might be possible to find the acceleration on each wheel with respect to the body with a less sophisticated device than an accelerometer, which could then be added to the vehicle's acceleration to find the inertial space acceleration on the wheel. In either case, there is a requirement to develop the states of each wheel. On vehicle with several wheels (more than four) this will become quite hardware intensive.

The stochastic gradient identifier using the steepest descent algorithm was extremely susceptible to plateauing. The Lyapunov gain matrix produced better results, but the R(k) matrix was correlated with the H(k) matrix which would necessarily produce a bias. The extent of this bias was not analyzed.

The problem of plateauing observed in the last chapter, is unresolved. Plateauing is quite common in least squares identification, or the Kalman filter approach. This problem was also seen in the stochastic gradient algorithm to a much worse degree. The algorithm did obtain acceptable estimates for the parameters that were not corrupted by road noise. In the case of correlated inputs, the scheme was able to identify the rear wheel parameters. If the center of gravity of the vehicle is known, and the dynamic elements of the front and rear wheels are the same, it should be possible to obtain the front wheel parameters

from the rear wheel elements. In order to remove the correlate between the inputs to the front and rear wheels. it was necessary to allow the time interval between samples to vary. This variation in sampling interval allowed the system to take measurements when when the rear wheel is over a point where a front wheel measurement was taken. The problem of state measurement noise bias worsened with the increase in state size. This is not surprising, but presents a severe limitation to this form of identification. Two possible solutions to be investigated are presented in the next section.

The scheme used to remove the correlation between the input and the state was successful. The implementation of this scheme might prove difficult. Whenever there is a requirement for an object to be over exactly one location at the proper time there is not much room for error. The question of exactly how accurate this timing must be. to obtain acceptable results. might best be answered on a real vehicle.

The Kalman filter identifier possessed the best characteristic of the three identification schemes chosen for analysis in this thesis, but it has a very limited capability to perform in the scenario described in Chapter I. Under different conditions it might perform much better. If for instance. several of the vehicles dynamic elements were known. but some others were likely to change very slowly throughout their operating life time. the number of parameters to be identified would decrease. This decrease would allow the Kalman filter to function more effectively.

The least squares approach works very well when there are large initial conditions, even in the presence of substantial measurement noise. The Kalman filter identifier, being a recursive least squares algorithm, should perform as well with the proper initial conditions. Thus, if the identifier is activated after large initial conditions have been imparted to the vehicle, it should produce much better results. This procedure would improve both the plateauing problem and the state measurement noise bias.

C. FURTHER STUDY

This thesis dealt with a methodology and did not attempt to use the parameters of an actual vehicle or the limitations of actual measurement equipment. Before any experimentation is done on an actual vehicle with real measurement equipment, several aspects of the work done in this thesis deserve further investigation.

The present analysis shows that the signal to noise ratio of the vehicle state sensors must be approximately 40 dB. This is an rather large signal to noise ratio for this type of sensor. There are two possible solution to this problem which should be investigated. The first is the extended Kalman filter identifier [Ref. 6:pp. 422-23]. Theoretically this approach estimates the state and identifies the parameters. There are no perpetual motion machines, however, and the price for this more complicated identifier will certainly be slower convergence. If on the

other hand, the slower convergence is not excessive, this approach would then be an excellent prospect for a vehicle identifier.

The other possible solution deals with the stochastic gradient approach. This approach was unsuccessful using the Venter gain matrix, and the Lyapunov gain matrix approach was abandoned because of theoretical inconsistencies. The resulting effect of this inconsistency was not investigated. This approach may perform somewhat better than the Kalman filter, even though theoretically flawed, when moderate state measurement noise is present.

The investigation of parameter tracking was not completed in this thesis. Although this area of analysis is easily interpolated from the previous study, no real experimentation was done. Therefore, the study of the parameter tracking problem, and which approach to use remains to be completed.

In chapter VI, the problem of plateauing was identified, but no analysis was done to try to reduce the probability of its occurrence. The use of a reinitialization scheme, such as the one discussed in Chapter IV, may allow the identifier to break through plateaus. This would reduce this serious problem to one of identification of plateaus, and decisions on how wide a plateau must be before reinitialization is preformed.

LIST OF REFERENCES

1. Adam, J. A., "Technology '86: Aerospace and Military," IEEE Spectrum, pp. 76-81, January 1986.

2. Nitao, J. J. and Parodi, A. M., "A Real-Time Reflexive Pilot for an Autonomous Land Vehicle," IEEE Control Systems, pp.14-23, February 1986.

3. Horgan, J., "Roboticists Aim to Ape Nature," IEEE Spectrum, pp. 66-71, February 1986.

4. Mendel, J. M., **Discrete Techniques of Parameter Estimation: The Equation Error Formulation**, Mercel Dekker, Inc., 1973.

5. Gelb, A., **Optimal Estimation**, The M.I.T. Press, 1974.

6. Ljung, L. and Soderstrom, T., **Theory and Practice of Recursive Identification**, The M.I.T. Press, 1986.

7. Goodwin, G. C. and Sin, K. S., **Adaptive Filtering Prediction and Control**, Prentice-Hall, Inc., 1984.

8. Cooper, G. R. and McGillem, C. D., **Probabilistic Methods of Signal and System Analysis**, Holt, Rinehart and Winston, Inc., 1971.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943-5002	2
2.	Chairman, Code 62 Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, California 93943-5000	1
3.	Prof. Robert B. McGhee, Code 52MZ Department of Computer Science Naval Postgraduate School Monterey, California 93943-5000	1
4.	Prof. Srbijanka Turajlic, Code 62TC Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, California 93943-5000	1
5.	Prof. Roberto Cristi, Code 62CX Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, California 93943-5000	1
6.	Prof. Harold Titus, Code 62TS Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, California 93943-5000	1
7.	Mr. Scott Harmon, Code 442 Naval Ocean Systems Center San Diego, California 92152	1
8.	Dr. Andrew Chang FMC Corporation Central Engineering Laboratories Santa Clara, California 95092	1

9.	Dr. James W. Lowrie Martin Marietta Denver Aerospace P.O. Box 179 Denver, Colorado 80201	1
10.	Dr. William Isler Defense Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, Va 22209	1
11.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
12.	Capt. James Kessler 13287 Kitten Ct. Woodbridge, Virginia 22193	1