COMBINED STATE AND PARAMETER ESTIMATION FOR ON-LINE APPLICATIONS

by

PETER S. MAYBECK

S. B., Massachusetts Institute of Technology

(1968)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1972

Signature of Author: _____

Department of Aeronautics and Astronautics
December 15, 1971

Certified by: _____

Thesis Supervisor

_____

Thesis Supervisor

_____

Thesis Supervisor

Accepted by: _____

Chairman, Departmental Graduate Committee

Schweppe, F. C. (1965), "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Transactions on Information Theory, Vol. IT-11, No. 1, pp. 61-70, January.

Schweppe, F. C. (1970), Course Notes for "Uncertain Dynamic Systems," printed at Massachusetts Institute of Technology, Cambridge, Mass.

Shellenbarger, J. C. (1966), "Estimation of Covariance Parameters for an Adaptive Kalman Filter," Proceedings of the National Electronics Conference, Vol. 22, pp. 298-302, Chicago, Illinois, October.

Sims, F. L. and Lainiotis, D. G. (1969), "Recursive Algorithm for the Calculation of the Adaptive Kalman Filter Weighting Coefficients," IEEE Transactions on Automatic Control, Vol. AC-14, No. 2, pp. 215-217, April.

Speyer, J. L. and Desai, M. N. (1970), "An Algorithm for Propagating the Square Root of the Covariance Matrix," MIT Instrumentation Laboratory ATC Memo No. 1-70, Cambridge, Mass., June.

Steiglitz, K. and McBride, L. E. (1965), "A Technique for the Identification of Linear Systems," IEEE Transactions on Automatic Control, Vol. AC-10, No. 4, pp. 461-464, October.

Widnall, W. S. (1968), Applications of Optimal Control Theory to Computer Controller Design, Research Monograph No. 48, The M.I.T. Press, Cambridge, Mass.

Wilks, S. S. (1963), Mathematical Statistics, John Wiley & Sons, Inc., New York, New York.

BIOGRAPHY

Peter S. Maybeck was born February 9, 1947 in Queens, New York. While attending Baldwin High School in Baldwin, New York, he won the Alicia Patterson Scholarship, an annual award presented to the outstanding senior on Long Island. He entered M.I.T. and received the Bachelor of Science Degree in June of 1968. During this time, he became a member of Sigma Chi Fraternity and Sigma Gamma Tau honorary, and also served as President of Tau Beta Pi. Industrial experience has included summer employment with Kaman Avidyne and two summers at the M.I.T. Instrumentation Laboratory, now the Charles Stark Draper Laboratory. Partially on the basis of papers which won the James Means Memorial Prize and the Luis de Florez Prize, he was permitted by the Department of Aeronautics and Astronautics to enter directly into the doctoral program. His graduate education has been supported financially by the Fannie and John Hertz Foundation through its national fellowship program.

In June of 1968, he and Miss Beverly J. Bedsaul were married, and they and their daughter, Kristen, currently reside in Cambridge, Massachusetts. After receiving his degree, Mr. Maybeck will serve as an engineer in the Flight Control Division of the Air Force Flight Dynamics Laboratory.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Professor Wallace E. Vander Velde, Professor John J. Deyst, Jr., and Doctor William S. Widnall. Not only have they provided invaluable assistance and support for this research, but have also served as an inspirational standard of excellence: the two most outstanding professors I have had the privelege to study under and an engineer who is a professional in the truest sense of the word, three remarkable individuals to know as people. I am especially indebted to Professor Vander Velde for the painstaking care with which he read and suggested improvements for the manuscript.

I would also like to thank Professor James E. Potter and Professor Frederick C. Schweppe for their assistance during the formative stages of this research. Throughout this study, I have enjoyed stimulating discussions with my friends and associates M. Edward Womble, Jr., William R. Killingsworth, Jr., James C. Deckert, and J. Edwin Jones, Jr. Doctor Donald C. Fraser generously made the facilities of the Digital Autopilot Group of the Charles Stark Draper Laboratory available for this research, and also arranged for the preparation of the final draft and figures by the Technical Publications Group under Robert E. Weatherbee. I am very grateful to the Fannie and John Hertz Foundation for providing financial support for my entire graduate education.

My sincerest appreciation goes to my wife Beverly, who not only typed all drafts of this document with care and precision, but also proofread the thesis to ensure clarity of presentation at a layman's level. Moreover, her unending support and encouragement were essential to the development of early, rudimentary ideas into the final realization of this document.

ACKNOWLEDGMENTS (cont)

This report was prepared under DSR Project 55-23890, sponsored by the Manned Spacecraft Center of the National Aeronautics and Space Administration through Contract NAS 9-4065.

The publication of this report does not constitute approval by the Charles Stark Draper Laboratory or the National Aeronautics and Space Administration of the findings or the conclusions contained herein. It is published only for the exchange and stimulation of ideas.

COMBINED STATE AND PARAMETER ESTIMATION FOR ON-LINE APPLICATIONS

by

PETER S. MAYBECK

Submitted to the Department of Aeronautics and Astronautics on December 15, 1971, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

ABSTRACT

This research develops an effective and practical means of simultaneously estimating the states and parameters of a linear dynamic system model, based upon noisy measurements of the actual system outputs. The accuracy of the resulting state estimates far surpasses that of a filter that does not simultaneously estimate the uncertain parameters. Unlike previously suggested techniques for providing such capability, this method of estimation directly and systematically incorporates the physical knowledge that the parameters are more slowly varying than the states. Moreover, it does not require complete apriori parameter statistics, but can utilize any such information that is available. Thus, it is unique in its ease of application.

Under very general assumptions, the parameter estimator is shown to be consistent and asymptotically unbiased, efficient, and normally distributed about the true value. The behavior of the state estimator converges to that of the optimum linear state estimator based upon the true parameter values. Further performance analysis is provided by the concept of an ambiguity function, which also serves as a useful design tool to ensure adequate performance from the estimators generated in this research.

As means to provide on-line applicability, approximating the iterative solution procedure, including only the most significant terms in the estimator equations, and precomputing various required quantities are all shown to yield very satisfactory performance. The computational feasibility is further enhanced by methods that do not inherently involve approximations: using an advantageous canonical state space form, exploiting symmetry, and modifying the incorporation of measurements. Extensive computational results substantiate the theoretical developments and demonstrate the ability of the proposed technique to produce a viable on-line estimator.

Thesis Advisors:

Wallace E. Vander Velde, Sc.D.
Title: Professor of Aeronautics and Astronautics

John J. Deyst, Jr., Sc.D.
Title: Assistant Professor of Aeronautics and Astronautics

William S. Widnall, Sc.D.
Title: Senior Staff Engineer, Intermetrics, Inc.

TABLE OF CONTENTS

Chapter		Page
	NOTATION AND TABLE OF SYMBOLS	x
I	INTRODUCTION	1
	1.1 Problem Statement and Motivation	1
	1.2 Choice of Estimation Method	2
	1.3 Previous Research	6
	1.4 Summary of Thesis	9
II	STATE ESTIMATION WITHOUT SIMULTANEOUS ESTIMATION OF PARAMETERS	12
	2.1 Maximum Likelihood Estimation of the State Alone	12
	2.1.1 The System Model	13
	2.1.2 Derivation of the State Estimator	15
	2.1.3 Statistics of the Estimation Errors	25
	2.1.4 Statistics of the Measurement Information and Residuals	26
	2.1.5 Inverse Covariance Form	28
	2.2 Error Sensitivity of Optimal State Estimators	38
III	COMBINED STATE AND PARAMETER ESTIMATION	48
	3.1 Ability to Perform the Estimation	54
	3.2 Derivations of the Likelihood Equations	59
	3.2.1 Likelihood Functions of the Form $\ell n\ f(\underline{x},\underline{Z}\|\underline{a})$	65
	3.2.2 Likelihood Functions of the Form $\ell n\ f(\underline{x},\underline{a}\|\underline{Z})$	78
	3.2.3 Likelihood Functions of the Form $\ell n\ f(\underline{Z}\|\underline{x},\underline{a})$	85
	3.2.4 Likelihood Functions of the Form $\ell n\ f(\underline{x}\|\underline{Z},\underline{a})$	91
	3.2.5 Perfect Measurements of Part of the State - A Special Case	93
	3.3 Iterative Solution to the Likelihood Equation	98
	3.4 Estimate Computations	109

Chapter		Page
IV	PERFORMANCE ANALYSIS	122
	4.1 Asymptotic Properties	123
	4.1.1 Regularity Conditions	129
	4.1.2 Asymptotic Distribution of the Score	132
	4.1.3 Convergence of the Maximum Likelihood Estimator	133
	4.1.4 Asymptotic Distribution of the Maximum Likelihood Estimator; Consistency	136
	4.1.5 Asymptotic Efficiency of the Maximum Likelihood Estimator	141
	4.1.6 Summary of ML Parameter Estimator Asymptotic Properties	142
	4.1.7 Combined State and Parameter Estimation	144
	4.2 Ambiguity Functions	149
V	METHODS OF ATTAINING ON-LINE APPLICABILITY	166
	5.1 Approximations to the Solution Algorithm	169
	5.2 On-Line Conceptualizations	175
	5.3 Approximations to the Likelihood Equations	186
	5.3.1 Inclusion of Only Weighted Least Squares Terms	187
	5.3.2 Towards a Desirable Form	194
	5.4 Feasible Precomputations	197
	5.5 Advantageous State Space Representations	202
	5.5.1 Standard Observable Phase Variables; Scalar Measurements	204
	5.5.2 Modified Jordan Canonical Form; Scalar Measurements	205
	5.5.3 Vector Measurements	213
	5.6 Exploiting Symmetry	215
	5.7 Modified Measurement Incorporation	218
VI	COMPUTATIONAL RESULTS	225
	6.1 Second Order System with Uncertain Parameters in $\underline{\Phi}$	225

Chapter			Page
		6.1.1 Prediction of Performance	229
		6.1.2 Full-Scale Estimator	233
		6.1.3 Effects of Apriori Statistical Information	243
		6.1.4 On-Line Technique; Parameter Estimate Every N Samples	245
		6.1.5 On-Line Technique; More Frequent Parameter Estimation	250
		6.1.6 Inclusion of Weighted Least Squares Terms Only	252
		6.1.7 Precomputations	256
		6.1.8 Time Varying System Parameter	259
	6.2	Second Order System with Uncertain Parameters in $\underline{B}$	262
		6.2.1 Prediction of Performance	265
		6.2.2 Full-Scale Estimator	265
		6.2.3 On-Line Techniques	270
	6.3	Apollo Thrust Vector Controller	273
		6.3.1 The Need for Parameter Estimation	281
		6.3.2 Full-Scale Estimator Performance	282
		6.3.3 On-Line Estimation	290
		6.3.4 Canonical State Space Form	293
		6.3.5 Comparison to Extended Kalman Filter	297
VII	CONCLUSION		305
	7.1	Summary of Results	305
	7.2	Recommendations for Future Investigations	307
	APPENDICES		
	A	Estimator Equations that Allow Uncertainties in $\underline{H}(i)$	310
	B	Discrete Representation of Continuous Dynamic Systems	313
	C	Approximate One-Step Recursions for Fixed-Length Memory State Estimators	317
	D	Recursive State Estimation Equations for $\ell n\ f(\underline{Z}\|\underline{x},\underline{a})$ or $\ell n\ f(\underline{Z}_N\|\underline{x},\underline{a})$ with $\underline{Q} = \underline{0}$	325
	E	Evaluation of $\underline{J}\left[i, \underline{a}_t\right]$ for the Likelihood Function $\ell n\ f(\underline{Z}(i)\|\underline{a})$	332
	F	Computational Requirements	337
	G	An Approximate Closed Form Estimator of Uncertain Entries in $\underline{\Phi}$ or $\underline{B}$	345
	REFERENCES		354

NOTATION

1) Matrices are represented by upper case letters with underscoring; vectors appear as lower case letters with underscoring; scalar quantities are not underscored. The one notable exception to this convention is $\underline{Z}$, a composite vector made up of a time history of $\underline{z}$ vectors. The k-th component of a general vector $\underline{v}$ is denoted by v_k; the element appearing in the k-th row and ℓ-th column of a general matrix $\underline{M}$ is denoted by $M_{k\ell}$.

2) $\underline{I}$ is the identity matrix; $\underline{0}$ is either a matrix or a vector composed entirely of zeroes, the differentiation being clear from context.

3) $\underline{M}^T$ is the transpose of the matrix $\underline{M}$; $\underline{M}^{-1}$ is the inverse of a square, nonsingular matrix $\underline{M}$; $|\underline{M}|$ or det $\underline{M}$ represents the determinant of $\underline{M}$; tr $\underline{M}$ is the trace of $\underline{M}$ - the sum of its diagonal elements. The time derivative of $\underline{M}$ is denoted by $\dot{\underline{M}}$, and the partial derivative of $\underline{M}$ with respect to the scalar s is $\partial\underline{M}/\partial s$. The partial derivative of a scalar s with respect to a vector $\underline{v}$ is defined to be a vector transpose; a recurrent example in the sequel is $\partial L/\partial \underline{a} = \underline{s}^T$.

4) The notation $\underline{M} \geq \underline{0}$ means the matrix $\underline{M}$ is positive semidefinite; and $\underline{M} > \underline{0}$ conveys that $\underline{M}$ is positive definite. If $\underline{M}$ and $\underline{N}$ are both matrices, then $\underline{M} \geq \underline{N}$ means that the matrix $(\underline{M}-\underline{N})$ is positive semidefinite, and similarly for $\underline{M} > \underline{N}$.

5) The indices i and j are time indices: the i-th and j-th measurement instants; k may also be used as a time index.

6) When a quantity is expressed as $x(i;a,b,c)$, the variables to the right of the semicolon are included in the notation to clarify the appropriate terms with which to evaluate the quantity $x(i)$: one example would be $\hat{\underline{x}}(i;\underline{a}^*(i-1))$.

7) The notations $f(\cdot)$ and $F(\cdot)$ represent probability density and distribution functions respectively. Conditional probability densities are expressed as $f(\cdot|\cdot)$. Similarly the expectation and conditional expectation operators are written as $E\{\cdot\}$ and $E\{\cdot|\cdot\}$ respectively; an alternate notation for $E\{\cdot\}$ is $\{\bar{\cdot}\}$. $P\{A\}$ denotes the probability of the event A.

8) Exponentials and natural logarithms are written as $\exp(\cdot)$ and $\ell n(\cdot)$ respectively.

9) Arg $(\max_{\underline{x}} \underline{A})$ is the value of $\underline{x}$ that maximizes $\underline{A}$ as a function of $\underline{x}$.

TABLE OF SYMBOLS

Symbol	Defined or First Used
$\underline{A}$	(2.2.5)
$\underline{\tilde{A}}$	(3.2.38)
$\underline{a}$	(2.1.10)
a_ι	Sec. 3.2.2
$\underline{a}_t$	(2.2.1)
a_u	Sec. 3.2.2
$\underline{a}^*$	(3.2.11)
$\underline{a}_*$	Sec. 3.4
$\underline{\hat{a}}$	Sec. 4.1.3
$\underline{\hat{a}}_x$	Sec. 4.1.7
$\underline{a}^\circ$	Sec. 4.1.4
$\underline{B}$	(2.1.1)
$\underline{B}(t)$	(3.4)
$\underline{B}'$	(5.5.8)
$\underline{B}'_a$	(5.5.36)
$\underline{B}'_b$	(5.5.36)
$\tilde{B}_{k\ell}$	(G-10)
b	(4.2.5)
$\underline{C}$	(2.2.27)
$\underline{\mathfrak{C}}$	(C-8)
$\underline{D}$	(3.2.70)
$\underline{D}(t,t_i)$	(B-10)
$\underline{d}$	(4.2.23)
$d_{k\ell}$	(G-14)
$d'_{k\ell}$	(G-20)

Symbol	Defined or First Used
$\partial\underline{A}/\partial a_\ell$	(3.4.20)
$\partial\underline{A}_j/\partial a_\ell$	(5.7.21)
$\partial\tilde{\underline{A}}/\partial a_\ell$	(3.2.37)
$\partial\underline{B}/\partial a_\ell$	(3.4.21)
$\partial\underline{B}'/\partial a_\ell$	(5.5.38)
$\partial\underline{B}'_a/\partial a_\ell$	(5.5.38)
$\partial\underline{D}/\partial a_\ell$	(3.2.69)
$\partial\underline{H}/\partial a_\ell$	App. A
$\partial\underline{H}'/\partial a_\ell$	Sec. 5.5.3
$\partial\underline{M}/\partial a_\ell$	(3.4.19)
$\partial\underline{M}_j/\partial a_\ell$	(5.7.20)
$\partial\underline{M}'/\partial a_\ell$	(5.5.40)
$\partial\underline{M}'_{aa}/\partial a_\ell$	(5.5.41)
$\partial\underline{M}'_{ab}/\partial a_\ell$	(5.5.41)
$\partial\underline{M}'_{bb}/\partial a_\ell$	(5.5.41)
$\partial\underline{P}/\partial a_\ell$	(3.4.33)
$\partial\underline{P}'_{aa}/\partial a_\ell$	(5.5.41)
$\partial\underline{P}'_{ab}/\partial a_\ell$	(5.5.41)
$\partial\underline{P}'_{bb}/\partial a_\ell$	(5.5.41)
$\partial\tilde{\underline{P}}/\partial a_\ell$	(3.2.37)
$\partial\bar{\underline{x}}/\partial a_\ell$	(3.4.21)
$\partial\bar{\underline{x}}'/\partial a_\ell$	(5.5.39)
$\partial\hat{\underline{x}}/\partial a_\ell$	(3.4.34)
$\partial\hat{\underline{x}}_j/\partial a_\ell$	(5.7.19)
$\partial\hat{\underline{x}}'_a/\partial a_\ell$	(5.5.39)
$\partial\tilde{\underline{x}}^-/\partial a_\ell$	(3.2.37)
$\partial\underline{x}_*/\partial a_\ell$	(3.2.69)
$\partial\underline{\Phi}/\partial a_\ell$	(3.4.19)

Symbol	Defined or First Used
$\partial \underline{\Phi}'/\partial a_\ell$	(5.5.37)
$\partial \underline{\Phi}_a'/\partial a_\ell$	(5.5.37)
$\underline{e}$, $\underline{e}(i^+)$	(2.1.45)
$\underline{e}(i^-)$	(2.2.14)
e_k	(3.3.15)
$\underline{F}(t)$	(3.4)
$\underline{\mathcal{F}}$	(4.2.26)
$\underline{\mathcal{F}}_j$	(4.2.30)
$f(\underline{a})$	(3.2.47)
$f_{k\ell}$	(G-12)
$f'_{k\ell}$	(G-18)
$\underline{G}$	(2.1.1)
$\underline{G}'$	(5.5.9)
$\underline{G}(t)$	(3.4)
$\underline{\mathcal{G}}$	(5.3.12)
$g_{k\ell}$	(G-13)
$\underline{H}$	(2.1.2)
$\underline{H}_j$	(5.7.5)
$\underline{H}'$	(5.5.10)
$\underline{h}^T$	(5.5.11)
$\underline{h}_j^T$	(5.7.16)
$\underline{h}'^T$	(5.5.29)
$\underline{h}_k'^T$	(5.5.30,31)
$\hbar_k$	Sec. 4.1.1
$\hbar'_{k\ell}$	Sec. 4.1.1
$\hbar''_{k\ell}$	Sec. 4.1.1
$I(\underline{\delta},\epsilon)$	Sec. 4.1.3
$I(\underline{\delta},\epsilon,i_\circ)$	Sec. 4.1.4
$\underline{\mathcal{J}}(i,1)$	(2.1.74)
$\underline{\mathcal{J}}(i,i-N+1)$	(2.1.90)
$i_\circ$	Sec. 4.1.4
$\underline{J}(i,\underline{a}_t)$	(4.1.4)
$\underline{J}(i,\underline{a}_t,\underline{a})$	(4.1.11)
$\underline{J}(i,\underline{\Theta}_*)$	(3.3.7)
$\underline{J}(\underline{Z},\underline{a})$	Sec. 5.1
$\underline{J}^1(i,\underline{a}_t)$	(4.1.13)
$\underline{J}^1(i,\underline{a}_t,\underline{a})$	(4.1.9)
$\underline{J}^1(\underline{Z},\underline{a})$	Sec. 5.1
$\tilde{\underline{J}}$	(5.2.2)
K	(4.1.65)
$\underline{K}$	(2.1.35)
$\underline{K}_j$	(5.7.9)
$\tilde{\underline{K}}$	(3.2.31)
$\underline{k}_j$	(5.7.17)
L	(2.1.11)
ℓ	(2.1.10)
$\underline{M}$	(2.1.14)
$\underline{M}_j$	(5.7.11)
$\underline{M}_\circ^{-1}$	(2.1.61)
$\underline{M}_t$	(2.2.16)
$M(\underline{Z})$	(4.1.64)
$\tilde{\underline{M}}$	(3.2.32)
$\tilde{\underline{M}}^*$	(C-5)
m	Sec. 2.1.1
N	(2.1.86)
N(,)	(3.2.34)

Symbol	Defined or First Used
$\underline{N}(t,t_i)$	(B-17)
$N_r(\underline{\theta}^*)$	(3.3.17)
n	Sec. 2.1.1
$\underline{P}$	(2.1.16)
$\underline{P}_o$	(2.1.9)
$\underline{P}_t$	(2.2.17)
$\underline{P}'_{aa}$	(5.5.41)
$\underline{P}'_{ab}$	(5.5.41)
$\tilde{\underline{P}}$	(3.2.33)
$\tilde{\underline{P}}_o(i-N)$	(3.2.28)
$\tilde{\underline{P}}^*$	(C-6)
$\underline{P}_*^{-1}$	(3.2.24)
p	(3.3.23)
$\underline{Q}$	(2.1.5)
$\underline{Q}(t)$	(B-4)
$\underline{q}_k$	(2.2.13)
$\underline{R}$	(2.1.6)
$\underline{R}_j$	(5.7.6)
r	Sec. 2.1.1
$\underline{r}$	(4.1.4)
$\tilde{\underline{r}}$	(4.2.22)
$\underline{S}_k$	(5.5.21)
s	Sec. 2.1.1
$\underline{s}(\underline{Z},\underline{a})$	(4.1.7)
$\underline{s}^{-1}(\underline{Z},\underline{a})$	(4.1.6)
$\tilde{\underline{s}}$	(5.2.1)
$\underline{T}$	(5.5.3)
$\underline{T}_1$	(5.5.18)
$\underline{T}_2$	(5.5.23)
$\underline{T}_{2k}$	(5.5.26,28)
$\underline{T}_{2k\ell}$	(5.5.27)
$t_{2k\ell}$	(5.5.28)
$\underline{u}$	(2.1.1)
$\underline{u}(t)$	(3.4)
$\underline{u}^*(i-N)$	(2.1.93)
V	(2.2.30)
$\underline{v}$	(2.1.2)
$\underline{v}_j$	(5.7.5)
$\underline{W}$	(5.3.3)
W_k	(G-5)
$\mathcal{W}$	(2.2.30)
$\underline{w}$	(2.1.1)
$\underline{w}(t)$	(3.5)
$\underline{x}$	(2.1.1)
$\underline{x}(t)$	(3.4)
$\underline{x}'$	(5.5.3)
$\underline{x}_o$	Sec. 3
$\bar{\underline{x}}$	(2.1.13)
$\hat{\underline{x}}$	(2.1.15)
$\hat{\underline{x}}\big\|_{\underline{a}\to\underline{a}^*}$	(3.2.13)
$\hat{\underline{x}}_j$	(5.7.7)
$\hat{\underline{x}}_o$	(2.1.8)
$\hat{\underline{x}}'_a$	(5.5.39)
$\underline{x}^*$	(2.1.26)
$\underline{x}_*$	(3.2.24)

Symbol	Defined or First Used
$\underline{\tilde{x}}$	(2.1.78)
$\tilde{x}_k$	(G-8)
$\underline{\tilde{x}}_\circ(i-N)$	(3.2.27)
$\underline{\tilde{x}}^-$	(3.2.29)
$\underline{\tilde{x}}^+$	(3.2.30)
$\underline{x}^\circ$	(4.2.16)
Y	(4.1.48)
$\underline{\bar{Y}}$	(2.1.66)
$\underline{\hat{Y}}$	(2.1.67)
$\underline{Z}$	Sec. 2.1.2
$\underline{Z}_N$	(2.1.86)
$\underline{Z}^\circ$	(4.2.17)
$\underline{z}$	(2.1.2)
$\underline{z}_j$	(5.7.5)
$\underline{z}^*(i)$	(D-13)
$\underline{z}^*(i-N)$	(D-13)
$\underline{z}_*$	(3.2.58)
$\underline{\hat{z}}_*$	(3.2.24)
$\underline{z}^\circ$	(4.2.15)
α	Sec. 5.5.2
β	Sec. 5.5.2
$\underline{\beta}(t)$	(3.4)
$\underline{\Gamma}_{ij}$	(4.2.30)
$\underline{\gamma}(\underline{Z},\underline{a})$	(3.4.7)
$\underline{\delta}$	Sec. 4.1.3
δ_{jk}	(4.1.16)
$\delta(t-\tau)$	(B-4)

Symbol	Defined or First Used
$\Delta\underline{a}^*$	(5.3.22)
$\Delta\underline{B}$	(2.2.12)
$\Delta V/\Delta a_\ell$	(2.2.31)
$\Delta V'/\Delta a_\ell$	(2.2.32)
$\Delta\underline{\Phi}$	(2.2.11)
ϵ	(4.1.33)
ζ	Sec. 4.1
$\zeta_\circ$	Sec. 4.1.4
ζ'	Sec. 4.2
$\underline{\eta}$	(3.4.17)
Θ	Sec. 4.1.1
$\tilde{\Theta}$	Sec. 4.1.1
$\underline{\theta}$	(3.3.1)
$\underline{\theta}_k$	(3.3.13)
$\underline{\theta}_t$	Sec. 4.1
$\underline{\theta}(\underline{Z})$	Sec. 4.1
$\underline{\theta}^*$	(3.3.2)
$\underline{\theta}_*$	(3.3.3)
$\underline{\Lambda}$	(3.2.94)
$\underline{\Lambda}'$	(3.2.97)
$\underline{\lambda}$	(3.2.94)
$\underline{\lambda}'$	(3.2.97)
λ_k	(5.5.22)
$\lambda_\circ$	(3.2.94)
$\lambda_\circ'$	(3.2.97)
$\lambda(\theta)$	(3.3.11)
μ_k	Sec. 5.5.2

Symbol	Defined or First Used
$\underline{\nu}$	(2.2.10)
σ_k	(5.5.21)
$\underline{\sigma}^1$	(4.1.30)
$\underline{\bar{\sigma}}$	(4.1.31)
$\underline{\Phi}$	(2.1.1)
$\underline{\Phi}'$	(5.5.7)
$\underline{\Phi}'_a$	(5.5.35)
$\underline{\Phi}'_b$	(5.5.35)
$\underline{\Phi}'_k$	(5.5.20,22)
$\underline{\Phi}^{-1}$	(C-13)
$\underline{\Phi}^{-2}$	(C-21)
$\underline{\Phi}^{-T}$	(C-13)
$\tilde{\Phi}_{k\ell}$	(G-9)
$\underline{\phi}^T$	(3.2.83)
$\underline{\phi}'^T$	(5.5.14)
$\phi(\Theta)$	(3.3.11)
$\underline{\chi}$	(2.1.62)
Ψ_i	(4.2.1)
ψ_i	(4.2.31,41)
$\psi_{i,N}$	(4.2.42)
ψ'_i	(4.2.44)
$\tilde{\psi}'_i$	(4.2.45)
$\tilde{\psi}_{i,N}$	(4.2.43)
$\underline{\Omega}$	(3.4.18)
Ω_j	(5.7.21)
ω_k	(5.5.21)

CHAPTER I

INTRODUCTION

1.1 Problem Statement and Motivation

At the present time, optimal linear estimation and control theory is finding progressively wider realms of application. Various modelling techniques are used to produce an adequate system description in the form of a linear differential or difference equation, driven by known inputs and "white" noises, with which it is possible to develop an optimal state estimator and/or controller. However, the optimality of these devices is dependent upon complete knowledge of the parameters which define the system dynamics. In any practical application, these quantities are known only with some uncertainty, and the performance degradation that results from improperly assigned values can be severe.

The extent of the model uncertainty and the sensitivity of filter performance to such uncertainty varies substantially from one problem to another. For example, in thrust vector control applications, the physics involved are well understood and can lead to a very accurate linear system model, but certain key parameters, as bending mode natural frequency, are typically known imprecisely; an optimal filter-controller combination is sensitively tuned to this parameter, and even a small deviation in its value can cause inadequate, or even unstable, control. On the other hand, the model is less well defined for the dynamics involved in many process control applications. Distillation towers would be a prime example: a wide range of nonlinear dynamics occur within each tower, but practical experience has shown that adequate control is possible by assuming an approximate model of a second order system plus a time delay. Although the towers are

qualitatively similar, the appropriate parameters for the linear model differ for each particular tower, and vary slowly in time as well. Apriori information about these parameters would be, at best, a range of physically admissible values.

Thus, in order to improve the quality of the state estimates, it would be desirable in many instances to estimate a number of uncertain parameters in the dynamics model. A number of methods have been suggested for developing such capacity, as using an extended Kalman filter to solve the nonlinear estimation problem that results from treating the parameters as additional state variables. However, these techniques usually depend upon apriori parameter statistical information, or require a difficult interpretation of physical knowledge about the parameters (i.e., that they are more slowly varying than the states) into the specification of appropriate noise strengths to drive the dynamics model. The objective of this research is to provide a feasible state and parameter estimator that 1) does not require complete apriori parameter statistics, but can utilize any such information that is available, 2) allows the engineer to use his knowledge that the parameters are slowly varying in a direct and physically meaningful manner, and 3) provides, or yields to approximations that provide, both on-line capability and adequate performance.

1.2 Choice of Estimation Method

Having described the general estimation problem to be considered, there are various means of specifying exactly what is meant by an "optimal" estimator, and thus a number of different methods for solving the "optimal estimation problem." The particular choice of the maximum likelihood

technique resulted from many considerations, which will now be enumerated.

First of all, certain properties of a maximum likelihood estimator make it especially attractive. The following theorems have been proven by Cramér (1946) for the case of independent, identically distributed measurements, and Chapter IV will develop the corresponding generalizations to the context of the current problem. If an efficient estimate exists (i.e., if there exists an unbiased estimate with finite covariance such that no other unbiased estimate has a lower covariance), it can always be found through maximum likelihood methods. Further, if an efficient estimate exists, the likelihood equation will have a unique solution that equals the efficient estimate. Under rather general conditions, the likelihood equation has a solution which converges in probability to the true value of the variables as the number of sample elements goes to infinity; i.e., it is consistent. This solution is an asymptotically normal and asymptotically efficient estimate. Kerr (1965) further asserts these additional properties. If any single sufficient statistic for the estimated variable exists, the maximum likelihood estimate will be sufficient, and under very general conditions, it will be at least asymptotically sufficient and unbiased. Even though the estimate will generally be biased for small samples, it will provide the unique minimum attainable variance estimate under the existence of sufficient statistics, attaining the Cramér-Rao lower bound if this is possible. Without sufficient statistics, this optimal behavior cannot be proven for small samples, but it will still be asymptotically optimal and usually a good small-sample estimator.

With regard to the bias, maximum likelihood estimates tend to have the true value of the estimated variable near the center of their distributions, so that the bias is often

negligible. Levin (1964) further states that if the measurement noise variance is small with respect to the actual system output, then the bias in the estimate of the pulse transfer function parameters of the system model will be negligible compared to the standard deviation of the estimate. He also demonstrates that, in the absence of the Gaussian assumption on the noises, the maximum likelihood estimate will at least provide a generalized least squares fit to data.

Besides these desirable properties, other considerations tend to favor maximum likelihood techniques as well. Once a particular conditional density has been propagated, the various possible estimates, as the mode, median, or mean, will often be very close, especially for unimodal densities concentrated about their mean. Therefore, it is logical to choose the estimate that is easiest to compute or approximate. Schweppe (1970) and others have noted that estimates based on the peak value of the density often have this characteristic. Furthermore, the mean and median estimate computations are made complex by truncating the density, as by imposing an admissible range of parameter values, whereas the maximum likelihood estimate is not affected at all (see section 3.2.2). In his work toward simultaneously estimating the system state and noise statistics, Abramson (1968) chose the maximum likelihood method partially because of severe analytical difficulties in a minimum variance formulation. Especially since one objective of this research is to provide on-line algorithms, the rationale of using the estimator which yields the simplest implementation is important, so the use of maximum likelihood techniques is further substantiated.

A potentially simpler technique might be the method of weighted least squares. With the "proper" choice of weighting factors, this method can derive the Kalman filter and other powerful estimation results. However, the appropriate

choice of the weighting factors is usually based upon considerable hindsight gained from comparison to other estimation techniques. The primary conceptual disadvantage to this method is that it does not attempt to propagate statistical or probabilistic information in time, and the knowledge of certain conditional densities provides a complete, rational basis of estimation. Although the available statistical information can often be incorporated into the weighting factors, the estimation technique itself does not reveal how to do so.

Bayesian estimation is conceptually satisfying in that it propagates the conditional density of the variables to be estimated, conditioned upon the values assumed by the sample elements that are actually observed. However, this requires an apriori specification of a parameter density function, and sufficient statistical information to define such a density adequately is quite often lacking in real applications. A maximum likelihood formulation need not suffer from this drawback. The classical likelihood function is the conditional density of the measurements, conditioned upon the value of the uncertain parameters, and the estimation is then a form of hypothesis testing to find the values which maximize the probability of the events (the measurements) that have actually occurred. However, there are also more general forms of likelihood functions, in the form of appropriately defined conditional densities, which can exploit as much, or as little, of the apriori statistical information that is available. In this respect, the maximum likelihood approach is more generally applicable than Bayesian methods. Similarly, Kashyap (1970) has asserted that maximum likelihood methods are more generally applicable than least squares, the instrumental variable method, stochastic approximation, or other methods commonly suggested for system identification purposes.

Perhaps the single greatest disadvantage to maximum likelihood estimation is the lack of theoretical knowledge about the behavior of the estimates for small sample sizes. Much is known about the asymptotic behavior as the number of samples becomes infinite, however, and a considerable amount of practical experience with the method instills confidence in its viability.

1.3 Previous Research

There is a substantial amount of literature in the fields of linear state estimation and system identification, and clearly this work is, in part, a product of the insights gained from previous research. Of the many sources of maximum likelihood techniques, those actively employed were Wilks (1963), Cramér (1946), Rao (1968), Deutsch (1968), Schweppe (1965), and Rauch, Tung and Striebel (1965). Rao (1968) also introduced the method of scoring for parameters as a means of solving the likelihood equations, which will be adopted in the sequel. Other sources of estimation concepts and solution procedures include Schweppe (1970), Jazwinski (1970), Sage (1968), Aoki (1967), and Lee (1964). Finally, Galiana (1969) presented a useful overview of the objectives and methods of system identification.

The various interrelationships of maximum likelihood, weighted least squares, regressions, and stochastic approximation methods of estimation have been developed by Geise and McGhee (1965), Kerr (1965), Levin (1964), Mowery (1965), Ho (1962), and Fagin (1964). In addition, Levin (1964) and Kerr (1965) considered bounds on errors, asymptotic properties, sufficient statistics, and generalized likelihood functions. Sensitivity analyses of optimal estimators that do not adapt to parameter changes were presented by Jazwinski (1970), Aoki (1967), and Neal (1967).

The maximum likelihood criterion was applied by Rogers and Steiglitz (1967) to the estimation of parameters in a rational transfer function, obtaining a nonlinear regression equation to be solved iteratively by a gradient-seeking procedure. Levin (1964) converted maximum likelihood identification of a system pulse transfer function into generalized least squares fit to nonoverlapping sets of data. Then Steiglitz and McBride (1965) combined his approach with an alteration of Kalman's (1958) least squares regression (based on minimizing $\hat{d}y - \hat{n}u$, where $\hat{n}$ and $\hat{d}$ represent estimates of the numerator and denominator of the transfer function, u is the input, and y is the observed output) to obtain iterative solutions to the identification problem. Subsequently, Schulz (1968) applied quasilinearization methods to obtain quadratic convergence characteristics for the same basic approach.

Quasilinearization methods were applied to least squares identification techniques by Kumar and Sridhar (1964). The least squares identification recursions were also developed by Lee (1964), whose ideas were expanded by Ho and Lee (1965) to consider stochastic convergence. Further use of least squares, combined with approximations and invariant imbedding, was developed by Cox (1964), Detchmendy and Sridhar (1966), Sage and Masters (1966), and Sage and Ellis (1966). Here, the approach was to augment the state vector with the parameters and treat the result as a nonlinear estimation problem.

Stochastic approximation techniques and properties of the resulting parameter estimates were developed by Sakrison (1967) and Albert and Gardiner (1967). The relationship of parameter identification to minimum norm solutions for a set of overdetermined equations and to spanning sets of vectors in state space was explored by Pearson (1967) and Ho and Whalen (1963).

To attain a feasible solution to the combined state estimation - parameter identification problem, Magill (1965) and Sims and Lainiotis (1969) considered those processes whose parameter vector values could be modelled as being obtained from a set of apriori known values (an application would be the detection of jet failures in an attitude controller). Essentially, the result was a set of elemental estimators for each value of the parameter vector, with weighting coefficients for each such that the coefficient for the correct parameter vector's estimator would converge to unity, and the others converge to zero.

The combined recursive estimation of state and parameter variables was also considered by Farison (1964) and Farison, Graham, and Shelton (1967), who separated the state estimation and identification of time-correlated Gauss-Markov random parameters. Each "learning" calculation was used to adjust the system model to be used for state estimation. However, like the state augmentation methods, it required an apriori stochastic description of the parameters, an undesirable feature for real applications.

Similarly, Sage and Husa (1969) approached estimation of the state and noise statistics by assuming a form for the apriori densities and applying Bayesian filtering techniques, incorporating the discrete maximum principle and invariant imbedding to reach a solution. Abramson (1968) demonstrated the validity of the maximum likelihood approach, without the requirement for apriori densities for all estimated variables, to the same problem. (Shellenbarger (1966) had previously used maximum likelihood concepts to estimate noise covariances, but considered only approximate solutions and used measurement residuals one at a time.) This motivated a search for a means of estimating states and parameters, without requiring apriori parameter statistical information, by means of maximum likelihood methods. Moreover, it was

desired not to restrict attention to time invariant cases as in Abramson (1968), but to slowly varying or time invariant parameters.

The concept of basing the parameter estimate on only a prespecified number of the most recent data samples was partially motivated by the "finite memory" estimators developed by Schweppe (1964) and Jazwinski (1968) for state estimation. In both of these developments, a necessary assumption was that no dynamic driving noise entered the system. This thesis will develop the corresponding state estimator that allows dynamic driving noise, but of greater importance is the fundamental use of only the N most recent measurements to process the parameter estimate, combined with the assumption that the parameters are constant over this interval of time. Such a basis for parameter estimation is original to this research.

1.4 Summary of Thesis

Chapter II introduces the fundamental concepts of maximum likelihood estimation by deriving the optimal linear state estimator. By presenting an error sensitivity analysis of such a state estimator, it further substantiates the need for simultaneous estimation of uncertain parameters.

In Chapter III, the likelihood equations that correspond to appropriately chosen conditional densities are derived. The solution to these equations yields the mode of the density as a function of the variables being estimated, thereby defining the optimal state and parameter estimates. After investigating the relative merits of each particular formulation, an efficient iterative solution technique is introduced, and the full-scale estimation algorithms developed in detail.

Chapter IV provides an analysis of the performance to be expected from the estimators. First it delineates

the asymptotic properties of the estimates: the behavior exhibited as the number of measurements processed grows without bound. Then the concept of an ambiguity function is used to generate a global performance analysis, one product of which is the Cramér-Rao lower bound on the parameter estimate error variance. Ambiguity functions have been developed previously, as in Schweppe (1970), but the convenient evaluation procedure described in the chapter is original.

Since the full-scale estimator requires considerable computation time, Chapter V examines methods of attaining on-line applicability. Various approximations to the solution of the likelihood equations, as well as approximations to the equations themselves, are considered, noting both the benefits and disadvantages of each. Also, the feasibility of precomputations is investigated as a further approximation to the estimation technique. Subsequently, methods that do not inherently involve approximations are discussed: the use of advantageous state space representations, exploitation of symmetry, and incorporation of a vector measurement as an equivalent succession of scalar measurements. To the knowledge of the author, the explicit means of specifying the unique modified Jordan canonical form for the state transition matrix and measurement matrix, and solving for the transformation without requiring an eigenvector evaluation, is original to this research. Moreover, the idea of replacing a vector measurement update with scalar updates has been suggested in the past, but the proofs of equivalency have depended upon rather complex algebraic manipulation; the proof in Chapter V is an original one, whose simplicity is gained by considering the inverse covariances rather than the covariances directly.

The methods of the thesis are applied to three practical problems in Chapter VI, and the computational results sub-

stantiate the concepts and performance predictions of the previous chapters. Finally, Chapter VII presents the conclusions of the research and suggestions for future investigations.

CHAPTER II

STATE ESTIMATION WITHOUT SIMULTANEOUS ESTIMATION OF PARAMETERS

This chapter will introduce the maximum likelihood conceptualization for solving the problem of estimating unknown quantities from samples of noise-corrupted data. Section 2.1 will develop the idea of deriving appropriate likelihood functions based upon the amount of available statistical information, and generating from these functions the likelihood equations whose solutions provide the desired estimates. Various forms of maximum likelihood state estimators will then be developed under the assumption that all parameters of the system structure are known precisely. Subsequently, section 2.2 will describe the detrimental effects that uncertainty in some of these parameters can impose upon the performance of the estimator. The results of this error sensitivity analysis, combined with the fact that parameter uncertainties do in fact prevail in most actual applications, serve to substantiate the need for parameter estimation. The following chapter will then develop the maximum likelihood estimate of both state variables and critical parameter values, providing a state estimate accuracy beyond the capabilities of a filter that estimates the state alone.

2.1 Maximum Likelihood Estimation of the State Alone

This section is meant to serve a number of purposes. First, the maximum likelihood rationale will be used to derive certain forms of "optimal" state estimators. These results will be important in and of themselves because they will motivate the need for, and relate to the structure of, the combined state and parameter estimators to be derived later. Furthermore, the solution to this simpler problem will introduce some concepts that will be instrumental in

the later developments. One particularly significant concept is the nonuniqueness of the likelihood function - that by appropriate choice of the likelihood function, the designer can incorporate as much information into the problem formulation as he desires. Moreover, certain choices lend themselves to simple or efficient computational forms more readily than others, and are thus preferable in many applications.

2.1.1 The System Model

It will be assumed that modelling techniques have produced an adequate system description in the form of a linear difference equation, driven by a combination of known inputs and white noise. A discrete time description of the plant dynamics is chosen because of the ultimate goal of implementing the estimation scheme on-line by a digital computer. Thus, the actual dynamics might well be a continuous-time process, modelled in discrete time by means of the state transition matrix conceptualization. Measurements are made upon the actual system, and these are corrupted by white noise. It is desired to combine the measurement data with the information provided by the system model so as to provide an optimal estimate of the state of the system, in this context "optimal" being defined in a maximum likelihood sense.

Mathematically, the system can be defined in the following manner. The system state satisfies the discrete time, linear difference equation

$$\underline{x}(i+1) = \underline{\Phi}(i+1,i)\underline{x}(i) + \underline{B}(i)\underline{u}(i) + \underline{G}(i)\underline{w}(i) \qquad (2.1.1)$$

and the noise-corrupted linear observations on the system at time instant i can be described by

$$\underline{z}(i) = \underline{H}(i)\underline{x}(i) + \underline{v}(i) \tag{2.1.2}$$

in which are defined the vector variables

$\underline{x}(i)$ = n-dimensional state vector at instant i
$\underline{u}(i)$ = r-dimensional deterministic input
$\underline{w}(i)$ = s-dimensional driving white noise sequence
$\underline{z}(i)$ = m-dimensional measurement vector
$\underline{v}(i)$ = m-dimensional measurement white noise sequence

and the system matrices

$\underline{\Phi}(i+1,i)$ = n-by-n state transition matrix
$\underline{B}(i)$ = n-by-r deterministic input matrix
$\underline{G}(i)$ = n-by-s noise input matrix
$\underline{H}(i)$ = m-by-n measurement matrix

For this derivation, it will be assumed that $\underline{w}(i)$ and $\underline{v}(i)$ are independent zero mean white noise sequences, each having a normal density with known covariance:

$$E\left\{\underline{w}(i)\right\} = \underline{0} \tag{2.1.3}$$

$$E\left\{\underline{v}(i)\right\} = \underline{0} \tag{2.1.4}$$

$$E\left\{\underline{w}(i)\underline{w}(j)^T\right\} = \begin{cases} \underline{Q}(i) & i=j \\ \underline{0} & i\neq j \end{cases} \tag{2.1.5}$$

$$E\left\{\underline{v}(i)\underline{v}(j)^T\right\} = \begin{cases} \underline{R}(i) & i=j \\ \underline{0} & i\neq j \end{cases} \tag{2.1.6}$$

$$E\left\{\underline{w}(i)\underline{v}(i)^T\right\} = \underline{0} \tag{2.1.7}$$

where $\underline{Q}(i)$ is a positive semidefinite s-by-s matrix and $\underline{R}(i)$ is a positive definite m-by-m matrix (all components of the measurement vector are corrupted by white noise).

The dynamic relation, equation (2.1.1), is valid for all $i \geq 0$, once an initial condition, $\underline{x}(0)$, is specified. For any particular operation of the real system, the initial

state assumes a specific value. However, because this value is not precisely known, it will be modelled as a random variable (over the ensemble of all possible initial conditions) with a normal probability density parameterized by a mean $\hat{\underline{x}}_o$ and a covariance $\underline{P}_o$. It is assumed that $\hat{\underline{x}}_o$ and $\underline{P}_o$ can be predetermined, and that the mean $\hat{\underline{x}}_o$ (which is also the maximum likelihood value because of the unimodality and symmetry of the assumed density) will be used as the apriori estimate of the initial state:

$$E\left\{ \underline{x}(0) \right\} = \hat{\underline{x}}_o \tag{2.1.8}$$

$$E\left\{ \left[\underline{x}(0)-\hat{\underline{x}}_o \right] \left[\underline{x}(0)-\hat{\underline{x}}_o \right]^T \right\} = \underline{P}_o \tag{2.1.9}$$

where $\underline{P}_o$ is a positive semidefinite n-by-n matrix. It is further assumed that $\underline{x}(0)$, $\underline{w}(i)$, and $\underline{v}(j)$ are independent for all i and j.

Generalizations of this description are possible, such as to allow correlated noises, semidefinite or null $\underline{R}$ matrices, or a control vector $\underline{u}$ that is a function of the measurements taken and thus a function of a set of random variables. Note also that certain terms may not be present in some applications: there may be no input $\underline{u}$ or noise $\underline{w}$, but the results will not be notably altered. However, equations (2.1.1) to (2.1.9) will serve as the initial basis of definition.

2.1.2 Derivation of the State Estimator

Let $\underline{Z}(i)$ denote the vector of realized values of the measurements $\underline{z}(1)$, ... $\underline{z}(i)$; thus, $\underline{Z}(i)$ is a vector of dimension $(i \cdot m)$. In order to obtain a maximum likelihood estimate of the system state, an appropriate likelihood function must be defined as a function relating these measurements (whose values are known), the state variables (the unknowns to be estimated), and any other pertinent parameters. One con-

venient likelihood function would be the conditional probability density of the state, given the measurements and values of the parameters:

$$\ell [\, \underline{x}(i), \underline{a}, \underline{Z}(i)] \quad = \quad f(\underline{x}(i) \mid \underline{Z}(i), \underline{a}) \qquad (2.1.10)$$

where $\underline{a}$ represents, in this derivation, the known sequence of parameter values $\underline{a}(0)$, ... $\underline{a}(i)$. The motivation for making the dependence on $\underline{a}$ explicit (the likelihood function will also be implicitly dependent upon all elements of $\underline{\Phi}$, $\underline{B}$, $\underline{G}$, $\underline{H}$, $\underline{R}$, $\underline{Q}$, and $\underline{P}_o$) is that eventually it will be desired to estimate certain system parameters as well as the state.

To specify a maximum likelihood estimate, then, the value $\underline{x}^*(i)$ is sought that will yield the greatest possible value of the likelihood function. In this particular case, $\underline{x}^*(i)$ is the value that will yield the peak of the conditional density function $f(\underline{x}(i) \mid \underline{Z}(i), \underline{a})$ as a function of $\underline{x}$; i.e., maximize the probability of the variables to be estimated, conditioned upon the events actually observed. Rather than directly maximize equation (2.1.10), it is more convenient to find the maximum of the natural logarithm of $\ell [\, \underline{x}(i), \underline{a}, \underline{Z}(i)]$; this is a valid step because for any function f, f and $\ell n(f)$ attain their maxima at the same point. Thus, the log-likelihood function is defined as

$$L [\, \underline{x}(i), \underline{a}, \underline{Z}(i)] \quad = \quad \ell n \; f(\underline{x}(i) \mid \underline{Z}(i), \underline{a}) \qquad (2.1.11)$$

in order to evaluate the desired state estimate.

First of all, Bayes' Rule is used to put the conditional density in equation (2.1.11) into a more convenient form:

$$f(\underline{x}(i) \mid \underline{Z}(i), \underline{a}) = \frac{f(\underline{x}(i), \underline{Z}(i), \underline{a})}{f(\underline{Z}(i), \underline{a})}$$

$$= \frac{f(\underline{z}(i)|\underline{Z}(i-1),\underline{x}(i),\underline{a})\quad f(\underline{Z}(i-1),\underline{x}(i),\underline{a})}{f(\underline{z}(i)|\underline{Z}(i-1),\underline{a})\qquad\quad f(\underline{Z}(i-1),\underline{a})}$$

$$= \left[\frac{f(\underline{z}(i)|\underline{Z}(i-1),\underline{x}(i),\underline{a})}{f(\underline{z}(i)|\underline{Z}(i-1),\underline{a})}\right] f(\underline{x}(i)|\underline{Z}(i-1),\underline{a}) \qquad (2.1.12)$$

By means of this equation, the conditional density for $\underline{x}(i)$ given all measurements through $\underline{z}(i)$, is directly related to the density conditioned on all measurements up to, but not including, $\underline{z}(i)$. It will be used to relate the estimate of $\underline{x}(i)$ just after the measurement $\underline{z}(i)$ to the estimate just before $\underline{z}(i)$. The component terms of equation (2.1.12) will now be evaluated.

Let the mean of the density $f(\underline{x}(i)|\underline{Z}(i-1),\underline{a})$ be denoted as $\overline{\underline{x}}(i)$, and denote the corresponding conditional covariance as $\underline{M}(i)$:

$$\overline{\underline{x}}(i) = E\left\{\underline{x}(i)\middle|\underline{Z}(i-1),\underline{a}\right\} \qquad (2.1.13)$$

$$\underline{M}(i) = E\left\{[\underline{x}(i)-\overline{\underline{x}}(i)]\ [\underline{x}(i)-\overline{\underline{x}}(i)]^T\middle|\underline{Z}(i-1),\underline{a}\right\} \qquad (2.1.14)$$

Similarly, denote the mean of $f(\underline{x}(i)|\underline{Z}(i),\underline{a})$ by $\hat{\underline{x}}(i)$ and the corresponding conditional covariance by $\underline{P}(i)$:

$$\hat{\underline{x}}(i) = E\left\{\underline{x}(i)\middle|\underline{Z}(i),\underline{a}\right\} \qquad (2\ 1.15)$$

$$\underline{P}(i) = E\left\{[\underline{x}(i)-\hat{\underline{x}}(i)]\ [\underline{x}(i)-\hat{\underline{x}}(i)]^T\middle|\underline{Z}(i),\underline{a}\right\} \qquad (2.1.16)$$

These expectations are ensemble averages over all possible initial conditions and both driving and measurement noise sequences, all conditioned on the values of the parameters $\underline{a}$ and the measurements.

The density $f(\underline{x}(i-1)|\underline{Z}(i-1),\underline{a})$ will be assumed normal, and based on this assumption, $f(\underline{x}(i)|\underline{Z}(i),\underline{a})$ will be shown to be normal for any $i \geq 1$; note that $f(\underline{x}(0))$ is assumed to be normal, or possibly an impulse, so that this becomes an inductive proof. Because the system is linear and driven by a white Gaussian noise sequence and/or a deterministic input, if $f(\underline{x}(i-1)|\underline{Z}(i-1),\underline{a})$ is a normal density, then the conditional probability density function of $\underline{x}(i)$ before the measurement at instant i is also normal. Thus,

$$f(\underline{x}(i)|\underline{Z}(i-1),\underline{a}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\underline{M}(i)|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{ -\tfrac{1}{2}\left[\underline{x}(i)-\overline{\underline{x}}(i)\right]^{T} \underline{M}^{-1}(i)\left[\underline{x}(i)-\overline{\underline{x}}(i)\right] \right\} \qquad (2.1.17)$$

Since this density is unimodal and symmetric about the mean $\overline{\underline{x}}(i)$, the maximum likelihood estimate is

$$\arg\left(\max_{\underline{x}}\left[\ell n\ f(\underline{x}(i)|\underline{Z}(i-1),\underline{a})\right]\right) = \overline{\underline{x}}(i) \qquad (2.1.18)$$

Having specified one of the terms in equation (2.1.12), investigate the term $f(\underline{z}(i)|\underline{Z}(i-1),\underline{x}(i),\underline{a})$. The observation at time i is

$$\underline{z}(i) = \underline{H}(i)\underline{x}(i) + \underline{v}(i) \qquad (2.1.2)$$

The $\underline{v}(i)$ sequence has been assumed to be normally distributed and independent of $\underline{x}(i)$. Therefore, conditioned on a particular value of $\underline{x}(i)$, $\underline{z}(i)$ is a normal random variable with conditional mean

$$E\left\{\underline{z}(i)|\underline{Z}(i-1),\underline{x}(i),\underline{a}\right\} = E\left\{\underline{z}(i)|\underline{x}(i),\underline{a}\right\}$$

$$= \underline{H}(i)\underline{x}(i) \qquad (2.1.19)$$

and conditional covariance

$$E\left\{ \left[\underline{z}(i)-\underline{H}(i)\underline{x}(i)\right] \left[\underline{z}(i)-\underline{H}(i)\underline{x}(i)\right]^{T} \middle| \underline{Z}(i-1),\underline{x}(i),\underline{a} \right\} =$$

$$= E\left\{ \underline{v}(i)\underline{v}(i)^{T} \right\} = \underline{R}(i) \qquad (2.1.20)$$

so that the conditional probability density of $\underline{z}(i)$ given $\underline{Z}(i-1)$, $\underline{x}(i)$, and $\underline{a}$ is

$$f(\underline{z}(i)\middle|\underline{Z}(i-1),\underline{x}(i),\underline{a}) = \frac{1}{(2\pi)^{\frac{m}{2}} \left|\underline{R}(i)\right|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}\left[\underline{z}(i)-\underline{H}(i)\underline{x}(i)\right]^{T} \underline{R}^{-1}(i)\left[\underline{z}(i)-\underline{H}(i)\underline{x}(i)\right]\right\} \qquad (2.1.21)$$

Finally, look at $f(\underline{z}(i)\middle|\underline{Z}(i-1),\underline{a})$. The output of the dynamic equation is a Gauss-Markov sequence; in other words, $\underline{x}(i)$ is a Gaussian random variable. Furthermore, linear operations on Gaussian random variables produce Gaussian random variables, so $\underline{H}(i)\underline{x}(i)$ is normally distributed. The sum of two independent normal random variables, i.e., $\underline{H}(i)\underline{x}(i)$ and $\underline{v}(i)$, yields a normal variable, so $f(\underline{z}(i)\middle|\underline{Z}(i-1),\underline{a})$ is a normal density, characterized by conditional mean and covariance:

$$E\left\{ \underline{z}(i)\middle|\underline{Z}(i-1),\underline{a}\right\} = \underline{H}(i)E\left\{\underline{x}(i)\middle|\underline{Z}(i-1),\underline{a}\right\} + E\left\{\underline{v}(i)\middle|\underline{Z}(i-1),\underline{a}\right\}$$
$$= \underline{H}(i)\overline{\underline{x}}(i) \qquad (2.1.22)$$

$$E\left\{ \left[\underline{z}(i)-\underline{H}(i)\overline{\underline{x}}(i)\right] \left[\underline{z}(i)-\underline{H}(i)\overline{\underline{x}}(i)\right]^{T} \middle| \underline{Z}(i-1),\underline{a} \right\} =$$
$$= \underline{H}(i)\underline{M}(i)\underline{H}^{T}(i) + \underline{R}(i) \qquad (2.1.23)$$

so that it is possible to define

$$f(\underline{z}(i)\middle|\underline{Z}(i-1),\underline{a}) = \frac{1}{(2\pi)^{\frac{m}{2}} \left|\underline{H}(i)\underline{M}(i)\underline{H}^{T}(i) + \underline{R}(i)\right|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}\left[\underline{z}(i)-\underline{H}(i)\overline{\underline{x}}(i)\right]^{T}\left[\underline{H}(i)\underline{M}(i)\underline{H}^{T}(i)+\underline{R}(i)\right]^{-1}\cdot\right.$$

$$\left.\cdot\ \left[\underline{z}(i)-\underline{H}(i)\overline{\underline{x}}(i)\right]\right\} \qquad (2.1.24)$$

and note that this is not a function of $\underline{x}(i)$.

By incorporating equations (2.1.17), (2.1.21), and (2.1.24) into the expression for $f(\underline{x}(i)|\underline{Z}(i),\underline{a})$ given by (2.1.12) and taking the natural logarithm, the log-likelihood function is obtained as

$$L[\underline{x}(i),\underline{a},\underline{Z}(i)] =$$

$$= \ell n\left\{(2\pi)^{-\frac{n}{2}}\left|\underline{H}(i)\underline{M}(i)\underline{H}^{T}(i)+\underline{R}(i)\right|^{\frac{1}{2}}\left|\underline{R}(i)\right|^{-\frac{1}{2}}\left|\underline{M}(i)\right|^{-\frac{1}{2}}\right\}$$

$$-\tfrac{1}{2}\left\{\left[\underline{x}(i)-\overline{\underline{x}}(i)\right]^{T}\underline{M}^{-1}(i)\left[\underline{x}(i)-\overline{\underline{x}}(i)\right]\right\}$$

$$-\tfrac{1}{2}\left\{\left[\underline{z}(i)-\underline{H}(i)\underline{x}(i)\right]^{T}\underline{R}^{-1}(i)\left[\underline{z}(i)-\underline{H}(i)\underline{x}(i)\right]\right\}$$

$$+\tfrac{1}{2}\left\{\left[\underline{z}(i)-\underline{H}(i)\overline{\underline{x}}(i)\right]^{T}\left[\underline{H}(i)\underline{M}(i)\underline{H}^{T}(i)+\underline{R}(i)\right]^{-1}\cdot\right.$$

$$\left.\cdot\left[\underline{z}(i)-\underline{H}(i)\overline{\underline{x}}\ (i)\right]\right\} \qquad (2.1.25)$$

where the first and last terms do not involve $\underline{x}(i)$.

To obtain the maximum likelihood estimate of $\underline{x}(i)$, the equation

$$\left.\frac{\partial L[\underline{x}(i),\underline{a},\underline{Z}(i)]}{\partial\underline{x}(i)}\right|_{\underline{x}(i)\rightarrow\underline{x}^{*}(i)} = \underline{0}^{T} \qquad (2.1.26)$$

is solved, where $\underline{x}^{*}(i)$ denotes the maximum likelihood estimate of $\underline{x}(i)$, which causes (2.1.26) to be satisfied. Substituting (2.1.25) into this equation yields

$$\left.\underline{M}^{-1}(i)\left[\underline{x}(i)-\overline{\underline{x}}(i)\right]\ -\ \underline{H}^{T}(i)\underline{R}^{-1}(i)\left[\underline{z}(i)-\underline{H}(i)\underline{x}(i)\right]\right|_{\underline{x}(i)\rightarrow\underline{x}^{*}(i)} = \underline{0} \qquad (2.1.27)$$

The solution is then

$$\underline{x}^*(i) = [\underline{M}^{-1}(i)+\underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)]^{-1} \cdot$$

$$\cdot [\underline{M}^{-1}(i)\overline{\underline{x}}(i)+\underline{H}^T(i)\underline{R}^{-1}(i)\underline{z}(i)] \qquad (2.1.28)$$

Instead of solving the likelihood equation, (2.1.26), directly by using equation (2.1.25), algebraic manipulations (completing squares, applying the inversion lemma) can first be performed upon the log-likelihood function to obtain the form

$$L[\underline{x}(i),\underline{a},\underline{Z}(i)] = \ell n\left\{(2\pi)^{-\frac{n}{2}}|\underline{P}(i)|^{-\frac{1}{2}}\right\}$$

$$-\tfrac{1}{2}[\underline{x}(i)-\hat{\underline{x}}(i)]^T \underline{P}^{-1}(i)[\underline{x}(i)-\hat{\underline{x}}(i)] \qquad (2.1.29)$$

where

$$\hat{\underline{x}}(i) = [\underline{M}^{-1}(i)+\underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)]^{-1} \cdot$$

$$\cdot [\underline{M}^{-1}(i)\overline{\underline{x}}(i)+\underline{H}^T(i)\underline{R}^{-1}(i)\underline{z}(i)] \qquad (2.1.30)$$

$$\underline{P}^{-1}(i) = \underline{M}^{-1}(i) + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i) \qquad (2.1.31)$$

This demonstrates that $f(\underline{x}(i)|\underline{Z}(i),\underline{a})$ is in fact a normal density, with mean $\hat{\underline{x}}(i)$ and covariance $\underline{P}(i)$:

$$f(\underline{x}(i)|\underline{Z}(i),\underline{a}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\underline{P}(i)|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}[\underline{x}(i)-\hat{\underline{x}}(i)]^T \underline{P}^{-1}(i)\ [\underline{x}(i)-\hat{\underline{x}}(i)]\right\} \qquad (2.1.32)$$

As expected, due to the symmetry of this density, the maximum likelihood estimate and the mean of the density coincide. The matrix inversion lemma can be used to put this estimate

into a better computational form, requiring inversion of an m-by-m matrix instead of a n-by-n (m is typically smaller than n, and since measurements can be incorporated one at a time, the inversion can become a simple division). For $\underline{M}(i)$ and $\underline{R}(i)$ positive definite, the lemma yields

$$(\underline{M}^{-1} + \underline{H}^T\underline{R}^{-1}\underline{H})^{-1} = \underline{M}-\underline{M}\underline{H}^T(\underline{H}\underline{M}\underline{H}^T + \underline{R})^{-1}\underline{H}\underline{M} \qquad (2.1.33)$$

so that the state estimate can be written as

$$\hat{\underline{x}}(i) = \bar{\underline{x}}(i) + \underline{K}(i)[\underline{z}(i) - \underline{H}(i)\bar{\underline{x}}(i)] \qquad (2.1.34)$$

$$\underline{K}(i) = \underline{M}(i)\underline{H}^T(i)[\underline{H}(i)\underline{M}(i)\underline{H}^T(i) + \underline{R}(i)]^{-1} \qquad (2.1.35)$$

which can be recognized as the form of the Kalman filter. The update of the covariance matrix from just before the i-th measurement to just after is given by (2.1.31), re-written as

$$\underline{P}(i) = [\underline{M}^{-1}(i) + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)]^{-1} \qquad (2.1.36)$$

which, although it adds the measurement information in a simple way, requires an n-by-n matrix inversion. Equation (2.1.33) yields

$$\underline{P}(i)=\underline{M}(i)-\underline{M}(i)\underline{H}^T(i)[\underline{H}(i)\underline{M}(i)\underline{H}^T(i)+\underline{R}(i)]^{-1}\underline{H}(i)\underline{M}(i) \qquad (2.1.37)$$

$$=[\underline{I}-\underline{K}(i)\underline{H}(i)]\,\underline{M}(i) \qquad (2.1.38)$$

Although these forms involve m-by-m inversions, they do have some undesirable characteristics. Equation (2.1.37) is sometimes a small difference of large numbers (especially if the measurements are very accurate) and thus is subject to numerical precision problems; equation (2.1.38) does not assure positive definiteness or symmetry, and can thus lead

to numerical difficulties also. Now, if the estimate equation, (2.1.34), is rewritten as

$$\hat{\underline{x}}(i) = \left[\underline{I}-\underline{K}(i)\underline{H}(i)\right]\bar{\underline{x}}(i) + \underline{K}(i)\underline{z}(i) \qquad (2.1.39)$$

then it can readily be shown that an equivalent expression for $\underline{P}(i)$ would be

$$\underline{P}(i) = \left[\underline{I}-\underline{K}(i)\underline{H}(i)\right]\underline{M}(i)\left[\underline{I}-\underline{K}(i)\underline{H}(i)\right]^{T} + \underline{K}(i)\underline{R}(i)\underline{K}^{T}(i) \qquad (2.1.40)$$

This is the sum of two symmetric, positive definite matrices, so that numerical computations based upon this form will be better conditioned, better assuring the symmetry and positive definiteness of $\underline{P}(i)$. Furthermore it is insensitive,to first order, to small errors $\underline{\delta K}$ in the computed filter gain, whereas the error in $\underline{P}(i)$ due to $\underline{\delta K}(i)$ in (2.1.38) would be

$$\underline{\delta P}(i) = - \underline{\delta K}(i)\underline{H}(i)\underline{M}(i) \qquad (2.1.41)$$

Furthermore, as demonstrated by Fraser (1967), the update given by (2.1.40) is less sensitive to arithmetic truncation than the other forms: especially in cases where the measurement noise is small, $(\underline{I}-\underline{KH})$ and $(\underline{M}-\underline{MH}^{T}\underline{A}^{-1}\underline{HM})$ will have first order truncation errors whereas equation (2.1.40) will only be affected to second order. This becomes a crucial consideration for on-line applications in which the minimum computer wordlength that achieves adequate performance is sought. Equation (2.1.40) thus seems to be the most desirable, even though it requires a considerable amount of computation. For on-line applications where time and wordlength constraints are critical, symmetry can be exploited by propagating only lower triangular forms, and precision can be improved for a short wordlength by using a square-root formulation. (See section 5.6.)

In order to propagate the estimates between measurements, it has been shown that the probability densities involved are normal, so the propagation of the entire density function can be specified by the time history of its mean and covariance. Thus,

$$\begin{aligned}\bar{\underline{x}}(i+1) &= E\left\{\underline{x}(i+1)\middle|\, \underline{Z}(i),\underline{a}\right\} \\ &= \underline{\Phi}(i+1,i)\,\hat{\underline{x}}(i) + \underline{B}(i)\underline{u}(i)\end{aligned} \tag{2.1.42}$$

$$\begin{aligned}\underline{M}(i+1) &= E\left\{\left[\underline{x}(i+1)-\bar{\underline{x}}(i+1)\right]\left[\underline{x}(i+1)-\bar{\underline{x}}(i+1)\right]^{T}\middle|\, \underline{Z}(i),\underline{a}\right\} \\ &= E\left\{\left[\underline{\Phi}(i+1,i)\left\{\underline{x}(i)-\hat{\underline{x}}(i)\right\} + \underline{G}(i)\underline{w}(i)\right]\cdot\right. \\ &\quad \left.\cdot\left[\underline{\Phi}(i+1,i)\left\{\underline{x}(i)-\hat{\underline{x}}(i)\right\} + \underline{G}(i)\underline{w}(i)\right]^{T}\middle|\, \underline{Z}(i),\underline{a}\right\} \\ &= \underline{\Phi}(i+1,i)\underline{P}(i)\underline{\Phi}^{T}(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^{T}(i)\end{aligned} \tag{2.1.43}$$

To summarize, at a measurement, the estimate is updated using:

$$\hat{\underline{x}}(i) = \bar{\underline{x}}(i) + \underline{K}(i)\left[\underline{z}(i)-\underline{H}(i)\bar{\underline{x}}(i)\right] \tag{2.1.34}$$

$$\underline{P}(i) = \left[\underline{I}-\underline{K}(i)\underline{H}(i)\right]\underline{M}(i)\left[\underline{I}-\underline{K}(i)\underline{H}(i)\right]^{T} + \underline{K}(i)\underline{R}(i)\underline{K}^{T}(i) \tag{2.1.40}$$

where

$$\underline{K}(i) = \underline{M}(i)\underline{H}^{T}(i)\left[\underline{H}(i)\underline{M}(i)\underline{H}^{T}(i) + \underline{R}(i)\right]^{-1} \tag{2.1.35}$$

and, to propagate the estimate up to the time of the next measurement, the relations are

$$\bar{\underline{x}}(i+1) = \underline{\Phi}(i+1,i)\hat{\underline{x}}(i) + \underline{B}(i)\underline{u}(i) \tag{2.1.42}$$

$$\underline{M}(i+1) = \underline{\Phi}(i+1,i)\underline{P}(i)\underline{\Phi}^{T}(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^{T}(i) \tag{2.1.43}$$

$\underline{K}(i)$ can also be expressed as

$$\underline{K}(i) = \underline{P}(i)\underline{H}^{T}(i)\ \underline{R}^{-1}(i) \tag{2.1.44}$$

(This can be expanded using the expression for $\underline{P}(i)$ and shown equal to the previous evaluation of $\underline{K}(i)$.) Although it is a simpler expression and has the same appearance as the continuous-time Kalman gain, it is not as convenient to use in recursions (see Ho, 1962').

2.1.3 Statistics of the Estimation Errors

Now investigate the statistics of the error committed by the estimate immediately after the measurement at time i. The error is

$$\underline{e}(i) = \underline{x}(i) - \hat{\underline{x}}(i) \tag{2.1.45}$$

Since both $\underline{x}(i)$ and $\hat{\underline{x}}(i)$ are Gaussian, $\underline{e}(i)$ must also have a Gaussian distribution, and thus its density function is completely specified by its mean and covariance, written as

$$E\left\{\underline{e}(i)\,|\,\underline{Z}(i),\underline{a}\right\} = \hat{\underline{x}}(i) - \hat{\underline{x}}(i) = \underline{0} \tag{2.1.46}$$

$$E\left\{\underline{e}(i)\underline{e}^{T}(i)\,|\,\underline{Z}(i),\underline{a}\right\} = \underline{P}(i) \tag{2.1.47}$$

Thus, the desired density can be written as

$$f(\underline{e}(i)\,|\,\underline{Z}(i),\underline{a}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\underline{P}(i)|^{\frac{1}{2}}} \exp\left\{-\tfrac{1}{2}\underline{e}^{T}(i)\underline{P}^{-1}(i)\underline{e}(i)\right\} \tag{2.1.48}$$

Since $\underline{P}(i)$ can be propagated independent of the particular sequence of measurements, using equations (2.1.40) and (2.1.43), and since $\underline{e}(i)$ is a dummy variable in equation (2.1.48), it can be seen that $f(\underline{e}(i)\,|\,\underline{Z}(i),\underline{a})$ is in fact

equivalent to $f(\underline{e}(i)|\underline{a})$. In other words, $\underline{e}(i)$ is statistically dependent upon the $\underline{R}(j)$'s but is independent of the particular $\underline{Z}(i)$.

2.1.4 Statistics of the Measurement Information and Residuals

Consider the statistics of the new information introduced at each measurement. Looking at equation (2.1.34), this information is portrayed by

$$\hat{\underline{x}}(i) - \bar{\underline{x}}(i) = \underline{K}(i)\left[\underline{z}(i)-\underline{H}(i)\bar{\underline{x}}(i)\right] \qquad (2.1.49)$$

where the term in brackets is the measurement residual at time i: it is the difference between the measurement $\underline{z}(i)$ and the best prediction of $\underline{z}(i)$ based upon $\underline{Z}(i-1)$. The residual itself can also be expressed as

$$\underline{z}(i) - \underline{H}(i)\bar{\underline{x}}(i) = \underline{H}(i)\left[\underline{x}(i)-\bar{\underline{x}}(i)\right] + \underline{v}(i) \qquad (2.1.50)$$

The correction to the state estimate is proportional to the residual, the gain $\underline{K}(i)$ serving as proportionality constant. By arguments of linear operations on Gaussian random variables, the residual itself is shown to be a Gaussian random variable, with mean and covariance given by:

$$\begin{aligned} &E\left\{\left[\underline{z}(i)-\underline{H}(i)\bar{\underline{x}}(i)\right] \middle| \underline{Z}(i-1),\underline{a}\right\} = \\ &\quad = E\left\{\left[\underline{z}(i)-E\{\underline{z}(i)\,|\,\underline{Z}(i-1),\underline{a}\}\right] \middle| \underline{Z}(i-1),\underline{a}\right\} \\ &\quad = \underline{0} \end{aligned} \qquad (2.1.51)$$

$$E\left\{\left[\underline{z}(i)-\underline{H}(i)\bar{\underline{x}}(i)\right]\left[\underline{z}(i)-\underline{H}(i)\bar{\underline{x}}(i)\right]^{T} \middle| \underline{Z}(i-1),\underline{a}\right\} =$$

$$= E\left\{\left(\underline{H}(i)\left[\underline{x}(i)-\overline{\underline{x}}(i)\right] + \underline{v}(i)\right) \cdot \right.$$

$$\left. \cdot\left(\underline{H}(i)\left[\underline{x}(i)-\overline{\underline{x}}(i)\right] + \underline{v}(i)\right)^{T} \,\middle|\, \underline{Z}(i-1), \underline{a}\right\}$$

$$= \underline{H}(i)\underline{M}(i)\underline{H}^{T}(i) + \underline{R}(i) \qquad (2.1.52)$$

Before these expressions are used to derive the statistics of the information added by each measurement, observe that this new information can be written as

$$\hat{\underline{x}}(i)-\overline{\underline{x}} = \underline{K}(i)\left[\underline{H}(i)\underline{x}(i)-\underline{H}(i)\overline{\underline{x}}(i) + \underline{v}(i)\right]$$

$$= \underline{K}(i)\left[\underline{H}(i)\underline{\Phi}(i,i-1)\left\{\underline{x}(i-1)-\hat{\underline{x}}(i-1)\right\}\right.$$

$$\left. + \underline{H}(i)\underline{\Phi}(i,i-1)\underline{G}(i-1)\underline{w}(i-1) + \underline{v}(i)\right] \qquad (2.1.53)$$

where the term in brackets is the sum of three independent terms, each independent of $\underline{Z}(i-1)$. Consequently, the value of $\left[\hat{\underline{x}}(i)-\overline{\underline{x}}(i)\right]$ is independent of $\underline{Z}(i-1)$, and is therefore also independent of $\left[\hat{\underline{x}}(j)-\overline{\underline{x}}(j)\right]$ for all j less than i, since these are dependent upon $\underline{Z}(i-1)$. This also demonstrates the whiteness of the residual sequence. Thus, the statistics of the information added by a measurement can be expressed as:

$$E\left\{\left[\hat{\underline{x}}(i)-\overline{\underline{x}}(i)\right] \,\middle|\, \underline{Z}(i-1), \underline{a}\right\} = E\left\{\left[\hat{\underline{x}}(i)-\overline{\underline{x}}(i)\right] \,\middle|\, \underline{a}\right\}$$

$$= \underline{0} \qquad (2.1.54)$$

$$E\left\{\left[\hat{\underline{x}}(i)-\overline{\underline{x}}(i)\right]\left[\hat{\underline{x}}(i)-\overline{\underline{x}}(i)\right]^{T} \,\middle|\, \underline{Z}(i-1), \underline{a}\right\}$$

$$= E\left\{\left[\hat{\underline{x}}(i)-\overline{\underline{x}}(i)\right]\left[\hat{\underline{x}}(i)-\overline{\underline{x}}(i)\right]^{T} \,\middle|\, \underline{a}\right\}$$

$$= \underline{K}(i)\left[\underline{H}(i)\underline{M}(i)\underline{H}^{T}(i)+\underline{R}(i)\right]\underline{K}^{T}(i)$$

$$= \underline{M}(i)\underline{H}^T(i)\left[\underline{H}(i)\underline{M}(i)\underline{H}^T(i)+\underline{R}(i)\right]^{-1}\underline{H}(i)\underline{M}(i)$$

$$= \underline{K}(i)\underline{H}(i)\underline{M}(i) \qquad (2.1.55)$$

which is the term subtracted from the estimate error covariance at each measurement (see equation (2.1.38)).

2.1.5 Inverse Covariance Form

In certain circumstances, the apriori statistical information about the state may not be complete; i.e., there is no information about the state initial conditions in some or all directions of state space. This can be visualized as the limiting case of certain rows and columns in $\underline{P}(0)$ (or the result of $\underline{P}(0)$ under a similarity transformation) going to infinity, or, of the corresponding terms in $\underline{P}^{-1}(0)$ going to zero. Because it remains finite, $\underline{P}^{-1}(i)$ is more desirable to utilize. The matrix "$\underline{P}^{-1}(0)$" is singular and has no inverse; until $\underline{P}^{-1}(i)$ attains full rank, a unique estimate of the full state cannot be made, but the following development allows a viable startup procedure. In practice, it will also avoid numerical problems associated with certain elements of $\underline{P}_o$ being many orders of magnitude larger than others.

The usual recursion relations for the covariance matrix can be written as

$$\underline{M}(i+1) = \underline{\Phi}(i+1,i)\underline{P}(i)\underline{\Phi}^T(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^T(i) \qquad (2.1.56)$$

$$\underline{P}^{-1}(i) = \underline{M}^{-1}(i) + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i) \qquad (2.1.57)$$

Applying the matrix inversion lemma to equation (2.1.57) yields the familiar Kalman equations. Instead, apply the lemma, which states that for $\underline{X}$ and $\underline{Y}$ both (n x k) matrices,

$$(\underline{A}+\underline{X}^T\underline{Y})^{-1} = \underline{A}^{-1} - \underline{A}^{-1}\underline{X}^T(\underline{I}+\underline{Y}\underline{A}^{-1}\underline{X}^T)^{-1}\underline{Y}\underline{A}^{-1} \qquad (2.1.58)$$

to equation (2.1.56) by identifying

$$\underline{A} = \underline{\Phi}(i+1,i)\underline{P}(i)\underline{\Phi}^T(i+1,i), \quad \underline{X}^T = \underline{G}(i)\underline{Q}(i), \quad \underline{Y} = \underline{G}^T(i) \qquad (2.1.59)$$

to yield, for $\underline{Q}$ nonsingular

$$\underline{M}^{-1}(i+1) = \underline{M}_\circ^{-1}(i+1) - \underline{M}_\circ^{-1}(i+1)\,\underline{G}(i) \cdot \left[\underline{G}^T(i)\underline{M}_\circ^{-1}(i+1)\underline{G}(i)+\underline{Q}^{-1}(i)\right]^{-1}\underline{G}^T(i)\underline{M}_\circ^{-1}(i+1) \qquad (2.1.60)$$

where

$$\underline{M}_\circ^{-1}(i+1) = \underline{\Phi}^T(i,i+1)\underline{P}^{-1}(i)\underline{\Phi}(i,i+1) \qquad (2.1.61)$$

(Note that if $\underline{Q} = \underline{0}$, $\underline{M}^{-1}(i+1) = \underline{M}_\circ^{-1}(i+1)$.)

Although equation (2.1.60) is correct, it is expressed in terms of the difference of two large quantities. In the Kalman formulation, it was shown that the form $\{[\underline{I}-\underline{K}\underline{H}]\,\underline{M}[\underline{I}-\underline{K}\underline{H}]^T + \underline{K}\underline{R}\underline{K}^T\}$ is to be preferred to $\{[\underline{I}-\underline{K}\underline{H}]\,\underline{M}\}$. An analogous development in this case yields the following result (other forms, as described by Fraser (1967) and Abramson (1968), can be shown to be equivalent to this new, more suitable form). Define the matrix

$$\underline{\chi}(i) = \underline{M}_\circ^{-1}(i+1)\underline{G}(i)\left[\underline{G}^T(i)\underline{M}_\circ^{-1}(i+1)\underline{G}(i) + \underline{Q}^{-1}(i)\right]^{-1} \qquad (2.1.62)$$

with which the covariance propagation can be written as:

$$\underline{M}^{-1}(i+1) = \left[\underline{I}-\underline{\chi}(i)\underline{G}^{T}(i)\right] \underline{M}_{o}^{-1}(i+1)\left[\underline{I}-\underline{\chi}(i)\underline{G}^{T}(i)\right]^{T} + \underline{\chi}(i)\underline{Q}^{-1}(i)\underline{\chi}^{T}(i) \quad (2.1.63)$$

$$\underline{P}^{-1}(i) = \underline{M}^{-1}(i) + \underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{H}(i) \quad (2.1.64)$$

from the initial condition

$$\underline{P}^{-1}(0) = \underline{P}_{o}^{-1} = \text{given} \quad (2.1.65)$$

The equivalence of equations (2.1.63) and (2.1.60) can readily be shown by expansion of (2.1.63).

For the "state" estimate define

$$\overline{\underline{Y}}(i) = \underline{M}^{-1}(i)\overline{\underline{x}}(i) \quad (2.1.66)$$

$$\hat{\underline{Y}}(i) = \underline{P}^{-1}(i)\hat{\underline{x}}(i) \quad (2.1.67)$$

which, for the finite time until $\underline{M}^{-1}$ and $\underline{P}^{-1}$ attain full rank, cannot determine a unique $\overline{\underline{x}}$ or $\hat{\underline{x}}$. The recursions for $\hat{\underline{Y}}$ are then generated as

$$\overline{\underline{Y}}(i+1) = \left[\underline{I}-\underline{\chi}(i)\underline{G}^{T}(i)\right] \underline{\Phi}^{T}(i,i+1)\left[\hat{\underline{Y}}(i)+\underline{P}^{-1}\underline{\Phi}(i,i+1)\underline{B}(i)\underline{u}(i)\right] \quad (2.1.68)$$

$$\hat{\underline{Y}}(i) = \overline{\underline{Y}}(i) + \underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{z}(i) \quad (2.1.69)$$

$$\hat{\underline{Y}}(0) = \underline{P}_{o}^{-1}\hat{\underline{x}}_{o} \quad (2.1.70)$$

If there is no driving noise ($\underline{Q} = \underline{0}$), then equations (2.1.63) and (2.1.68) are replaced by

$$\underline{M}^{-1}(i+1) = \underline{M}_{o}^{-1}(i+1) \quad (2.1.71)$$

$$\bar{\underline{y}}(i+1) = \underline{\Phi}^T(i,i+1)\left[\hat{\underline{y}}(i)+\underline{P}^{-1}(i)\underline{\Phi}(i,i+1)\underline{B}(i)\underline{u}(i)\right] \tag{2.1.72}$$

Once $\underline{P}^{-1}(i)$ becomes nonsingular, then its inverse can be taken to obtain $\underline{P}(i)$, and then the optimal state estimate can be expressed as

$$\hat{\underline{x}}(i) = \underline{P}(i)\hat{\underline{y}}(i) \tag{2.1.73}$$

From this time forward, it is possible to revert to the more familiar recursions of $\bar{\underline{x}}$, $\hat{\underline{x}}$, $\underline{M}$, and $\underline{P}$, or continue using equations (2.1.61) through (2.1.70). It is conceivable that the latter may be more convenient in certain situations, even when $\underline{P}_\circ^{-1}$ is nonsingular.

A concept related to the inverse covariance is the Fisher Information Matrix, which is a measure of the certainty of the state estimate due to measurement data alone; i.e., the apriori information provided by the initial conditions $\hat{\underline{x}}_\circ$ and $\underline{P}_\circ$ is disregarded. The Information Matrix is defined as

$$\underline{\mathcal{J}}(i,1) = \sum_{j=1}^{i} \underline{\Phi}^T(j,i)\underline{H}^T(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i) \tag{2.1.74}$$

where $\underline{\Phi}(j,i)$ for $j<i$ is the transition matrix for propagating the state backward in time. In order to relate this concept directly to the previous derivations, it must be assumed that the dynamics are noise free - there is no $\underline{w}(j)$ sequence, or equivalently, all $\underline{Q}(j) = \underline{0}$. Under this assumption

$$\underline{\mathcal{J}}(i,1) = \underline{P}^{-1}(i) - \underline{\Phi}^T(0,i)\underline{P}_\circ^{-1}\underline{\Phi}(0,i) \tag{2.1.75}$$

If there were no apriori information about the state, or formally, if $\underline{P}_\circ^{-1} = \underline{0}$, then the Information Matrix is the inverse of the corresponding estimate error covariance. Expressing the definition of the Information Matrix for time instants i and (i+1), and equating like terms, yields the

following recursion for $\underline{\mathcal{L}}$:

$$\underline{\mathcal{L}}(i+1,1) = \underline{\Phi}^T(i,i+1)\underline{\mathcal{L}}(i,1)\underline{\Phi}(i,i+1)+\underline{H}^T(i+1)\underline{R}^{-1}(i+1)\underline{H}(i+1) \tag{2.1.76}$$

From this relation, it can be seen that the "information" contained in the single measurement at time i is $\underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)$. Through matrix manipulations, equation (2.1.30) can be transformed into the following equation when $\underline{Q} = \underline{0}$:

$$\begin{aligned}\hat{\underline{x}}(i) = &\left[\underline{\Phi}^T(0,i)\underline{P}_o^{-1}\underline{\Phi}(0,1) + \underline{\mathcal{L}}(i,1)\right]^{-1} \cdot \\ &\cdot \Bigg[\sum_{j=1}^{i} \underline{\Phi}^T(j,i)\underline{H}^T(j)\underline{R}^{-1}(j)\Big\{\underline{z}(j) \\ &+ \underline{H}(j)\sum_{k=j+1}^{i} \underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\Big\} \\ &+ \underline{\Phi}^T(0,i)\underline{P}_o^{-1}\Big\{\hat{\underline{x}}_o + \sum_{j=1}^{i} \underline{\Phi}(0,j)\underline{B}(j-1)\underline{u}(j-1)\Big\}\Bigg]\end{aligned} \tag{2.1.77}$$

thereby separating the effects of initial conditions and measurements upon the estimate. Note that if no apriori information is available, this reduces to the classical solution for the linear, unbiased, minimum variance estimate of $\underline{x}$ (assuming $\underline{v}$ is white but not necessarily Gaussian) given by the Gauss-Markov theorem:

$$\begin{aligned}\tilde{\underline{x}}(i) = &\underline{\mathcal{L}}^{-1}(i,1)\sum_{j=1}^{i} \underline{\Phi}^T(j,i)\underline{H}^T(j)\underline{R}^{-1}(j)\Big\{\underline{z}(j) \\ &+ \underline{H}(j)\sum_{k=j+1}^{i} \underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\Big\}\end{aligned} \tag{2.1.78}$$

If apriori state information is not available, there is a more appropriate log-likelihood function than that used previously. Consider

$$L[\underline{x}(i),\underline{a},\underline{Z}(i)] = \ell n\ f(\underline{Z}(i)|\underline{x}(i),\underline{a}) \qquad (2.1.79)$$

For a specific value of $\underline{a}$, this is the "classical" likelihood function, used to estimate $\underline{x}(i)$ by maximizing the probability of occurrence of the events that have in fact occurred (the measurement history) by varying the value of $\underline{x}(i)$. By repeatedly applying Bayes' Rule, the density in (2.1.79) can be transformed into a product of separate densities:

$$\begin{aligned} f(\underline{Z}(i)|\underline{x}(i),\underline{a}) &= f(\underline{z}(i)|\underline{Z}(i-1),\underline{x}(i),\underline{a})\ \ f(\underline{Z}(i-1)|\underline{x}(i),\underline{a}) \\ &= f(\underline{z}(i)|\underline{Z}(i-1),\underline{x}(i),\underline{a})\cdot \\ &\quad \cdot f(\underline{z}(i-1)|\underline{Z}(i-2),\underline{x}(i),\underline{a})\ \ f(\underline{Z}(i-2)|\underline{x}(i),\underline{a}) \\ &\ \ \vdots \\ &= \prod_{j=1}^{i} f(\underline{z}(j)|\underline{Z}(j-1),\underline{x}(i),\underline{a}) \qquad (2.1.80) \end{aligned}$$

For the case of $\underline{Q} = \underline{0}$, conditioning on a value of $\underline{x}(i)$ completely determines these densities independent of $\underline{Z}(j-1)$, and an explicit expression for the separate densities can be obtained, so that the logarithm of (2.1.80) becomes

$$\begin{aligned} L[\underline{x}(i),\underline{a},\underline{Z}(i)] &= -\frac{im}{2}\ \ell n\ (2\pi) - \tfrac{1}{2}\sum_{j=1}^{i} \ell n|\underline{R}(j)| \\ &-\tfrac{1}{2}\sum_{j=1}^{i}\left[\underline{z}(j)-\underline{H}(j)\left\{\underline{\Phi}(j,i)\underline{x}(i)-\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\right]^{T}\underline{R}^{-1}(j)\cdot \\ &\cdot\left[\underline{z}(j)-\underline{H}(j)\left\{\underline{\Phi}(j,i)\underline{x}(i)-\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\right] \qquad (2.1.81) \end{aligned}$$

Taking the derivative with respect to $\underline{x}(i)$ yields the likelihood equation:

$$\sum_{j=1}^{i} \left[\underline{z}(j)-\underline{H}(j)\underline{\Phi}(j,i)\underline{x}(i)+\sum_{k=j+1}^{i} \underline{H}(j)\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right]^{T}\cdot$$

$$\cdot\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i)\Big|_{\underline{x}(i)\to\underline{x}^{*}(i)} = \underline{0}^{T} \qquad (2.1.82)$$

for which the solution is

$$\begin{aligned}\underline{x}^{*}(i) &= \tilde{\underline{x}}(i)\\ &= \underline{\mathcal{Q}}^{-1}(i,1)\Big[\sum_{j=1}^{i} \underline{\Phi}^{T}(j,i)\underline{H}^{T}(j)\underline{R}^{-1}(j)\Big\{\underline{z}(j)\\ &\quad + \underline{H}(j)\sum_{k=j+1}^{i} \underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\Big\}\Big] \qquad (2.1.83)\end{aligned}$$

(When j reaches i, the second sum is $\sum_{k=i+1}^{i}$, which by convention is zero.) This agrees with the previous results. In order to form a completely defined estimate of the state, at least i_o measurements must be made, where i_o is the smallest integer that causes the Information Matrix to become positive definite (of rank n) and thus to have a unique inverse. Previous to the i_o-th measurement, the Information Matrix is singular, and certain linear combinations of the state vector components cannot be determined. The existence of such an i_o is, in fact, the complete observability condition. Until the i_o-th stage of the process is reached, the startup procedures described by equations (2.1.61) to (2.1.72) can be employed, using the initial conditions $\underline{P}^{-1}(0) = \underline{0}$, $\hat{\underline{y}}(0) = \underline{0}$. Other similar startup procedures are also possible, such as waiting until i_o measurements are made and using the $\underline{\mathcal{Q}}^{-1}(i_o,1)$ as an "initial" $\underline{P}_o$ to define the usual recursions (Schweppe, 1964), or making i_o fictitious measurements to define $\underline{P}_o$ and then removing their effect algebraically once i_o real measurements have been made (Fagin, 1964).

It can be shown through straightforward algebra that the solution (2.1.83) can be expressed recursively as:

$$\tilde{\underline{x}}(i) = \underline{\Phi}(i,i-1)\tilde{\underline{x}}(i-1) + \underline{B}(i-1)\underline{u}(i-1)$$
$$+ \underline{\mathcal{P}}^{-1}(i,1)\underline{H}^{T}(i)\underline{R}^{-1}(i)\cdot$$
$$\cdot\left[\underline{z}(i)-\underline{H}(i)\left\{\underline{\Phi}(i,i-1)\tilde{\underline{x}}(i-1) + \underline{B}(i-1)\underline{u}(i-1)\right\}\right] \tag{2.1.84}$$

where $\underline{\mathcal{P}}(i,1)$ satisfies

$$\underline{\mathcal{P}}(i,1) = \underline{\Phi}^{T}(i-1,i)\underline{\mathcal{P}}(i-1,1)\underline{\Phi}(i-1,i) + \underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{H}(i) \tag{2.1.85}$$

Now suppose it is desired to estimate the state using only the N most recent measurements, rather than all measurements up to the current time. In view of the result produced by the likelihood function given by (2.1.79), it would appear that an appropriate choice for this case would be

$$L\left[\underline{x}(i),\underline{a},\underline{Z}_{N}(i)\right] = \ell n\ f(\underline{Z}_{N}(i)\,|\,\underline{x}(i),\underline{a}) \tag{2.1.86}$$

where $\underline{Z}_{N}(i)$ is the vector of the N most recent measurements, $\underline{z}(i-N+1)$, $\underline{z}(i-N+2)$, ... $\underline{z}(i-1)$, $\underline{z}(i)$. Repeated application of Bayes' Rule yields

$$f(\underline{Z}_{N}(i)\,|\,\underline{x}(i),\underline{a}) = \prod_{j=i-N+1}^{i} f(\underline{z}(j)\,|\,\underline{Z}_{j-[i-N+1]}(j-1),\underline{x}(i),\underline{a}) \tag{2.1.87}$$

where $\underline{Z}_{j-[i-N+1]}(j-1)$ is the most recent $j-[i-N+1]$ measurements at time (j-1) and $\underline{Z}_{o}(k)$ denotes no measurements for any k. Again, it is possible to write the terms in this equation explicitly for the instance of $\underline{Q}(j) = \underline{0}$ for all j:

$$L\left[\underline{x}(i),\underline{a},\underline{Z}_N(i)\right] = -\frac{Nm}{2}\ln(2\pi) - \tfrac{1}{2}\sum_{j=i-N+1}^{i}\ln\left|\underline{R}(j)\right|$$

$$-\tfrac{1}{2}\sum_{j=i-N+1}^{i}\left[\underline{z}(j)-\underline{H}(j)\left\{\underline{\Phi}(j,i)\underline{x}(i)-\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\right]^{T}\cdot$$

$$\cdot\,\underline{R}^{-1}(j)\left[\underline{z}(j)-\underline{H}(j)\left\{\underline{\Phi}(j,i)\underline{x}(i)-\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\right] \tag{2.1.88}$$

The solution to the likelihood equation then becomes

$$\underline{x}^{*}(i) = \underline{\mathcal{J}}^{-1}(i,i-N+1)\Big[\sum_{j=i-N+1}^{i}\underline{\Phi}^{T}(j,i)\underline{H}^{T}(j)\underline{R}^{-1}(j)\cdot$$

$$\cdot\left\{\underline{z}(j)+\underline{H}(j)\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\Big] \tag{2.1.89}$$

where the N-step Information Matrix is defined as

$$\underline{\mathcal{J}}(i,i-N+1) = \sum_{j=i-N+1}^{i}\underline{\Phi}^{T}(j,i)\underline{H}^{T}(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i) \qquad (i\geq N) \tag{2.1.90}$$

The startup procedure described previously is applicable here up to the N-th instant. It is necessary that $\underline{\mathcal{J}}(i,i-N+1)$ be of rank n for all $i\geq N$ in order to define the state estimate completely. In other words, each successive set of N measurements at a time must provide information in all directions of state space. It is not sufficient for one such set to satisfy this condition because, unlike $\underline{\mathcal{J}}(i,1)$, $\underline{\mathcal{J}}(i,i-N+1)$ does not have a rank that is a nondecreasing function of i. From the definition of the N-step Information Matrix, the recursion that propagates its value can be expressed as

$$\underline{\mathcal{J}}(i,i-N+1) = \underline{\Phi}^{T}(i-1,i)\underline{\mathcal{J}}(i-1,i-N)\underline{\Phi}(i-1,i)+\underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{H}(i)$$

$$-\underline{\Phi}^{T}(i-N,i)\underline{H}^{T}(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\ \underline{\Phi}(i-N,i) \tag{2.1.91}$$

This equation simply states that to obtain $\underline{\mathcal{Q}}(i,i-N+1)$ from the one-step-propagated $\underline{\mathcal{Q}}(i-1,i-N)$, add the information obtained from $\underline{z}(i)$ and subtract the information due to $\underline{z}(i-N)$. After considerable algebra, the state estimate (2.1.89) can be put into recursive form:

$$\underline{x}^*(i) = \underline{\Phi}(i,i-1)\underline{x}^*(i-1) + \underline{B}(i-1)\underline{u}(i-1) + \underline{\mathcal{Q}}^{-1}(i,i-N+1) \cdot$$

$$\cdot \Big\{ \underline{H}^T(i)\underline{R}^{-1}(i)\big[\underline{z}(i)-\underline{H}(i)\underline{\Phi}(i,i-1)\underline{x}^*(i-1)-\underline{H}(i)\underline{B}(i-1)\underline{u}(i-1)\big]$$

$$-\underline{\Phi}^T(i-N,i)\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\big[\underline{z}(i-N)-\underline{H}(i-N)\underline{\Phi}(i-N,i-1)\underline{x}^*(i-1)$$

$$+\underline{H}(i-N)\sum_{k=i-N+1}^{i-1}\underline{\Phi}(i-N,k)\underline{B}(k-1)\underline{u}(k-1)\big]\Big\} \qquad (2.1.92)$$

This can be interpretted as follows. The expected value of $\underline{x}(i)$, given the value of $\underline{x}^*(i-1)$, or in other words just before $\underline{z}(i)$ is incorporated, is

$$\{\underline{\Phi}(i,i-1)\underline{x}^*(i-1) + \underline{B}(i-1)\underline{u}(i-1)\}$$

whereas the expected value of $\underline{x}^*(i-N)$, given $\underline{x}^*(i-1)$ is

$$\Big\{\underline{\Phi}(i-N,i-1)\underline{x}^*(i-1) - \sum_{k=i-N+1}^{i-1}\underline{\Phi}(i-N,k)\underline{B}(k-1)\underline{u}(k-1)\Big\}$$

so that the two terms within the braces in (2.1.92) are the values of the residuals at times i and (i-N) as propagated from the best state estimate at time (i-1). The summation in the last term can be replaced by a more convenient recursion by defining, for $i \geq N$,

$$\underline{u}^*(i-N) = \sum_{k=i-N+1}^{i-1}\underline{\Phi}(i-N,k)\underline{B}(k-1)\underline{u}(k-1) \qquad (2.1.93)$$

which can be generated by:

$$\underline{u}^*(i-N) = \underline{\Phi}(i-N,i-N-1)\underline{u}^*(i-N-1) - \underline{B}(i-N-1)\underline{u}(i-N-1)$$
$$+ \underline{\Phi}(i-N,i-1)\underline{B}(i-2)\underline{u}(i-2) \qquad (2.1.94)$$

thereby transforming the last bracketted expression in (2.1.92) into

$$\left[\underline{z}(i-N) - \underline{H}(i-N)\underline{\Phi}(i-N,i-1)\underline{x}^*(i-1) + \underline{H}(i-N)\underline{u}^*(i-N)\right]$$

2.2 Error Sensitivity of Optimal State Estimators

In many instances of applying linear Kalman filter techniques for state estimation, a number of parameters in the system dynamics are not completely determined apriori. Modelling techniques can often produce an adequate system description (adequate for the requirements of the particular application, as navigation or control to within specified limits) in the form of a linear difference equation driven by known inputs and white noise, but with certain entries in the state transition matrix or input matrix not known exactly. Sensitivity analyses have shown in the past that the uncertainty in these parameters can cause estimation errors that are unacceptably large for system requirements. For example, if the filter is cascaded with an optimal linear controller to provide an optimal stochastic controller for linear systems, the errors committed due to the parameter uncertainties may well be sufficient to cause instability when applied to a real situation. Or, the performance of a navigation system may deteriorate beyond bounds imposed by mission specifications because of such uncertainties. Therefore, it is often desired to be able to implement an on-line technique for simultaneous estimation of both the state variables and those system parameters which sensitivity analyses have revealed to be insufficiently well known.

The error analysis to be presented in this section has a twofold purpose. First, it will determine whether the estimator performance is adequate without incorporating any parameter estimates - if it is, then the considerations of computational speed and simplicity (with its inherent benefit to reliability) would dictate against the use of parameter estimation. Secondly, if the performance is not suitable, the error analysis will delineate which parameters are most crucial to estimate. Furthermore, it will be shown in section 4.3 that these state estimator error analyses also provide the data to determine how accurately each of the parameters can be estimated. Thus, the designer will be able to concentrate upon the critical uncertain parameters and determine whether the performance improvement afforded by estimating one or a number of them will yield an estimator that meets specifications.

To perform the error analysis, it will be assumed that a Kalman form of estimator based upon parameter values $\underline{a}$ is operating on a system which in fact is characterized by parameter values $\underline{a}_t$. In other words, the true system satisfies

$$\underline{x}(i+1;\underline{a}_t) = \underline{\Phi}(i+1,i;\underline{a}_t)\underline{x}(i;\underline{a}_t) + \underline{B}(i;\underline{a}_t)\underline{u}(i) + \underline{G}(i)\underline{w}(i) \tag{2.2.1}$$

$$\underline{z}(i;\underline{a}_t) = \underline{H}(i)\underline{x}(i;\underline{a}_t) + \underline{v}(i) \tag{2.2.2}$$

whereas the filter is described by

$$\overline{\underline{x}}(i+1;\underline{a}) = \underline{\Phi}(i+1,i;\underline{a})\hat{\underline{x}}(i;\underline{a}) + \underline{B}(i;\underline{a})\underline{u}(i) \tag{2.2.3}$$

$$\hat{\underline{x}}(i;\underline{a}) = \overline{\underline{x}}(i;\underline{a}) + \underline{M}(i;\underline{a})H^T(i)\underline{A}^{-1}(i;\underline{a}) \cdot \cdot \left[\underline{z}(i;\underline{a}_t) - \underline{H}(i)\overline{\underline{x}}(i;\underline{a})\right] \tag{2.2.4}$$

where, for convenience, $\underline{A}(i;\underline{a})$ has been defined as:

$$\underline{A}(i;\underline{a}) = \underline{H}(i)\underline{M}(i;\underline{a})\underline{H}^{T}(i) + \underline{R}(i) \qquad (2.2.5)$$

and where the estimated covariances propagate as

$$\underline{M}(i+1;\underline{a}) = \underline{\Phi}(i+1,i;\underline{a})\underline{P}(i;\underline{a})\underline{\Phi}^{T}(i+1,i;\underline{a}) + \underline{G}(i)\underline{Q}(i)\underline{G}^{T}(i) \qquad (2.2.6)$$

$$\underline{P}(i;\underline{a}) = \underline{M}(i;\underline{a}) - \underline{M}(i;\underline{a})\underline{H}^{T}(i)\underline{A}^{-1}(i;\underline{a})\underline{H}(i)\underline{M}(i;\underline{a}) \qquad (2.2.7)$$

$$= \left[\underline{I} - \underline{K}(i;\underline{a})\underline{H}(i)\right]\underline{M}(i;\underline{a})\left[\underline{I} - \underline{K}(i;\underline{a})\underline{H}(i)\right]^{T} + \underline{K}(i;\underline{a})\underline{R}(i)\underline{K}^{T}(i;\underline{a}) \qquad (2.2.8)$$

where

$$\underline{K}(i;\underline{a}) = \underline{M}(i;\underline{a})\underline{H}^{T}(i)\underline{A}^{-1}(i;\underline{a}) \qquad (2.2.9)$$

Note that this formulation delineates the effects of uncertain parameters in $\underline{\Phi}$ and $\underline{B}$. The method can readily be extended to account for errors in the value of $\underline{H}$, $\underline{Q}$, or $\underline{R}$ used by the filter. An additional bias or random sequence term could also be added to (2.2.1) in the form

$$\underline{x}(i+1;\underline{a}_t) = \underline{\Phi}(i+1,i;\underline{a}_t)\underline{x}(i;\underline{a}_t) + \underline{B}(i;\underline{a}_t)\underline{u}(i) + \underline{G}(i)\underline{w}(i) + \underline{\nu}(i) \qquad (2.2.10)$$

to approximate the effect of neglecting nonlinearities, using a reduced dimensional system, or other omissions necessary to achieve a practical model of a real physical system. It is also assumed here that the control input is completely deterministic or that it is based upon feedback of calculated variables, and thus $\underline{u}(i)$ is known exactly to the filter; the more general case, which allows the $\underline{u}(i)$ actually entering the system and that assumed by the filter to differ, can also be developed using the techniques of this section. However,

it is currently desired to concentrate on the effects of uncertainties in the $\underline{\Phi}$ and $\underline{B}$ matrices exclusively.

For convenience, define

$$\Delta\underline{\Phi}(i+1,i) = \underline{\Phi}(i+1,i;\underline{a}_t) - \underline{\Phi}(i+1,i;\underline{a}) \tag{2.2.11}$$

$$\Delta\underline{B}(i) = \underline{B}(i;\underline{a}_t) - \underline{B}(i;\underline{a}) \tag{2.2.12}$$

The analysis equations will be developed in terms of a general $\Delta\underline{\Phi}$ and $\Delta\underline{B}$. In practical applications, parameter variations would be investigated one at a time, so that one of these terms may in fact be zero for all a_ℓ (this need not be the case since if the real system is continuous-time, a single system parameter could affect a number of entries in both $\underline{\Phi}$ and $\underline{B}$). For the particularly simple case of only one uncertain entry, for example the entry in the k-th row and ℓ-th column of $\underline{\Phi}$,

$$\Delta\underline{\Phi}(i+1,i) = (a_t-a)\underline{q}_k\underline{q}_\ell^T\,,\ \Delta\underline{B}(i) = \underline{0} \tag{2.2.13}$$

where $\underline{q}_j$ is an n-vector composed of all zeroes except for a one as its j-th component.

The objective of this error analysis is to obtain an expression for the actual estimation error covariance matrix, since this will indicate the real filter performance under the influence of the uncertain parameter values. Define the true estimation errors, before and after incorporation of the i-th measurement (for particular choice of $\underline{a}$ and $\underline{a}_t$), as

$$\underline{e}(i^-;\underline{a}_t,\underline{a}) = \underline{x}(i;\underline{a}_t) - \bar{\underline{x}}(i;\underline{a}) \tag{2.2.14}$$

$$\underline{e}(i^+;\underline{a}_t,\underline{a}) = \underline{x}(i;\underline{a}_t) - \hat{\underline{x}}(i;\underline{a}) \tag{2.2.15}$$

and the corresponding actual error covariances as

$$\underline{M}_t(i;\underline{a}_t,\underline{a}) = E\left\{\underline{e}(i^-;\underline{a}_t,\underline{a})\underline{e}^T(i^-;\underline{a}_t,\underline{a})\,\middle|\,\underline{a}_t\right\} \qquad (2.2.16)$$

$$\underline{P}_t(i;\underline{a}_t,\underline{a}) = E\left\{\underline{e}(i^+;\underline{a}_t,\underline{a})\underline{e}^T(i^+;\underline{a}_t,\underline{a})\,\middle|\,\underline{a}_t\right\} \qquad (2.2.17)$$

where the expectation is taken over the values of the initial state uncertainty and the noise sequences. It is the value of $\underline{M}_t(i;\underline{a}_t,\underline{a})$ and $\underline{P}_t(i;\underline{a}_t,\underline{a})$, which will henceforth be denoted simply as $\underline{M}_t(i)$ and $\underline{P}_t(i)$, that will be used both to depict this error analysis and also to evaluate the ambiguity functions in section 4.3 for a measure of parameter estimation performance.

From the defining relations,

$$\begin{aligned}
\underline{e}(i+1^-) &= \underline{x}(i+1;\underline{a}_t) - \bar{\underline{x}}(i+1;\underline{a}) \\
&= \left[\underline{\Phi}(i+1,i;\underline{a}) + \Delta\underline{\Phi}(i+1,i)\right]\underline{x}(i;\underline{a}_t) \\
&\quad + \left[\underline{B}(i;\underline{a}) + \Delta\underline{B}(i)\right]\underline{u}(i) + \underline{G}(i)\underline{w}(i) \\
&\quad - \underline{\Phi}(i+1,i;\underline{a})\hat{\underline{x}}(i;\underline{a}) - \underline{B}(i;\underline{a})\underline{u}(i) \\
&= \underline{\Phi}(i+1,i;\underline{a})\underline{e}(i^+) + \Delta\underline{\Phi}(i+1,i)\underline{x}(i;\underline{a}_t) \\
&\quad + \Delta\underline{B}(i)\underline{u}(i) + \underline{G}(i)\underline{w}(i)
\end{aligned} \qquad (2.2.18)$$

where $\underline{e}(i^+)$ is obtained as

$$\begin{aligned}
\underline{e}(i^+) &= \underline{x}(i;\underline{a}_t) - \hat{\underline{x}}(i;\underline{a}) \\
&= \underline{x}(i;\underline{a}_t) - \left[\underline{I} - \underline{K}(i;\underline{a})\underline{H}(i)\right]\bar{\underline{x}}(i;\underline{a}) \\
&\quad - \underline{K}(i;\underline{a})\left[\underline{H}(i)\underline{x}(i;\underline{a}_t) + \underline{v}(i)\right] \\
&= \left[\underline{I} - \underline{K}(i;\underline{a})\underline{H}(i)\right]\underline{e}(i^-) - \underline{K}(i;\underline{a})\underline{v}(i)
\end{aligned} \qquad (2.2.19)$$

These expressions can be used directly to derive the desired covariances. Thus,

$$\underline{M}_t(i+1) = \underline{\Phi}(i+1,i;\underline{a})\underline{P}_t(i)\underline{\Phi}^T(i+1,i;\underline{a}) + \underline{G}(i)\underline{Q}(i)\underline{G}^T(i)$$

$$+ \Delta\underline{\Phi}(i+1,i)\left[\overline{\underline{x}(i)\underline{x}^T(i)}\right]\Delta\underline{\Phi}^T(i+1,i)$$

$$+ \underline{\Phi}(i+1,i;\underline{a})\left[\overline{\underline{e}(i^+)\underline{x}^T(i)}\right]\Delta\underline{\Phi}^T(i+1,i)$$

$$+ \Delta\underline{\Phi}(i+1,i)\left[\overline{\underline{x}(i)\underline{e}^T(i^+)}\right]\underline{\Phi}^T(i+1,i;\underline{a})$$

$$+ \left\{\underline{\Phi}(i+1,i;\underline{a})\left[\overline{\underline{e}(i^+)\underline{u}^T(i)}\right]\Delta\underline{B}^T(i)\right.$$

$$+ \Delta\underline{B}(i)\left[\overline{\underline{u}(i)\underline{e}^T(i^+)}\right]\underline{\Phi}^T(i+1,i;\underline{a})$$

$$+ \Delta\underline{\Phi}(i+1,i)\left[\overline{\underline{x}(i)\underline{u}^T(i)}\right]\Delta\underline{B}^T(i)$$

$$+ \Delta\underline{B}(i)\left[\overline{\underline{u}(i)\underline{x}^T(i)}\right]\Delta\underline{\Phi}^T(i+1,i)$$

$$\left. + \Delta\underline{B}(i)\left[\overline{\underline{u}(i)\underline{u}^T(i)}\right]\Delta\underline{B}^T(i)\right\} \qquad (2.2.20)$$

where the bar notation denotes expectation. Note that the term in braces is nonzero only when the uncertain parameters are present in the $\underline{B}$ matrix. The covariance after the measurement at instant i is

$$\underline{P}_t(i) = \left[\underline{I} - \underline{K}(i;\underline{a})\underline{H}(i)\right]\underline{M}_t(i)\left[\underline{I} - \underline{K}(i;\underline{a})\underline{H}(i)\right]^T$$

$$+ \underline{K}(i;\underline{a})\underline{R}(i)\underline{K}^T(i;\underline{a}) \qquad (2.2.21)$$

The initial condition for this recursion is

$$\underline{P}_t(0) = \underline{P}(0) = \underline{P}_o \qquad (2.2.22)$$

To completely specify these equations, the recursion for the matrix $\left[\overline{\underline{x}(i)\underline{x}^T(i)}\right]$ can be described by:

$$\left[\overline{\underline{x}(i+1)\underline{x}^T(i+1)}\right] = \underline{\Phi}(i+1,i;\underline{a}_t)\left[\overline{\underline{x}(i)\underline{x}^T(i)}\right]\underline{\Phi}^T(i+1,i;\underline{a}_t)$$
$$+ \underline{B}(i;\underline{a}_t)\left[\overline{\underline{u}(i)\underline{u}^T(i)}\right]\underline{B}^T(i;\underline{a}_t)$$
$$+ \underline{\Phi}(i+1,i;\underline{a}_t)\left[\overline{\underline{x}(i)\underline{u}^T(i)}\right]\underline{B}^T(i;\underline{a}_t)$$
$$+ \underline{B}(i;\underline{a}_t)\left[\overline{\underline{u}(i)\underline{x}^T(i)}\right]\underline{\Phi}^T(i+1,i;\underline{a}_t)$$
$$+ \underline{G}(i)\underline{Q}(i)\underline{G}^T(i) \qquad (2.2.23)$$

from the initial condition

$$\left[\overline{\underline{x}(0)\underline{x}^T(0)}\right] = \underline{P}_t(0) + \hat{\underline{x}}(0)\hat{\underline{x}}^T(0) = \underline{P}_o + \hat{\underline{x}}_o\hat{\underline{x}}_o^T$$

and also the recursion

$$\left[\overline{\underline{e}(i+1^-)\underline{x}^T(i+1)}\right] = \underline{\Phi}(i+1,i;\underline{a})\left[\overline{\underline{e}(i^+)\underline{x}^T(i)}\right]\underline{\Phi}^T(i+1,i;\underline{a}_t)$$
$$+ \underline{\Phi}(i+1,i;\underline{a})\left[\overline{\underline{e}(i^+)\underline{u}^T(i)}\right]\underline{B}^T(i;\underline{a}_t)$$
$$+ \Delta\underline{\Phi}(i+1,i)\left[\overline{\underline{x}(i)\underline{x}^T(i)}\right]\underline{\Phi}^T(i+1,i;\underline{a}_t)$$
$$+ \Delta\underline{\Phi}(i+1,i)\left[\overline{\underline{x}(i)\underline{u}^T(i)}\right]\underline{B}^T(i;\underline{a}_t)$$
$$+ \underline{G}(i)\underline{Q}(i)\underline{G}^T(i)$$
$$+ \left\{\Delta\underline{B}(i)\left[\overline{\underline{u}(i)\underline{x}^T(i)}\right]\underline{\Phi}^T(i+1,i;\underline{a}_t)\right.$$
$$\left. + \Delta\underline{B}(i)\left[\overline{\underline{u}(i)\underline{u}^T(i)}\right]\underline{B}^T(i;\underline{a}_t)\right\} \qquad (2.2.24)$$

$$\left[\overline{\underline{e}(i^+)\underline{x}^T(i)}\right] = \left[\underline{I} - \underline{K}(i;\underline{a})\underline{H}(i)\right]\left[\overline{\underline{e}(i^-)\underline{x}^T(i)}\right] \qquad (2.2.25)$$

from the initial condition:

$$\left[\overline{\underline{e}(0^+)\underline{x}^T(0)}\right] = \underline{P}_t(0) = \underline{P}_o \qquad (2.2.26)$$

where again the term in braces in (2.2.24) denotes the contribution due to uncertain parameters appearing in $\underline{B}$.

Finally, the expectations involving $\underline{u}(i)$ must be evaluated. If the $\underline{u}(i)$ sequence is precomputed, then all terms of the form $E\{\underline{u}\cdot\}$ simply become $\underline{u}\,E\{\cdot\}$ and $E\{\underline{u}\underline{u}^T\}$ becomes $\underline{u}\underline{u}^T$. However, if the control input is a feedback of calculated variables,

$$\underline{u}(i) = -\underline{C}(i;\underline{a})\hat{\underline{x}}(i;\underline{a}) \qquad (2.2.27)$$

then the required expectations do not separate, but can conveniently be expressed in terms of $\underline{P}_t(i)$, $\left[\overline{\underline{x}(i)\underline{x}^T(i)}\right]$, $\left[\overline{\underline{e}(i^+)\underline{x}^T(i)}\right]$ and its transpose $\left[\overline{\underline{x}(i)\underline{e}^T(i^+)}\right]$, for which recursions have already been obtained. Thus, by noting that

$$\underline{P}_t = E\{[\underline{x}-\hat{\underline{x}}][\underline{x}-\hat{\underline{x}}]^T\} = E\{\underline{x}\underline{x}^T-\underline{x}\hat{\underline{x}}^T-\hat{\underline{x}}\underline{x}^T+\hat{\underline{x}}\hat{\underline{x}}^T\}$$

$$\left[\overline{\underline{e}^+\underline{x}^T}\right] = E\{[\underline{x}-\hat{\underline{x}}]\underline{x}^T\} = E\{\underline{x}\underline{x}^T-\hat{\underline{x}}\underline{x}^T\}$$

$$\left[\overline{\underline{x}\underline{e}^{+T}}\right] = E\{\underline{x}[\underline{x}-\hat{\underline{x}}]^T\} = E\{\underline{x}\underline{x}^T-\underline{x}\hat{\underline{x}}^T\}$$

$$\left[\overline{\underline{x}\underline{x}^T}\right] = E\{\underline{x}\underline{x}^T\}$$

the desired matrices can be written as

$$\begin{aligned}\overline{\underline{u}(i)\underline{u}^T(i)} &= \underline{C}(i;\underline{a})\left[\overline{\hat{\underline{x}}(i;\underline{a})\hat{\underline{x}}^T(i;\underline{a})}\right]\underline{C}^T(i;\underline{a}) \\ &= \underline{C}(i;\underline{a})\Big\{\underline{P}_t(i) - \left[\overline{\underline{e}(i^+)\underline{x}^T(i)}\right] - \left[\overline{\underline{x}(i)\underline{e}^T(i^+)}\right] \\ &\quad + \left[\overline{\underline{x}(i)\underline{x}^T(i)}\right]\Big\}\underline{C}^T(i;\underline{a}) \qquad (2.2.28)\end{aligned}$$

$$\overline{[\underline{e}(i^+)\underline{u}^T(i)]} = -\overline{[\underline{e}(i^+)\hat{\underline{x}}^T(i)]}\,\underline{C}^T(i;\underline{a})$$
$$= \left\{\underline{P}_t(i) - \overline{[\underline{e}(i^+)\underline{x}^T(i)]}\right\}\underline{C}^T(i;\underline{a})$$
$$\overline{[\underline{x}(i)\underline{u}^T(i)]} = -\overline{[\underline{x}(i)\hat{\underline{x}}^T(i)]}\,\underline{C}^T(i;\underline{a})$$
$$= \left\{\overline{[\underline{x}(i)\underline{e}^T(i^+)]} - \overline{[\underline{x}(i)\underline{x}^T(i)]}\right\}\underline{C}^T(i;\underline{a}) \qquad (2.2.29)$$

The evaluation of $\underline{M}_t(i)$ and $\underline{P}_t(i)$ for various values of $\underline{a}$ and $\underline{a}_t$ will indicate the expected degradation in performance due to the variation of individual parameter values. In order to determine the range of each parameter value that yields acceptable performance, and also to compare the effects of the different parameters, it would be convenient to derive a scalar "cost functional" to serve as a measure of filter performance. In any given application, there may be certain state variables that must be estimated much more accurately than others to achieve filter system objectives. Using engineering judgment, the designer can define a (diagonal) weighting matrix $\underline{\mathcal{W}}$ to reflect these relative importances, the more crucial the state variable, the heavier the corresponding diagonal term of $\underline{\mathcal{W}}$. For instance, if the system estimates $\underline{x} = [x_1, x_2]^T$, and only the effectiveness of estimating x_1 is vital, $\underline{\mathcal{W}}$ would be $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Using the matrix $\underline{\mathcal{W}}$, the "cost" would be written as

$$V(i;\underline{a},\underline{a}_t,\underline{\mathcal{W}}) = \mathrm{tr}\left(\underline{\mathcal{W}}\underline{P}_t(i;\underline{a}_t,\underline{a})\right) \qquad (2.2.30)$$

An associated form that would help determine the critical parameters would be

$$\frac{\Delta V}{\Delta a_\ell}(i;\underline{a},\underline{a}_t,\underline{\mathcal{W}}) = \frac{\mathrm{tr}\left(\underline{\mathcal{W}}[\underline{P}_t(i;\underline{a}_t,\underline{a}) - \underline{P}_t(i;\underline{a}_t,\underline{a}_t)]\right)}{|a_{t\ell} - a_\ell|} \qquad (2.2.31)$$

By evaluating this term for each parameter a_ℓ, and knowing the magnitude of the possible variation in each uncertain parameter value, those parameters with the most detrimental effect can be isolated. It may also be convenient to express this as a percentage increase over the best achievable performance when the parameter value is known exactly:

$$\frac{\Delta V'}{\Delta a_\ell}(i;\underline{a},\underline{a}_t,\underline{W}) = \frac{\mathrm{tr}\left(\underline{W}[\underline{P}_t(i;\underline{a}_t,\underline{a})-\underline{P}_t(i;\underline{a}_t,\underline{a}_t)]\right)}{|a_{t\ell}-a_\ell|\ \mathrm{tr}[\underline{W}\underline{P}_t(i;\underline{a}_t,\underline{a}_t)]} \tag{2.2.32}$$

Note that the evaluation of $\underline{P}_t(i)$ depends upon the level of the system state activity. In practice, a sufficiently high level of operations, representative of the expected range of operational environments, would be used. Since the various system modes would be excited differently according to the frequency content of the inputs, a number of representative (or "worst case") trials would often be incorporated into the analysis.

Expressions similar to (2.2.31) and (2.2.32), but with $\underline{P}_t(i;\underline{a}_t,\underline{a}_t)$ replaced by $\underline{P}(i;\underline{a})$, can be used to indicate how erroneously the computed $\underline{P}$ represents the actual state estimate error covariance due to faulty knowledge of $\underline{a}$, thereby providing a measure of state estimator reliability. With these various performance measures, it is possible to decide whether parameter estimation need be incorporated into a proposed filter system, and if so, which parameters should be estimated.

CHAPTER III

COMBINED STATE AND PARAMETER ESTIMATION

This chapter will address the problem of estimating the state variables of a linear dynamic system optimally when a number of parameters in the system equation are known only with some uncertainty. In order to refine the state estimate under these conditions, the uncertain parameters will be estimated simultaneously by means of maximum likelihood methods.

To formulate the problem fully, it will be assumed that modelling techniques have produced an adequate system description in the form of a linear difference equation driven by known inputs and white noise sequences, on which noise-corrupted measurements are taken. As mentioned previously, a discrete-time representation of the plant dynamics has been chosen to facilitate the eventual implementation by an on-line digital computer. This model is as given by equations (2.1.1) to (2.1.9). However, certain parameters in the state transition matrix $\underline{\Phi}(i+1,i)$ or input matrix $\underline{B}(i)$ are not determined completely. To conform with a realistic situation, it will further be assumed that the designer will often lack sufficient information to develop valid statistical models or probability density functions for these parameters. Instead, he will be able to provide, at best, a range of possible values and a most probable value for each parameter by examining the physics involved in the particular problem. These parameters will be distinguished from the state variables in that they will vary signficantly more slowly than the states, and may in fact be time-invariant.

Uncertainties in the measurement matrix $\underline{H}(i)$ will not be emphasized, as mentioned in Chapter I. When the state space model is expressed in physical variables, the measurement matrices are generally better known in most problems than

$\underline{\Phi}(i+1,i)$, $\underline{B}(i)$, or $\underline{G}(i)$, and thus it is usually not as necessary to estimate parameters in $\underline{H}(i)$. Furthermore, uncertainties in $\underline{H}(i)$ often cannot be distinguished from uncertainties in $\underline{\Phi}(i,i-1)$ or $\underline{B}(i)$ on the basis of the measurement and control time histories, so a convenient choice of state space representation can usually avoid uncertain parameters in $\underline{H}(i)$. For example, in the simple case of a scalar state system with no control inputs or driving noise, the i-th measurement is $[h(i)\phi(i,0)x_o+v(i)]$, from which an estimate can be made of the product $h(i)\phi(i,0)$, but not of $h(i)$ or $\phi(i,0)$ separately. Appendix A will present the estimator formulation that allows uncertain parameters in $\underline{H}(i)$, especially useful if canonical variables are employed (see section 5.5). Note that uncertainties in $\underline{G}(i)$ are not included, but this case can be treated equivalently as uncertain parameters in the covariance matrix $\underline{Q}(i)$, which has already been investigated in the literature (see Abramson (1968)).

The simplest problem is that of estimating a single entry in the $\underline{\Phi}(i+1,i)$ or $\underline{B}(i)$ matrix simultaneously with the state. Consider the case of one entry in the $\underline{\Phi}(i+1,i)$ matrix to be estimated. If it is located in the k-th row and the ℓ-th column, then $\underline{\Phi}(i+1,i)$ can be written as

$$\underline{\Phi}(i+1,i) = \begin{bmatrix} \phi_{11}(i+1,i) & \cdots & \phi_{1\ell}(i+1,i) & \cdots & \phi_{1n}(i+1,i) \\ \vdots & & \vdots & & \vdots \\ \phi_{k1}(i+1,i) & \cdots & a(i) & \cdots & \phi_{kn}(i+1,i) \\ \vdots & & \vdots & & \vdots \\ \phi_{n1}(i+1,i) & \cdots & \phi_{n\ell}(i+1,i) & \cdots & \phi_{nn}(i+1,i) \end{bmatrix} \tag{3.1}$$

where $a(i)$ is the parameter to be estimated. This can be expressed somewhat more conveniently as

$$\underline{\Phi}(i+1,i) = \underline{\Phi}_o(i+1,i) + a(i)\underline{q}_k\underline{q}_\ell^T \tag{3.2}$$

where $\underline{\Phi}_o(i+1,i)$ is the known part of the state transition matrix, with a zero as its k-ℓ-th entry, and $\underline{q}_j$ is an n-vector composed of all zeroes except for one as its j-th component. Using this relation, a term such as $[\partial\underline{\Phi}(i+1,i)/\partial a]\ \hat{\underline{x}}(i)$ readily becomes

$$[\partial\underline{\Phi}(i+1,i)/\partial a]\ \hat{\underline{x}}(i) = \underline{q}_k\ \underline{q}_\ell^T\ \hat{\underline{x}}(i)$$

$$= \underline{q}_k\ \hat{x}_\ell(i) \qquad (3.3)$$

Recognition of such convenient forms will yield more efficient programming of the final algorithms. These results can easily be generalized to the case of multiple entries of $\underline{\Phi}(i+1,i)$ and/or $\underline{B}(i)$ to be estimated.

However, the uncertain parameters need not be explicit entries in the $\underline{\Phi}$ or $\underline{B}$ matrices. If the actual system is a continuous time system, as described by the equation

$$d\underline{x}(t) = \underline{F}(t)\underline{x}(t)dt + \underline{B}(t)\underline{u}(t)dt + \underline{G}(t)d\underline{\beta}(t) \qquad (3.4)$$

or the (mathematically less precise) relation

$$\dot{\underline{x}}(t) = \underline{F}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) + \underline{G}(t)\underline{w}(t) \qquad (3.5)$$

then a single uncertain parameter in $\underline{F}(t)$ will generally affect both $\underline{\Phi}(i+1,i)$ and $\underline{B}(i)$ in the equivalent discrete representation (see Appendix B for the continuous-discrete relationships). In the most general case, the matrices $\partial\underline{\Phi}(i+1,i)/\partial a_\ell$ and $\partial\underline{B}(i)/\partial a_\ell$ (where a_ℓ is the ℓ-th component of $\underline{a}$) will be functions of the time instant i, the parameter value $\underline{a}(i)$, and in some infrequent cases, the value of the state vector $\underline{x}(i)$. These matrices would then be evaluated using the state and parameter estimates. In many foreseeable applications, though, $\underline{\Phi}$ and $\underline{B}$ can adequately be modelled as time-invariant matrices whose entries are linear functions

of the parameters: the actual functional relationships can be precomputed and then curve-fitted. Even if linear relationships are not adequate, piecewise-linear relations are probably preferable to higher order curve fitting for on-line implementations, since a computer can perform a "branch" based on $\underline{\hat{a}}(i)$ and a linear function evaluation more easily than a higher order function evaluation. For these applications, the matrices $\partial\underline{\Phi}(i+1,i)/\partial a_\ell$ and $\partial\underline{B}(i)/\partial a_\ell$ are constants that can be precomputed and stored in the computer memory, to be used in various propagations and also to evaluate the new value of $\underline{\Phi}(i+1,i)$ and $\underline{B}(i)$ from the estimate $\underline{\hat{a}}(i)$.

The mathematical model employed for the uncertain parameters has been designed for general applicability and ease of usage. To incorporate as much of the physical characterization of the parameters as possible, the model should reflect the fact that the time variation of parameters is usually slow with respect to the time variation of state variables; the uncertain parameters may well be unknown constants. An "optimal" estimation of the parameters should exploit the knowledge that their values will be more consistent from sample to sample than will any of the noise sources entering into the system dynamics or measurements.

Assume that an adequate model is that the parameters are essentially constant over any given interval of N sample periods. That is, at a given time, i, the parameters are modelled as remaining constant over the samples (i-N+1) to i. At the next instant of time, the parameters are again assumed to be constants, though possibly of different magnitude, over the samples (i-N+2) to (i+1). In a more general extension of this idea, a polynomial of fixed order might be used to model the parameters over fixed-duration intervals, but use of a constant can usually be justified as the best apriori approximation, as well as the assumed form that yields the least complex estimation equations.

This model is more appropriate than a stochastic descrip-

tion for a number of reasons. Stochastic models require an apriori specification of the probability densities for the range of admissible parameter values, a difficult requirement to satisfy in most realistic applications. A solution would be to assume a form for the densities, as to consider the parameters as random variables with Gaussian distributions or uniform distributions between physically determined bounds, or to model them as the outputs of integrators driven by white Gaussian noise. Claims of overall system "optimality" must then be viewed as estimation that optimizes a performance criterion, provided that the initial, rather arbitrary, assumption on the stochastic nature of the parameters was correct. Because of the central limit theorem, a Gaussian process (white Gaussian noise through a linear shaping filter) is often a good model of physical noise driving the system state, but there is no such physical motivation for a stochastic model for the parameters. Moreover, an engineer would find it difficult to transform his physical knowledge of the problem into an appropriate covariance matrix for such a model. On the other hand, he would usually be able to determine, though admittedly in a subjective manner, an adequate value for N from the physics of the particular problem at hand.

Based upon the "constant over N steps" model for the parameters, one is naturally led to considering a fixed-length memory type of estimator for the parameters. Whereas the usual recursive filters "remember" all observations back to the initial time, a fixed-length memory would retain only the N most recent data points. Conceptually, then, one can ask, "What constant values for the parameters will fit these N samples best in a maximum likelihood sense?" It will be shown in section 3.2 that, by proper choice of the likelihood function, such a parameter estimate can be combined with a state estimate that depends either on the most recent N samples of data or on all data taken. The choice of the likelihood function for eventual implementation will depend

upon tractability of the resulting equations, physical sensibility of the likelihood function, and overall performance of the estimator.

Much can be gained by such fixed-length memory estimators based on the parameter model just described. They are less susceptible to errors due to poor quality measurement data than are systems that perform an independent parameter estimate at each sample; they also reflect the physical expectation that the parameter values are somewhat consistent over a certain interval of time. If a recursive filter were based upon the assumption that the parameter values are unknown constants throughout the entire period of estimation, this assumption might be invalid, and the estimator would become less sensitive to parameter changes occurring later in the period than to those that occur earlier. If such an assumption were not incorporated into recursive estimators, their computational complexity might well be prohibitive for on-line applications. Thus, the motivation for considering fixed-length memory methods would be to provide an on-line estimator that would remain sensitive to parameter variations while not being overly sensitive to bad data.

State estimation using a limited memory length can be advantageous in certain applications as well. Jazwinski (1968) has shown it to be an effective means of removing estimator error divergence caused by system nonlinearities: the system is assumed to be adequately modelled as linear over N samples, and therefore the data previous to an N-step interval is discarded. However, his and Schweppe's (1964) analytical results require the assumption of no dynamic noise entering the system. This chapter will develop fixed-length memory state estimators that do not impose this restriction, and these results can be generalized to the case of simultaneously estimating the uncertain system parameters.

Apparent drawbacks of the fixed-length memory formulation, as compared to the recursive form, would be the necessity to

remember N sets of measurement data at all times and the foreseeable use of "batch processing" of the data collected over an interval each time an estimate is made. Both objections can be allayed by realizing that as the system passes from time (i-1) to i, the information provided by the (i-N)-th set of data would be removed, and replaced by the current information from the i-th measurement. This suggests storing data in the form of the variables that appear directly in the estimate updating equations, which, as will be shown subsequently, infers storing $N/2\ (p^2 + 3p)$ variables, or Np if certain precomputations are used. Thus, the memory requirements could be satisfied by a moderate amount of "inexpensive" read-write storage, using core memory index-changing methods to overwrite the newest data into the locations previously occupied by the oldest data. Furthermore, another general form of estimate implementation will be developed in which the data incorporation can be performed with no greater computational difficulty than a recursive formulation. Finally, the slowly varying nature of the parameters can be exploited by estimating them less frequently than the state variables. These concepts will be developed in more detail in Chapter V.

Section 3.1 will discuss the conditions that the system must satisfy in order to perform the state and parameter estimation. Section 3.2 will then derive the likelihood equations appropriate to the various assumptions and amounts of available statistical information in different applications. The resulting equations will generally not have closed form analytical solutions, and therefore section 3.3 investigates iterative solutions, and section 3.4 demonstrates the full-scale implementation of the method deemed most suitable, called "scoring for parameters."

3.1 Ability to Perform the Estimation

In order to establish the conditions necessary for

state and parameter estimation to be possible, the purposes and means of such estimation must be investigated. A fundamental assumption is that a linear model has been developed to represent the behavior of a physical system adequately, but that certain parameters in this linear description are not known exactly. Equations depict only those aspects of a physical process that are most significant to the user's purposes: there are no n-dimensional processes in nature, only models of processes with such a state space description. Thus, one cannot say he is trying to estimate the "true" values of the states and parameters of the system, but can attempt to find the values which when substituted into the model, yield a model output behavior that best duplicates the actual system performance (which can be measured only with a degree of uncertainty). Moreover, this optimal replication is achieved for a single set of input and output sequences, with no guarantee that the same parameter values provide an optimal system representation for other sequences as well.

The ability to perform the estimation will thus be determined by conditions upon the mathematical model employed as an adequate system representation. If the estimation is impossible with the originally formulated problem, one may need to incorporate different measurements, additional measurements, or a modified system model, in order to satisfy these conditions.

In the case of estimating only the state for a linear, time invariant system model whose parameters are known explicity, the condition for being able to improve the estimates of all state variables with the measurement data, observability, is readily verifiable. Unfortunately, the necessary conditions that follow for the case of estimating both states and parameters are not as practical to apply.

Consider first the problem of estimating the state for a given set of parameter values. A linear system (either

unforced or driven by a deterministic input) is termed observable if it is possible to determine the entire state at the initial time instant, $\underline{x}(0)$, exactly from only the noise-free measurements $\underline{z}(1)$, ..., $\underline{z}(i)$ of the system output over some finite interval of time. Because of uniqueness of solutions to linear dynamic equations, knowledge of $\underline{x}(0)$ also determines the entire state trajectory from $\underline{x}(0)$ to $\underline{x}(i)$. It is easily shown (see Meditch (1969)) that a necessary and sufficient condition for observability is that, for some finite $i > 0$,

$$\sum_{j=1}^{i} \underline{\Phi}^T(j,0;\underline{a})\underline{H}^T(j)\underline{H}(j)\underline{\Phi}(j,0;\underline{a}) > \underline{0} \qquad (3.1.1)$$

If this process is viewed as uniquely determining an arbitrary vector in n-dimensional state space from a set of i m-dimensional vectors, it is apparent that at least n/m (or the least integer greater than this number) measurements must be taken before the system is completely observable. Physically, an observable system has a measuring device such that all of the modes of the system response become visible to it in a finite interval of time.

When noise enters into the measurements, it is no longer appropriate to speak of determining the state exactly from the measurements, but rather of finding the conditional density of the state, given the measurements. A completely observable system then is one in which the measurements provide information in all n directions of state space. In this context, the system model is said to be completely observable with respect to the measurements $\underline{z}(k)$, ..., $\underline{z}(i)$ if and only if

$$\sum_{j=k}^{i} \underline{\Phi}^T(j,i;\underline{a})\underline{H}^T(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i;\underline{a}) > \underline{0} \qquad (3.1.2)$$

where the matrix summation is recognized as $\underline{\mathcal{J}}(i,k;\underline{a})$. If $\underline{a}_t(j)$ is the (unknown) value of the uncertain parameters which yields the most accurate linear model of the real

system, then for state estimation purposes, the model should be completely observable with respect to $\underline{z}(1), \ldots, \underline{z}(i)$, for some i, for $\underline{a}(j) = \underline{a}_t(j)$. However, since $\underline{a}_t(j)$ is not known apriori, this condition would require verification for all admissible sequences of $\underline{a}(j)$, a very difficult task in practice.

Since the system model is assumed time-invariant over N steps, the observability criterion can also be written as

$$\text{rank}\left(\left[\underline{H}^T : \underline{\Phi}^T\underline{H}^T : \cdots : \underline{\Phi}^{T(N-1)}\underline{H}^T\right]\right) = n \qquad (3.1.3)$$

or, if N is greater than n, the equivalent condition except that the last partition becomes $(\underline{\Phi}^{T(n-1)}\underline{H}^T)$. (The equivalence is demonstrated by means of cyclic generators; see Gantmacher (1959).) However, this condition would have to be tested for all admissible constant values of the parameters, which unfortunately is still not very practical in most applications.

To investigate the ability to perform the parameter estimation, it is convenient to transform to canonical variables and thereby separate system modes. Each uncertain parameter affects a subset of these modes. Of the modes that a particular parameter affects, at least one must be:

1) observable;
2) excited by the initial conditions, or controllable with respect to the points of entry of the dynamic noise, or controllable with respect to the points of entry of the control input which is not identically zero;
3) such that $\underline{\Phi}$ and $\underline{B}$ do not assume identical values over a range of parameter values (lack of sensitivity to parameter variation) or for a number of parameter values (as, the aliasing phenomenon, in which a high frequency oscillation can appear to have a much lower frequency of oscillation to a sampling device);

over the admissable range of parameter values. Conditions (1) and (2) imply verifying observability and controllability (or nonzero magnitude of initial state conditions) for each mode partition. Statement (2) can also be expressed as the state initial conditions, or dynamic noise, or control, have nontrivial projections onto the eigenvector(s) that correspond to a system mode that the uncertain parameter affects. If the parameter is confined to $\underline{B}$, then statement (2) is altered to include only the last of its three conditions. With regard to (3), $\underline{\Phi}$ and $\underline{B}$ would rarely be functionally independent of a parameter over a finite range if the parameter were critical enough to estimate, but a low sensitivity to the parameter value would be of importance practically; this aspect is deferred to section 4.2, in which such sensitivity can be demonstrated explicitly. Also related to (3) is the fact that $\underline{\Phi}$ and $\underline{B}$ (or, in the particular case of maximum likelihood estimation, the log-likelihood function) must not assume the same values for greater than one set of values of the parameters.

The above conditions admit the estimation of a parameter in a mode that is observable but uncontrollable, if the initial conditions excite the mode. If this mode is stable, its energy level would eventually decay to a point beyond which the parameter estimation would be very inaccurate, if not altogether impossible; but at this point the effect of the mode could be ignored.

The choice of control input can significantly affect the parameter estimation. For instance, if an uncertain parameter corresponded to an oscillation frequency in the system dynamics, it would be advantageous to introduce considerable control energy at this (uncertain) frequency to excite the mode; however, if the admissible range of frequencies is large, incorporating control energy over the entire range may well be prohibitive. Usually the control input has other purposes than facilitating parameter

estimation, but Schweppe (1970) has suggested a procedure for determining a desirable control history by means of minimizing the Cramér-Rao lower bound on parameter estimate error variance as a function of the control sequence (a formidable task).

The ambiguity function concept to be discussed in Chapter IV will also indicate whether the estimation of particular parameters is feasible, and do so in a practical manner. Moreover, it will predict the accuracy with which each parameter can be estimated.

The previous discussion concerned conditions for the ability to estimate the parameter values. In certain instances, one is interested only in the ability to predict input/output behavior, and does not necessarily require that a "correct" parameter value be estimated. Practically, an ambiguity function analysis and simulation results would be employed to ascertain the adequacy of performance.

3.2 Derivations of the Likelihood Equations

For any specified log-likelihood function $L[\underline{\xi}(i),\underline{Z}(i)]$, where $\underline{\xi}(i)$ is the vector of variables to be estimated and $\underline{Z}(i)$ is the set of realized values of the measurements to be used as data, the objective of maximum likelihood estimation is to find that value $\underline{\xi}^*(i)$ which maximizes $L[\underline{\xi}(i),\underline{Z}(i)]$ as a function of $\underline{\xi}(i)$. When $L[\underline{\xi}(i),\underline{Z}(i)]$ is differentiable with respect to $\underline{\xi}(i)$ and the maximum actually lies within the admissable range of parameter values, this maximum can be defined by the solution of the likelihood equation:

$$\left.\frac{\partial L[\underline{\xi}(i),\underline{Z}(i)]}{\partial \underline{\xi}(i)}\right|_{\underline{\xi}(i)\to\underline{\xi}^*(i)} = \underline{0}^T \qquad (3.2.1)$$

When such a solution $\underline{\xi}^*(i)$ exists, it is the maximum likelihood estimate of $\underline{\xi}(i)$.

The likelihood functions for which results have been obtained appear in Table 3.2.1. The functions in the left column utilize all data back to the initial time, whereas those on the right use only the most recent N samples of data. Thus, the functions on the left also serve as starting procedures for the corresponding functions on the right, before N samples have been accumulated.

Of these possible likelihood functions, the one which incorporates the most available information relevant to state and parameter estimation is the most beneficial. Furthermore, a form of solution that can eventually lend itself to on-line computations is desired: being not too complex and admitting prospective simplifications, while allowing adequate accuracy in the estimation.

Consider the first entry in Table 3.2.1. From the Bayesian viewpoint, it is the most logical function to use: it is the log of the conditional density of the variables to be estimated, conditioned on the values of the observed data. Notice that it is the only function of the table in which the parameter vector $\underline{a}$ appears to the left of the conditioning sign. As such, it is the only likelihood function which requires apriori specification of a probability density for the parameters. In many physically motivated problems, such a density would be difficult, if not impossible, to assess. Moreover, only a restricted class of densities, as uniform or normal, could be handled conveniently in the framework of the current mathematical models. In the case of a uniform density between two boundary values, it will be shown that the same results can be achieved more readily by using the fourth entry in the table and limiting the estimates to the specified range. Thus, a problem formulation that does not impose the need for a somewhat arbitrary apriori parameter density may well be preferable.

Table 3.2.1

Likelihood Functions

1) $\ell n\ f(\underline{x}(i),\underline{a}\|\underline{Z}(i))$	1) $\ell n\ f(\underline{x}(i),\underline{a}\|\underline{Z}_N(i))$
2) $\ell n\ f(\underline{Z}(i)\|\underline{x}(i),\underline{a})$	2a) $\ell n\ f(\underline{Z}_N(i)\|\underline{x}(i),\underline{a})$
	2b) $\ell n\ f(\underline{Z}_N(i)\|\underline{Z}(i-N),\underline{x}(i),\underline{a})$
3) $\ell n\ f(\underline{x}(i)\|\underline{Z}(i),\underline{a})$	3) $\ell n\ f(\underline{x}(i)\|\underline{Z}_N(i),\underline{a})$
4) $\ell n\ f(\underline{x}(i),\underline{Z}(i)\|\underline{a})$	4a) $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)\|\underline{a})$
	4b) $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)\|\underline{Z}(i-N),\underline{a})$

where

$\underline{x}(i)$ = state vector at time instant i

$\underline{a}$ = parameter vector

$\underline{Z}(i)$ = measurement history $\underline{z}(1)$, $\underline{z}(2)$, ... $\underline{z}(i)$

$\underline{Z}_N(i)$ = most recent N measurements $\underline{z}(i-N+1)$, $\underline{z}(i-N+2)$, ... $\underline{z}(i)$

$\underline{Z}(i-N)$ = measurement history $\underline{z}(1)$, $\underline{z}(2)$, ... $\underline{z}(i-N)$

When the densities are instead conditioned upon a given value of the parameters, their Gaussian characterization is readily maintained. Conceptually, when the derivatives of such log-likelihood functions with respect to $\underline{a}$ are set equal to zero, one is saying, "Set up an entire range of densities, each conditioned on a different value of the parameter vector, and then choose the density whose magnitude evaluated at the point, $\underline{Z}$ = the actually observed measurement values, is the greatest of all."

Now consider the second likelihood function. For a specified value of $\underline{a}$, this is the "classical" likelihood function for state estimation. The density $f(\underline{Z}(i)|\underline{x}(i),\underline{a})$ is the conditional density of the known data, conditioned upon the variables to be estimated: those values of $\underline{x}(i)$ and $\underline{a}$ that maximize the probability of the events that did in fact occur (the $\underline{Z}(i)$ or $\underline{Z}_N(i)$ histories) thus define a logical choice for a means of estimation. However, through use of Bayes' Rule (using the left column of Table 3.2.1 for sake of argument)

$$f(\underline{Z}(i)|\underline{x}(i),\underline{a})\ f(\underline{x}(i)|\underline{a}) = f(\underline{x}(i),\underline{Z}(i)|\underline{a}) \qquad (3.2.2)$$

In other words, the information contained in the density $f(\underline{Z}(i)|\underline{x}(i),\underline{a})$ is also contained in $f(\underline{x}(i),\underline{Z}(i)|\underline{a})$, but the latter's dependence upon the propagation of the apriori state statistics, given by $f(\underline{x}(i)|\underline{a})$, has been removed from $f(\underline{Z}(i)|\underline{x}(i),\underline{a})$. Usually such apriori state information is available and is a valuable asset, and thus it would be desirable to incorporate it into the filter formulation. These remarks also pertain to function number (2b) in the right column of Table 3.2.1, for which the apparently odd use of the measurement history in the density will be discussed in conjunction with number (4b).

As shown in section 2.1, the third likelihood function in the table can be used, for a specified value of $\underline{a}$, to

generate the Kalman filter through maximum likelihood techniques. The difference between this function and the last entry in the table can be displayed by means of Bayes' Rule:

$$f(\underline{x}(i)|\underline{Z}(i),\underline{a})\ f(\underline{Z}(i)|\underline{a}) = f(\underline{x}(i),\underline{Z}(i)|\underline{a}) \qquad (3.2.3)$$

It can be shown that, under the assumptions made on the system, equation (3.2.3) is the product of Gaussian densities, expressed as

$$f(\underline{x}(i)|\underline{Z}(i),\underline{a}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\underline{P}(i)|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}[\underline{x}(i)-\hat{\underline{x}}(i)]^{T}\,\underline{P}^{-1}(i)[\underline{x}(i)-\hat{\underline{x}}(i)]\right\} \qquad (3.2.4)$$

$$f(\underline{Z}(i)|\underline{a}) = \prod_{j=1}^{i} \frac{1}{(2\pi)^{\frac{m}{2}}|\underline{A}(j)|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]^{T}\,\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]\right\} \qquad (3.2.5)$$

Thus, both $f(\underline{x}(i)|\underline{Z}(i),\underline{a})$ and $f(\underline{x}(i),\underline{Z}(i)|\underline{a})$ are convenient to work with mathematically. The difference between them, $f(\underline{Z}(i)|\underline{a})$ is noteworthy in that, if it were desired to estimate only the parameters and not the state, and if the apriori statistics on $\underline{a}$ were not complete, then $f(\underline{Z}(i)|\underline{a})$ would be the natural density to use for deriving the likelihood function. Because $f(\underline{Z}(i)|\underline{a})$ is not an explicit function of $\underline{x}(i)$, it does not affect the state estimate: the derivative with respect to $\underline{x}$ of $\ell n\, f(\underline{x}(i)|\underline{Z}(i),\underline{a})$ and $\ell n\, f(\underline{x}(i),\underline{Z}(i)|\underline{a})$ yield the same result. But there is a difference in the parameter estimates, and since the distinction is caused by exactly those terms which naturally produce estimates of the parameters alone, one would suspect that valuable information is contained in the extra terms (i.e., yielding a higher

sensitivity to differences in parameter values). This will, in fact, be substantiated by the performance analysis provided by ambiguity functions in section 4.3. Furthermore, the additional information is in a particularly useful form: the likelihood equations will involve terms of $[\underline{z}(j)-\underline{H}(j)\overline{\underline{x}}(j)]$, which, for a particular realization of the $\underline{z}(j)$ history, are just the measurement residuals that can easily be evaluated on-line by feedback means. Thus, $\ell n\ f(\underline{x}(i),\underline{Z}(i)|\ \underline{a})$ would generally be preferable to $\ell n\ f(\underline{x}(i)|\ \underline{Z}(i),\underline{a})$ for the filter derivations.

Finally, the distinction between (4a) and (4b) in the table will be determined. Conceptually, (4a) will treat the estimation problem as though, at any time i greater than N, only the most recent N measurements have actually been taken: it will "forget" the information provided by the previous measurements. That is, the statistical quantities $\hat{\underline{x}}(i-N)$ and $\underline{P}(i-N)$ needed to initialize the N-step propagations are calculated from the values of $\hat{\underline{x}}_o$ and $\underline{P}_o$ as though no measurements had been taken up through time (i-N). In contrast, (4b) is conditioned upon the realization of $\underline{Z}(i-N)$, and therefore $\hat{\underline{x}}(i-N)$ and $\underline{P}(i-N)$ depend upon the measurements $\underline{z}(1)$, ... $\underline{z}(i-N)$ as well as the dynamic model and initial conditions $\hat{\underline{x}}_o$ and $\underline{P}_o$. It is evident that this latter form will be less subject to large propagated errors in $\hat{\underline{x}}(i-N)$ and $\underline{P}(i-N)$ (and thereby in the final estimates) due to erroneous values of $\hat{\underline{x}}_o$ or $\underline{P}_o$.

This section will derive the likelihood equations resulting from the likelihood functions in Table 3.2.1. For the reasons cited, emphasis will be concentrated upon the developments using the fourth entry in the table, and then the results for the other functions will be presented with less intermediate detail.

3.2.1 Likelihood Functions of the Form $\ell n\ f(\underline{x},\underline{Z}|\underline{a})$

One convenient choice of likelihood function form would be $\ell n\ f(\underline{x}(i),\underline{Z}(i)|\underline{a})$ and its derivative N-measurement forms, $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{a})$ and $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})$. As will be shown, these functions will most readily treat the case in which apriori information is available for the state, but not for the parameters. Moreover, section (3.2.2) will demonstrate the desirability of these forms even when some apriori parameter information can be incorporated.

To begin, investigate $\ell n\ f(\underline{x}(i),\underline{Z}(i)|\underline{a})$. The resulting likelihood equations will utilize all measurement data from the initial time to the current instant, $\underline{Z}(i)$, and will thus provide both a growing-length-memory estimator and the start-up procedure for the fixed-length (N-measurement) types. Bayes' Rule can be applied repeatedly to the conditional density $f(\underline{x}(i),\underline{Z}(i)|\underline{a})$ to obtain

$$\begin{aligned} f(\underline{x}(i),\underline{Z}(i)|\underline{a}) &= f(\underline{x}(i)|\underline{Z}(i),\underline{a})\ f(\underline{Z}(i)|\underline{a}) \\ &= f(\underline{x}(i)|\underline{Z}(i),\underline{a})\ f(\underline{z}(i)|\underline{Z}(i-1),\underline{a})\ f(\underline{Z}(i-1)|\underline{a}) \\ &\vdots \\ &= f(\underline{x}(i)|\underline{Z}(i),\underline{a}) \prod_{j=1}^{i} f(\underline{z}(j)|\underline{Z}(j-1),\underline{a}) \end{aligned} \tag{3.2.6}$$

where, since $\underline{z}(1)$ is the first measurement, the term in the product for j=1 is $f(\underline{z}(1)|\underline{a})$. Using the notation developed in section 2.1, each of the separate densities in equation (3.2.6) can be written out explicitly:

$$\begin{aligned} f(\underline{x}(i)|\underline{Z}(i),\underline{a}) &= \frac{1}{(2\pi)^{\frac{n}{2}}|\underline{P}(i)|^{\frac{1}{2}}}\ \exp\ \{\cdot\} \\ \{\cdot\} &= \left\{-\tfrac{1}{2}[\underline{x}(i)-\hat{\underline{x}}(i)]^T\ \underline{P}^{-1}(i)[\underline{x}(i)-\hat{\underline{x}}(i)]\right\} \end{aligned} \tag{3.2.7}$$

where $\hat{\underline{x}}(i)$ and $\underline{P}(i)$ are implicitly dependent upon the parameter values $\underline{a}$, and

$$f(\underline{z}(j)|\underline{Z}(j-1),\underline{a}) = \frac{1}{(2\pi)^{\frac{m}{2}}|\underline{A}(j)|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]^{T}\,\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]\right\} \tag{3.2.8}$$

where

$$\underline{A}(j) = \underline{H}(j)\underline{M}(j)\underline{H}^{T}(j) + \underline{R}(j) \tag{3.2.9}$$

and again $\bar{\underline{x}}(j)$, $\underline{M}(j)$, and $\underline{A}(j)$ are implicitly functions of $\underline{a}$. By substituting these expressions into equation (3.2.6), the log-likelihood function can be written as

$$\begin{aligned}
\ell n\, f(\underline{x}(i),\underline{Z}(i)|\,\underline{a}) &= \ell n\, f(\underline{x}(i)|\,\underline{Z}(i),\underline{a}) + \sum_{j=1}^{i} \ell n\, f(\underline{z}(j)|\,\underline{Z}(j-1),\underline{a}) \\
&= -\frac{n+im}{2}\,\ell n\,(2\pi) - \tfrac{1}{2}\ell n\{|\underline{P}(i)|\} - \tfrac{1}{2}\sum_{j=1}^{i}\ell n\,\{|\underline{A}(j)|\} \\
&\quad - \tfrac{1}{2}\,[\underline{x}(i) - \hat{\underline{x}}(i)]^{T}\,\underline{P}^{-1}(i)\,[\underline{x}(i) - \hat{\underline{x}}(i)] \\
&\quad - \tfrac{1}{2}\sum_{j=1}^{i}\,[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]^{T}\,\underline{A}^{-1}(j)[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]
\end{aligned} \tag{3.2.10}$$

Note that the alternative expression for $\ell n\, f(\underline{x}(i)|\,\underline{Z}(i),\underline{a})$, equation (2.1.25), could have been used to obtain an equivalent relation, but for the purposes of this derivation, the likelihood function is most conveniently expressed as in (3.2.10).

The desired likelihood equations result from setting the derivatives of $\ell n\, f(\underline{x}(i),\underline{Z}(i)|\,\underline{a})$ with respect to $\underline{x}(i)$ and $\underline{a}$ to zero simultaneously:

$$\left.\frac{\partial \ell n\, f(\underline{x}(i),\underline{Z}(i)|\,\underline{a})}{\partial \underline{x}(i)}\right|_{\substack{\underline{x}(i)\to\underline{x}^*(i)\\ \underline{a}\to\underline{a}^*(i)}} = \underline{0}^T \qquad (3.2.11a)$$

$$\left.\frac{\partial \ell n\, f(\underline{x}(i),\underline{Z}(i)|\,\underline{a})}{\partial \underline{a}}\right|_{\substack{\underline{x}(i)\to\underline{x}^*(i)\\ \underline{a}\to\underline{a}^*(i)}} = \underline{0}^T \qquad (3.2.11b)$$

Using equation (3.2.10), the first of these becomes

$$\left.-[\underline{x}(i)-\hat{\underline{x}}(i)]^T\, \underline{P}^{-1}(i)\right|_{\substack{\underline{x}(i)\to\underline{x}^*(i)\\ \underline{a}\to\underline{a}^*(i)}} = \underline{0}^T \qquad (3.2.12)$$

for which the solution is

$$\underline{x}^*(i) = \hat{\underline{x}}(i)\Big|_{\underline{a}\to\underline{a}^*(i)} \qquad (3.2.13)$$

In other words, the maximum likelihood estimate of the system state $\underline{x}(i)$, given the measurements $\underline{Z}(i)$, is obtained by using the maximum likelihood state estimator derived in section 2.1, with the estimate $\underline{a}^*(i)$ replacing the value of $\underline{a}$ for the required propagations.

Now consider the partial derivative in (3.2.11b). Since $\underline{a}^*(i)$ will comprise uncertain parameters in $\underline{\Phi}$ or $\underline{B}$ of the system model, $\underline{a}$ will affect $\overline{\underline{x}}$, $\hat{\underline{x}}$, $\underline{M}$, and $\underline{P}$; it explicitly will not affect $\underline{x}$, $\underline{z}$, $\underline{H}$, $\underline{R}$, or $\underline{Q}$ in the expansion of equation (3.2.10). From this expansion, it can be seen that the typical forms to appear will be the partial

derivative of the log of a determinant and also the derivative of a quadratic form involving an inverse matrix. For any general matrix $\underline{X}$ which is a function of a parameter, a_ℓ, these forms can be expressed as

$$\frac{\partial \ell n |\underline{X}|}{\partial a_\ell} = \frac{\partial \ell n |\underline{X}|}{\partial |\underline{X}|} \frac{\partial |\underline{X}|}{\partial a_\ell}$$

$$= \frac{1}{|\underline{X}|} \frac{\partial |\underline{X}|}{\partial a_\ell}$$

$$= \text{tr} \left\{ \underline{X}^{-1} \frac{\partial \underline{X}}{\partial a_\ell} \right\} \qquad (3.2.14)$$

$$\frac{\partial \underline{X}^{-1}}{\partial a_\ell} = - \underline{X}^{-1} \frac{\partial \underline{X}}{\partial a_\ell} \underline{X}^{-1} \qquad (3.2.15)$$

Using these relations, equation (3.2.10) can be used to determine the desired partial derivative as:

$$\begin{aligned} -2 \frac{\partial \ell n\, f(\underline{x}(i), \underline{Z}(i) | \underline{a})}{\partial a_\ell} &= \text{tr} \left\{ \underline{P}^{-1}(i) \frac{\partial \underline{P}(i)}{\partial a_\ell} \right\} \\ &-2 \frac{\partial \hat{\underline{x}}^T(i)}{\partial a_\ell} \underline{P}^{-1}(i) [\underline{x}(i) - \hat{\underline{x}}(i)] \\ &-[\underline{x}(i) - \hat{\underline{x}}(i)]^T \underline{P}^{-1}(i) \frac{\partial \underline{P}(i)}{\partial a_\ell} \underline{P}^{-1}(i) [\underline{x}(i) - \hat{\underline{x}}(i)] \\ &+ \sum_{j=1}^{i} \text{tr} \left\{ \underline{A}^{-1}(j) \frac{\partial \underline{A}(j)}{\partial a_\ell} \right\} \\ &-2 \sum_{j=1}^{i} \frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell} \underline{H}^T(j) \underline{A}^{-1}(j) [\underline{z}(j) - \underline{H}(j) \bar{\underline{x}}(j)] \\ &- \sum_{j=1}^{i} [\underline{z}(j) - \underline{H}(j) \bar{\underline{x}}(j)]^T \underline{A}^{-1}(j) \frac{\partial \underline{A}(j)}{\partial a_\ell} \underline{A}^{-1}(j) [\underline{z}(j) - \underline{H}(j) \bar{\underline{x}}(j)] \end{aligned} \qquad (3.2.16)$$

From equation (3.2.13), $\underline{x}(i)$ is simultaneously being set to $\hat{\underline{x}}(i)\big|_{\underline{a}\to\underline{a}^*(i)}$, so this knowledge can be incorporated into equation (3.2.16) by setting the terms $[\underline{x}(i)-\hat{\underline{x}}(i)]$ to zero. Noting that for general vectors $\underline{f}$ and $\underline{g}$,

$$\underline{f}^T\underline{g} = \text{tr}\{\underline{f}\underline{g}^T\} = \text{tr}\{\underline{g}\underline{f}^T\} \tag{3.2.17}$$

the likelihood equation for each parameter, a_ℓ, can be written as

$$\text{tr}\left\{\underline{P}^{-1}(i)\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\} - 2\sum_{j=1}^{i}\frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell}\underline{H}^T(j)\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]$$
$$+\sum_{j=1}^{i}\text{tr}\left\{\left[\underline{A}^{-1}(j)-\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]\,[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]^T\underline{A}^{-1}(j)\right]\cdot\right.$$
$$\left.\cdot\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}\Bigg|_{\underline{a}\to\underline{a}^*(i)} = 0 \tag{3.2.18}$$

Unfortunately, there is no general closed form solution to this equation, and so an iterative procedure will have to be employed to determine $\underline{a}^*(i)$. This will be described in section 3.3.

For the particular case of no driving noise in the dynamics, or $\underline{Q} = \underline{0}$, there is another means of manipulating the log-likelihood function that is advantageous. Instead of using equation (3.2.6), Bayes' Rule can generate

$$f(\underline{x}(i),\underline{Z}(i)\big|\underline{a}) = f(\underline{x}(i)\big|\underline{a})\; f(\underline{Z}(i)\big|\underline{x}(i),\underline{a})$$
$$= f(\underline{x}(i)\big|\underline{a})\;\prod_{j=1}^{i} f(\underline{z}(j)\big|\underline{Z}(j-1),\underline{x}(i),\underline{a}) \tag{3.2.19}$$

In the case of $\underline{Q} = \underline{0}$, these terms can be evaluated separately as:

$$f(\underline{x}(i)|\underline{a}) = \frac{1}{(2\pi)^{\frac{n}{2}}\left|\underline{\Phi}(i,0)\underline{P}_o\underline{\Phi}^T(i,0)\right|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}\left[\underline{x}(i)-\underline{\Phi}(i,0)\hat{\underline{x}}_o - \sum_{j=1}^{i}\underline{\Phi}(i,j)\underline{B}(j-1)\underline{u}(j-1)\right]^T\cdot\right.$$

$$\cdot\left[\underline{\Phi}(i,0)\underline{P}_o\underline{\Phi}^T(i,0)\right]^{-1}\cdot$$

$$\left.\cdot\left[\underline{x}(i)-\underline{\Phi}(i,0)\hat{\underline{x}}_o - \sum_{j=1}^{i}\underline{\Phi}(i,j)\underline{B}(j-1)\underline{u}(j-1)\right]\right\} \qquad (3.2.20)$$

and since knowledge of $\underline{x}(i)$ completely specifies the density for $\underline{z}(j)$ for the case of $\underline{Q} = \underline{0}$, which is to say that conditioning on $\underline{Z}(j-1)$ as well is superfluous,

$$f(\underline{z}(j)|\underline{Z}(j-1),\underline{x}(i),\underline{a}) = f(\underline{z}(j)|\underline{x}(i),\underline{a})$$

$$= \frac{1}{(2\pi)^{\frac{m}{2}}\left|\underline{R}(j)\right|^{\frac{1}{2}}} \exp\left\{\cdot\right\}$$

$$\left\{\cdot\right\} = \left\{-\tfrac{1}{2}\left[\underline{z}(j)-\underline{H}(j)\underline{\Phi}(j,i)\underline{x}(i)+\underline{H}(j)\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right]^T\cdot\right.$$

$$\left.\cdot\,\underline{R}^{-1}(j)\left[\underline{z}(j)-\underline{H}(j)\underline{\Phi}(j,i)\underline{x}(i)+\underline{H}(j)\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right]\right.$$

(3.2.21)

Substituting these into (3.2.19) allows evaluation of the likelihood equations, the first of which, equation (3.2.11a), yields a solution of

$$\underline{x}^*(i) = \hat{\underline{x}}(i) \;\Big|\; \begin{array}{l}\underline{Q} = \underline{0} \\ \underline{a} \rightarrow \underline{a}^*(i)\end{array}$$

$$= \left[\underline{\Phi}^T(0,i)\underline{P}_\circ^{-1}\underline{\Phi}(0,i) + \underline{\mathcal{Q}}(i,1)\right]^{-1} \cdot$$

$$\cdot\left[\sum_{j=1}^{i}\underline{\Phi}^T(j,i)\underline{H}^T(j)\underline{R}^{-1}(j)\left\{\underline{z}(j)+\underline{H}(j)\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\right.$$

$$\left. + \underline{\Phi}^T(0,i)\underline{P}_\circ^{-1}\left\{\hat{\underline{x}}_\circ + \sum_{k=1}^{i}\underline{\Phi}(0,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\right]\Bigg|_{\underline{a}\to\underline{a}^*(i)}$$

(3.2.22)

which can be recognized as the form of equation (2.1.77) with $\underline{a}^*(i)$ replacing $\underline{a}$ in the development. After considerable algebra, the likelihood equation for one component of the parameter vector, a_ℓ, can be expressed as

$$\mathrm{tr}\left\{\left[\underline{\Phi}^T(0,i) - \underline{P}_*^{-1}(i)[\hat{\underline{x}}(i) - \underline{x}_*(i)] \cdot \right.\right.$$

$$\left.\left. \cdot\left[\hat{\underline{x}}(i) - \sum_{k=1}^{i}\underline{\Phi}(i,k)\underline{B}(k-1)\,\underline{u}(k-1)\right]^T\underline{\Phi}^T(0,i)\right]\frac{\partial\underline{\Phi}^T(i,0)}{\partial a_\ell}\right\}$$

$$-\sum_{j=1}^{i}\left[\hat{\underline{x}}^T(i)\,\frac{\partial\underline{\Phi}^T(j,i)}{\partial a_\ell} + \sum_{k=j+1}^{i}\underline{u}^T(k-1)\,\frac{\partial[\underline{B}^T(k-1)\,\underline{\Phi}^T(j,k)]}{\partial a_\ell}\right]\cdot$$

$$\cdot\,\underline{H}^T(j)\underline{R}^{-1}(j)\,[\underline{z}(j) - \hat{\underline{z}}_*(j)]$$

$$-\sum_{k=1}^{i}\underline{u}^T(k-1)\,\frac{\partial[\underline{B}^T(k-1)\,\underline{\Phi}^T(i,k)]}{\partial a_\ell}\,\underline{P}_*^{-1}(i)[\hat{\underline{x}}(i) - \underline{x}_*(i)]\Bigg|_{\underline{a}\to\underline{a}^*(i)} = \underline{0}$$

(3.2.23)

where

$$\underline{x}_*(i) = \underline{\Phi}(i,0)\hat{\underline{x}}_\circ + \sum_{k=1}^{i}\underline{\Phi}(i,k)\underline{B}(k-1)\underline{u}(k-1)$$

$$\hat{\underline{z}}_*(j) = \underline{H}(j)\underline{\Phi}(j,i)\hat{\underline{x}}(i) - \underline{H}(j)\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)$$

$$\underline{P}_*^{-1}(i) = \underline{\Phi}^T(0,i)\underline{P}_\circ^{-1}\underline{\Phi}(0,i)$$

(3.2.24)

(Setting all $\underline{u}$'s to zero yields the results for the case of no deterministic inputs.)

The preceding developments pertain to a "growing-length memory": the estimator "remembers" the effects of all data back to the initial time, as reflected by the ever-increasing number of terms in the summations. As discussed in the beginning of this chapter, there may be numerous advantages in restricting these sums to the effect of only the most recent N data measurements. For this reason, consider the log-likelihood function $\ell n\, f(\underline{x}(i), \underline{Z}_N(i) | \underline{a})$, where $\underline{Z}_N(i)$ is defined by:

$$\underline{Z}_N(i) = \underline{z}(i-N+1),\ \underline{z}(i-N+2),\ \ldots\ \underline{z}(i-1),\ \underline{z}(i) \tag{3.2.25}$$

and is thus the most recent N measurements at the time instant i. From Bayes' Rule (for $i \geq N$),

$$\begin{aligned} f(\underline{x}(i), \underline{Z}_N(i) | \underline{a}) &= f(\underline{x}(i) | \underline{Z}_N(i), \underline{a})\ f(\underline{Z}_N(i) | \underline{a}) \\ &= f(\underline{x}(i) | \underline{Z}_N(i), \underline{a}) \prod_{j=i-N+1}^{i} f(\underline{z}(j) | \underline{Z}_{j-[i-N+1]}(j-1), \underline{a}) \end{aligned} \tag{3.2.26}$$

where $\underline{Z}_{j-[i-N+1]}(j-1)$ is the most recent $\{j-[i-N+1]\}$ measurements at time $(j-1)$, $\underline{Z}_0(k)$ being defined as no measurements for any k. The first term in (3.2.26) is the probability density of $\underline{x}(i)$ conditioned on the value of $\underline{a}$ and the measurement sequence $\underline{z}(i-N+1), \ldots \underline{z}(i)$. Now, the expected value of $\underline{x}(i-N)$, given that no measurements are taken, is simply $f(\underline{x}(i-N) | \underline{a})$, which is a Gaussian density characterized by a mean

$$\tilde{\underline{x}}_o(i-N) = \underline{\Phi}(i-N, 0)\hat{\underline{x}}_o + \sum_{j=1}^{i-N} \underline{\Phi}(i-N, j)\underline{B}(j-1)\underline{u}(j-1) \tag{3.2.27}$$

and covariance

$$\tilde{\underline{P}}_{\circ}(i-N) = \underline{\Phi}(i-N,0)\underline{P}_{\circ}\underline{\Phi}^{T}(i-N,0) + \sum_{j=1}^{i-N} \underline{\Phi}(i-N,j)\underline{G}(j-1)\underline{Q}(j-1)\underline{G}^{T}(j-1)\underline{\Phi}^{T}(i-N,j) \tag{3.2.28}$$

Subsequently, measurements are taken for N steps of the propagation. This can be treated simply as an N-step dynamic process with measurements, starting from the "initial conditions" given by (3.2.27) and (3.2.28). Using the same arguments about a Markov process emanating from an apriori Gaussian density as used previously, it can be concluded that defining the propagation

$$\tilde{\underline{x}}^{-}(j+1) = \underline{\Phi}(j+1,j)\tilde{\underline{x}}^{+}(j) + \underline{B}(j)\underline{u}(j) \tag{3.2.29}$$

$$\tilde{\underline{x}}^{+}(j) = \tilde{\underline{x}}^{-}(j) + \tilde{\underline{K}}(j)[\underline{z}(j) - \underline{H}(j)\tilde{\underline{x}}^{-}(j)] \tag{3.2.30}$$

where

$$\tilde{\underline{K}}(j) = \tilde{\underline{M}}(j)\underline{H}^{T}(j)[\underline{H}(j)\tilde{\underline{M}}(j)\underline{H}^{T}(j) + \underline{R}(j)]^{-1} \tag{3.2.31}$$

$$\tilde{\underline{M}}(j+1) = \underline{\Phi}(j+1,j)\tilde{\underline{P}}(j)\underline{\Phi}^{T}(j+1,j) + \underline{G}(j)\underline{Q}(j)\underline{G}^{T}(j) \tag{3.2.32}$$

$$\tilde{\underline{P}}(j) = [\underline{I} - \tilde{\underline{K}}(j)\underline{H}(j)]\tilde{\underline{M}}(j)[\underline{I} - \tilde{\underline{K}}(j)\underline{H}(j)]^{T} + \tilde{\underline{K}}(j)\underline{R}(j)\tilde{\underline{K}}^{T}(j) \tag{3.2.33}$$

for j = (i-N+1), (i-N+2), ... (i), from the "initial conditions" specified by (3.2.27) and (3.2.28), determines the values of $\tilde{\underline{x}}^{+}(i)$ and $\tilde{\underline{P}}(i)$. These, in turn, completely define the Gaussian density:

$$f(\underline{x}(i)\,|\,\underline{Z}_N(i),\underline{a}) = N[\tilde{\underline{x}}^+(i),\ \tilde{\underline{P}}(i)] \tag{3.2.34}$$

The notation $N[\underline{\mu}, \underline{\Omega}]$ denotes a normal density with mean $\underline{\mu}$ and covariance $\underline{\Omega}$. Similarly, the second term in equation (3.2.26) can be written as

$$f(\underline{Z}_N(i)\,|\,\underline{a}) = \prod_{j=i-N+1}^{i} N[\underline{H}(j)\tilde{\underline{x}}^-(j),\ \underline{H}(j)\tilde{\underline{M}}(j)\underline{H}^T(j) + \underline{R}(j)] \tag{3.2.35}$$

Notice that the same expressions can be obtained formally by letting $\underline{R}(j) = \infty\underline{I}$ for all $j \leq (i-N)$ in the equations previously obtained for the log-likelihood function $\ell n\ f(\underline{x}(i),\underline{Z}(i)|\underline{a})$. Equations (3.2.34) and (3.2.35) are substituted into (3.2.26) to define the log-likelihood function $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\,\underline{a})$, the derivatives with respect to $\underline{x}(i)$ and $\underline{a}$ set equal to zero, and a development paralleling the previous derivation yields

$$\underline{x}^*(i) = \tilde{\underline{x}}^+(i)\Big|_{\underline{a}\to\underline{a}^*(i)} \tag{3.2.36}$$

That is, the maximum likelihood estimate of the state is given by equations (3.2.29) to (3.2.33), with the estimated value $\underline{a}^*(i)$ used for the propagation. The components of $\underline{a}^*(i)$ are the solutions to the likelihood equations

$$\begin{aligned} &\mathrm{tr}\left\{\tilde{\underline{P}}^{-1}(i)\frac{\partial\tilde{\underline{P}}(i)}{\partial a_\ell}\right\} - 2\sum_{j=i-N+1}^{i}\frac{\partial\tilde{\underline{x}}^{-T}(j)}{\partial a_\ell}\underline{H}^T(j)\tilde{\underline{A}}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\tilde{\underline{x}}^-(j)] \\ &+\sum_{j=i-N+1}^{i}\mathrm{tr}\left\{\left[\tilde{\underline{A}}^{-1}(j)-\tilde{\underline{A}}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\tilde{\underline{x}}^-(j)]\cdot\right.\right. \\ &\left.\left.\cdot\,[\underline{z}(j)-\underline{H}(j)\tilde{\underline{x}}^-(j)]^T\tilde{\underline{A}}^{-1}(j)\right]\frac{\partial\tilde{\underline{A}}(j)}{\partial a_\ell}\right\}\Bigg|_{\underline{a}\to\underline{a}^*(i)} = 0 \end{aligned} \tag{3.2.37}$$

where

$$\tilde{\underline{A}}(j) = \underline{H}(j)\tilde{\underline{M}}(j)\underline{H}^{T}(j) + \underline{R}(j) \qquad (3.2.38)$$

Again, an iterative solution is required, but here the number of terms in the likelihood equation remains fixed.

Although this result is conceptually straightforward and can be defined by an N-step propagation each time the system progresses from one sample time to the next, the amount of computation required is prohibitive for on-line applications. Nor does the form of the state estimate readily yield to approximating recursions suggested by the Fisher Information Matrix propagation: i.e., replace the N-step state and covariance propagations with an approximate one-step propagation of the previous evaluations, plus a term due to the information from $\underline{z}(i)$, minus a corresponding term due to the data from $\underline{z}(i-N)$. (Appendix C presents a further development of these concepts.) Moreover, the rejection of measurement data previous to $\underline{z}(i-N)$ often has serious drawbacks for state estimation, such as the degradation of $\tilde{\underline{x}}(i-N)$ and $\tilde{\underline{P}}(i-N)$ due to incorrect values for $\hat{\underline{x}}_{\circ}$ and $\underline{P}_{\circ}$ as mentioned previously. There are certain applications in which a fixed-length memory state estimator is in fact advantageous (for instance, the orbit determination problem, as in Jazwinski (1970)), but very often the N-step validity of the mathematical model pertains more critically to the assumption that the uncertain parameters are essentially constant over this interval. Thus an approximate likelihood function is sought to yield a parameter estimate based upon the most recent N measurements, but a state estimate which incorporates the information afforded by apriori statistics and the entire measurement history.

Consider the conditional density $f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$. Through use of Bayes' Rule for conditional densities:

$$f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a}) = f(\underline{x}(i)|\underline{Z}_N(i),\underline{Z}(i-N),\underline{a})\ f(\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})$$

$$= f(\underline{x}(i)|\underline{Z}_N(i),\underline{Z}(i-N),\underline{a}) \cdot \frac{f(\underline{Z}_N(i),\underline{Z}(i-N)|\underline{a})}{f(\underline{Z}(i-N)|\underline{a})}$$

$$= f(\underline{x}(i)|\underline{Z}(i),\underline{a}) \cdot \frac{f(\underline{Z}(i)|\underline{a})}{f(\underline{Z}(i-N)|\underline{a})} \qquad (3.2.39)$$

It can readily be shown that each of the densities in this expression is Gaussian:

$$f(\underline{x}(i)|\underline{Z}(i),\underline{a}) = N[\hat{\underline{x}}(i),\underline{P}(i)]$$

$$f(\underline{Z}(i)|\underline{a}) = \prod_{j=1}^{i} N[\underline{H}(j)\bar{\underline{x}}(j),\ \underline{A}(j)] \qquad (3.2.40)$$

$$f(\underline{Z}(i-N)|\underline{a}) = \prod_{j=1}^{i-N} N[\underline{H}(j)\bar{\underline{x}}(j),\ \underline{A}(j)]$$

so that the log-likelihood function $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})$ can be expressed as

$$\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a}) = \ell n\ f(\underline{x}(i)|\underline{Z}(i),\underline{a})$$

$$+ \sum_{j=i-N+1}^{i} \ell n\ f(\underline{z}(j)|\underline{Z}(j-1),\underline{a})$$

$$= -\frac{n+Nm}{2}\ \ell n\ (2\pi) - \tfrac{1}{2}\ \ell n\{|\underline{P}(i)|\} - \tfrac{1}{2}\sum_{j=i-N+1}^{i} \ell n\{|\underline{A}(j)|\}$$

$$- \tfrac{1}{2}[\underline{x}(i) - \hat{\underline{x}}(i)]^T\ \underline{P}^{-1}(i)[\underline{x}(i) - \hat{\underline{x}}(i)]$$

$$- \tfrac{1}{2}\sum_{j=i-N+1}^{i} [\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]^T\ \underline{A}^{-1}(j)[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]$$

(3.2.41)

This is the same as the log-likelihood function (3.2.10) developed from the density $f(\underline{x}(i),\underline{Z}(i)|\underline{a})$, except that the lower limit on the summation terms is (i-N+1) instead of 1.

Therefore, by a parallel derivation it is concluded that the maximum likelihood estimate of the state is given by

$$\underline{x}^*(i) = \hat{\underline{x}}(i) \Big|_{\underline{a} \to \underline{a}^*(i)} \tag{3.2.42}$$

i.e., the form of state estimator derived in section 2.1, with the value of $\underline{a}$ being replaced by $\underline{a}^*(i)$ for the recursions. In this case, the components a_ℓ of the parameter vector satisfy the likelihood equation

$$\begin{aligned} &\mathrm{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\} - 2\sum_{j=i-N+1}^{i} \frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell}\,\underline{H}^T(j)\underline{A}^{-1}(j)[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)] \\ &+ \sum_{j=i-N+1}^{i} \mathrm{tr}\left\{\left[\underline{A}^{-1}(j) - \underline{A}^{-1}(j)[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]\cdot\right.\right. \\ &\qquad \left.\left.\cdot\,[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]^T\,\underline{A}^{-1}(j)\right]\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}\Bigg|_{\underline{a} \to \underline{a}^*(i)} = 0 \end{aligned} \tag{3.2.43}$$

where

$$\underline{A}(j) = \underline{H}(j)\underline{M}(j)\underline{H}^T(j) + \underline{R}(j) \tag{3.2.44}$$

For the case of a fixed-length memory filter with $\underline{Q}(j) = \underline{0}$, the simplification can be made that

$$\begin{aligned} f(\underline{z}(j) \,|\, \underline{Z}_{j-[i-N+1]}(j-1), \underline{x}(i), \underline{a}) &= f(\underline{z}(j) \,|\, \underline{Z}(j-1), \underline{x}(i), \underline{a}) \\ &= f(\underline{z}(j) \,|\, \underline{x}(i), \underline{a}) \end{aligned} \tag{3.2.45}$$

so instead of expanding $f(\underline{x}(i), \underline{Z}_N(i) \,|\, \underline{a})$ as in (3.2.26) it is more appropriate to expand as:

$$f(\underline{x}(i), \underline{Z}_N(i) \,|\, \underline{a}) = f(\underline{x}(i) \,|\, \underline{a})\; f(\underline{Z}_N(i) \,|\, \underline{x}(i), \underline{a})$$

$$= f(\underline{x}(i)|\underline{a}) \prod_{j=i-N+1}^{i} f(\underline{z}(j)|\underline{Z}_{j-[i-N+1]}(j-1),\underline{x}(i),\underline{a})$$

$$= f(\underline{x}(i)|\underline{a}) \prod_{j=i-N+1}^{i} f(\underline{z}(j)|\underline{x}(i),\underline{a}) \qquad (3.2.46)$$

Carrying through the development yields a result identical to that of equations (3.2.22) to (3.2.24), except that the summations over j have a lower limit of (i-N+1) instead of 1, and $\underline{\mathcal{J}}(i,1)$ is replaced by $\underline{\mathcal{J}}(i,i-N+1)$ as defined by equation (2.1.91).

3.2.2 Likelihood Functions of the Form $\ell n\ f(\underline{x},\underline{a}|\underline{Z})$

If apriori statistical information is available for both the state and parameters, a suitable choice of log-likelihood function is $\ell n\ f(\underline{x}(i),\underline{a}|\underline{Z}(i))$. Bayes' Rule can be employed to manipulate the corresponding density into

$$\begin{aligned} f(\underline{x}(i),\underline{a}|\underline{Z}(i)) &= f(\underline{x}(i)|\underline{Z}(i),\underline{a})\ f(\underline{a}|\underline{Z}(i)) \\ &= f(\underline{x}(i)|\underline{Z}(i),\underline{a})\ \frac{f(\underline{Z}(i)|\underline{a})\ f(\underline{a})}{f(\underline{Z}(i))} \\ &= f(\underline{x}(i),\underline{Z}(i)|\underline{a})\ \frac{f(\underline{a})}{f(\underline{Z}(i))} \end{aligned} \qquad (3.2.47)$$

The estimation problem is then solved by finding the maximum of the log-likelihood function

$$\ell n\ f(\underline{x}(i),\underline{a}|\underline{Z}(i)) = \ell n\ f(\underline{x}(i),\underline{Z}(i)|\underline{a})$$

$$+ \ell n\ f(\underline{a}) - \ell n\ f(\underline{Z}(i)) \qquad (3.2.48)$$

This can be recognized as the log-likelihood function used in the previous section, plus two additional terms. Since the last term in (3.2.48) is not an explicit function of either $\underline{x}(i)$ or $\underline{a}$, it will have no effect on the likelihood equations.

If, as is generally the case, $f(\underline{a})$ is not a function of $\underline{x}(i)$, then the likelihood equations for the state estimate are:

$$\left.\frac{\partial\, \ell n\, f(\underline{x}(i), \underline{a} | \underline{Z}(i))}{\partial \underline{x}(i)}\right|_{\substack{\underline{x}(i)\to\underline{x}^*(i) \\ \underline{a}\to\underline{a}^*(i)}} =$$

$$= \left.\frac{\partial\, \ell n\, f(\underline{x}(i), \underline{Z}(i) | \underline{a})}{\partial \underline{x}(i)}\right|_{\substack{\underline{x}(i)\to\underline{x}^*(i) \\ \underline{a}\to\underline{a}^*(i)}} = \underline{0}^T \quad (3.2.49)$$

so that the maximum likelihood state estimate is of the same structural form as in the last section:

$$\underline{x}^*(i) = \hat{\underline{x}}(i) \Big|_{\underline{a}\to\underline{a}^*(i)} \quad (3.2.50)$$

The likelihood equations to be solved for $\underline{a}^*(i)$ are altered by the addition of a term:

$$\underline{0}^T = \left.\frac{\partial\, \ell n\, f(\underline{x}(i), \underline{a} | \underline{Z}(i))}{\partial \underline{a}}\right|_{\substack{\underline{x}(i)\to\underline{x}^*(i) \\ \underline{a}\to\underline{a}^*(i)}}$$

$$= \left\{ \frac{\partial \ell n\ f(\underline{x}(i),\underline{Z}(i)|\underline{a})}{\partial \underline{a}} + \frac{\partial \ell n\ f(\underline{a})}{\partial \underline{a}} \right\} \Bigg|_{\substack{\underline{x}(i)\to\underline{x}^*(i) \\ \underline{a}\to\underline{a}^*(i)}} \tag{3.2.51}$$

Thus, the general equation for the component a_ℓ is:

$$\begin{aligned} &\mathrm{tr}\left\{\underline{P}^{-1}(i)\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\} - 2\sum_{j=1}^{i}\frac{\partial \overline{\underline{x}}^T(j)}{\partial a_\ell}\underline{H}^T(j)\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\overline{\underline{x}}(j)] \\ &+ \sum_{j=1}^{i}\mathrm{tr}\left\{\left[\underline{A}^{-1}(j) - \underline{A}^{-1}(j)[\underline{z}(j) - \underline{H}(j)\overline{\underline{x}}(j)]\cdot\right.\right. \\ &\left.\left.\cdot\,[\underline{z}(j) - \underline{H}(j)\overline{\underline{x}}(j)]^T\,\underline{A}^{-1}(j)\right]\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\} \\ &- 2\frac{\partial \ell n\ f(\underline{a})}{\partial a_\ell}\Bigg|_{\underline{a}\to\underline{a}^*(i)} = 0 \end{aligned} \tag{3.2.52}$$

When the parameters that compose $\underline{a}$ are independent, the last term of this equation becomes

$$-2\frac{\partial \ell n\ f(\underline{a})}{\partial a_\ell} \longrightarrow -2\frac{\partial \ell n\ f(a_\ell)}{\partial a_\ell} \tag{3.2.53}$$

Similar developments can be made for the case of a fixed number (N) of measurements being used to form the estimate by writing

$$f(\underline{x}(i),\underline{a}|\underline{Z}_N(i)) = f(\underline{x}(i),\underline{Z}_N(i)|\underline{a})\ \frac{f(\underline{a})}{f(\underline{Z}_N(i))} \tag{3.2.54}$$

$$f(\underline{x}(i),\underline{a}|\underline{Z}(i)) = f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})\ f(\underline{Z}(i-N)|\underline{a})\cdot \cdot\frac{f(\underline{a})}{f(\underline{Z}(i))} \tag{3.2.55}$$

to yield the same form of results as in section 3.2.1 but with the term (3.2.53) added to the likelihood equations for the parameters.

The current formulation allows the incorporation of apriori parameter information, in the form of the density $f(\underline{a})$, into the problem. However, this has been accomplished at the expense of additional calculations, as compared to the corresponding formulation without such apriori information. A tradeoff evaluation must be made, especially for on-line applications, to decide which form to use.

It will now be demonstrated that if the apriori density $f(\underline{a})$ is a uniform density between two bounds, which is to say that the designer can only predetermine a range of admissible parameter values, then it is better not to incorporate it into the likelihood equations. A more efficient procedure would be to solve the likelihood equations without the effect of $f(\underline{a})$, and then to limit the solution to lie within the acceptable boundaries. To show this, the behavior of the log-likelihood function and its derivative, both with and without the influence of $f(\underline{a})$, will be examined.

If an assumed apriori density $f(\underline{a})$ that specifies the bounds of acceptable parameter values were to be incorporated into the likelihood function, i.e., an $f(\underline{a})$ that is zero outside a certain range, as by truncation, it would disallow any parameter values outside that range from being a solution to the likelihood equations. As $f(\underline{a})$ goes abruptly to zero, $\ell n\, f(\underline{a})$ goes abruptly to negative infinity, so that at that point the value of the log-likelihood function would be $(-\infty)$, no matter what the other terms contribute (it is assumed that the other contributions will not be $+\infty$). Furthermore, if the assumed $f(\underline{a})$ were uniform in between the bounds, it would not alter the shape of the density associated with the likelihood function, as shown in figure 3.2.1. Unless the bounds of $f(\underline{a})$ are inappropriate, the peak of the density should always lie within the specified range, so the maximum

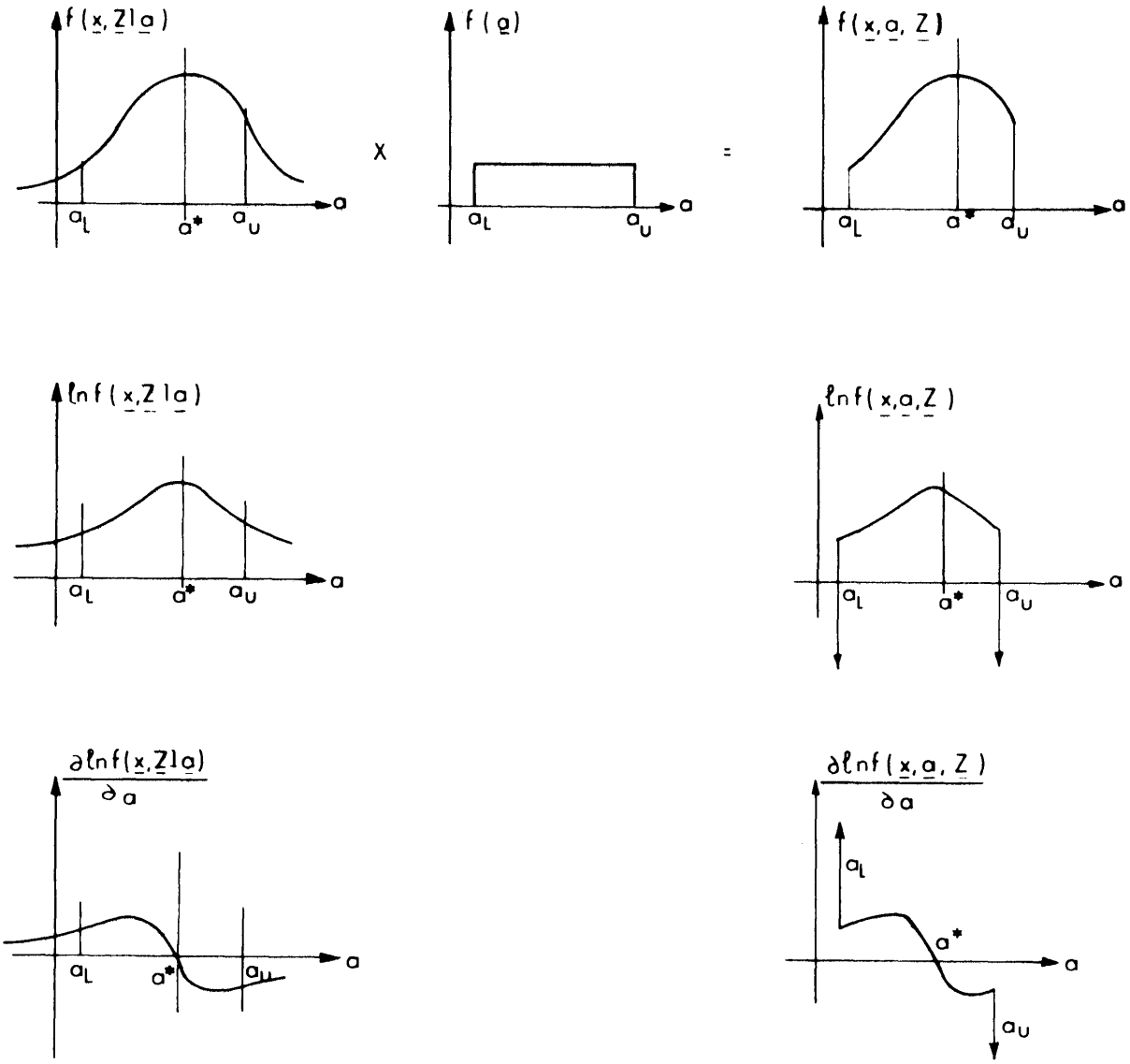


Figure 3.2.1 Incorporating f(a) into the Likelihood Function

likelihood estimate would not be affected by incorporating the uniform $f(\underline{a})$. It should be noted that other estimates, such as the conditional mean, would, in contrast, generally be affected.

Figure 3.2.1 also displays the log-likelihood function, and its derivative, corresponding to the densities $f(\underline{x}, \underline{Z}|\underline{a})$ and $f(\underline{x},\underline{a},\underline{Z})$. For the latter, the derivative must be defined as a limit from the right at a_L, and as a limit from the left at a_u. Since the limits do not exist from both sides at these points, and since the derivative is not defined outside the range of a_L to a_u, there may well be numerical difficulty in applying iterative techniques to solving the likelihood equation. This is especially true if $\underline{a}^*$ is actually close to the bounding values. The derivative that does not involve $f(\underline{a})$ effects is not subject to such problems.

Ill-conditioned problems arise when the peak of the density not incorporating $f(\underline{a})$ falls directly on or outside the boundaries. This is illustrated by figure 3.2.2. From this figure, it can be seen that use of $f(\underline{x},\underline{Z}|\underline{a})$ would readily lead to the estimate a^*, whereupon a_u would be chosen as the closest acceptable value. If $f(\underline{x},\underline{a},\underline{Z})$ had been used instead, the derivative function depicted in the figure reveals that an iterative method would have extreme difficulty in attaining an estimate.

Unless $f(\underline{a})$ is known with some certainty, it may be preferable not to enter it into the likelihood function. Boundaries of admissible regions can be treated with less required computation time and fewer numerical problems by the method described above. Furthermore, if $f(\underline{a})$ is not well known, an arbitrary selection of a reasonable density may in fact deteriorate performance if its peak differs significantly from the true value of $\underline{a}$. Since an engineer's best estimate of $\underline{a}$ would serve as the initial point for the parameter iterations using $f(\underline{x},\underline{Z}|\underline{a})$, and since this formulation can include admissible ranges, more would have to be known about $f(\underline{a})$ than its peak and bounds before its use in the likelihood

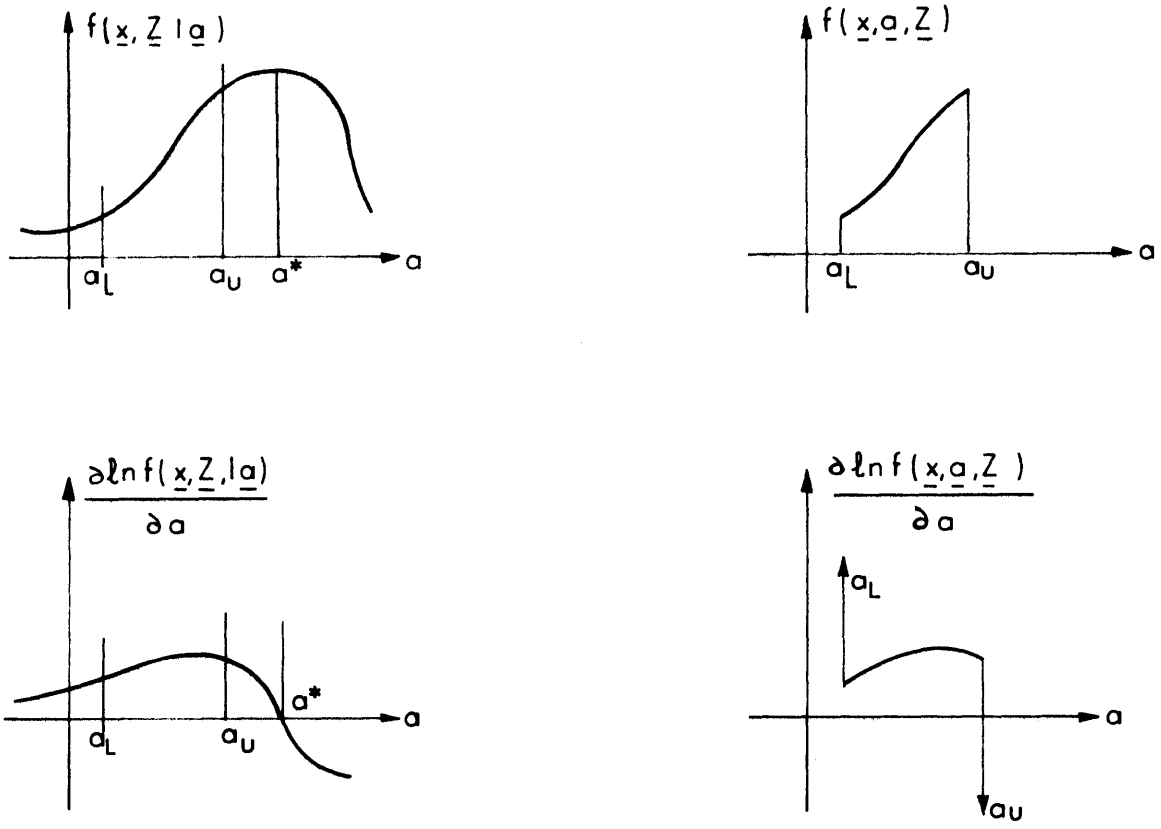


Figure 3.2.2 Poorly Conditioned Problems

equations would be warranted. Thus, although there will be some applications in which using apriori parameter information is justified, it is expected that the development of section 3.2.1 will be more useful in practice.

3.2.3 Likelihood Functions of the Form $\ell n\, f(\underline{Z}|\underline{x},\underline{a})$

Since $f(\underline{Z}(i)|\underline{x}(i),\underline{a})$ is conditioned on both $\underline{x}(i)$ and $\underline{a}$, this density is appropriate for developing a maximum likelihood estimator that does not depend on apriori statistics for either the state or parameters. It will yield practical results for the case of no dynamic noise, or $\underline{Q} = \underline{0}$, but no useful new results are obtained for those systems in which dynamic noise is present. The problem of insufficient data to define an estimate for the early sample times will prevail, and a startup procedure as described by equations (2.1.61) to (2.1.72) must be provided.

Under the assumption that $\underline{Q} = \underline{0}$, a convenient expansion of $f(\underline{Z}(i)|\underline{x}(i),\underline{a})$ is

$$f(\underline{Z}(i)|\underline{x}(i),\underline{a}) = \prod_{j=1}^{i} f(\underline{z}(j)|\underline{Z}(j-1),\underline{x}(i),\underline{a}) \qquad (3.2.56)$$

since, as explained in section 3.2.1, each of the densities in the product can be written, for this case, as

$$f(\underline{z}(j)|\underline{Z}(j-1),\underline{x}(i),\underline{a}) = f(\underline{z}(j)|\underline{x}(i),\underline{a})$$

$$= \frac{1}{(2\pi)^{\frac{m}{2}}|\underline{R}(j)|^{\frac{1}{2}}} \exp\left\{-\tfrac{1}{2}[\underline{z}(j)-\underline{z}_*(j)]^T \underline{R}^{-1}(j)[\underline{z}(j)-\underline{z}_*(j)]\right\} \qquad (3.2.57)$$

where

$$\underline{z}_*(j) = \underline{H}(j)\left\{\underline{\Phi}(j,i)\underline{x}(i) - \sum_{k=j+1}^{i} \underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\} \qquad (3.2.58)$$

Taking the derivative with respect to $\underline{x}(i)$ and setting it to zero yields, for $\underline{Q} = \underline{0}$,

$$\underline{x}^*(i) = \underline{\mathcal{Q}}^{-1}(i,1)\sum_{j=1}^{i}\underline{\Phi}^T(j,i)\underline{H}^T(j)\underline{R}^{-1}(j)\cdot$$
$$\cdot\left\{\underline{z}(j)+\underline{H}(j)\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\}\Bigg|_{\underline{a}\to\underline{a}^*(i)} \tag{3.2.59}$$

where $\underline{\mathcal{Q}}^{-1}(i,1)$, $\underline{\Phi}(j,k)$ and $\underline{B}(k-1)$ are all (potential) functions of $\underline{a}$. As expected, this is the same as equation (3.2.22) with $\underline{P}_o^{-1} = \underline{0}$, i.e., with no apriori statistical knowledge of the state. Setting the derivative with respect to $\underline{a}$ equal to zero produces the likelihood equations

$$\sum_{j=1}^{i}\left[\hat{\underline{x}}^T(i)\frac{\partial\underline{\Phi}^T(j,i)}{\partial a_\ell}+\sum_{k=j+1}^{i}\underline{u}^T(k-1)\frac{\partial[\underline{B}^T(k-1)\underline{\Phi}^T(j,k)]}{\partial a_\ell}\right]\cdot$$
$$\cdot\,\underline{H}^T(j)\underline{R}^{-1}(j)[\underline{z}(j)-\hat{\underline{z}}_*(j)]\Bigg|_{\underline{a}\to\underline{a}^*(i)} = 0 \tag{3.2.60}$$

where

$$\underline{z}_*(j) = \underline{H}(j)\left\{\underline{\Phi}(j,i)\hat{\underline{x}}(i)-\sum_{k=j+1}^{i}\underline{\Phi}(j,k)\underline{B}(k-1)\underline{u}(k-1)\right\} \tag{3.2.61}$$

As done previously, the result is expressed in a manner that can readily be simplified for the case of no control inputs merely by setting the $\underline{u}$'s or $\underline{B}$'s to zero.

The same arguments applied to the function $\ell n\, f(\underline{Z}_N(i)|\underline{x}(i),\underline{a})$ yield a "fixed-length memory" for both state and parameters, being of the same form as (3.2.59) and (3.2.60), but with the lower limit on the j summations changed to (i-N+1), and $\underline{\mathcal{Q}}(i,1)$ replaced by $\underline{\mathcal{Q}}(i,i-N+1)$. For on-line applications, recursions

for the state estimate are desirable, and appendix D shows that the relations given by (2.1.84) and (2.1.85), or (2.1.91) and (2.1.92), can be adapted to this situation. Finally, equation (3.2.60) with the lower limit on the j summation set to (i-N+1) is compatible with the unaltered (3.2.59), so that a "fixed-length memory" parameter estimate can be combined with the "growing memory" recursions for the state.

When dynamic noise does enter the system, it is no longer possible to invoke equation (3.2.57). Instead, the density $f(\underline{Z}(i)|\underline{x}(i),\underline{a})$ can be transformed by Bayes' Rule into

$$f(\underline{Z}(i)|\underline{x}(i),\underline{a}) = \frac{f(\underline{x}(i),\underline{Z}(i)|\underline{a})}{f(\underline{x}(i)|\underline{a})} \tag{3.2.62}$$

(This can also be obtained by expressing the individual product terms in (3.2.56) as

$$f(\underline{z}(j)|\underline{Z}(j-1),\underline{x}(i),\underline{a}) = \frac{f(\underline{x}(i)|\underline{Z}(j),\underline{a})\ f(\underline{z}(j)|\underline{Z}(j-1),\underline{a})}{f(\underline{x}(i)|\underline{Z}(j-1),\underline{a})} \tag{3.2.63}$$

i.e., in terms of three normal densities, and performing the multiplication.) If it is assumed that there is no apriori knowledge about the state, then (3.2.62) cannot be evaluated. Therefore, assume that $\hat{\underline{x}}_o$ and $\underline{P}_o$ are known, and then look at the results for limiting the case of $\underline{P}_o \rightarrow \infty \underline{I}$ or $\underline{P}_o^{-1} \rightarrow \underline{0}$.

Under the assumptions of known $\hat{\underline{x}}_o$ and $\underline{P}_o$,

$$\begin{aligned} f(\underline{x}(i),\underline{Z}(i)|\underline{a}) &= f(\underline{x}(i)|\underline{Z}(i),\underline{a})\ f(\underline{Z}(i)|\underline{a}) \\ &= f(\underline{x}(i)|\underline{Z}(i),\underline{a}) \prod_{j=1}^{i} f(\underline{z}(j)|\underline{Z}(j-1),\underline{a}) \\ &= N[\hat{\underline{x}}(i),\underline{P}(i)] \prod_{j=1}^{i} N[\underline{H}(j)\overline{\underline{x}}(j),\underline{A}(j)] \end{aligned} \tag{3.2.64}$$

$$f(\underline{x}(i)|\,\underline{a}) = N[\underline{x}_*(i),\ \underline{\Phi}(i,0)\underline{P}_0\underline{\Phi}^T(i,0) + \sum_{k=1}^{i} \underline{\Phi}(i,k)\underline{G}(k)\underline{Q}(k)\underline{G}^T(k)\underline{\Phi}^T(i,k)] \qquad (3.2.65)$$

where

$$\underline{x}_*(i) = \underline{\Phi}(i,0)\hat{\underline{x}}_0 + \sum_{k=1}^{i} \underline{\Phi}(i,k)\underline{B}(k-1)\underline{u}(k-1) \qquad (3.2.66)$$

Putting these equations into the log-likelihood function and differentiating with respect to $\underline{x}(i)$ yields the likelihood equations for the state,

$$\left[\underline{x}(i)-\underline{x}_*(i)\right]^T \left[\underline{\Phi}(i,0)\underline{P}_0\underline{\Phi}^T(i,0) + \sum_{k=1}^{i} \underline{\Phi}(i,k)\underline{G}(k)\underline{Q}(k)\underline{G}^T(k)\underline{\Phi}^T(i,k)\right]^{-1} - \left[\underline{x}(i)-\hat{\underline{x}}(i)\right]^T \underline{P}^{-1}(i) \Bigg|_{\substack{\underline{x}(i)\to\underline{x}^*(i) \\ \underline{a}\to\underline{a}^*(i)}} = \underline{0}^T \qquad (3.2.67)$$

to which the solution is

$$\underline{x}^*(i) = \left[\underline{P}^{-1}(i) - \left\{\underline{\Phi}(i,0)\underline{P}_0\underline{\Phi}^T(i,0) + \sum_{k=1}^{i} \underline{\Phi}(i,k)\underline{G}(k)\underline{Q}(k)\underline{G}^T(k)\underline{\Phi}^T(i,k)\right\}^{-1}\right]^{-1} \cdot \left[\underline{P}^{-1}(i)\hat{\underline{x}}(i) - \left\{\underline{\Phi}(i,0)\underline{P}_0\underline{\Phi}^T(i,0) + \sum_{k=1}^{i} \underline{\Phi}(i,k)\underline{G}(k)\underline{Q}(k)\underline{G}^T(k)\underline{\Phi}^T(i,k)\right\}^{-1}\underline{x}_*(i)\right] \Bigg|_{\underline{a}\to\underline{a}^*(i)} \qquad (3.2.68)$$

Similarly, the likelihood equation for a_ℓ, the ℓ-th component of $\underline{a}$, is:

$$\text{tr}\left\{\left[\underline{P}^{-1}(i)-\underline{P}^{-1}(i)[\underline{x}^*(i)-\hat{\underline{x}}(i)][\underline{x}^*(i)-\hat{\underline{x}}(i)]^T\underline{P}^{-1}(i)\right]\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\}$$

$$-\ 2\ \frac{\partial \hat{\underline{x}}^T(i)}{\partial a_\ell}\ \underline{P}^{-1}(i)[\underline{x}^*(i)-\hat{\underline{x}}(i)]$$

$$-\ \text{tr}\left\{\left[\underline{D}^{-1}(i)-\underline{D}^{-1}(i)[\underline{x}^*(i)-\underline{x}_*(i)][\underline{x}^*(i)-\underline{x}_*(i)]^T\underline{D}^{-1}(i)\right]\frac{\partial \underline{D}(i)}{\partial a_\ell}\right\}$$

$$+\ 2\ \frac{\partial \underline{x}_*^T(i)}{\partial a_\ell}\ \underline{D}^{-1}(i)[\underline{x}^*(i)-\underline{x}_*(i)]$$

$$+\ \sum_{j=1}^{i}\text{tr}\left\{\left[\underline{A}^{-1}(j)-\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]\cdot\right.\right.$$

$$\left.\left.\cdot[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]^T\underline{A}^{-1}(j)\right]\ \frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}$$

$$-\ 2\ \sum_{j=1}^{i}\frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell}\ \underline{H}^T(j)\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)]\Bigg|_{\underline{a}\to\underline{a}^*(i)} = 0 \qquad (3.2.69)$$

where $\underline{D}(i) = \underline{\Phi}(i,0)\underline{P}_o\underline{\Phi}^T(i,0) + \sum_{k=1}^{i}\underline{\Phi}(i,k)\underline{G}(k)\underline{Q}(k)\underline{G}^T(k)\underline{\Phi}^T(i,k)$ (3.2.70)

Now, in the limit as $\underline{P}_o \to \infty\underline{I}$, (3.2.68) becomes

$$\underline{x}^*(i) = \left[\underline{P}^{-1}(i)-\underline{0}\right]^{-1}\left[\underline{P}^{-1}(i)\hat{\underline{x}}(i)-\underline{0}\right]$$

or $$\underline{P}^{-1}(i)\underline{x}^*(i) = \underline{P}^{-1}(i)\hat{\underline{x}}(i) \qquad (3.2.71)$$

and, once the matrix "$\underline{P}^{-1}(i)$" becomes of full rank, this relation can become

$$\underline{x}^*(i) = \underline{P}(i)\underline{P}^{-1}(i)\hat{\underline{x}}(i) = \hat{\underline{x}}(i) \qquad (3.2.72)$$

With respect to the parameter likelihood equation, as $\underline{P}_o \to \infty \underline{I}$, $\underline{D}^{-1}(i) \to \underline{0}$. Once "$\underline{P}^{-1}(i)$" becomes nonsingular, it is valid to assert that the terms $[\underline{x}^*(i) - \hat{\underline{x}}(i)] \to \underline{0}$ and that $\underline{A}^{-1} (= \underline{R}^{-1} - \underline{R}^{-1}\underline{H}[\underline{P}^{-1}]^{-1}\underline{H}^T\underline{R}^{-1})$ and $\frac{\partial \underline{A}}{\partial a_\ell}$ can in fact be defined, so that the likelihood equations become:

$$\operatorname{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\} - 2\sum_{j=1}^{i}\frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell}\,\underline{H}^T(j)\underline{A}^{-1}(j)[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]$$

$$+ \sum_{j=1}^{i}\operatorname{tr}\left\{\left[\underline{A}^{-1}(j) - \underline{A}^{-1}(j)[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]\cdot\right.\right.$$

$$\left.\left.\cdot[\underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j)]^T\,\underline{A}^{-1}(j)\right]\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}\Bigg|_{\underline{a}\to\underline{a}^*(i)} = 0 \qquad (3.2.73)$$

which is the same result as obtained by equation (3.2.18) using $\ell n\, f(\underline{x}(i), \underline{Z}(i) | \underline{a})$. These results dictate the use of the startup procedure of equations (2.1.61) to (2.1.72) with $\underline{P}_o^{-1} = \underline{0}$ and an initial best estimate of $\underline{a}$ to define the propagations, until $[\underline{P}^{-1}(i)]$ becomes nonsingular, at which point the estimation procedure derived in section 3.2.1 for $\ell n\, f(\underline{x}(i), \underline{Z}(i) | \underline{a})$ is utilized.

Similar developments can be made for $\ell n\, f(\underline{Z}_N(i) | \underline{x}(i), \underline{a})$ to relate $\underline{x}^*(i)$ to $\tilde{\underline{x}}(i)$ defined for $\ell n\, f(\underline{x}(i), \underline{Z}_N(i) | \underline{a})$ by noting that

$$f(\underline{Z}_N(i) | \underline{x}(i), \underline{a}) = \frac{1}{f(\underline{x}(i) | \underline{a})}\, f(\underline{x}(i), \underline{Z}_N(i) | \underline{a}) \qquad (3.2.74)$$

To define the estimates properly, N must be large enough to allow $\tilde{\underline{P}}^{-1}(N)$ to be nonsingular and such that the rank of $\tilde{\underline{P}}^{-1}(i)$ is maintained at n. Also, to relate to the $\ell n\, f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$ results, it is possible to write

$$f(\underline{Z}_N(i) | \underline{Z}(i-N), \underline{x}(i), \underline{a}) = \frac{1}{f(\underline{x}(i) | \underline{Z}(i-N), \underline{a})}\, f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a}) \qquad (3.2.75)$$

where

$$f(\underline{x}(i)|\underline{Z}(i-N),\underline{a}) = N\Big[\underline{\Phi}(i,i-N)\hat{\underline{x}}(i-N) + \sum_{j=i-N+1}^{i}\underline{\Phi}(i,j)\underline{B}(j-1)\underline{u}(j-1),$$

$$\underline{\Phi}(i,i-N)\underline{P}(i-N)\underline{\Phi}^{T}(i,i-N)$$

$$+\sum_{j=i-N+1}^{i}\underline{\Phi}(i,j)\underline{G}(j)\underline{Q}(j)\underline{G}^{T}(j)\underline{\Phi}^{T}(i,j)\Big] \qquad (3.2.76)$$

However, these yield rather unweildly results and will not be pursued. Either a large $\underline{P}_o$ can be used in practice with $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\ \underline{a})$, or the aforementioned startup procedure (or large $\underline{P}_o$) used to properly initialize the recursions for $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\ \underline{Z}(i-N),\underline{a})$.

3.2.4 Likelihood Functions of the Form $\ell n\ f(\underline{x}|\underline{Z},\underline{a})$

As shown in section 3.2.1, $\ell n\ f(\underline{x}(i)|\ \underline{Z}(i),\underline{a})$ can be written as

$$\ell n\ f(\underline{x}(i)|\ \underline{Z}(i),\underline{a}) = -\frac{n}{2}\ \ell n\ (2\pi) - \tfrac{1}{2}\ \ell n\ \{|\underline{P}(i)|\}$$

$$-\tfrac{1}{2}\left[\underline{x}(i)-\hat{\underline{x}}(i)\right]^{T}\underline{P}^{-1}(i)\left[\underline{x}(i)-\hat{\underline{x}}(i)\right] \qquad (3.2.77)$$

so that the likelihood equations for this function are

$$-\left[\underline{x}(i)-\hat{\underline{x}}(i)\right]^{T}\underline{P}^{-1}(i)\ \Big|_{\substack{\underline{x}(i)\to\underline{x}^{*}(i)\\ \underline{a}\to\underline{a}^{*}(i)}} = \underline{0}^{T} \qquad (3.2.78)$$

$$\text{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\} - 2\,\frac{\partial \hat{\underline{x}}^{T}(i)}{\partial a_\ell}\,\underline{P}^{-1}(i)\,[\underline{x}(i)-\hat{\underline{x}}(i)]$$

$$- [\underline{x}(i)-\hat{\underline{x}}(i)]^{T}\,\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_\ell}\,\underline{P}^{-1}(i)\,[\underline{x}(i)-\hat{\underline{x}}(i)]\Big|_{\substack{\underline{x}(i)\to\underline{x}^*(i)\\ \underline{a}\to\underline{a}^*(i)}} = 0 \qquad (3.2.79)$$

The solution to equation (3.2.78) is

$$\underline{x}^*(i) = \hat{\underline{x}}(i)\Big|_{\underline{a}\to\underline{a}^*(i)} \qquad (3.2.80)$$

Imposing this upon the solution of equation (3.2.79) causes the second and third terms to vanish, yielding a parameter likelihood equation of

$$\text{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\}\Big|_{\underline{a}\to\underline{a}^*(i)} = 0 \qquad (3.2.81)$$

However, this is equivalent to

$$\frac{\partial}{\partial a_\ell}\left[\ell n\left\{|\underline{P}(i)|\right\}\right]\Big|_{\underline{a}\to\underline{a}^*(i)} = 0 \qquad (3.2.82)$$

i.e., minimizing $\ell n\left\{|\underline{P}(i)|\right\}$ as a function of the uncertain parameters. However, $\underline{P}(i)$ does not depend upon the actual measurements taken. Therefore, this log-likelihood function cannot define a valid parameter esti-

mate, since the measurements are the sole source of information for correcting the estimated parameter values.

3.2.5 Perfect Measurements of Part of the State - A Special Case

The previous sections have all assumed that the measurements are noise-corrupted. Although this is a realistic assumption, there are certain instances in which "perfect" measurements can be made upon part of the state, which is to say that they are so accurate that a noise-free model of the measuring device is suitable. For such cases, the concept of a state observer, as proposed by Luenberger (1966), would replace the state estimator. This section will briefly describe a viable technique for obtaining a parameter estimate in conjunction with the output of a state observer. The most general case will not be developed, but instead a convenient state space representation will be utilized. The formulation used is in part motivated by previous work by Schweppe (1970).

Assume that the system under investigation is a single input-single output dynamic system described by

$$\begin{bmatrix} \underline{x}(i+1) \end{bmatrix} = \begin{bmatrix} \underline{0} & \vdots & \underline{I} \\ \cdots & \vdots & \cdots \\ \underline{\phi}^{T}(i+1,i) & & \end{bmatrix} \begin{bmatrix} \underline{x}(i) \end{bmatrix} + \begin{bmatrix} \underline{0} \\ \cdots \\ b(i) \end{bmatrix} u(i) + \begin{bmatrix} \underline{0} \\ \cdots \\ 1 \end{bmatrix} w(i) \qquad (3.2.83)$$

$$z(i) = \begin{bmatrix} \underline{0}^{T} & \vdots & 1 \end{bmatrix} \underline{x}(i) \qquad (3.2.84)$$

where w(i) is a white, Gaussian, zero mean sequence with covariance Q(i). Looking at the last row of equation (3.2.83), this system can be expressed equivalently as:

$$y(i+1) = \underline{\phi}^T(i+1,i)\underline{x}(i) + b(i)u(i) + w(i)$$

$$= \phi_0\, y(i) + \ldots + \phi_{n-1} y(i-n+1) + b(i)u(i) + w(i) \qquad (3.2.85)$$

by defining the vectors $\underline{\phi}(i+1,i)$ and $\underline{x}(i)$ as

$$\underline{\phi}^T(i+1,i) = \left[\phi_{n-1} \cdot \cdot \cdot \phi_1 \quad \phi_0\right] \qquad (3.2.86a)$$

$$\underline{x}^T(i) = \left[y(i-n+1) \cdot \cdot \cdot y(i-1)\ y(i)\right] \qquad (3.2.86b)$$

The parameter to be identified will be all of $\underline{\phi}(i)$ and $b(i)$, which are assumed to be functionally independent, and essentially constant over an interval of N samples. The value of the control $u(i)$ will not depend on these parameter values. For convenience, it is further assumed that $u(0)=0$ and $\underline{x}(0)=\underline{0}$.

An appropriate choice of log-likelihood function for this problem is $\ell n\ f(\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})$. Applying Bayes' Rule to this function yields

$$\ell n\ f(\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a}) = \sum_{j=i-N+1}^{i} \ell n\ f(z(j)|\underline{Z}(j-1),\underline{a}) \qquad (3.2.87)$$

Each of the terms within the summation is evaluated as follows. Since $w(j)$ is a zero mean, white Gaussian sequence, the statistics of the measured output can be expressed as

$$E\left\{z(j)\middle|\underline{Z}(j-1),\underline{a}\right\} = E\left\{y(j)\middle|\underline{Z}(j-1),\underline{a}\right\}$$

$$= E\Big\{\phi_0 z(j-1) + \ldots + \phi_{n-1} z(j-n) + b(j-1)u(j-1)$$

$$+ w(j-1)\Big|\underline{Z}(j-1),\underline{a}\Big\}$$

$$= \underline{\phi}^T(j,j-1)\underline{Z}_n(j-1) + b(j-1)u(j-1) \qquad (3.2.88)$$

where $\underline{Z}_n(j-1)$ is the most recent n measurements at time (j-1): z(j-1), ..., z(j-n). Note that $\underline{Z}_n(0)$ is simply $\underline{0}$. The covariance is

$$E\left\{\left[z(j) - \underline{\phi}^T(j,j-1)\underline{Z}_n(j-1) - b(j-1)u(j-1)\right]^2 \middle| \underline{Z}(j-1),\underline{a}\right\} =$$

$$= E\left\{\left[w(j)\right]^2 \middle| \underline{Z}(j-1),\underline{a}\right\}$$

$$= Q(j)$$

Thus, the log-likelihood function can be written as

$$\ell n\ f(\underline{Z}_N(i)|\ \underline{Z}(i-N),\underline{a}) = -\frac{N}{2}\ \ell n(2\pi) - \tfrac{1}{2}\sum_{j=i-N+1}^{i} \ell n[Q(j)]$$

$$- \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\left[z(j) - \underline{\phi}^T(j,j-1)\underline{Z}_n(j-1) - b(j-1)u(j-1)\right]^2$$

(3.2.90)

Taking the derivatives with respect to $\underline{\phi}$ and b, using the facts that these are not functionally dependent, nor is the control a function of either $\underline{\phi}$ or b, yields the likelihood equations

$$\underline{0}^T = \left.\frac{\partial\ \ell n\ f(\underline{Z}_N(i)|\ \underline{Z}(i-N),\underline{a})}{\partial\underline{\phi}}\right|_{\underline{a}\to\underline{a}^*(i)}$$

$$= \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\left[z(j) - \underline{\phi}^{*T}(j,j-1)\underline{Z}_n(j-1) - b^*(j-1)u(j-1)\right]\underline{Z}_n(j-1)$$

(3.2.91)

$$0 = \left.\frac{\partial \ell n\ f(\underline{Z}_N(i)\,|\,\underline{Z}(i-N),\underline{a})}{\partial b}\right|_{\underline{a}\rightarrow\underline{a}^*(i)}$$

$$= \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\left[\underline{z}(j)-\underline{\phi}^{*T}(j,j-1)\underline{Z}_n(j-1)-b^*(j-1)u(j-1)\right]u(j-1) \tag{3.2.92}$$

These can be written as a single equation

$$\begin{bmatrix} \underline{\Lambda} & \vdots & \underline{\lambda} \\ \cdots & \vdots & \cdots \\ \underline{\lambda}^T & \vdots & \lambda_o \end{bmatrix} \begin{bmatrix} \underline{\phi}^*(i,i-1) \\ \cdots \\ b^*(i-1) \end{bmatrix} = \begin{bmatrix} \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\,\underline{Z}_n(j-1)z(j) \\ \cdots \\ \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\,u(j-1)z(j) \end{bmatrix} \tag{3.2.93}$$

where the matrix $\underline{\Lambda}$, vector $\underline{\lambda}$, and scalar λ_o are defined by

$$\left.\begin{aligned} \underline{\Lambda} &= \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\,\underline{Z}_n(j-1)\underline{Z}_n^T(j-1) \\ \underline{\lambda} &= \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\,\underline{Z}_n(j-1)u(j-1) \\ \lambda_o &= \sum_{j=i-N+1}^{i} \frac{1}{Q(j)}\,u^2(j-1) \end{aligned}\right\} \tag{3.2.94}$$

The fact that $\underline{\phi}(j,j-1)$ and $b(j-1)$ are assumed to be constant over N sample periods has been used explicitly in obtaining equation (3.2.93). This relation can be solved for the parameter estimates, provided that the required inverse exists, as:

$$\begin{bmatrix} \underline{\phi}^* \\ \cdots \\ b^* \end{bmatrix} = \begin{bmatrix} \underline{\Delta} & \vdots & \underline{\lambda} \\ \cdots & \cdots & \cdots \\ \underline{\lambda}^T & \vdots & \lambda_o \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=i-N+1}^{i} \frac{1}{Q(j)} \underline{Z}_n(j-1)z(j) \\ \cdots \\ \sum_{j=i-N+1}^{i} \frac{1}{Q(j)} u(j-1)z(j) \end{bmatrix} \tag{3.2.95}$$

If the noise covariance Q(j) is time invariant, i.e., if Q(j) equals Q for all j, then equation (3.2.95) is simplified to

$$\begin{bmatrix} \underline{\phi}^* \\ \cdots \\ b^* \end{bmatrix} = \begin{bmatrix} \underline{\Delta}' & \vdots & \underline{\lambda}' \\ \cdots & \cdots & \cdots \\ \underline{\lambda}'^T & \vdots & \lambda'_o \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=i-N+1}^{i} \underline{Z}_n(j-1)z(j) \\ \cdots \\ \sum_{j=i-N+1}^{i} u(j-1)z(j) \end{bmatrix} \tag{3.2.96}$$

where

$$\left.\begin{aligned} \underline{\Delta}' &= \sum_{j=i-N+1}^{i} \underline{Z}_n(j-1)\underline{Z}_n^T(j-1) \\ \underline{\lambda}' &= \sum_{j=i-N+1}^{i} \underline{Z}_n(j-1)u(j-1) \\ \lambda'_o &= \sum_{j=i-N+1}^{i} u^2(j-1) \end{aligned}\right\} \tag{3.2.97}$$

In this case it is also possible to estimate the value of Q by maximum likelihood means. Setting the Q-partial derivative of the log-likelihood function, equation (3.2.90), equal to zero, the estimate is obtained as

$$Q^* = \frac{1}{N} \sum_{j=i-N+1}^{i} \left[z(j) - \underline{\phi}^T(j,j-1)\underline{Z}_n(j-1) - b(j-1)u(j-1) \right]^2 \tag{3.2.98}$$

To provide recursive estimates efficiently, the inverse

matrices appearing in equations (3.2.95) and (3.2.96) can be expressed in terms of the corresponding inverse matrix at the previous sample time. Either the matrix inversion lemma can be applied, or the approximation used, for $\underline{X}^{-1}(i)$ defined as the required inverse, that

$$\delta\underline{X}^{-1}(i) \cong -\underline{X}^{-1}(i-1)\delta\underline{X}(i)\underline{X}^{-1}(i-1) \qquad (3.2.99)$$

which is especially valid for large N.

3.3 Iterative Solution to the Likelihood Equation

The derivations of section 3.2 have led to a set of likelihood equations, the solution to which will yield the desired estimates of system states and parameters. However, these equations do not have a general closed form solution, so it is necessary to consider iterative techniques.

One means of obtaining the solution to the likelihood equations is to employ the Newton-Raphson iteration. If $\underline{\theta}$ is defined as the vector of all quantities to be estimated,

$$\underline{\theta}(i) = \begin{bmatrix} \underline{x}(i) \\ \underline{a}(i) \end{bmatrix} \qquad (3.3.1)$$

then the likelihood equations can be written as

$$\left.\frac{\partial L[\underline{\theta}, \underline{Z}(i)]}{\partial \underline{\theta}}\right|_{\underline{\theta} \to \underline{\theta}^*(i)} = \underline{0}^T \qquad (3.3.2)$$

where $\underline{\theta}^*(i)$ is the maximum likelihood estimate of $\underline{\theta}$. The philosophy of the Newton-Raphson procedure is to approximate the log-likelihood function as a quadratic in $\underline{\theta}$:

$$L[\underline{\theta}^*(i),\underline{Z}(i)] \cong L[\underline{\theta}_*(i),\underline{Z}(i)]$$

$$+ \frac{\partial L[\underline{\theta}_*(i),\underline{Z}(i)]}{\partial \underline{\theta}} \left[\underline{\theta}^*(i) - \underline{\theta}_*(i)\right]$$

$$+ \tfrac{1}{2} \left[\underline{\theta}^*(i) - \underline{\theta}_*(i)\right]^T \frac{\partial^2 L[\underline{\theta}_*(i),\underline{Z}(i)]}{\partial \underline{\theta}^2} \left[\underline{\theta}^*(i) - \underline{\theta}_*(i)\right] \tag{3.3.3}$$

where $\underline{\theta}_*(i)$ is some admissable value of the vector $\underline{\theta}$, and the notation $\partial L/\partial \underline{\theta}^T[\underline{\theta}_*(i),\underline{Z}(i)]$ refers more explicitly to the vector quantity $\partial L/\partial \underline{\theta}^T[\underline{\theta},\underline{Z}(i)]$, evaluated at the point $\underline{\theta} = \underline{\theta}_*(i)$. Equivalently, the derivative of the log-likelihood function is approximated as a linear function of $\underline{\theta}$:

$$\frac{\partial L[\underline{\theta}^*(i),\underline{Z}(i)]^T}{\partial \underline{\theta}} \cong \frac{\partial L[\underline{\theta}_*(i),\underline{Z}(i)]^T}{\partial \underline{\theta}}$$

$$+ \frac{\partial^2 L[\underline{\theta}_*(i),\underline{Z}(i)]}{\partial \underline{\theta}^2} \left[\underline{\theta}^*(i) - \underline{\theta}_*(i)\right] \tag{3.3.4}$$

If this equation were satisfied exactly, the likelihood equation would be soluble in one step. For $\underline{\theta}_*(i)$ as any trial solution, then, equation (3.3.4) could be set equal to zero to obtain the maximum likelihood estimate as

$$\underline{\theta}^*(i) = \underline{\theta}_*(i) - \left[\frac{\partial^2 L[\underline{\theta}_*(i),\underline{Z}(i)]}{\partial \underline{\theta}^2}\right]^{-1} \frac{\partial L[\underline{\theta}_*(i),\underline{Z}(i)]^T}{\partial \underline{\theta}} \tag{3.3.5}$$

It is evident that a necessary condition for the form of this

equation to exist is that the Jacobian matrix,

$$\frac{\partial^2 L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}^2}$$

be of full rank.

If equation (3.3.4) were not actually an exact equality, but rather a Taylor series about an assumed solution denoted by $\underline{\theta}_*(i)$, in which only terms to first order in $[\underline{\theta}^*(i) - \underline{\theta}_*(i)]$ are retained, then equation (3.3.5) can be considered a first order correction to the estimate $\underline{\theta}_*(i)$. To converge on the solution, it is possible to apply (3.3.5) repeatedly at each step until the corrections become negligibly small: a $\underline{\theta}_*(i)$ is chosen and used to calculate $\underline{\theta}^*(i)$, this $\underline{\theta}^*(i)$ then being used as the $\underline{\theta}_*(i)$ value in the next iteration, and so on. This is a process known as local relinearization, but it is probably impractical for any on-line applications.

Unfortunately, even if only one iteration of (3.3.5) were to be processed at each time step, the computation of an analytic expression for the required Jacobian matrix is complicated and would put a prohibitive load on the computer if on-line operation is desired. Rao (1968) has suggested an approximation called "scoring" which simplifies the computations substantially while maintaining accuracy over large samples. The approximation made is that

$$\frac{\partial^2 L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}^2} \quad \cong \quad - \underline{J}[\, i \,, \underline{\theta}_*(i)] \qquad (3.3.6)$$

where the matrix

$$\underline{J}[\, i \,, \underline{\theta}_*(i)] = E\left\{ \frac{\partial L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}}^{T} \; \frac{\partial L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}} \,\middle|\, \underline{\theta}_*(i) \right\} \qquad (3.3.7)$$

is termed the augmented conditional information matrix by Rao. This can be motivated by the following equality (the proof of which can be found in section 4.1.1):

$$\frac{\partial^2}{\partial \underline{\theta}^2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\underline{Z}(i) \mid \underline{\theta}_*(i)) d\underline{Z}(i) = \frac{\partial^2}{\partial \underline{\theta}^2} (1)$$

$$= 0 = E\left\{ \frac{\partial L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}}^T \frac{\partial L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}} \,\middle|\, \underline{\theta}_*(i) \right\}$$

$$+ E\left\{ \frac{\partial^2 L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}^2} \,\middle|\, \underline{\theta}_*(i) \right\} \qquad (3.3.8)$$

for $L[\underline{\theta}_*(i), \underline{Z}(i)] = \ell n\, f(\underline{Z}(i) \mid \underline{\theta}_*(i))$. For this case, then, the approximation is that the second partial derivative of $L[\underline{\theta}_*(i), \underline{Z}(i)]$ for a particular realization of $\underline{Z}(i)$ can be adequately represented by its average over all possible $\underline{Z}(i)$ sequences. Incorporating this into equation (3.3.5), the first order correction equation becomes

$$\underline{\theta}^*(i) = \underline{\theta}_*(i) + \underline{J}[i, \underline{\theta}_*(i)]^{-1} \frac{\partial L[\underline{\theta}_*(i), \underline{Z}(i)]}{\partial \underline{\theta}}^T \qquad (3.3.9)$$

As Rao has shown, the error committed by using the approximation of equation (3.3.6) is of order 1/N for large samples. This is one of the numerous factors that dictate that the size of the fixed-length interval of data, N samples long, should be as large as modelling accuracy (the parameters are assumed to be essentially constant over the interval) and computer capacity (only a finite amount of memory space and calculation time is available for any practical on-line device)

will allow. It can also be shown that, as N is allowed to become larger and once a stabilized solution of $\underline{\theta}^*(i)$ has been achieved, the asymptotic estimation error can be described statistically by a normal distribution with mean zero and a conditional covariance of $\underline{J}^{-1}[\,i\,,\,\underline{\theta}_t(i)]$, $\underline{\theta}_t(i)$ being the true value of $\underline{\theta}(i)$. Thus, $\underline{J}^{-1}[\,i\,,\,\underline{\theta}^*(i)]$, which is calculable, is a very good approximation to this covariance.

Other iterative methods are feasible, but investigations have revealed the "scoring" technique to be superior in numerous respects. To make a theoretical comparison of the various methods, it is useful to generalize the derivation of an iterative procedure so as to encompass all of the specific forms, thereby providing a uniform frame of reference. For convenience, the scalar case of equation (3.3.2) will be considered, but the results are readily applicable to estimating a vector of parameters. Thus, the equation to be solved is

$$\left.\frac{\partial L[\theta, \underline{Z}(i)]}{\partial \theta}\right|_{\theta \to \theta^*(i)} = 0 \qquad (3.3.10)$$

where the true solution, $\theta^*(i)$, is assumed to exist. The standard solution procedure is to define $\phi(\theta)$ to be a differentiable function of θ which has no zero in the neighborhood of the root $\theta^*(i)$ of the likelihood equation. Then the function $\lambda(\theta)$ can be defined as

$$\lambda(\theta) = \theta - \phi(\theta)\ \frac{\partial L}{\partial \theta} \qquad (3.3.11)$$

Therefore, the solution to the original equation, (3.3.10), can be found by solving the equivalent relation

$$\theta^*(i) = \lambda[\theta^*(i)] \qquad (3.3.12)$$

from which the most evident iterative scheme is

$$\theta_{k+1}(i) = \lambda[\theta_k(i)]$$

$$= \theta_k(i) - \phi[\theta_k(i)] \left[\frac{\partial L}{\partial \theta} \quad [\theta_k(i)]\right] \qquad (3.3.13)$$

The choice of the function $\phi(\theta)$ distinguishes the various forms of iterations. The second order Newton-Raphson technique results from

$$\phi(\theta) = \left[\frac{\partial^2 L}{\partial \theta^2}\right]^{-1} \qquad (3.3.14)$$

The gradient methods do not attempt to approximate this value, but instead use an "appropriate" weighting value, such as a constant. In the case of a vector of parameters, rather than attempt to evaluate the matrix corresponding to (3.3.14), they seek an appropriate scalar ϕ, as by performing a local optimization on ϕ so as to proceed along the current gradient direction, $\partial L/\partial \underline{\theta}^T[\underline{\theta}_k(i)]$ until the next gradient direction, $\partial L/\partial \underline{\theta}^T[\underline{\theta}_{k+1}(i)]$, is orthogonal to it. However, as shown by equations (3.3.6) and (3.3.7), the scoring method does attempt to evaluate the inverse Jacobian, using first derivative information.

To have confidence in the use of such iterative solutions, their convergence properties must be determined. Considerable theory about this aspect has been developed by Deutsch (1968), Kale (1961), Householder (1953), Isaacson and Keller (1966), and Kantorovich and A kilov (1964). If the estimation error at the k-th iteration is defined as the absolute value (or appropriate norm for a vector of parameters) of the difference between the k-th iterate and the true value $\underline{\theta}^*(i)$,

$$e_k(i) = |\theta_k(i) - \theta^*(i)| \qquad (3.3.15)$$

then the iteration described by equation (3.3.13) will converge to $\theta^*(i)$ if the estimation error decreases as k increases and tends to zero as k grows without bound:

$$\left.\begin{aligned} & e_{k+1} < e_k \\ & e_k \longrightarrow 0 \qquad \text{as} \qquad k \longrightarrow \infty \end{aligned}\right\} \quad (3.3.16)$$

Householder (1953) has proven that the following two conditions are sufficient to satisfy (3.3.16):

1) There exists an r-neighborhood of $\theta^*(i)$, $N_r[\theta^*(i)]$, such that if θ^I and θ^{II} are both contained within $N_r[\theta^*(i)]$, then for some c, $0 \leq c < 1$, it is true that

$$\left| \lambda(\theta^I) - \lambda(\theta^{II}) \right| \leq c \qquad (3.3.17)$$

If $\lambda(\theta)$ is differentiable, then the existence of $N_r[\theta^*(i)]$ is assured if

$$\left| \frac{\partial \lambda}{\partial \theta} \left[\theta^*(i)\right] \right| < 1 \qquad (3.3.18)$$

2) The initial estimate, $\theta_0(i)$, is contained within $N_r[\theta^*(i)]$.

Kale (1961) demonstrated that the Newton-Raphson process satisfies Householder's conditions, and that it has second order convergence (as shown also by Kantorovich and Akilov (1964). He also showed that the scoring procedure satisfies these conditions if $J[\,i\,,\,\theta^*(i)]$ is differentiable, which is not very restrictive. Furthermore, scoring is less rapidly convergent than the Newton-Raphson method near the solution point, but more so than the first order gradient techniques. A gradient method employing a constant weighting value, $\phi(\theta) = \nu$ = constant, can be shown to converge for

$$0 < \nu \frac{\partial^2 L}{\partial \theta^2} \left[\theta^*(i)\right] < 2 \qquad (3.3.19)$$

However, it is very difficult to find an appropriate value ν, such that the convergence proceeds, but not too slowly.

With regard to the size of the neighborhood about $\theta^*(i)$ containing the admissible initial conditions, it can be said that the higher the rate of convergence of the iteration technique, the better the initial estimate must be to insure convergence. However, the exact size of the neighborhood is difficult to determine, and obviously the solution point, necessary to position the neighborhood, would not be known apriori even if the size of $N_r[\theta^*(i)]$ could be determined. Thus, it is usually not possible to prespecify a set of initial conditions for which convergence is assured in an actual application.

For the Newton-Raphson procedure, there is a sufficient condition for convergence which can be explicitly verified without knowledge of the solution point. Since it is also valid for scoring, it is presented here, although the amount of computation it requires might make it impractical compared to the alternative of extensive simulation analysis. The proof can be found in Isaacson and Keller (1966). Assume the initial condition on the iteration $\underline{\theta}_\circ(i)$, determines a Jacobian matrix $\partial^2 L/\partial \underline{\theta}^2 \left[\underline{\theta}_\circ(i), \underline{Z}(i)\right]$ which has an inverse with norm bounded by

$$\left\| \frac{\partial^2 L^{-1}}{\partial \underline{\theta}^2} \left[\underline{\theta}_\circ(i), \underline{Z}(i)\right] \right\| \leq a \qquad (3.3.20)$$

Further assume that the difference between the first two iterates is bounded by:

$$\|\underline{\theta}_1(i) - \underline{\theta}_0(i)\| = \left\| \left[\frac{\partial^2 L}{\partial \underline{\theta}^2}\left[\underline{\theta}_\circ(i), \underline{Z}(i)\right]\right]^{-1} \frac{\partial L}{\partial \underline{\theta}}^T \left[\underline{\theta}_\circ(i), \underline{Z}(i)\right] \right\| \leq b \tag{3.3.21}$$

The vector norms, as in (3.3.21), are understood to be the maximum norms, and the matrix norms, as in (3.3.20), are the corresponding induced natural norms, defined as

$$\|\underline{\alpha}\| = \max_i \; |\alpha_i| \tag{3.3.22}$$

$$\|\underline{A}\| = \max_i \; \left\{ \sum_{j=1}^{p} |a_{ij}| \right\} \tag{3.3.23}$$

where p is the dimension of the parameter vector. Finally assume that the components of $(\partial L/\partial \underline{\theta})^T$ have continuous second derivatives which satisfy

$$\sum_{\ell=1}^{p} \left| \frac{\partial^3 L[\underline{\theta}_k(i), \underline{Z}(i)]}{\partial \theta_i \partial \theta_j \partial \theta_\ell} \right| \leq \frac{c}{p} \tag{3.3.24}$$

for i and j, taking on all values 1, 2, ... p and for all $\underline{\theta}_k(i)$ within $\|\underline{\theta}_k(i) - \underline{\theta}_\circ(i)\| \leq 2b$. If the constants a, b, and c satisfy the relation

$$abc \leq \tfrac{1}{2} \tag{3.3.25}$$

then the Newton-Raphson iterates are uniquely defined and lie within the 2b-sphere about $\underline{\theta}_\circ(i)$, defined by

$$\|\underline{\theta}_k(i) - \underline{\theta}_\circ(i)\| \leq 2b \tag{3.3.26}$$

and the iterates converge to some vector, $\underline{\theta}_\infty(i)$,

$$\lim_{n\to\infty} \underline{\theta}_k(i) = \underline{\theta}_\infty(i) \tag{3.3.27}$$

for which

$$\frac{\partial L}{\partial \underline{\theta}}^T \left[\underline{\theta}_\infty(i)\right] = \underline{0} \tag{3.3.28}$$

$$\left\| \underline{\theta}_k(i) - \underline{\theta}_\infty(i) \right\| \leq \frac{2b}{2^n} \tag{3.3.29}$$

Besides providing an adequately convergent algorithm, it is essential that an iterative process be as efficient and simple as possible for on-line usage. The Newton-Raphson technique is attractive because it converges rapidly (quadratically) near the optimum solution, unlike a gradient procedure where rate of convergence is usually very slow near the optimum, but it requires considerable computation. Scoring reduces the calculations, at the expense of some rate of convergence, but because it does approximate the inverse Jacobian its rate is superior to that of gradient methods. As more samples of data are taken, the approximation becomes better, and the convergence rate approaches that of the Newton-Raphson method. Computationally, scoring is not as simple as the simplest gradient techniques, but it does not require any local optimizations as the more sophisticated gradient methods do.

There are disadvantages to the scoring method as well. Whereas a gradient method will always converge at least to a local optimum, scoring will converge only if the initially assumed values are sufficiently close to the optimal values.

In practice, the initial evaluations of $\underline{J}[\,i\,,\,\underline{\theta}_*(i)]$ yield a matrix with a small magnitude (and not a good approximation to the negative Jacobian), so that its inverse is very large. (Even if the Jacobian were well approximated, the resulting Newton-Raphson iteration would generally exhibit initial convergence behavior inferior to that of a gradient technique.) Thus, a two-mode mechanization might be desirable: until $\underline{J}\,[\,i\,,\,\underline{\theta}_*(i)]$ reaches a minimal size, a precomputed $\underline{J}^{-1}$ matrix (the value that $\underline{J}^{-1}$ might converge to in simulation studies for an average value for $\underline{\theta}$) could be substituted into the recursion.

Another disadvantage is the need to calculate the value and inverse of $\underline{J}\,[\,i\,,\,\underline{\theta}_*(i)]$. However, repeated evaluations are not necessary in practice because $\underline{J}\,[\,i\,,\,\underline{\theta}_*(i)]$ will converge after a number of steps. It has been found that it is adequate to retain a fixed value of $\underline{J}^{-1}$ after a certain stage, periodically verifying that the corresponding $\underline{J}$ does not differ substantially from $\underline{J}\,[\,i\,,\,\underline{\theta}_*(i)]$; the stage at which this is initiated and the difference magnitude that dictates the re-evaluation of the inverse can be determined from simulation studies to insure adequate performance. Investigation upon the examples described in Chapter VI reveal that the two-mode idea mentioned previously can successfully be combined with the present concept: use a precomputed average value of $\underline{J}^{-1}$ until a certain stage, and then test periodically for the need to re-evaluate the inverse. In fact, for many on-line applications, adequate performance can be achieved by using the precomputed $\underline{J}^{-1}$ throughout the estimation process: this becomes somewhat similar conceptually to a fixed weighting value gradient technique, but it retains the structure of incorporating an approximate inverse Jacobian, rather than a more arbitrary scalar, into the iteration.

There are other iterative methods besides scoring that approximate the inverse Jacobian with first derivative information. For instance, the Davidon method constructs the required inverse directly from first order information, rather

than computing a matrix and then its inverse. However, it entails more computations than does scoring, requires considerable storage, uses a local scalar optimization, and is rather susceptible to single precision accuracy problems. Furthermore, the routine generates all of the quantities needed to evaluate $\underline{J}\left[i, \underline{\theta}_*(i)\right]$ directly. Therefore, especially since the inversion of $\underline{J}$ is not a crucial problem as shown previously, the scoring method is preferable. The fact is that scoring has provided very acceptable performance in all cases examined, and will be the procedure employed in solving the likelihood equations in the next section.

3.4 Estimate Computations

This section will develop the recursions to yield a full-scale solution to the state and parameter estimation problem. They are not intended to provide on-line capability, but rather to establish the best performance achievable from maximum likelihood methods. In Chapter V, various concepts will be developed to furnish the desired on-line applicability, and the performance of the resulting estimators can be compared to the full-scale solutions of this section. Thus, a valid tradeoff of estimation accuracy and computational efficiency can be made. Furthermore, investigations to depict the prime determinants of the behavior of the full-scale estimators will, in fact, suggest some of the simplifications to be used to attain an on-line algorithm.

For the sake of clarity, one representative estimator will be developed. The most useful form is that developed from a log-likelihood function of $\ell n\, f(\underline{x}(i), \underline{Z}_N(i) \mid \underline{Z}(i-N), \underline{a})$. Treatment of this estimator will encompass most of the aspects embodied in the other forms, so the results will be easily generalized. From equations (3.2.42) and (3.2.43), the state estimate is:

$$\underline{x}^*(i) = \hat{\underline{x}}(i) \Big|_{\underline{a} \to \underline{a}^*(i)} \tag{3.4.1}$$

where the calculation of $\hat{\underline{x}}(i)$ has been described in section 2.1, and where the parameter estimate used to propagate $\hat{\underline{x}}(i)$ is to be found from solving the likelihood equations

$$\operatorname{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\} - 2\sum_{j=i-N+1}^{i} \frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell}\,\underline{H}^T(j)\underline{A}^{-1}(j)\underline{r}(j)$$
$$+ \sum_{j=i-N+1}^{i} \operatorname{tr}\left\{\left[\underline{A}^{-1}(j) - \underline{A}^{-1}(j)\underline{r}(j)\underline{r}^T(j)\underline{A}^{-1}(j)\right]\,\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}\Bigg|_{\underline{a} \to \underline{a}^*(i)} = 0 \tag{3.4.2}$$

where the measurement residual $\underline{r}(j)$ and matrix $\underline{A}(j)$ are defined as

$$\underline{r}(j) = \underline{z}(j) - \underline{H}(j)\bar{\underline{x}}(j) \tag{3.4.3}$$

$$\underline{A}(j) = \underline{H}(j)\underline{M}(j)\underline{H}^T(j) + \underline{R}(j) \tag{3.4.4}$$

To implement the scoring iteration for equations of the form (3.4.2), it is necessary to generate the score $\partial L/\partial \underline{a}^T[\underline{x}_*(i), \underline{a}_*(i), \underline{Z}(i)]$ and the conditional information matrix, $\underline{J}[\,i\,,\,\underline{x}_*(i)\,,\,\underline{a}_*(i)\,]$. The score is a p-dimensional vector whose components are of the form of equation (3.4.2) times $(-\frac{1}{2})$, but in which the terms are evaluated with the parameter estimate $\underline{a}_*(i)$ rather than the actual, but unknown maximum likelihood estimate $\underline{a}^*(i)$. Recursions can be developed for the required terms, and for this purpose it is convenient to decompose the score into the sum of single measurement scores, $\underline{s}^1[\underline{Z}(j), \underline{a}_*(i)]$, and the term $\underline{g}[\underline{Z}(i), \underline{a}_*(i)]$. For the ℓ-th component,

$$\frac{\partial L}{\partial a_\ell}\left[\underline{x}_*(i),\underline{a}_*(i),\underline{Z}(i)\right] = \gamma_\ell\left[\underline{Z}(i),\underline{a}_*(i)\right] + \sum_{j=i-N+1}^{i} s^1{}_\ell\left[\underline{Z}(j),\underline{a}_*(i)\right] \tag{3.4.5}$$

where

$$s^1{}_\ell\left[\underline{Z}(j),\underline{a}\right] = \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell}\,\underline{H}^T(j)\underline{A}^{-1}(j)\underline{r}(j) - \tfrac{1}{2}\,\mathrm{tr}\left\{\left[\underline{A}^{-1}(j)-\underline{A}^{-1}(j)\underline{r}(j)\underline{r}^T(j)\underline{A}^{-1}(j)\right]\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\} \tag{3.4.6}$$

and, from equation (3.2.16),

$$\gamma_\ell\left[\underline{Z}(i),\underline{a}\right] = \frac{\partial \underline{\hat{x}}^T(i)}{\partial a_\ell}\,\underline{P}^{-1}(i)\left[\underline{x}(i)-\underline{\hat{x}}(i)\right] - \tfrac{1}{2}\,\mathrm{tr}\left\{\left[\underline{P}^{-1}(i)-\underline{P}^{-1}(i)\left[\underline{x}(i)-\underline{\hat{x}}(i)\right]\left[\underline{x}(i)-\underline{\hat{x}}(i)\right]\underline{P}^{-1}(i)\right]\frac{\partial \underline{P}(i)}{\partial a_\ell}\right\} \tag{3.4.7}$$

It should be realized that, although $s^1{}_\ell\left[\underline{Z}(j),\underline{a}\right]$ explicitly contains only the residual at time instant j, this single measurement score is in fact a function of the entire measurement history, $\underline{Z}(j)$. In equation (3.4.7), the solution form yields $\left[\underline{x}(i) - \underline{\hat{x}}(i)\right]\rightarrow\underline{0}$, so that only the leading term in the trace remains in the score evaluation.

With regard to the conditional information matrix, it is possible to obtain a closed-form expression for the case of the parameter value assuming its true value, $\underline{a}_t$. This expression will then be evaluated using $\underline{a}_*(i)$ as a suitable recursive approximation. In section 4.1, it will be shown that, by employing the independence of successive residuals for a linear filter using the true parameter values, the

component of the conditional information matrix due to the single measurement scores can be expressed as

$$E\left\{ \sum_{j=i-N+1}^{i} \underline{s}^{1}[\underline{Z}(j),\underline{a}_t] \cdot \sum_{n=i-N+1}^{i} \underline{s}^{1^T}[\underline{Z}(n),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} =$$

$$= E\left\{ \sum_{j=i-N+1}^{i} \underline{s}^{1}[\underline{Z}(j),\underline{a}_t]\ \underline{s}^{1^T}[\underline{Z}(j),\underline{a}_t] \,\middle|\, \underline{a}_t \right\}$$

$$= \sum_{j=i-N+1}^{i} E\left\{ \underline{s}^{1}[\underline{Z}(j),\underline{a}_t]\ \underline{s}^{1^T}[\underline{Z}(j),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} \tag{3.4.8}$$

Similarly, since the error in the state estimate for such a filter is a zero-mean random variable independent of the current and previous residuals, the total conditional information matrix can be decomposed into a matrix whose k-ℓ-th entry is

$$J_{k\ell}[\, i\,,\, \underline{x}_*(i)\,,\, \underline{a}_t\,] = E\left\{ r_k[\underline{Z}(i),\underline{a}_t]\ r_\ell[\underline{Z}(i),\underline{a}_t] \,\middle|\, \underline{a}_t \right\}$$

$$+ \sum_{j=i-N+1}^{i} E\left\{ s^{1}{}_{k}[\underline{Z}(j),\underline{a}_t]\, s^{1}{}_{\ell}[\underline{Z}(j),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} \tag{3.4.9}$$

Appendix E demonstrates that each term in the sum term can be evaluated as

$$E\left\{ s^{1}{}_{k}[\underline{Z}(j),\underline{a}_t]\ s^{1}{}_{\ell}[\underline{Z}(j),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} =$$

$$= \tfrac{1}{2}\, \mathrm{tr}\left\{ \underline{A}^{-1}(j)\, \frac{\partial \underline{A}(j)}{\partial a_k}\, \underline{A}^{-1}(j)\, \frac{\partial \underline{A}(j)}{\partial a_\ell} \right.$$

$$\left. + 2\, \underline{A}^{-1}(j)\underline{H}(j) E\left[\frac{\partial \underline{\bar{x}}(j)}{\partial a_k}\ \frac{\partial \underline{\bar{x}}^{T}(j)}{\partial a_\ell} \,\middle|\, \underline{a}_t \right] \underline{H}^{T}(j) \right\} \tag{3.4.10}$$

By similar reasoning, it can be shown that

$$E\left\{ r_k[\underline{Z}(i),\underline{a}_t]\ r_\ell[\underline{Z}(i),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} =$$

$$= \tfrac{1}{2}\ \mathrm{tr}\left\{ \underline{P}^{-1}(i)\frac{\partial \underline{P}(i)}{\partial a_k}\underline{P}^{-1}(i)\ \frac{\partial \underline{P}(i)}{\partial a_\ell} + 2\ \underline{P}^{-1}(i)\ E\left[\frac{\partial \hat{\underline{x}}(i)}{\partial a_k}\ \frac{\partial \hat{\underline{x}}^T(i)}{\partial a_\ell} \,\middle|\, \underline{a}_t\right]\right\} \qquad (3.4.11)$$

Equations (3.4.10) and (3.4.11) are exact: no approximations were used to derive them. The approximation to be made is that the same expressions can be used for parameter values other than $\underline{a}_t$.

Equations (3.4.5) and (3.4.9) provide explicit evaluations of the score and conditional information matrix, expressed in terms of quantities which can be defined recursively. The parameter estimator seeks the value of the parameter vector, constant over the most recent N sample times, that best matches the system model to measurements taken from the real system. Therefore, the parameter value will be allowed to vary only within the N-step interval in the computations. In other words, if the estimate is being made at time i, then $\hat{\underline{x}}(i-N)$ and $\underline{P}(i-N)$ are considered to be immutable; thus, the initial conditions of the N-step recursions are

$$\left.\begin{aligned} &\hat{\underline{x}}(i-N) = \text{previously computed} \\ &\underline{P}(i-N) = \text{previously computed} \\ &\frac{\partial \hat{\underline{x}}(i-N)}{\partial a_\ell} = \underline{0} \end{aligned}\right\} \qquad (3.4.12)$$

$$\left.\begin{aligned} &\frac{\partial \underline{P}(i-N)}{\partial a_\ell} = \underline{0} \\ &E\left\{ \frac{\partial \hat{\underline{x}}(i-N)}{\partial a_k} \; \frac{\partial \hat{\underline{x}}^T(i-N)}{\partial a_\ell} \;\middle|\; \underline{a}_*(i) \right\} = \underline{0} \end{aligned}\right\} \quad (3.4.12)$$

Thus, the likelihood equations, (3.4.1) and (3.4.2), can be solved by the scoring iteration in the following manner. Assume for the present that a state and parameter estimate is required at every time step. Let the time instant be i, and let $\underline{a}_*(i)$ be the parameter estimate achieved at the previous sample time, (i-1). The state estimate is obtained as the N-step propagation from $\hat{\underline{x}}(i-N)$ and $\underline{P}(i-N)$ to $\hat{\underline{x}}(i)$, using the value $\underline{a}_*(i)$ for the computations: $\hat{\underline{x}}(i\,;\,\underline{a}_*(i))$. The terms needed for the parameter estimate are evaluated simultaneously, so that at each step, j, for j = (i-N+1), (i-N+2), ... (i-1),i, $\hat{\underline{x}}(j;\underline{a}_*(i))$ and $\underline{P}(j;\underline{a}_*(i))$ are computed, $\underline{s}^1[\underline{Z}(j),\underline{a}_*(i)]$ calculated and added to the running sum for the score, and $E\{\underline{s}^1[\underline{Z}(j),\underline{a}_*(i)] \cdot \cdot \underline{s}^{1T}[\underline{Z}(j),\underline{a}_*(i)] \,|\, \underline{a}_*(i)\}$ evaluated and added to the running sum comprising the conditional information matrix. At the end of the N-step propagation, the vector $\underline{\gamma}[\underline{Z}(i),\underline{a}_*(i)]$ is evaluated and added to the score running sum, and similarly, $E\{\underline{\gamma}[\underline{Z}(i),\underline{a}_*(i)]\,\underline{\gamma}^T[\underline{Z}(i),\underline{a}_*(i)] \,|\, \underline{a}_*(i)\}$ is evaluated and added to the conditional information matrix running sum. Finally the new parameter estimate is made by solving

$$\underline{a}^*(i) = \underline{a}_*(i) + \underline{J}[\,i\,,\,\underline{x}_*(i)\,,\,\underline{a}_*(i)\,]^{-1} \frac{\partial L^T}{\partial \underline{a}}\,[\underline{x}_*(i),\underline{a}_*(i),\underline{Z}(i)] \quad (3.4.13)$$

The entire process could be iterated locally by setting $\underline{a}_*(i)$ equal to this value of $\underline{a}^*(i)$, and recalculating the entire procedure. When satisfactory convergence is obtained, the resulting $\underline{a}^*(i)$ is considered to be an adequate representation

of the maximum likelihood estimate of $\underline{a}$. Local iterations require considerable amounts of computation and are probably restricted from on-line applications. In fact, since the parameters are expected to vary slowly, it may well be acceptable for on-line usage to re-evaluate the parameter estimate only periodically instead of every sample.

The recursions will now be described in detail. To propagate from just after the measurement at time (j-1), to just before the measurement at time j (for j = i-N+1, i-N+2, ..., i-1,i), the relations are, where it is implicit that $\underline{a}_*(i)$ is used for the computations:

I) state related equations:

$$\bar{\underline{x}}(j) = \underline{\Phi}(j,j-1)\hat{\underline{x}}(j-1) + \underline{B}(j-1)\underline{u}(j-1) \qquad (3.4.14)$$

$$\underline{M}(j) = \underline{\Phi}(j,j-1)\underline{P}(j-1)\underline{\Phi}^T(j,j-1) + \underline{G}(j-1)\underline{Q}(j-1)\underline{G}^T(j-1) \qquad (3.4.15)$$

$$\underline{A}(j) = \underline{H}(j)\underline{M}(j)\underline{H}^T(j) + \underline{R}(j) \qquad (3.4.16)$$

$$\underline{\eta}(j) = \underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j)] \qquad (3.4.17)$$

These are the usual optimal linear state estimate propagation equations, but the $\underline{\Phi}$ and $\underline{B}$ matrices are evaluated using the most recent parameter estimate, $\underline{a}_*(i)$. The initial conditions of $\hat{\underline{x}}(i-N)$ and $\underline{P}(i-N)$ are obtained from the computations at the previous sample time, (i-1). The vector $\underline{\eta}(j)$ is defined to decrease subsequent computations.

II) score equations:

$$\underline{\Omega}(j) = \underline{A}^{-1}(j) - \underline{\eta}(j)\underline{\eta}^T(j) \qquad (3.4.18)$$

$$\frac{\partial \underline{M}(j)}{\partial a_\ell} = \frac{\partial \underline{\Phi}(j,j-1)}{\partial a_\ell} \underline{P}(j-1)\underline{\Phi}^T(j,j-1) + \underline{\Phi}(j,j-1)\underline{P}(j-1) \frac{\partial \underline{\Phi}^T(j,j-1)}{\partial a_\ell}$$
$$+ \underline{\Phi}(j,j-1) \frac{\partial \underline{P}(j-1)}{\partial a_\ell} \underline{\Phi}^T(j,j-1) \qquad (3.4.19)$$

$$= \underline{0} \text{ if } a_\ell = \text{a parameter in } \underline{B} \qquad (3.4.19')$$

$$\frac{\partial \underline{A}(j)}{\partial a_\ell} = \underline{H}(j) \frac{\partial \underline{M}(j)}{\partial a_\ell} \underline{H}^T(j) \qquad (3.4.20)$$

$$= \underline{0} \text{ if } a_\ell = \text{a parameter in } \underline{B} \qquad (3.4.20')$$

$$\frac{\partial \bar{\underline{x}}(j)}{\partial a_\ell} = \underline{\Phi}(j,j-1) \frac{\partial \hat{\underline{x}}(j-1)}{\partial a_\ell} + \frac{\partial \underline{\Phi}(j,j-1)}{\partial a_\ell} \hat{\underline{x}}(j-1)$$
$$+ \frac{\partial \underline{B}(j-1)}{\partial a_\ell} \underline{u}(j-1) \qquad (3.4.21)$$

$$s^1{}_\ell[\underline{Z}(j), \underline{a}_*(i)] = \frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell} \underline{H}^T(j)\underline{\eta}(j) - \tfrac{1}{2} \operatorname{tr}\left\{\underline{\Omega}(j) \frac{\partial \underline{A}(j)}{\partial a_\ell}\right\} \qquad (3.4.22)$$

The matrix $\underline{\Omega}(j)$ is defined to decrease the number of computations. If a_ℓ is the α-β-th entry in the $\underline{\Phi}$ matrix and no other entries are functionally related to it, then

$$\frac{\partial \underline{\Phi}}{\partial a_\ell} = \underline{q}_\alpha \underline{q}_\beta{}^T \qquad (3.4.23)$$

where $\underline{q}_\alpha$ is a vector of all zeroes except a one in the α-th component. A similar relation is true for $\partial \underline{B}/\partial a_\ell$ with a_ℓ an entry in $\underline{B}$, and such forms can lead to computational efficiency. If a_ℓ is a parameter in the dynamics of a continuous plant, then both $\partial \underline{\Phi}/\partial a_\ell$ and $\partial \underline{B}/\partial a_\ell$ are in general nonzero, unless a_ℓ is in the continuous-time control input

matrix, for which $\partial\underline{\Phi}/\partial a_\ell$ is zero generally. Note that no term of the form $\partial\underline{u}/\partial a_\ell$ appears: when at time i, the control inputs $\underline{u}(i-N)$, ... $\underline{u}(i-1)$ have already been applied, so even if the control is of a feedback nature, $\partial\underline{u}(j)/\partial a_\ell = \underline{0}$ for all j less than i. The initial conditions for these recursions are that $\partial\underline{P}(i-N)/\partial a_\ell$ is the zero matrix, and $\partial\hat{\underline{x}}(i-N)/\partial a_\ell$ is the zero vector.

III) conditional information matrix computations:

$$E\left\{\frac{\partial\bar{\underline{x}}(j)}{\partial a_k}\ \frac{\partial\bar{\underline{x}}^T(j)}{\partial a_\ell}\ \middle|\ \underline{a}_*(i)\right\} =$$

$$= \underline{\Phi}(j,j-1)\ E\left\{\frac{\partial\hat{\underline{x}}(j-1)}{\partial a_k}\ \frac{\partial\hat{\underline{x}}^T(j-1)}{\partial a_\ell}\ \middle|\ \underline{a}_*(i)\right\}\ \underline{\Phi}^T(j,j-1)$$

$$+ \frac{\partial\underline{\Phi}}{\partial a_k}\ E\{\hat{\underline{x}}\hat{\underline{x}}^T|\underline{a}_*(i)\}\ \frac{\partial\underline{\Phi}^T}{\partial a_\ell} + \frac{\partial\underline{B}}{\partial a_k}\ E\{\underline{u}\underline{u}^T|\underline{a}_*(i)\}\ \frac{\partial\underline{B}^T}{\partial a_\ell}$$

$$+ \underline{\Phi}\ E\left\{\frac{\partial\hat{\underline{x}}}{\partial a_k}\ \hat{\underline{x}}^T\middle|\underline{a}_*(i)\right\}\ \frac{\partial\underline{\Phi}^T}{\partial a_\ell} + \frac{\partial\underline{\Phi}}{\partial a_k}\ E\left\{\hat{\underline{x}}\ \frac{\partial\hat{\underline{x}}^T}{\partial a_\ell}\middle|\underline{a}_*(i)\right\}\ \underline{\Phi}^T$$

$$+ \underline{\Phi}\ E\left\{\frac{\partial\hat{\underline{x}}}{\partial a_k}\ \underline{u}^T\middle|\underline{a}_*(i)\right\}\ \frac{\partial\underline{B}^T}{\partial a_\ell} + \frac{\partial\underline{B}}{\partial a_k}\ E\left\{\underline{u}\ \frac{\partial\hat{\underline{x}}^T}{\partial a_\ell}\middle|\underline{a}_*(i)\right\}\ \underline{\Phi}^T$$

$$+ \frac{\partial\underline{\Phi}}{\partial a_k}\ E\{\hat{\underline{x}}\underline{u}^T|\underline{a}_*(i)\}\ \frac{\partial\underline{B}^T}{\partial a_\ell} + \frac{\partial\underline{B}}{\partial a_k}\ E\{\underline{u}\hat{\underline{x}}^T|\underline{a}_*(i)\}\ \frac{\partial\underline{\Phi}^T}{\partial a_\ell} \qquad (3.4.24)$$

$$E\{\bar{\underline{x}}(j)\bar{\underline{x}}^T(j)|\underline{a}_*(i)\} = \underline{\Phi}(j,j-1)\ E\{\hat{\underline{x}}(j-1)\hat{\underline{x}}^T(j-1)|\underline{a}_*(i)\}\ \underline{\Phi}^T(j,j-1)$$

$$+ \underline{\Phi}\ E\{\hat{\underline{x}}\underline{u}^T|\underline{a}_*(i)\}\ \underline{B}^T + \underline{B}\ E\{\underline{u}\hat{\underline{x}}^T|\underline{a}_*(i)\}\ \underline{\Phi}^T$$

$$+ \underline{B}\ E\{\underline{u}\underline{u}^T|\underline{a}_*(i)\}\ \underline{B}^T \qquad (3.4.25)$$

$$E\left\{\frac{\partial \bar{\underline{x}}(j)}{\partial a_\ell}\,\bar{\underline{x}}^T(j)\,\middle|\,\underline{a}_*(i)\right\} = \underline{\Phi}(j,j-1)E\left\{\frac{\partial \hat{\underline{x}}(j-1)}{\partial a_\ell}\,\hat{\underline{x}}^T(j-1)\,\middle|\,\underline{a}_*(i)\right\}\underline{\Phi}^T(j,j-1)$$
$$+\frac{\partial \underline{\Phi}}{\partial a_\ell}E\{\hat{\underline{x}}\hat{\underline{x}}^T|\underline{a}_*(i)\}\,\underline{\Phi}^T+\frac{\partial \underline{B}}{\partial a_\ell}E\{\underline{u}\underline{u}^T|\underline{a}_*(i)\}\,\underline{B}^T$$
$$+\frac{\partial \underline{\Phi}}{\partial a_\ell}E\{\hat{\underline{x}}\underline{u}^T|\underline{a}_*(i)\}\,\underline{B}^T+\frac{\partial \underline{B}}{\partial a_\ell}E\{\underline{u}\hat{\underline{x}}^T|\underline{a}_*(i)\}\,\underline{\Phi}^T$$
$$+\underline{\Phi}\,E\left\{\frac{\partial \hat{\underline{x}}}{\partial a_\ell}\,\underline{u}^T\,\middle|\,\underline{a}_*(i)\right\}\underline{B}^T \qquad (3.4.26)$$

$$E\left\{s^1{}_k\,[\underline{Z}(j),\underline{a}_*(i)]\;\;s^1{}_\ell\,[\underline{Z}(j),\underline{a}_*(i)]\;\middle|\;\underline{a}_*(i)\right\} =$$
$$=\tfrac{1}{2}\,\mathrm{tr}\left\{\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_k}\,\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_\ell}\right.$$
$$\left.+\,2\,\underline{A}^{-1}(j)\underline{H}(j)E\left[\frac{\partial \bar{\underline{x}}(j)}{\partial a_k}\,\frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell}\,\middle|\,\underline{a}_*(i)\right]\underline{H}^T(j)\right\} \qquad (3.4.27)$$

Notice that equations (3.4.25) and (3.4.26) are needed to define certain terms in (3.4.24), and that once (3.4.26) is computed, both it and its transpose will be used. The initial conditions on (3.4.24) and (3.4.26) are the zero matrix, and $E\{\hat{\underline{x}}(i-N)\hat{\underline{x}}^T(i-N)\}$ is obtained from the calculations at the previous sample instant. The expectations involving $\underline{u}$ are evaluated in one of two ways. If the $\underline{u}(j)$ history is completely precomputed (as especially zero), then $E\{\underline{u}\;(\;)^T|\,\underline{a}_*(i)\}$ becomes $\underline{u}\;E\{(\;)^T|\,\underline{a}_*(i)\}$, and thus the following recursions must be defined:

$$E\{\hat{\underline{x}}(j)|\,\underline{a}_*(i)\} = \underline{x}^\circ(j;\hat{\underline{x}}(i-N))$$
$$=\underline{\Phi}(j,j-1)E\{\hat{\underline{x}}(j-1)|\,\underline{a}_*(i)\}+\underline{B}(j-1)\underline{u}(j-1) \qquad (3.4.28)$$

$$E\left\{\frac{\partial \hat{\underline{x}}(j)}{\partial a_\ell}\middle|\underline{a}_*(i)\right\} = \left[\underline{I}-\underline{K}(j)\underline{H}(j)\right]\left[\underline{\Phi}(j,j-1)E\left\{\frac{\partial \hat{\underline{x}}(j-1)}{\partial a_\ell}\middle|\underline{a}_*(i)\right\}\right.$$

$$\left. + \frac{\partial \underline{\Phi}(j,j-1)}{\partial a_\ell} E\{\hat{\underline{x}}(j-1)|\underline{a}_*(i)\} + \frac{\partial \underline{B}(j-1)}{\partial a_\ell}\underline{u}(j-1)\right] \tag{3.4.29}$$

If the control were instead computed in feedback form as $\underline{u}(j) = -\underline{C}(j)\hat{\underline{x}}(j)$, then the expectations of $E\{\underline{u}(\ )^T|\underline{a}_*(i)\}$ become $-\underline{C}\,E\{\hat{\underline{x}}(\ )^T|\underline{a}_*(i)\}$ and $E\{\underline{u}\underline{u}^T|\underline{a}_*(i)\}$ becomes $\underline{C}\,E\{\hat{\underline{x}}\hat{\underline{x}}^T|\underline{a}_*(i)\}\,\underline{C}^T$, for which the recursions have already been evaluated.

To incorporate the measurement at time instant j, the following relations are used:

I) state related equations:

$$\hat{\underline{x}}(j) = \bar{\underline{x}}(j) + \underline{M}(j)\underline{H}^T(j)\underline{\eta}(j) \tag{3.4.30}$$

$$\underline{K}(j) = \underline{M}(j)\underline{H}^T(j)\underline{A}^{-1}(j) \tag{3.4.31}$$

$$\underline{P}(j) = [\underline{I} - \underline{K}(j)\underline{H}(j)]\,\underline{M}(j)[\underline{I} - \underline{K}(j)\underline{H}(j)]^T + \underline{K}(j)\underline{R}(j)\underline{K}^T(j) \tag{3.4.32}$$

After the first time through these relations, the values of $\hat{\underline{x}}(i-N+1)$ and $\underline{P}(i-N+1)$ are saved in storage to serve as the initial conditions at the next sample time, (i+1).

II) score equations:

$$\frac{\partial \underline{P}(j)}{\partial a_\ell} = [\underline{I} - \underline{K}(j)\underline{H}(j)]\,\frac{\partial \underline{M}(j)}{\partial a_\ell}\,[\underline{I} - \underline{K}(j)\underline{H}(j)]^T \tag{3.4.33}$$

$$= \underline{0} \text{ if } a_\ell = \text{a parameter in } \underline{B} \tag{3.4.33'}$$

$$\frac{\partial \hat{\underline{x}}(j)}{\partial a_\ell} = \left[\underline{I} - \underline{K}(j)\underline{H}(j)\right]\left[\frac{\partial \bar{\underline{x}}(j)}{\partial a_\ell} + \frac{\partial \underline{M}(j)}{\partial a_\ell}\,\underline{H}^T(j)\underline{\eta}(j)\right] \qquad (3.4.34)$$

$$= \left[\underline{I} - \underline{K}(j)\underline{H}(j)\right]\left[\frac{\partial \bar{\underline{x}}(j)}{\partial a_\ell}\right] \quad \text{if } a_\ell = \text{a parameter in } \underline{B} \qquad (3.4.34')$$

It should be noted that, despite its simple appearance, equation (3.4.33) does account for the variation of $\underline{K}(j)$ with respect to a_ℓ; this can be verified by taking the partial of equation (3.4.32) with respect to a_ℓ, expressing $\partial \underline{K}(j)/\partial a_\ell$ in terms of $\partial \underline{M}(j)/\partial a_\ell$, and collecting like terms.

III) conditional information matrix computations:

$$E\left\{\frac{\partial \hat{\underline{x}}(j)}{\partial a_k}\,\frac{\partial \hat{\underline{x}}^T(j)}{\partial a_\ell}\,\middle|\,\underline{a}_*(i)\right\} = \left[\underline{I}-\underline{KH}\right]E\left\{\frac{\partial \bar{\underline{x}}(j)}{\partial a_k}\,\frac{\partial \bar{\underline{x}}^T(j)}{\partial a_\ell}\,\middle|\,\underline{a}_*(i)\right\}\left[\underline{I}-\underline{KH}\right]^T$$
$$+ \left[\underline{I}-\underline{KH}\right]\frac{\partial \underline{M}}{\partial a_k}\,\underline{H}^T\underline{A}^{-1}\underline{H}\,\frac{\partial \underline{M}}{\partial a_\ell}\left[\underline{I}-\underline{KH}\right]^T \qquad (3.4.35)$$

$$E\left\{\hat{\underline{x}}(j)\hat{\underline{x}}^T(j)\,\middle|\,\underline{a}_*(i)\right\} = E\left\{\bar{\underline{x}}(j)\bar{\underline{x}}^T(j)\,\middle|\,\underline{a}_*(i)\right\} + \underline{K}(j)\underline{A}(j)\underline{K}^T(j) \qquad (3.4.36)$$

$$E\left\{\frac{\partial \hat{\underline{x}}(j)}{\partial a_\ell}\,\hat{\underline{x}}^T(j)\,\middle|\,\underline{a}_*(i)\right\} = \left[\underline{I}-\underline{K}(j)\underline{H}(j)\right]\left[E\left\{\frac{\partial \bar{\underline{x}}(j)}{\partial a_\ell}\,\bar{\underline{x}}^T(j)\,\middle|\,\underline{a}_*(i)\right\}\right.$$
$$\left. + \frac{\partial \underline{M}(j)}{\partial a_\ell}\,\underline{H}^T(j)\underline{K}^T(j)\right] \qquad (3.4.37)$$

The second term in (3.4.35) and the second term within the brackets of equation (3.4.37) are zero if a_ℓ is a parameter in the $\underline{B}$ matrix.

At the end of the N-step recursion, $\underline{P}(i)$, $\partial \hat{\underline{x}}(i)/\partial a_\ell$, $\partial \underline{P}(i)/\partial a_\ell$, and $E\{\partial \hat{\underline{x}}(i)/\partial a_k \;\; \partial \hat{\underline{x}}^T(i)/\partial a_\ell \,|\, \underline{a}_*(i)\}$ will have been evaluated for all values of k and ℓ. These are precisely the terms required to form the final components of the score and conditional information matrix. Thus,

$$r_\ell[\underline{Z}(i),\underline{a}_*(i)] = -\tfrac{1}{2}\,\mathrm{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a}\right\} \tag{3.4.38}$$

and

$$E\left\{r_k[\underline{Z}(i),\underline{a}_*(i)]\;\; r_\ell[\underline{Z}(i),\underline{a}_*(i)]\,\middle|\,\underline{a}_*(i)\right\} =$$

$$= \tfrac{1}{2}\,\mathrm{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_k}\;\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_\ell} + 2\,\underline{P}^{-1}(i)\,E\left[\frac{\partial \hat{\underline{x}}(i)}{\partial a_k}\;\frac{\partial \hat{\underline{x}}^T(i)}{\partial a_\ell}\,\middle|\,\underline{a}_*(i)\right]\right\} \tag{3.4.39}$$

are computed and added to the running sums composing the score and conditional information matrix respectively.

The previous development pertained to an estimator based upon a likelihood function of $\ell n\, f(\underline{x}(i),\underline{Z}_N(i)\,|\,\underline{Z}(i\text{-}N),\underline{a})$. A duplicate procedure would be employed for an estimator derived from $\ell n\, f(\underline{x}(i),\underline{Z}_N(i)\,|\,\underline{a})$, with the only difference being that the initial conditions $\hat{\underline{x}}(i\text{-}N)$, $\underline{P}(i\text{-}N)$, and $E\{\hat{\underline{x}}(i\text{-}N)\hat{\underline{x}}^T(i\text{-}N)\}$ are replaced by $\tilde{\underline{x}}(i\text{-}N)$, $\tilde{\underline{P}}(i\text{-}N)$, and $\tilde{\underline{x}}(i\text{-}N)\tilde{\underline{x}}^T(i\text{-}N)$. The value of $\tilde{\underline{x}}(i\text{-}N)$ is defined as a one-step propagation from the previous value $\tilde{\underline{x}}(i\text{-}N\text{-}1)$ as

$$\tilde{\underline{x}}(i\text{-}N) = \underline{\Phi}(i\text{-}N,i\text{-}N\text{-}1;a_*(i))\;\tilde{\underline{x}}(i\text{-}N\text{-}1) + \underline{B}(i\text{-}N\text{-}1;a_*(i))\,\underline{u}(i\text{-}N\text{-}1) \tag{3.4.40}$$

for $i \geq (N+1)$, starting from the initial condition

$$\tilde{\underline{x}}(0) = \hat{\underline{x}}_\circ \tag{3.4.41}$$

Similarly, $\tilde{\underline{P}}(i\text{-}N)$ is obtained from

$$\tilde{\underline{P}}(i\text{-}N) = \underline{\Phi}(i\text{-}N,i\text{-}N\text{-}1;a_*(i))\;\tilde{\underline{P}}(i\text{-}N\text{-}1)\;\underline{\Phi}^T(i\text{-}N,i\text{-}N\text{-}1;a_*(i)) + \\ + \underline{G}(i\text{-}N\text{-}1)\;\underline{Q}(i\text{-}N\text{-}1)\;\underline{G}^T(i\text{-}N\text{-}1) \tag{3.4.42}$$

for $i \geq (N+1)$, from the initial condition $\tilde{\underline{P}}(0) = \underline{P}_\circ$.

CHAPTER IV

PERFORMANCE ANALYSIS

To be able to use this estimation technique with confidence, the user must be able to predict the performance achievable from its various implementations. This chapter provides two distinct performance analyses to satisfy this requirement.

Section 4.1 considers the asymptotic properties of a maximum likelihood state and parameter estimator: the behavior exhibited by the estimator as the number of measurements processed grows without bound. Although these properties can be strictly proven only as the number of data samples tends to infinity, they are not merely of theoretical significance. Rather, they delineate trends of estimator behavior as time progresses, and in most practical problems the actual behavior closely approximates that of the asymptotic properties well within the time interval of interest.

The concept of an ambiguity function is introduced in section 4.2. As the average value of the log-likelihood function, it provides both a global and local performance analysis: its shape indicates the ability of the log-likelihood function to produce a unique estimate, and the curvature at its peak value can be used to generate the Cramér-Rao lower bound on the estimate error covariance matrix. Moreover, it can be used as a design tool, since it can predict the sensitivity of the estimation accuracy to the length of the data interval N, the type and precision of measurements taken, assumed form of dynamics model, exclusion of certain terms from the likelihood equations, and other factors that the engineer can control to assure adequate performance. Finally, this versatile function can be evaluated entirely from the output of a state estimator sensitivity analysis, which would normally be performed before deciding upon parameter estimation to improve overall filter performance.

4.1 Asymptotic Properties

Most of the previous work concerning the properties of maximum likelihood estimators (as especially in the classical texts by Cramér (1946), Rao (1968), and Wilks (1963)) considers the case of measurements which are independent and identically distributed. Although these results are not directly applicable to the present situation, when they are combined with certain insights provided by Abramson (1968), they can serve to develop a valid description of our estimator's properties.

First some of the fundamental concepts used to describe an estimator will be delineated. These properties are not restricted to maximum likelihood estimators, and therefore they can be used to compare our estimator to others that may not depend on the maximum likelihood criterion. The reason for making this explicit is that one finds it difficult to evaluate the relative merits of maximum likelihood, Bayesian, weighted least squares, minimax, and other forms of estimators in a truly meaningful way.

Assume that it is desired to know the value of a certain parameter, θ. Further assume that its true value is θ_t, but that this value cannot be obtained with complete certainty. However, there is available a set of elements $(z_1, \ldots z_n)$ from a sample space ζ, which can be regarded as measurements upon the event or process under investigation. Suppose that an observable random variable, $\hat{\theta}(z_1, \ldots z_n)$ can be defined as a function of the sample elements $(z_1, \ldots z_n)$; that is, for every particular realization $(z_1, \ldots z_n)$ chosen out of the sample space ζ, the function $\hat{\theta}(z_1, \ldots z_n)$ assigns a real value. Further assume that the distribution of this random variable is "concentrated" about the true value θ_t in some sense. (The mean and covariance provide considerable information about the "concentration"; for normal random variables they specify the "concentration" completely.) Now, when the value of the random variable that results from a particular realization (set of measurements or observations), say $(z_1, \ldots z_n)$, is used to represent the true (unknown) value of the parameter, then this value, $\hat{\theta}(z_1, \ldots z_n)$ is called an estimator,

or point estimate, of θ_t.

One desirable property of an estimator is that it be unbiased. There are differing definitions of what unbiased means exactly, but one prevalent definition is that the conditional expectation of $\hat{\theta}$, given the true value θ_t, be θ_t itself. For $Z = (z_1, \ldots z_n)$, if

$$E\left\{\hat{\theta}|\theta_t\right\} = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \hat{\theta}(Z) f(Z|\theta_t) dZ = \theta_t \qquad (4.1.1)$$

then the random variable $\hat{\theta}$ is called an unbiased estimator of θ_t. This definition of unbiasedness is appropriate for an estimator that is dependent solely upon the measurements taken, and not upon any apriori information about the distribution of θ. Other definitions may well be more appropriate and natural in cases for which such apriori information is available, as in the case of Bayesian estimation for example. Eventually, the state and parameter estimation problem for the case of no apriori information about the parameters will be considered, so the definition given by equation (4.1.1) is suitable for the investigations into the bias of the parameter estimate. For the state estimate, the more familiar definition will be used: that an estimator is unbiased if, over an ensemble of trials, the expected value of the state estimate equals the expected value of the state.

An estimate $\hat{\theta}$ is termed consistent if it converges in probability to the true value θ_t, as the number of sample elements goes to infinity. If an estimate is unbiased and has a covariance which goes to zero as the number of sample elements goes to infinity, then it is a consistent estimate.

An efficient estimate $\hat{\theta}$ is an estimate of θ_t that satisfies the following conditions:

1) it is unbiased;
2) it has finite covariance;
3) there is no other unbiased estimate whose covariance is smaller than that of $\hat{\theta}$.

In general, an efficient estimate can only be formed under rather restrictive conditions (Cramér (1946)). However, asymptotically efficient estimates, that is, estimates which become efficient as

the number of sample elements goes to infinity (as, due to being asymptotically unbiased but biased for any finite number of sample elements) require only certain general regularity conditions (Cramér (1946)) to insure their existence. These conditions will be stated explicitly for the ML estimator later in this section.

This section will procede by proving certain properties of an estimator of the uncertain parameters which does not simultaneously estimate the state variables. These proofs were motivated by derivations by Wilks (for the case of estimating variables from independent, identically distributed samples) and by Abramson (for estimating noise covariances), both of whose work rely heavily upon the original proofs by Dugué in 1937. First, the behavior of a parameter estimator which uses all measurements from the initial time will be examined, assuming the parameters to be constant over the entire time interval of interest. Sections 4.1.2 to 4.1.5 present the pertinent derivations of asymptotic properties; each section is written in the format of a complete statement of the property, followed by the proof, so that the reader who wishes to omit the details of the proof may do so readily. Section 4.1.6 summarizes these characteristics. The results will then be related to an estimator which simultaneously solves for the state variables and the parameters. Finally, these will be used to determine the properties of a combined state-and-parameter estimator using a fixed number, N, of samples to provide a parameter estimate.

As shown in the previous chapter, by considering a log-likelihood function of

$$L[\underline{Z}(i),\underline{a}] = \ell n\ f(\underline{Z}(i)|\ \underline{a}) \tag{4.1.2}$$

the likelihood equations yield scalar equations for the estimates of a_ℓ, the ℓ-th component of the parameter vector, as

$$\sum_{j=1}^{i} \mathrm{tr}\left\{\left[\underline{A}^{-1}(j)-\underline{A}^{-1}(j)\underline{r}(j)\underline{r}^{T}(j)\underline{A}^{-1}(j)\right]\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\} - 2\sum_{j=1}^{i}\frac{\partial \underline{\bar{x}}^{T}(j)}{\partial a_\ell}\ \underline{H}^{T}(j)\underline{A}^{-1}(j)\underline{r}(j)\Bigg|_{\underline{a}\rightarrow\underline{\hat{a}}(i)} = 0 \tag{4.1.3}$$

where the residual $\underline{r}(j)$ and matrix $\underline{A}(j)$ have been defined as

$$\underline{r}(j) = \underline{z}(j) - \underline{H}(j)\overline{\underline{x}}(j) \tag{4.1.4}$$

$$\underline{A}(j) = \underline{H}(j)\underline{M}(j)\underline{H}^{T}(j) + \underline{R}(j) \tag{4.1.5}$$

Consistent with the conventional terminology, define the single measurement score, $\underline{s}^{1}[\underline{Z}(j),\underline{a}]$, as the vector composed of the components

$$s^{1}{}_{\ell}\,[\underline{Z}(j),\underline{a}] = \frac{\partial}{\partial a_{\ell}} \ell n\ f(\underline{z}(j)|\underline{Z}(j-1),\underline{a}) \tag{4.1.6a}$$

$$= -\tfrac{1}{2}\ \mathrm{tr}\left\{\left[\underline{A}^{-1}(j)-\underline{A}^{-1}(j)\underline{r}(j)\ \underline{r}^{T}(j)\underline{A}^{-1}(j)\right]\frac{\partial \underline{A}(j)}{\partial a_{\ell}}\right\}$$
$$+ \frac{\partial \overline{\underline{x}}^{T}(j)}{\partial a_{\ell}}\ \underline{H}^{T}(j)\underline{A}^{-1}(j)\underline{r}(j) \tag{4.1.6b}$$

It is a function of all measurements from the first up to, and including, the j-th measurement. Then the total measurement score at time instant i can be defined as $\underline{s}[\underline{Z}(i),\underline{a}]$, where

$$\underline{s}[\underline{Z}(i),\underline{a}] = \left[\frac{\partial}{\partial \underline{a}} \ell n\ f(\underline{Z}(i)|\underline{a})\right]^{T} \tag{4.1.7a}$$

$$= \sum_{j=1}^{i} \underline{s}^{1}[\underline{Z}(j),\underline{a}] \tag{4.1.7b}$$

Therefore, the likelihood equations, (4.1.3), can be written as

$$\underline{s}\,[\underline{Z}(i),\underline{a}] = \underline{0} \tag{4.1.8}$$

It is now possible to develop the conditional information matrix. Define the matrices $\underline{J}^{1}[\,j\,,\,\underline{a}_{t}\,,\,\underline{a}\,]$ and $\underline{J}[\,i\,,\,\underline{a}_{t}\,,\,\underline{a}\,]$ by means of their components as:

$$J^1_{k\ell}[\, j\,,\, \underline{a}_t\,,\, \underline{a}\,] = E\left\{ s^1_k[\underline{Z}(j),\underline{a}]\; s^1_\ell\,[\underline{Z}(j),\underline{a}] \,\middle|\, \underline{a}_t \right\} \tag{4.1.9}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} s^1_k[\underline{Z}(j),\underline{a}]\; s^1_\ell[\underline{Z}(j),\underline{a}]\; f(\underline{Z}(j)|\,\underline{a}_t)\; d\underline{Z}(j) \tag{4.1.10}$$

$$J_{k\ell}[\, i\,,\, \underline{a}_t\,,\, \underline{a}\,] = E\left\{ s_k\,[\underline{Z}(i),\underline{a}]\; s_\ell[\underline{Z}(i),\underline{a}] \,\middle|\, \underline{a}_t \right\} \tag{4.1.11}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} s_k[\underline{Z}(i),\underline{a}]\; s_\ell[\underline{Z}(i),\underline{a}]\; f(\underline{Z}(i)|\,\underline{a}_t)\; d\underline{Z}(i) \tag{4.1.12}$$

From these relations, the single measurement conditional information matrix, $\underline{J}^1[\, j\,,\, \underline{a}_t\,]$, can be defined as

$$\underline{J}^1[\, j\,,\, \underline{a}_t\,] = \underline{J}^1[\, j\,,\, \underline{a}_t\,,\, \underline{a}_t\,] \tag{4.1.13}$$

and the total measurement conditional information matrix, $\underline{J}[\, i\,,\, \underline{a}_t\,]$, as

$$\underline{J}[\, i\,,\, \underline{a}_t\,] = \underline{J}[\, i\,,\, \underline{a}_t\,,\, \underline{a}_t\,] \tag{4.1.14}$$

In order to utilize the results about asymptotic behavior from classical maximum likelihood estimation theory directly, independent, identically distributed measurements are required. Although the measurements themselves are not independent from sample to sample, linear filtering theory demonstrates that, if the parameter vector $\underline{a}$ assumes its true value $\underline{a}_t$, then the measurement residuals at different times are independent random variables with a normal distribution characterized by a zero mean and a covariance $\underline{A}(j)$:

$$E\left\{ \underline{r}(j) \,\middle|\, \underline{a}_t \right\} = E\left\{ [\underline{z}(j) - \underline{H}(j)\overline{\underline{x}}(j)] \,\middle|\, \underline{a}_t \right\} = \underline{0} \tag{4.1.15}$$

$$E\left\{ \underline{r}(j)\underline{r}^T(k) \middle| \underline{a}_t \right\} = E\left\{ \left[\underline{z}(j)-\underline{H}(j)\overline{\underline{x}}(j)\right]\left[\underline{z}(k)-\underline{H}(k)\overline{\underline{x}}(k)\right]^T \middle| \underline{a}_t \right\}$$

$$= \left[\underline{H}(j)\underline{M}(j;\underline{a}_t)\underline{H}^T(j) + \underline{R}(j)\right] \delta_{jk} \tag{4.1.16}$$

where $\underline{M}(j;\underline{a}_t)$ is written to demonstrate the explicit dependence of $\underline{M}(j)$ on the value $\underline{a}_t$, and where δ_{jk} is the Kronecker delta.

Using this information and properties of the linear stochastic system driven by Gaussian white noise, it can be shown (see Appendix E) that the expected values (conditioned on the true value of the parameters, $\underline{a}_t$) of the score and conditional information matrix for a single measurement are

$$E\left\{ s^1{}_\ell \left[\underline{Z}(j),\underline{a}_t\right] \middle| \underline{a}_t \right\} = 0 \tag{4.1.17}$$

$$E\left\{ s^1{}_k\left[\underline{Z}(i),\underline{a}_t\right] s^1{}_\ell \left[\underline{Z}(j),\underline{a}_t\right] \middle| \underline{a}_t \right\}$$

$$= \tfrac{1}{2} \operatorname{tr}\left[\underline{A}^{-1}(j) \frac{\partial \underline{A}(j)}{\partial a_k} \underline{A}^{-1}(j) \frac{\partial \underline{A}(j)}{\partial a_\ell} + 2\, \underline{A}^{-1}(j)\underline{H}(j) E\left\{ \frac{\partial \overline{\underline{x}}(j)}{\partial a_k} \frac{\partial \overline{\underline{x}}^T(j)}{\partial a_\ell} \middle| \underline{a}_t \right\} \underline{H}^T(j) \right] \delta_{ij} \tag{4.1.18}$$

where

$$\underline{A}(j) = \underline{H}(j)\underline{M}(j)\underline{H}^T(j) + \underline{R}(j) \tag{4.1.19}$$

and where k and ℓ take on all values 1, 2, ... p, where p is the dimension of the parameter vector. Now equation (4.1.18) reveals that $\underline{s}^1[\underline{Z}(j),\underline{a}_t]$ and $\underline{s}^1[\underline{Z}(i),\underline{a}_t]$ are independent for $i \neq j$. Therefore, to obtain the total measurement score and conditional information matrix, it is valid to add the single measurement scores and matricestogether:

$$E\left\{ \underline{s}\,[\underline{Z}(i),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} = \underline{0} \tag{4.1.20}$$

$$E\left\{ s_k\,[\underline{Z}(i),\underline{a}_t]\; s_\ell\,[\underline{Z}(i),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} = J_{k\ell}[\,i\,,\,\underline{a}_t\,] =$$

$$= \tfrac{1}{2} \sum_{j=1}^{i} \mathrm{tr}\left[\underline{A}^{-1}(j) \frac{\partial \underline{A}(j)}{\partial a_k} \underline{A}^{-1}(j) \frac{\partial \underline{A}(j)}{\partial a_\ell} + 2\,\underline{A}^{-1}(j)\underline{H}(j) E\left\{ \frac{\partial \underline{\tilde{x}}(j)}{\partial a_k} \frac{\partial \underline{\tilde{x}}^{T}(j)}{\partial a_\ell} \,\middle|\, \underline{a}_t \right\} \underline{H}^{T}(j) \right] \tag{4.1.21}$$

Thus, by conditioning the densities upon the true value of the parameter vector, $\underline{a}_t$, it has been possible to obtain independent measurement residuals. Although these are not identically distributed, Abramson (1968) first demonstrated the ability to manipulate this type of statistical description in order to obtain asymptotic properties very similar to those of the "classical" maximum likelihood estimators.

4.1.1 Regularity Conditions

The following properties of maximum likelihood estimators are dependent upon certain regularity conditions, which will now be made explicit. Suppose $\underline{Z}(i)$ is a vector random variable with a distribution function $F(\underline{Z}(i)|\,\underline{a})$, or a density function $f(\underline{Z}(i)|\,\underline{a})$, which is dependent upon a parameter vector $\underline{a}$. The parameter $\underline{a}$ can take on values in a parameter space Θ, and for each particular point $\underline{\tilde{a}}$ in the space Θ there is a corresponding $F(\underline{Z}(i)|\,\underline{\tilde{a}})$ or $f(\underline{Z}(i)|\,\underline{\tilde{a}})$. Of primary interest will be an open set $\tilde{\Theta}$ of admissible parameter values within the space Θ containing a particular value, $\underline{a}_t$, the true value of the parameter.

It is assumed that a density function does exist, and a basic property of such a density is that

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\underline{Z}(i)|\,\underline{a})\; d\underline{Z}(i) = 1 \tag{4.1.22}$$

"Regularity conditions" are sufficient conditions under which this equation can be differentiated under the integral one or more times with respect to components of $\underline{a}$. Regularity to a certain order is a weaker condition than the existence everywhere of derivatives of $f(\underline{Z}(i)|\underline{a})$ to that order, and regularity is all that is needed in the proofs to follow. If the differentiation were performed formally upon equation (4.1.22), then since the integral is equal to unity for all values of $\underline{a}$, its partial with respect to $\underline{a}$ will vanish. Combining this with the fact that $f = \exp(\ell n\ f)$,

$$\frac{\partial}{\partial a_k} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\underline{Z}(i)|\underline{a})\, d\underline{Z}(i) =$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial a_k} \ell n\ f(\underline{Z}(i)|\underline{a})\right] f(\underline{Z}(i)|\underline{a})\, d\underline{Z}(i)$$

$$= E\left\{\left[\frac{\partial}{\partial a_k} \ell n\ f(\underline{Z}(i)|\underline{a})\right] \middle| \underline{a}\right\} = 0 \qquad (4.1.23)$$

$$\frac{\partial^2}{\partial a_k \partial a_\ell} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\underline{Z}(i)|\underline{a})\, d\underline{Z}(i) =$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\frac{\partial^2}{\partial a_k \partial a_\ell} \ell n\ f(\underline{Z}(i)|\underline{a})\right] f(\underline{Z}(i)|\underline{a})\, d\underline{Z}(i)$$

$$+ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial a_k} \ell n\ f(\underline{Z}(i)|\underline{a})\right]\left[\frac{\partial}{\partial a_\ell} \ell n\ f(\underline{Z}(i)|\underline{a})\right] f(\underline{Z}(i)|\underline{a})\, d\underline{Z}(i)$$

$$= 0 \qquad (4.1.24)$$

One is interested in the validity of these two equations for k and ℓ equal to 1, 2, ... p, where p is the dimension of $\underline{a}$, and for all points in the open set Θ.

It can be shown that a sufficient condition for equation (4.1.23) to be valid is that $[\partial/\partial a_k\ \ell n\ f(\underline{Z}(i)|\underline{a})]$ be dominated by an integrable function $h_k(\underline{Z})$, for all k. Similarly, for (4.1.24)

to be valid, $[\partial^2/\partial a_k \partial a_\ell \ \ell n\ f(\underline{Z}(i)|\underline{a})]$ and $[\partial/\partial a_k \ \ell n\ f(\underline{Z}(i)|\underline{a})] \cdot [\partial/\partial a_\ell \ \ell n\ f(\underline{Z}(i)|\underline{a})]$ must be dominated by the integrable functions $h'_{k\ell}(\underline{Z})$ and $h''_{k\ell}(\underline{Z})$, respectively, for all k and ℓ. To be precise, a given measurable function $\phi(\underline{Z}(i)|\underline{a})$ is said to be dominated by the function $h(\underline{Z})$ if there exists a non-negative function $h(\underline{Z})$, measurable and having a finite mean value, such that

$$|\phi(\underline{Z}(i)|\underline{a})| < h(\underline{Z}) \tag{4.1.25}$$

The distribution $F(\underline{Z}(i)|\underline{a})$ is said to be regular with respect to its first partial $\underline{a}$-derivatives in the open parameter set $\tilde{\Theta}$ if equation (4.1.23) does hold:

$$E\left\{\left[\frac{\partial}{\partial a_k} \ell n\ f(\underline{Z}(i)|\underline{a})\right]\middle|\underline{a}\right\} = 0 \qquad k = 1, 2, \ldots p \tag{4.1.26}$$

where this is identifiable as the expectation of the score. The distribution is further termed regular with respect to its second $\underline{a}$-partial derivatives in $\tilde{\Theta}$ if $\| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} [\partial/\partial a_k \ \ell n\ f(\underline{Z}(i)|\underline{a})] \cdot [\partial/\partial a_\ell \ \ell n\ f(\underline{Z}(i)|\underline{a})]\ f(\underline{Z}(i)|\underline{a})\ d\underline{Z}(i) \|$ is finite and if equation (4.1.24) holds:

$$E\left\{\left[\frac{\partial}{\partial a_k} \ell n\ f(\underline{Z}(i)|\underline{a})\right]\left[\frac{\partial}{\partial a_\ell} \ell n\ f(\underline{Z}(i)|\underline{a})\right]\middle|\underline{a}\right\} + E\left\{\left[\frac{\partial^2}{\partial a_k \partial a_\ell} \ell n\ f(\underline{Z}(i)|\underline{a})\right]\middle|\underline{a}\right\} = 0 \qquad k = 1, 2, \ldots p \tag{4.1.27}$$

In all of the following proofs, the regularity of a distribution with respect to first and second $\underline{a}$-partial derivatives can be replaced by the more restrictive conditions that the corresponding density have finite first and second $\underline{a}$-partial derivatives everywhere, respectively.

4.1.2 Asymptotic Distribution of the Score

Let $(\underline{z}(1), \ldots \underline{z}(i))$ be a set of samples such that each $\underline{z}(j)$ is taken from the conditional probability density function $f(\underline{z}(j)|\underline{Z}(j-1),\underline{a}_t)$ for $j = 1, 2, \ldots i$, and let $F(\underline{z}(j)|\underline{Z}(j-1),\underline{a}_t)$ be regular with respect to its first $\underline{a}$-partial derivatives in the open set $\tilde{\Theta}$ within the entire admissible range of parameter values, Θ. If the matrix $\underline{J}[\,i\,,\underline{a}\,,\underline{a}\,]$ is positive for $\underline{a}$ in the range Θ, then the total measurement score $\underline{s}[\underline{Z}(i),\underline{a}_t]$ is asymptotically distributed for large i as a normal random variable with zero mean and covariance equal to $\underline{J}[\,i\,,\underline{a}_t\,]$:

$$f(\underline{s}[\underline{Z}(i),\underline{a}_t]\,|\,\underline{Z}(i),\underline{a}_t) \rightarrow N\left[\underline{0},\underline{J}[\,i\,,\underline{a}_t\,]\right] \qquad (4.1.28)$$

$$\text{as } i \rightarrow \infty$$

The proof is as follows. Equations (4.1.20) and (4.1.21) show that $\underline{s}[\underline{Z}(i),\underline{a}_t]$ is, for all i, a random variable with zero mean and covariance equal to $\underline{J}[\,i\,,\underline{a}_t\,]$. Now it is necessary to prove that the distribution becomes normal in the limit as i grows to infinity, and thus that the distribution becomes totally characterized by its mean and covariance. By its definition, given by equation (4.1.7), the total measurement score $\underline{s}[\underline{Z}(i),\underline{a}_t]$ is the sum of i independent single measurement scores. Under the assumption that no one of these individual terms takes on a large enough value with sufficient probability to dominate the sum, then since $\underline{J}[\,i\,,\underline{a}\,,\underline{a}\,]$ is assumed to be positive definite for $\underline{a}$ in Θ (a necessary condition for the following), then the Central Limit Theorem can be invoked in the limit as $i \rightarrow \infty$: that under certain general conditions, the distribution of the sum of independent random variables approaches a normal distribution as the number of variables increases, regardless of the distributions of the individual variables. One set of such conditions is that the sum of the variances of the individual variables becomes infinite and that for some $\alpha > 2$, $\int_{-\infty}^{\infty} x^{\alpha} f_i(x)dx < C$ = constant; these are not the most general conditions, but are very applicable. The

fourth assumption in the summary section, 4.1.6, is such a third moment condition. Papoulis (1965) points out that the normal curve is a good approximation even for moderate sample sizes if each of the densities is reasonably concentrated near its mean. For example, when uniform distributions are considered for each variable, the sum of three such variables is remarkably well represented by the normal approximation. Thus, it has been shown that the score $\underline{s}[\underline{Z}(i),\underline{a}_t]$ is asymptotically normally distributed as $i \rightarrow \infty$, and the proof is complete.

4.1.3 Convergence of the Maximum Likelihood Estimator

Let $(\underline{z}(1), \ldots \underline{z}(i))$ be a set of samples such that each $\underline{z}(j)$ is taken from the conditional probability density function $f(\underline{z}(j)|\underline{Z}(j-1),\underline{a}_t)$ for $j = 1, 2, \ldots i$, and let $F(\underline{z}(j)|\underline{Z}(j-1),\underline{a}_t)$ be regular with respect to its first $\underline{a}$-partial derivatives in the open set $\tilde{\Theta}$ within the admissible range of parameter values, Θ. Let each component of the single measurement score, $\underline{s}^1[\underline{Z}(i),\underline{a}]$, be a continuous function of $\underline{a}$ in Θ for all values of $\underline{Z}(i)$ (in real Euclidean space) except possibly for a set of probability zero. If, as i becomes infinitely large, the value of $\underline{J}^{-1}[\,i\,,\underline{a}_t\,]$ tends to the zero matrix:

$$\text{if } \underline{J}^{-1}[\,i\,,\underline{a}_t\,] \rightarrow \underline{0} \text{ as } i \rightarrow \infty \qquad (4.1.29)$$

then there exists a sequence of solutions to the likelihood equations, given by (4.1.8), which converges in probability to $\underline{a}_t$. If the solution is a unique vector, $\hat{\underline{a}}(i)$, for i greater than some i_o, then the sequence of vectors $\hat{\underline{a}}(i)$ converges in probability to $\underline{a}_t$ as i tends to infinity.

In order to prove this, it must be shown that, with probability tending to 1 as i becomes infinitely large, the likelihood equation (4.1.8) has a solution value $\hat{\underline{a}}(i)$, under the assumed conditions, that lies between the limits $(\underline{a}_t \pm \underline{\delta})$, no matter how small the positive vector quantity $\underline{\delta}$ is chosen to be. First, define the vector $\underline{g}^1[\,j\,,\underline{a}_t\,,\underline{a}\,]$ as

$$\underline{\sigma}^{1}[\,j\,,\,\underline{a}_t\,,\,\underline{a}\,] = E\left\{\,\underline{s}^{1}[\underline{Z}(j),\underline{a}]\,\middle|\,\underline{a}_t\right\}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\underline{s}^{1}[\underline{Z}(j),\underline{a}]\,f(\underline{Z}(j)|\,\underline{a}_t)\;d\underline{Z}(j) \tag{4.1.30}$$

and the vector $\bar{\underline{\sigma}}[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]$, which is the average of the individual $\underline{\sigma}^{1}[\,j\,,\,\underline{a}_t\,,\,\underline{a}\,]$ vectors, as

$$\bar{\underline{\sigma}}[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,] = \frac{1}{i}\sum_{j=1}^{i}\underline{\sigma}^{1}[\,j\,,\,\underline{a}_t\,,\,\underline{a}\,] \tag{4.1.31}$$

It follows from equation (4.1.17) and the assumptions of regularity of the distribution and continuity of the score, that the components of $\bar{\underline{\sigma}}[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]$ are continuous and strictly decreasing in $\underline{a}$ over some subrange $\tilde{\Theta}$ of Θ which includes $\underline{a}_t$ (see Wilks (1963)).

From the relation

$$\frac{1}{i}\,\underline{s}[\underline{Z}(i),\underline{a}] = \frac{1}{i}\sum_{j=1}^{i}\underline{s}^{1}[\underline{Z}(j),\underline{a}] \tag{4.1.32}$$

it is seen that $\frac{1}{i}\,\underline{s}\,[\underline{Z}(i),\underline{a}]$ is the sample mean of a sample of size i taken from a population whose mean is $\bar{\underline{\sigma}}\,[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]$. Using this fact and the assumption that $\underline{J}^{-1}\,[\,i\,,\,\underline{a}_t\,]$ tends to zero as i grows, the weak law of large numbers can be utilized to conclude that $\frac{1}{i}\,\underline{s}[\underline{Z}(i),\underline{a}]$converges in probability to $\bar{\underline{\sigma}}[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]$, which is to say, for arbitrary $\epsilon>0$, and k = 1, 2, ... p,

$$\lim_{i\to\infty} P\left\{\left|\frac{1}{i}\,s_k[\underline{Z}(i),\underline{a}] - \bar{\sigma}_k[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]\right| \geq \epsilon\right\} = 0 \tag{4.1.33}$$

Now, as shown above, $\bar{\underline{\sigma}}[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]$ is strictly decreasing over $\tilde{\Theta}$, which includes $\underline{a}_t$. Without any loss of generality, let $\tilde{\Theta}$ be the range $(\underline{a}_t-\underline{\delta},\ \underline{a}_t+\underline{\delta})$ for a positive $\underline{\delta}$. Since equation (4.1.20) yields

$$\underline{\bar{\sigma}}\,[\,i\,,\,\underline{a}_t\,,\,\underline{a}_t\,] \;=\; E\left\{\,\underline{s}[\underline{Z}(i),\underline{a}_t]\;\middle|\;\underline{a}_t\,\right\} \;=\; \underline{0} \qquad (4.1.34)$$

and since $\bar{\sigma}_k[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]$ is monotonically decreasing over $(\underline{a}_t-\underline{\delta},\underline{a}_t+\underline{\delta})$, then, on a componentwise basis,

$$\underline{\bar{\sigma}}[\,i\,,\,\underline{a}_t\,,\,\underline{a}_t-\underline{\delta}\,] \;>\; \underline{0} \qquad (4.1.35)$$

$$\underline{\bar{\sigma}}[\,i\,,\,\underline{a}_t\,,\,\underline{a}_t+\underline{\delta}\,] \;<\; \underline{0} \qquad (4.1.36)$$

Consequently, there exists a number $I\,(\underline{\delta},\epsilon)$, the value of which may depend on $\underline{\delta}$ and ϵ, such that the following inequalities are valid for all $i > I\,(\underline{\delta},\epsilon)$ with probability exceeding $(1-\epsilon)$:

$$\frac{1}{i}\,\underline{s}\,[\,\underline{Z}(i),\underline{a}\,] \;>\; \underline{0} \quad \text{componentwise if } \underline{a} = \underline{a}_t-\underline{\delta} \qquad (4.1.37)$$

$$\frac{1}{i}\,\underline{s}\,[\,\underline{Z}(i),\underline{a}\,] \;<\; \underline{0} \quad \text{componentwise if } \underline{a} = \underline{a}_t+\underline{\delta} \qquad (4.1.38)$$

Since it was assumed that each component of $\underline{s}^1[\underline{Z}(i),\underline{a}]$ is a continuous function of $\underline{a}$ in Θ (and thus in the subrange $(\underline{a}_t-\underline{\delta}$, $\underline{a}_t+\underline{\delta})$ certainly) for all values of $\underline{Z}(i)$ (in real Euclidean space) except possibly for a set of probability zero, then the same is true of $\frac{1}{i}\,\underline{s}[\underline{Z}(i),\underline{a}]$ because

$$\frac{1}{i}\,\underline{s}[\underline{Z}(i),\underline{a}] \;=\; \frac{1}{i}\sum_{j=1}^{i}\,\underline{s}^1[\underline{Z}(j),\underline{a}] \qquad (4.1.39)$$

is merely the sum of such continuous functions.

Using (4.1.37) and (4.1.38) combined with this continuity, it is concluded that for <u>some</u> $\underline{a}$ in $(\underline{a}_t\pm\underline{\delta})$ and for all $i > I(\underline{\delta},\epsilon)$, that

$$P\left\{\frac{1}{i}\,s_k[\underline{Z}(j),\underline{a}] \;=\; 0 \;\middle|\; \underline{a}_t\right\} \;>\; 1-\epsilon \qquad (4.1.40)$$

where k assumes the values 1, 2, ... p (dimension of $\underline{a}$). This implies that a sequence of solutions to the likelihood equations

exists which converges in probability to $\underline{a}_t$. If there is a unique solution $\hat{\underline{a}}(i)$ to the likelihood equations for i equal to some $i_o, i_o+1, \ldots$, then the sequence $\hat{\underline{a}}(i)$ for these values of i converges in probability to $\underline{a}_t$.

4.1.4 Asymptotic Distribution of the Maximum Likelihood Estimator; Consistency

Let $(\underline{z}(1), \ldots \underline{z}(i))$ be a set of samples such that each $\underline{z}(j)$ is taken from the conditional probability density function $f(\underline{z}(j)|\underline{Z}(j-1),\underline{a}_t)$ for $j = 1, 2, \ldots i$, and let $F(\underline{z}(j)|\underline{Z}(j-1),\underline{a}_t)$ be regular with respect to its first and second $\underline{a}$-partial derivatives in the open set $\tilde{\Theta}$ within the entire admissible range of parameter values, Θ. Further assume that the maximum likelihood estimator $\hat{\underline{a}}(i)$ for $\underline{a}_t$ is unique for time instant i greater than or equal to some i_o (and also that it is a measurable random variable). Then $\hat{\underline{a}}(i)$ is asymptotically distributed for large i as a normal random variable characterized by mean $\underline{a}_t$ and covariance $\underline{J}^{-1}[\,i\,,\underline{a}_t\,]$:

$$f(\hat{\underline{a}}(i)|\underline{Z}(i)) \underset{\text{as } i \to \infty}{\longrightarrow} N\left[\underline{a}_t,\ \underline{J}^{-1}[\,i\,,\underline{a}_t\,]\right] \qquad (4.1.41)$$

Furthermore, $\hat{\underline{a}}(i)$ is a consistent estimator of the parameters.

Conceptually, the proof will show that the vector $\underline{J}[\,i\,,\underline{a}_t\,]\ [\hat{\underline{a}}(i)-\underline{a}_t]$ converges in probability (componentwise) to the vector $\underline{s}[\underline{Z}(i),\underline{a}_t]$ as i tends to infinity, and thereby conclude that the vector $[\hat{\underline{a}}(i)-\underline{a}_t]$ is a vector-valued random variable whose asymptotic distribution is normal, characterized by mean zero and covariance $\underline{J}^{-1}[\,i\,,\underline{a}_t\,]$.

Since $\underline{s}^1[\underline{Z}(j),\underline{a}]$ has a derivative with respect to $\underline{a}$ everywhere in a subrange $\tilde{\Theta}$ of Θ including $\underline{a}_t$, for all $\underline{Z}(i)$ except possibly for a set of probability zero, the same statement can be made about $\underline{s}[\underline{Z}(i),\underline{a}]$. If $\hat{\underline{a}}(i)$ is unique

for i greater than or equal to some i_o, then from the result of section 4.1.3, $\underline{\hat{a}}(i)$ converges in probability to $\underline{a}_t$.

By specifying that $\underline{\hat{a}}(i)$ is a measurable random variable, it is possible to assert (see Wilks (1963)) that for arbitrary positive values of $\underline{\delta}$ and ϵ, there is an integer $I(\underline{\delta},\epsilon,i_o)$ and a set ξ_o in Euclidean space of dimension (im) defined by $|\hat{a}_k(i) - a_{tk}| < |\delta_k|$ such that

$$P\left\{ \underline{Z}(i) \in \xi_o \text{ for all } i > I(\underline{\delta},\epsilon,i_o) \,\middle|\, \underline{a}_t \right\} > 1-\epsilon \tag{4.1.42}$$

The statements in the proof will be confined to this ξ_o.

The components of the likelihood equation can be expanded through the mean value theorem as

$$0 = s_k[\underline{Z}(i),\underline{a}(i)] \tag{4.1.43}$$

$$= s_k[\underline{Z}(i),\underline{a}_t] + \frac{\partial s_k[\underline{Z}(i),\underline{a}^\circ]}{\partial \underline{a}} \left[\underline{\hat{a}}(i) - \underline{a}_t\right] \tag{4.1.44}$$

where $\underline{a}^\circ[\underline{Z}(i)]$ is a random variable satisfying

$$|\underline{a}_t - \underline{a}^\circ| \leq |\underline{a}_t - \underline{\hat{a}}(i)| \tag{4.1.45}$$

This implicitly assumes a solution $\underline{\hat{a}}(i)$, the existence of which was demonstrated in the last section. Note that the term $\partial s_k/\partial \underline{a}$ is a transposed vector of dimension p (p equals the number of uncertain parameters):

$$\frac{\partial s_k}{\partial \underline{a}} = \left[\frac{\partial s_k}{\partial a_1} \quad \cdots \quad \frac{\partial s_k}{\partial a_p} \right] \tag{4.1.46}$$

Now use the fact that $\underline{\hat{a}}(i)$ is a measurable random variable to strengthen the statement of equation (4.1.44).

Equations (4.1.42) and (4.1.44) together yield (see Wilks (1963)):

$$P\left\{ \underline{s}[\underline{Z}(i),\underline{a}_t] + \frac{\partial \underline{s}[\underline{Z}(i),\underline{a}^\circ]}{\partial \underline{a}} \left[\hat{\underline{a}}(i)-\underline{a}_t\right] = \underline{0} \right.$$

$$\left. \text{for all } i > I(\underline{\delta},\epsilon,i_\circ) \,\middle|\, \underline{a}_t \right\} \quad > \quad 1-\epsilon \tag{4.1.47}$$

which implies that $\underline{s}[\underline{Z}(i),\underline{a}_t]$ and $\partial \underline{s}[\underline{Z}(i),\underline{a}^\circ]/\partial \underline{a}\ [\hat{\underline{a}}(i)-\underline{a}_t]$ are vector valued random variables such that if either converges in distribution, then they both do and that their sum converges to the zero vector.

Now look at equation (4.1.47). If it can be shown that $\partial \underline{s}[\underline{Z}(i),\underline{a}^\circ]/\partial \underline{a}$ converges in probability to $\{-\underline{J}[\,i\,,\,\underline{a}_t\,]\}$ as i tends to infinity, then (4.1.47) would demonstrate that $\{(\partial \underline{s}[\underline{Z}(i),\underline{a}^\circ]/\partial \underline{a})\ [\hat{\underline{a}}(i)-\underline{a}_t]\}$, $\{-\underline{J}[\,i\,,\,\underline{a}_t\,]\ [\hat{\underline{a}}(i)-\underline{a}_t]\}$, and $\{-\underline{s}[\underline{Z}(i),\underline{a}_t]\}$ converge together in probability (and so also in distribution). Thus, $\{\underline{s}[\underline{Z}(i),\underline{a}_t] - \underline{J}[\,i\,,\,\underline{a}_t\,]\ [\hat{\underline{a}}(i)-\underline{a}_t]\}$ converges in probability to zero, which is the result that is sought. Therefore, it will now be shown that $\partial \underline{s}[\underline{Z}(i),\underline{a}^\circ]/\partial \underline{a}$ does in fact converge in probability to $\{-\underline{J}[\,i\,,\,\underline{a}_t\,]\}$ as $i \rightarrow \infty$.

As a first step, $\partial \underline{s}[\underline{Z}(i),\underline{a}_t]/\partial \underline{a}$ will be shown to converge to $\{-\underline{J}[\,i\,,\,\underline{a}_t\,]\}$. Toward this result, define the matrix $\underline{Y}[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,]$ as

$$\underline{Y}[\,i\,,\,\underline{a}_t\,,\,\underline{a}\,] \equiv E\left\{ \frac{1}{f(\underline{Z}(i)|\,\underline{a})} \quad \frac{\partial^2 f(\underline{Z}(i)|\,\underline{a})}{\partial \underline{a}^2} \,\middle|\, \underline{a}_t \right\} \tag{4.1.48}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{f(\underline{Z}(i)|\,\underline{a})} \quad \frac{\partial^2 f(\underline{Z}(i)|\,\underline{a})}{\partial \underline{a}^2} \ f(\underline{Z}(i)|\,\underline{a}_t)\, d\underline{Z}(i) \tag{4.1.49}$$

The usefulness of this definition lies in the fact that

$$\frac{1}{f(\underline{Z}(i)|\underline{a})}\frac{\partial^2 f(\underline{Z}(i)|\underline{a})}{\partial \underline{a}^2} = \frac{\partial^2 \ell n\, f(\underline{Z}(i)|\underline{a})}{\partial \underline{a}^2}$$

$$+ \frac{1}{[f(\underline{Z}(i)|\underline{a})]^2}\frac{\partial f(\underline{Z}(i)|\underline{a})^T}{\partial \underline{a}}\frac{\partial f(\underline{Z}(i)|\underline{a})}{\partial \underline{a}}$$

$$= \frac{\partial \underline{s}|\underline{Z}(i),\underline{a}|}{\partial \underline{a}} + \underline{s}[\underline{Z}(i),\underline{a}]\underline{s}^T[\underline{Z}(i),\underline{a}] \qquad (4.1.50)$$

and therefore $\underline{Y}[\,i\,,\underline{a}_t\,,\underline{a}_t\,]$ is equal to

$$\underline{Y}[\,i\,,\underline{a}_t\,,\underline{a}_t\,] = E\left\{\frac{\partial \underline{s}[\underline{Z}(i),\underline{a}_t]}{\partial \underline{a}}\,\middle|\,\underline{a}_t\right\} + \underline{J}[\,i\,,\underline{a}_t\,] \qquad (4.1.51)$$

But, assuming that differentiation with respect to $\underline{a}$ can be taken outside the integral, $\underline{Y}[\,i\,,\underline{a}_t\,,\underline{a}_t\,]$ can be calculated as

$$\underline{Y}[\,i\,,\underline{a}_t\,,\underline{a}_t\,] = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{\partial^2 f(\underline{Z}(i)|\underline{a}_t)}{\partial \underline{a}^2}\,d\underline{Z}(i)$$

$$= \frac{\partial^2}{\partial \underline{a}^2}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\underline{Z}(i)|\underline{a}_t)\,d\underline{Z}(i)$$

$$= \frac{\partial^2}{\partial \underline{a}^2}(1)$$

$$= \underline{0} \qquad (4.1.52)$$

Combining (4.1.51) and (4.1.52) yields:

$$E\left\{ \frac{\partial \underline{s}[\underline{Z}(i), \underline{a}_t]}{\partial \underline{a}} \,\middle|\, \underline{a}_t \right\} = -\underline{J}[\, i \,,\, \underline{a}_t \,] \tag{4.1.53}$$

By its definition,

$$\frac{\partial \underline{s}[\underline{Z}(i), \underline{a}_t]}{\partial \underline{a}} = \sum_{j=1}^{i} \frac{\partial \underline{s}^1[\underline{Z}(j), \underline{a}_t]}{\partial \underline{a}} \tag{4.1.54}$$

so that as i becomes large, the weak law of large numbers can be used to demonstrate the desired convergence in probability:

$$\frac{\partial \underline{s}[\underline{Z}(i), \underline{a}_t]}{\partial \underline{a}} \longrightarrow E\left\{ \frac{\partial \underline{s}[\underline{Z}(i), \underline{a}_t]}{\partial \underline{a}} \,\middle|\, \underline{a}_t \right\} = -\underline{J}[\, i \,,\, \underline{a}_t \,] \tag{4.1.55}$$

Now, in view of equation (4.1.45) and the fact that $\hat{\underline{a}}(i)$ converges in probability to $\underline{a}_t$, then $\underline{a}^\circ$ must also converge in probability to $\underline{a}_t$. (A general theorem to this effect can be found in Wilks (1963).) The fact that $\partial \underline{s}[\underline{Z}(i), \underline{a}_t]/\partial \underline{a}$ and $\underline{a}^\circ$ are sequences which converge respectively to $\{-\underline{J}[\, i \,,\, \underline{a}_t \,]\}$ and $\underline{a}_t$ (where $\underline{J}$ is continuous in $\underline{a}$) is sufficient (see Wilks for proof of theorem) to insure that

$$\frac{\partial \underline{s}[\underline{Z}(i), \underline{a}^\circ]}{\partial \underline{a}} \longrightarrow -\underline{J}[\, i \,,\, \underline{a}_t \,] \tag{4.1.56}$$

as $i \longrightarrow \infty$, which is the convergence in probability that has been sought.

The following convergence in probability has just been demonstrated:

$$\underline{s}[\underline{Z}(i), \underline{a}_t] - \underline{J}[\, i \,,\, \underline{a}_t \,]\, [\hat{\underline{a}}(i) - \underline{a}_t] \longrightarrow \underline{0} \tag{4.1.57}$$

But it has already been shown that $\underline{s}[\underline{Z}(i),\underline{a}_t]$ is asymptotically normally distributed with mean zero and covariance $\underline{J}[\,i\,,\,\underline{a}_t\,]$. Therefore, in order for (4.1.57) to be true, $[\hat{\underline{a}}(i)-\underline{a}_t]$ must be asymptotically normally distributed with mean zero and covariance $\underline{J}[\,i\,,\,\underline{a}_t\,]^{-1}$. Thus it is concluded that the asymptotic distribution of $\hat{\underline{a}}(i)$ is a normal distribution with mean $\underline{a}_t$ and covariance $\underline{J}[\,i\,,\,\underline{a}_t\,]^{-1}$.

It can be shown (see Brunk (1965)) that this is sufficient to prove $\hat{\underline{a}}(i)$ is also a consistent estimator of the parameters.

4.1.5 Asymptotic Efficiency of the Maximum Likelihood Estimator

Under the same assumptions as used in the previous section, the maximum likelihood estimator $\hat{\underline{a}}(i)$ has an asymptotic efficiency of 1 for estimating $\underline{a}_t$.

Let $\tilde{\underline{a}}(i)$ be a consistent estimator of $\underline{a}_t$, not necessarily a maximum likelihood estimator, which is to say that each component of $\tilde{\underline{a}}(i)$ is a consistent estimator of the corresponding component of $\underline{a}_t$. This allows $\tilde{\underline{a}}(i)$ to be biased, but as i tends to infinity, the bias goes to zero (which is indeed the case for a maximum likelihood estimator in general). Now let $[\tilde{\underline{a}}(i)-\underline{a}_t]$ be asymptotically normally distributed with mean $\underline{0}$ and covariance $\tilde{\underline{J}}[\,i\,,\,\underline{a}_t\,]^{-1}$ as $i\to\infty$. The Cramér-Rao Inequality can be used to show that the minimum covariance matrix that can be obtained in the limit as $i\to\infty$ is $\underline{J}[\,i\,,\,\underline{a}_t\,]^{-1}$:

$$\liminf_{i\to\infty}\left(\operatorname{cov}\,(\tilde{\underline{a}}(i)|\,\underline{a}_t)\right) \geq \underline{J}[\,i\,,\,\underline{a}_t\,]^{-1} \qquad (4.1.58)$$

Therefore, asymptotic efficiency of $\tilde{\underline{a}}(i)$ is defined to be

$$\text{leff}\ (\tilde{\underline{a}}(i)|\,\underline{a}_t) \;=\; \frac{\left|\tilde{\underline{J}}[\,i\,,\,\underline{a}_t\,]\right|}{\left|\underline{J}[\,i\,,\,\underline{a}_t\,]\right|} \qquad (4.1.59)$$

In the previous section, it was seen that the maximum likelihood estimator $\hat{\underline{a}}(i)$ has the property that $[\hat{\underline{a}}(i)-\underline{a}_t]$ is asymptotically normally distributed with mean zero and covariance $\underline{J}[\,i\,,\,\underline{a}_t\,]^{-1}$. Therefore, its asymptotic efficiency is:

$$\text{leff}\ (\hat{\underline{a}}(i)|\,\underline{a}_t) \;=\; \frac{\left|\underline{J}[\,i\,,\,\underline{a}_t\,]\right|}{\left|\underline{J}[\,i\,,\,\underline{a}_t\,]\right|} \;=\; 1 \qquad (4.1.60)$$

4.1.6 Summary of ML Parameter Estimator Asymptotic Properties

Assume the following general conditions on the likelihood function, $\ell n\ f(\underline{Z}(i)|\,\underline{a})$, are satisfied (they are sufficient for validity of the previous developments):

1) the derivatives

$$\frac{\partial\, \ell n\ f(\underline{Z}(i)|\,\underline{a})}{\partial a_j} \qquad j = 1,\ 2,\ \ldots\ p$$

$$\frac{\partial^2 \ell n\ f(\underline{Z}(i)|\,\underline{a})}{\partial a_j\, \partial a_k} \qquad j,k = 1,\ 2,\ \ldots\ p$$

$$\frac{\partial^3 \ell n\ f(\underline{Z}(i)|\,\underline{a})}{\partial a_j\, \partial a_k\, \partial a_\ell} \qquad j,k,\ell = 1,\ 2,\ \ldots\ p$$

exist in the range $\tilde{\Theta}$ of admissible $\underline{a}$ values, for all $\underline{Z}(i)$ except possibly a set of probability zero.

2) the following conditional expectations are zero:

$$E\left\{ \frac{1}{f(\underline{Z}(i)|\underline{a})} \frac{\partial f(\underline{Z}(i)|\underline{a})^T}{\partial \underline{a}} \,\middle|\, \underline{a}_t \right\} = \underline{0} \qquad (4.1.61)$$

$$E\left\{ \frac{1}{f(\underline{Z}(i)|\underline{a})} \frac{\partial^2 f(\underline{Z}(i)|\underline{a})}{\partial \underline{a}^2} \,\middle|\, \underline{a}_t \right\} = \underline{0} \qquad (4.1.62)$$

3) the following matrix is positive definite:

$$E\left\{ \frac{1}{[f(\underline{Z}(i)|\underline{a})]^2} \frac{\partial f(\underline{Z}(i)|\underline{a})^T}{\partial \underline{a}} \frac{\partial f(\underline{Z}(i)|\underline{a})}{\partial \underline{a}} \,\middle|\, \underline{a}_t \right\} > \underline{0} \qquad (4.1.63)$$

4) for every $\underline{a}$ in Θ, a bound on the third partial derivative of the likelihood function exists:

$$\frac{1}{i} \left| \frac{\partial^3 \ell n\, f(\underline{Z}(i)|\underline{a})}{\partial a_j\, a_k\, a_\ell} \right| < M[\underline{Z}(i)] \qquad (4.1.64)$$

for j, k, and ℓ taking on the values 1, 2, ... p, and where

$$E\left\{ M[\underline{Z}(i)] \,\middle|\, \underline{a}_t \right\} < K \qquad (4.1.65)$$

for some K which is independent of $\underline{a}$ and i.

Under these assumptions, the maximum likelihood estimator for the parameters $\underline{a}$ is:

1) consistent
2) asymptotically unbiased
3) asymptotically normally distributed according to $N\left[\underline{a}_t, \underline{J}[\,i\,,\,\underline{a}_t\,]^{-1}\right]$
4) asymptotically efficient.

4.1.7 Combined State and Parameter Estimation

This section will show that the asymptotic properties of the maximum likelihood estimator of both state variables and system parameters are the same as those of an estimator of the parameters alone.

The previous developments were based upon the set of likelihood equations

$$\left.\frac{\partial \ell n\ f(\underline{Z}(i)|\ \underline{a})^T}{\partial \underline{a}}\right|_{\underline{a} \rightarrow \hat{\underline{a}}(i)} = \underline{0} \qquad (4.1.66)$$

so that $\hat{\underline{a}}(i)$ as defined by the solution of these equations is the estimate of the parameter vector $\underline{a}$ determined independently of any estimate of the state. Now consider a simultaneous estimation of the state and parameter vectors by means of the likelihood equations

$$\left.\frac{\partial \ell n\ f(\underline{x}(i), \underline{Z}(i)|\ \underline{a})^T}{\partial \underline{x}}\right|_{\substack{\underline{x}(i) \rightarrow \hat{\underline{x}}(i) \\ \underline{a} \rightarrow \hat{\underline{a}}_x(i)}} = \underline{0} \qquad (4.1.67)$$

$$\left.\frac{\partial \ell n\ f(\underline{x}(i), \underline{Z}(i)|\ \underline{a})^T}{\partial \underline{a}}\right|_{\substack{\underline{x}(i) \rightarrow \hat{\underline{x}}(i) \\ \underline{a} \rightarrow \hat{\underline{a}}_x(i)}} = \underline{0} \qquad (4.1.68)$$

where $\hat{\underline{x}}(i)$ and $\hat{\underline{a}}_x(i)$ are defined to be the solution to the simultaneous set of equations (4.1.67) and (4.1.68). In other words, $\hat{\underline{a}}_x(i)$ is the estimate of the parameter vector $\underline{a}$ obtained simultaneously with the state estimate. It will be

shown that, as i becomes large, $\hat{\underline{a}}_x(i)$ will converge to $\hat{\underline{a}}(i)$, and thereby be concluded that the asymptotic behavior of the two estimators is the same.

To do this, expand equation (4.1.68) in a Taylor series through the second order, about the point $\underline{a} = \hat{\underline{a}}(i)$:

$$\left.\frac{\partial \ln f(\underline{x}(i),\underline{Z}(i)|\underline{a})^T}{\partial \underline{a}}\right|_{\substack{\underline{x}(i)\to\hat{\underline{x}}(i)\\ \underline{a}\to\hat{\underline{a}}_x(i)}} = \underline{0}$$

$$\cong \left.\frac{\partial \ln f(\underline{x}(i),\underline{Z}(i)|\underline{a})^T}{\partial \underline{a}}\right|_{\substack{\underline{x}(i)\to\hat{\underline{x}}(i)\\ \underline{a}\to\hat{\underline{a}}(i)}}$$

$$+ \left.\frac{\partial^2 \ln f(\underline{x}(i),\underline{Z}(i)|\underline{a})}{\partial \underline{a}^2}\right|_{\substack{\underline{x}(i)\to\hat{\underline{x}}(i)\\ \underline{a}\to\hat{\underline{a}}(i)}} \left[\hat{\underline{a}}_x(i)-\hat{\underline{a}}(i)\right]$$

$$= \left\{\frac{\partial \ln f(\underline{x}(i)|\underline{Z}(i),\underline{a})^T}{\partial \underline{a}} + \frac{\partial \ln f(\underline{Z}(i)|\underline{a})^T}{\partial \underline{a}}\right\}\Bigg|_{\substack{\underline{x}(i)\to\hat{\underline{x}}(i)\\ \underline{a}\to\hat{\underline{a}}(i)}}$$

$$+ \left\{\frac{\partial^2 \ln f(\underline{x}(i)|\underline{Z}(i),\underline{a})}{\partial \underline{a}^2} + \frac{\partial^2 \ln f(\underline{Z}(i)|\underline{a})}{\partial \underline{a}^2}\right\}\Bigg|_{\substack{\underline{x}(i)\to\hat{\underline{x}}(i)\\ \underline{a}\to\hat{\underline{a}}(i)}} \left[\hat{\underline{a}}_x(i)-\hat{\underline{a}}(i)\right] \qquad (4.1.69)$$

by application of Bayes' Rule to the density in the likelihood function. By definition, $\hat{\underline{a}}(i)$ is the solution to equation (4.1.66), so the second term inside the first set of brackets in (4.1.69) is zero. If the second term inside the other brackets were evaluated at $\underline{a}_t$ instead of $\hat{\underline{a}}(i)$, it would be approximately equal to $\{-\underline{J}[\,i\,,\,\underline{a}_t\,]\}$ for large enough i. If the estimate $\hat{\underline{a}}(i)$ is assumed sufficiently close to $\underline{a}_t$ for large i, then

$$\left.\frac{\partial^2 \ell n\ f(\underline{Z}(i)|\,\underline{a})}{\partial \underline{a}^2}\right|_{\underline{a}\rightarrow\hat{\underline{a}}(i)} \cong -\underline{J}[\,i\,,\,\underline{a}_t\,] \qquad (4.1.70)$$

The first term in the second brackets of equation (4.1.69) has been evaluated as

$$\left.\frac{\partial^2 \ell n\ f(\underline{x}(i)|\,\underline{Z}(i),\underline{a})}{\partial a_j \partial a_k}\right|_{\substack{\underline{x}(i)\rightarrow\hat{\underline{x}}(i)\\ \underline{a}\rightarrow\hat{\underline{a}}(i)}} =$$

$$= -\tfrac{1}{2}\,\mathrm{tr}\left\{\underline{P}^{-1}(i)\,\frac{\partial^2 \underline{P}(i)}{\partial a_j \partial a_k} - \underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_j}\,\underline{P}^{-1}(i)\,\frac{\partial \underline{P}(i)}{\partial a_k}\right\}$$

$$\left. - \frac{\partial \hat{\underline{x}}^T(i)}{\partial a_j}\,\underline{P}^{-1}(i)\,\frac{\partial \hat{\underline{x}}(i)}{\partial a_k}\right|_{\underline{a}\rightarrow\hat{\underline{a}}(i)} \qquad (4.1.71)$$

This is bounded except when $\underline{P}(i)$ becomes singular (in which case $\underline{P}^{-1}(i)$ and the density $f(\underline{x}(i)|\,\underline{Z}(i),\underline{a})$ itself do not exist) or possibly at points where the density or its first $\underline{a}$-partials are not continuous functions of the parameters (such continuity would be expected physically, and could be

assured by assuming bounded third partial derivatives, similar to equation (4.1.64)). Therefore, as i tends to infinity, the matrix composed of entries defined by equation (4.1.71) would be dominated by $\{-\underline{J}[\,i\,,\,\underline{a}_t\,]\}$.

Thus, equation (4.1.69) becomes

$$\left.\frac{\partial\, \ell n\ f(\underline{x}(i)|\,\underline{Z}(i),\underline{a})^T}{\partial \underline{a}}\right|_{\substack{\underline{x}(i)\to\hat{\underline{x}}(i)\\ \underline{a}\to\hat{\underline{a}}(i)}} - \underline{J}[\,i\,,\,\underline{a}_t\,]\,\left[\hat{\underline{a}}_x(i)-\hat{\underline{a}}(i)\right] \cong \underline{0} \tag{4.1.72}$$

Evaluating the first term results in

$$-\tfrac{1}{2}\left.\begin{bmatrix} \mathrm{tr}\left\{\underline{P}^{-1}(i)\,\dfrac{\partial \underline{P}(i)}{\partial \underline{a}_1}\right\} \\ \vdots \\ \mathrm{tr}\left\{\underline{P}^{-1}(i)\,\dfrac{\partial \underline{P}(i)}{\partial a_p}\right\} \end{bmatrix}\right|_{\underline{a}\to\hat{\underline{a}}(i)} - \underline{J}[\,i\,,\,\underline{a}_t\,]\left[\hat{\underline{a}}_x(i)-\hat{\underline{a}}(i)\right] \cong \underline{0} \tag{4.1.73}$$

Solving for $\left[\hat{\underline{a}}_x(i)-\hat{\underline{a}}(i)\right]$ yields

$$\left[\hat{\underline{a}}_x(i)-\hat{\underline{a}}(i)\right] = -\tfrac{1}{2}\,\underline{J}^{-1}[\,i\,,\,\underline{a}_t\,]\left.\begin{bmatrix} \mathrm{tr}\left\{\underline{P}^{-1}(i)\,\dfrac{\partial \underline{P}(i)}{\partial a_1}\right\} \\ \mathrm{tr}\left\{\underline{P}^{-1}(i)\,\dfrac{\partial \underline{P}(i)}{\partial a_p}\right\} \end{bmatrix}\right|_{\underline{a}\to\hat{\underline{a}}(i)} \tag{4.1.74}$$

as the first order difference between the parameter estimate obtained simultaneously with a state estimate and that obtained independent from a state estimate. It has been shown

that $\underline{J}^{-1}[\,i\,,\,\underline{a}_t\,]$ converges to $\underline{0}$ as i tends to infinity, so if the vector it premultiplies is bounded as i grows large (assured except for the cases cited below equation (4.1.71), $\hat{\underline{a}}_x(i)$ converges to $\hat{\underline{a}}(i)$ as i goes to infinity. Therefore, the asymptotic behaviors of the two parameter estimators are equivalent. That is to say that the maximum likelihood estimator of state and parameter values simultaneously yields a parameter estimate that is:

1) consistent
2) asymptotically unbiased
3) asymptotically normally distributed according to $N\left[\underline{a}_t, \underline{J}[\,i\,,\,\underline{a}_t\,]^{-1}\right]$
4) asymptotically efficient.

Furthermore, since the parameter estimate does converge to the true value $\underline{a}_t$, the corresponding state estimator, which uses $\hat{\underline{a}}_x(i)$ for the values of the parameters, converges to the behavior of the maximum likelihood state estimator (of Kalman form) that uses the true value, $\underline{a}_t$, for the parameters. These statements are conditioned upon the assumptions described previously and also upon the assumption that the model of the parameters as constants over the time interval of interest, i.e., over all time, is a valid one.

In other words, the asymptotic properties just described provide the following information. If all data from the initial time forward can be used, and if the parameters are indeed constant for all time, the behavior of the estimator will be as described above. If a fixed number, N, of measurements are used to process the parameter estimates, then the behavior of the estimator will be as described as N is allowed to increase to infinity, provided the constant parameter model remains valid. Thus, to be assured of estimator behavior consistent with that predicted theoretically, one should seek as large a value of N as allowed by the physics of the problem and by the constraints imposed by the computer capacity.

4.2 Ambiguity Functions

In order to analyze the performance to be expected from the estimators, a linearized error analysis can be generated to display the sensitivity of the estimates to variation in noise sequences, initial conditions, and other factors. However, the result is a local analysis, and an additional global view of the performance would be very desirable.

To accomplish this, the concept of an ambiguity function can be utilized (this is a substantial modification and extension of a presentation in Schweppe (1970)). Let $f(\underline{Z}(i);\underline{x},\underline{a})$ denote a general likelihood function and note that its natural logarithm, used to derive the maximum likelihood estimate of $\underline{x}$ and $\underline{a}$, is a random variable when viewed as a function of $\underline{Z}(i)$. The ambiguity function is then defined to be the average of the log-likelihood function. As an evident extension to previous literature, it can be written as:

$$\Psi_i\,[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\,[\ell n\, f(\underline{Z}(i);\underline{x},\underline{a})]\cdot \cdot f(\underline{Z}(i);\underline{x}_t,\underline{a}_t)\; d\underline{Z}(i) \qquad (4.2.1)$$

where $\underline{x}_t$ and $\underline{a}_t$ are the true, but unknown, values of the state and parameter vectors. The shape of this ambiguity function, as a function of $\underline{x}$ and $\underline{a}$, affords the designer considerable information about the performance he should expect from the estimator.

Before investigating the applications of the ambiguity function, this section will develop some of its properties. To simplify what follows, consider the simplest case of a single scalar measurement, the state known exactly, and a scalar parameter to be estimated. Possible generalizations will then be indicated. First consider:

$$\psi_1[\underline{x}_t, \underline{a}; \underline{x}_t, \underline{a}_t] = \int_{-\infty}^{\infty} [\ell n\ f(z;\underline{x}_t, a)]\ f(z;\underline{x}_t, a_t)\ dz \tag{4.2.2}$$

Let the z axis be divided arbitrarily into disjoint intervals, $(z_j, z_{j+1}]$ for j taking on all positive and negative integer values. Then (4.2.2) can be written as

$$\psi_1[\underline{x}_t, a; \underline{x}_t, a_t] = \sum_{j=-\infty}^{\infty} \int_{z_j}^{z_{j+1}} [\ell n\ f(z;\underline{x}_t, a)]\ f(z;\underline{x}_t, a_t)\, dz \tag{4.2.3}$$

By assuming that the distributions $F(z;\underline{x}_t, a)$ and $F(z;\underline{x}_t, a_t)$ are absolutely continuous with respect to each other, i.e., that there is no set ξ' in the z-space over which $f(z;\underline{x}_t, a)$ and $f(z;\underline{x}_t, a_t)$ are not simultaneously zero or simultaneously positive, the occurrence of terms in this sum that are equal to minus infinity is precluded. This condition is necessary for a finite value of $\psi_1[\underline{x}_t, a; \underline{x}_t, a_t]$ to exist at all points (a_t, a) in the Cartesian product space $[\tilde{\Theta} \times \tilde{\Theta}]$.

Wilks (1963) has further claimed that the assumption that the distribution $F(z;\underline{x}_t, a)$ is regular (see section 4.1) with respect to its first two a-derivatives is sufficient to prove that the ambiguity function has first partial derivatives with respect to a and a_t, and that for fixed a_t, the ambiguity function (as a function of a) reaches a maximum at $a = a_t$ if $f(z;\underline{x}, a) = f(z|a)$.

When a vector $\underline{z}$, or a number of such vectors, $\underline{Z}$, are considered, these results are easily extended by considering m-dimensional intervals (hypercubes). With respect to vectors of parameters, consider the components to be functionally independent, and look at component-wise relationships, such as partial derivatives with respect to each parameter.

The significance of the ambiguity function is that it provides both a global and local error analysis. It is the

average value of the log-likelihood function, where the average is taken over the possible state initial conditions, $\underline{x}(0)$, dynamic driving noise sequences, $\underline{w}(j)$, and measurement noise sequences, $\underline{v}(j)$. If it has multiple peaks, it indicates that individual log-likelihood functions realized by particular sets of $\underline{x}(0)$, $\underline{w}(j)$, and $\underline{v}(j)$ may well have multiple peaks, which can cause convergence to local (rather than global) maxima or failure of any convergence at all. (The name "ambiguity function" is derived from such a situation, since multiple peaks tend to cause ambiguities.) Furthermore, the "sharpness" of the peak of the ambiguity function in the immediate vicinity of the "true" values of the states and parameters (various chosen values of $\underline{a}_t$ and $\underline{x}_t$ sequences would be investigated in practice) conveys the preciseness with which a maximum likelihood estimate can be discerned. If the ambiguity function were to attain its maximum over a finite range of values, the log-likelihood function evaluated for a specific set of data may well demonstrate the same property, precluding a choice among that range by means of the maximum likelihood criterion. In fact, the curvature of the ambiguity function at the point of "true" values can be inversely related to the Cramér-Rao lower bound on the estimate error covariance matrix.

The ambiguity function is a versatile concept for performance analysis. Not only does it provide an absolute measure of a particular estimator's expected behavior, but it can also be used to relate the performance of estimates based upon different log-likelihood functions. Furthermore, it can determine the relative effectiveness of the individual terms of the likelihood equations in obtaining desired performance. (The latter consideration is important when the designer contemplates removing terms from the equations for expediency.) Ambiguity functions can also delineate the sensitivity of the various estimators' performance to

1) various sizes of the fixed-length memory (N)

2) form of the assumed dynamic system model (particular state space, model dimension, etc.)
3) types of measurements taken (the structure of the $\underline{H}(j)$ matrices) and their precision (the $\underline{R}(j)$ histories)
4) the magnitude and uncertainty in initial conditions
5) the dynamic driving noise intensities
6) control inputs

Thus, the ambiguity function can be an invaluable tool to insure adequate performance from an estimator as designed by the methods of this thesis.

The relationship between the ambiguity function concept and the Cramér-Rao lower bound on estimate error covariance will now be made explicit. For a scalar parameter, a, the Cramér-Rao bound is

$$E\left\{\left[\hat{a}(i)-a_t\right]^2\right\} \geq \frac{\left[1 + \left.\frac{\partial b}{\partial a}\right|_{a = at}\right]^2}{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial a} \ln f(\underline{Z}(i);\underline{x}_t,a_t)\right]^2 f(\underline{Z}(i);\underline{x}_t,a_t)d\underline{Z}(i)} \tag{4.2.4}$$

where b is the bias error of $\hat{a}$:

$$b(a_t,i) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(\hat{a} - a_t)\, f(\underline{Z}(i);\underline{x}_t,a_t)\, d\underline{Z}(i) \tag{4.2.5}$$

It can be shown (see section 4.1 on regularity conditions) that

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial a} \ln f(\underline{Z}(i);\underline{x}_t,a_t)\right]^2 f(\underline{Z}(i);\underline{x}_t,a_t)\, d\underline{Z}(i) =$$

$$= -\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[\frac{\partial^2}{\partial a^2} \ln f(\underline{Z}(i);\underline{x}_t,a_t)\right] f(\underline{Z}(i);\underline{x}_t,a_t)\, d\underline{Z}(i) \tag{4.2.6}$$

so that equation (4.2.4) can be written as:

$$E\left\{\left[\hat{a}(i)-a_t\right]^2\right\} \geq \frac{\left[1 + \left.\frac{\partial b}{\partial a}\right|_{a = a_t}\right]^2}{\left. - \frac{\partial^2}{\partial a^2} \psi_i[\underline{x}, a; \underline{x}_t, a_t]\right|_{\substack{a = a_t \\ \underline{x} = \underline{x}_t}}} \qquad (4.2.7)$$

Thus, the curvature of the ambiguity function at the point $a = a_t$ determines the lower bound on the variance of the estimation error. For a vector parameter, $\underline{a}$, the unbiased version of equation (4.2.7) would be

$$E\left\{\left[\hat{\underline{a}}(i)-\underline{a}_t\right]\left[\hat{\underline{a}}(i)-\underline{a}_t\right]^T\right\} \geq \left[\left. - \frac{\partial^2}{\partial \underline{a}^2} \psi_i[\underline{x}, \underline{a}; \underline{x}_t, \underline{a}_t]\right|_{\substack{\underline{a}=\underline{a}_t \\ \underline{x}=\underline{x}_t}}\right]^{-1} \qquad (4.2.8)$$

Note that, although biases are generally present in maximum likelihood estimates, they are often ignored in performance analyses for expediency. The covariance lower bound so derived serves the same purpose in practical design as any theoretically defined optimum cost function: as the highest attainable goal in performance, the standard of comparison for all practical, suboptimal implementations. It enables the designer to answer such questions as, "To obtain better performance, is it worthwhile to investigate more sophisticated implementations or larger samples of data (N), or is it in fact necessary to require more or better quality data from the measuring devices?"

It should be noted that the ambiguity function and the divergence distance measure of hypothesis testing theory are closely related. This can be exploited both in developing

ambiguity functions and in interpretting the information contained in them.

In applying this (and other) error analysis, it is assumed that, for some value of the parameters $\underline{a}$, the dynamic model

$$\underline{x}(i+1) = \underline{\Phi}(i+1,i)\underline{x}(i) + \underline{B}(i)\underline{u}(i) + \underline{G}(i)\underline{w}(i) \quad (4.2.9)$$

$$\underline{z}(i) = \underline{H}(i)\underline{x}(i) + \underline{v}(i) \quad (4.2.10)$$

accurately represents the actual physical system. In other words, such modelling errors as insufficient state dimension are not considered, but state and parameter estimate errors due to the uncertainty in initial conditions, driving noise sequences, and measurement noise sequences are fully studied.

In order to investigate the ambiguity functions associated with the various possible log-likelihood functions, it is convenient to consider the separate contributions due to:

1) $\ell n\ f(\underline{Z}(i)|\,\underline{a})$ or $\ell n\ f(\underline{Z}_N(i)|\,\underline{a})$ or $\ell n\ f(\underline{Z}_N(i)|\,\underline{Z}(i-N),\underline{a})$

2) $\ell n\ f(\underline{x}(i)|\,\underline{Z}(i),\underline{a})$ or $\ell n\ f(\underline{x}(i)|\,\underline{Z}_N(i),\underline{a})$

3) $\ell n\ f(\underline{a})$

Attention will be concentrated on the first of these, initially developing its contribution as described in previous literature on ambiguity functions, and then introducing an innovative method for calculating its value in a very convenient manner. Then the other contributions will be developed analogously, and finally the explicit forms of the ambiguity functions for the different log-likelihood functions will be presented.

The log-likelihood function $\ell n\ f(\underline{Z}(i)|\,\underline{a})$, under the assumptions made previously, can be written as

$$\ell n\ f(\underline{Z}(i)|\,\underline{a}) = -\frac{im}{2}\,\ell n\,(2\pi) - \tfrac{1}{2}\sum_{j=1}^{i}\ell n\left[\left|\underline{A}(j;\underline{a})\right|\right]$$

$$-\tfrac{1}{2}\sum_{j=1}^{i}\left[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j;\underline{a})\right]^T\underline{A}^{-1}(j;\underline{a})\left[\underline{z}(j)-\underline{H}(j)\bar{\underline{x}}(j;\underline{a})\right] \quad (4.2.11)$$

where

$$\underline{A}(j;\underline{a}) = \underline{H}(j)\underline{M}(j;\underline{a})\underline{H}^{T}(j) + \underline{R}(j) \qquad (4.2.12)$$

and where the explicit dependence upon $\underline{a}$ appears in the notation. Similarly,

$$\ell n\ f(\underline{Z}_N(i)|\,\underline{Z}(i-N),\underline{a}) = -\frac{Nm}{2}\,\ell n\ (2\pi) - \tfrac{1}{2}\sum_{j=i-N+1}^{i}\ell n\left[\left|\underline{A}(j;\underline{a})\right|\right] - \tfrac{1}{2}\sum_{j=i-N+1}^{i}[\underline{z}(j)-\underline{H}(j)\overline{\underline{x}}(j;\underline{a})]^{T}\ \underline{A}^{-1}(j;\underline{a})[\underline{z}(j)-\underline{H}(j)\overline{\underline{x}}(j;\underline{a})] \qquad (4.2.13)$$

and also

$$\ell n\ f(\underline{Z}_N(i)|\,\underline{a}) = -\frac{Nm}{2}\,\ell n\ (2\pi) - \tfrac{1}{2}\sum_{j=i-N+1}^{i}\ell n\left[\left|\tilde{\underline{A}}(j;\underline{a})\right|\right] - \tfrac{1}{2}\sum_{j=i-N+1}^{i}[\underline{z}(j)-\underline{H}(j)\tilde{\underline{x}}^{-}(j;\underline{a})]^{T}\ \tilde{\underline{A}}^{-1}(j;\underline{a})[\underline{z}(j)-\underline{H}(j)\tilde{\underline{x}}^{-}(j;\underline{a})] \qquad (4.2.14)$$

where the notation in equation (4.2.14) has been described previously. Because of the similarities of equations (4.2.11), (4.2.13), and (4.2.14), only equation (4.2.11) will be pursued, and then the implications for the other two forms will be indicated.

The first development is based upon a presentation by Schweppe (1970), but the algorithm for the $\underline{\mathcal{F}}(i,1;\underline{a})$ matrix and expression for $\overline{\underline{x}}(i;\underline{a})$ are original. It is analogous to the construction of the divergence distance measure of hypothesis testing.

First define $\underline{z}^{\circ}(i;\underline{a})$ as the mean value of the i-th measurement, $\underline{z}(i)$, as determined by the assumed state initial conditions, any known inputs, and a given value, $\underline{a}$, of the uncertain parameters:

$$\underline{z}^{\circ}(i;\underline{a}) = \underline{H}(j)\underline{x}^{\circ}(i;\underline{a}) \tag{4.2.15}$$

where

$$\underline{x}^{\circ}(i;\underline{a}) = \underline{\Phi}(i,0;\underline{a})\hat{\underline{x}}_{\circ} + \sum_{j=1}^{i} \underline{\Phi}(i,j;\underline{a})\underline{B}(j-1;\underline{a})\underline{u}(j-1) \tag{4.2.16}$$

Recognize specifically that $\underline{x}^{\circ}(i;\underline{a})$ is the expected value of the state at time i as propagated with parameter value $\underline{a}$ and without any conditioning on the previous measurements: it is completely deterministic and can be precomputed (unless $\underline{u}$ is computed in feedback form as a function of the estimated state). Define the time histories of the measurements and the mean values as

$$\underline{Z}(i;\underline{a}_t) = \begin{bmatrix} \underline{z}(1;\underline{a}_t) \\ \vdots \\ \underline{z}(i;\underline{a}_t) \end{bmatrix} \qquad \underline{Z}^{\circ}(i;\underline{a}) = \begin{bmatrix} \underline{z}^{\circ}(1;\underline{a}) \\ \vdots \\ \underline{z}^{\circ}(i;\underline{a}) \end{bmatrix} \tag{4.2.17}$$

Now define an (m x im) matrix, $\underline{\mathcal{F}}(i,1;\underline{a})$ as the linear system which operates on the zero mean component of $\underline{Z}(i;\underline{a})$, i.e. $[\underline{Z}(i;\underline{a}) - \underline{Z}^{\circ}(i;\underline{a})]$, to give the zero mean component of the conditional expectation of $\underline{z}(i+1;\underline{a})$ assuming that $\underline{Z}(i;\underline{a})$ is generated by

$$\underline{x}(i+1;\underline{a}) = \underline{\Phi}(i+1,i;\underline{a})\underline{x}(i;\underline{a}) + \underline{B}(i;\underline{a})\underline{u}(i) + \underline{G}(i)\underline{w}(i) \tag{4.2.18}$$

$$\underline{z}(i;\underline{a}) = \underline{H}(i)\underline{x}(i;\underline{a}) + \underline{v}(i) \tag{4.2.19}$$

with the given value of $\underline{a}$. Thus, the conditional expectation can be written:

$$\bar{\underline{z}}(i+1;\underline{a}) = \underline{z}^{\circ}(i+1;\underline{a}) + \underline{\mathcal{F}}(i,1;\underline{a})\left[\underline{Z}(i;\underline{a}_t) - \underline{Z}^{\circ}(i;\underline{a})\right] \tag{4.2.20}$$

and the corresponding residual can be decomposed into two parts:

$$\underline{z}(i+1;\underline{a}_t) - \bar{\underline{z}}(i+1;\underline{a}) = \tilde{\underline{r}}(i+1) + \underline{d}(i+1) \tag{4.2.21}$$

$$\tilde{\underline{r}}(i+1) = \underline{z}(i+1;\underline{a}_t) - \underline{z}^{\circ}(i+1;\underline{a}_t) - \underline{\mathcal{F}}(i,1;\underline{a})\left[\underline{Z}(i;\underline{a}_t) - \underline{Z}^{\circ}(i;\underline{a}_t)\right] \tag{4.2.22}$$

$$\underline{d}(i+1) = \underline{z}^{\circ}(i+1;\underline{a}_t) - \underline{z}^{\circ}(i+1;\underline{a}) - \underline{\mathcal{F}}(i,1;\underline{a})\left[\underline{Z}^{\circ}(i;\underline{a}_t) - \underline{Z}^{\circ}(i;\underline{a})\right] \tag{4.2.23}$$

where $\underline{d}(i+1)$ is the deterministic portion of the (i+1)-th residual, due to the difference between the assumed value $\underline{a}$ and the true value $\underline{a}_t$.

In order to determine the $\underline{\mathcal{F}}(i,1;\underline{a})$ matrix, the optimal estimate of the state $\underline{x}(i)$ before the i-th measurement, conditioned on the previous measurements and the assumed value of the parameters, $\bar{\underline{x}}(i;\underline{a})$, can be expressed as

$$\bar{\underline{x}}(1;\underline{a}) = \underline{x}^{\circ}(1;\underline{a})$$

$$\bar{\underline{x}}(2;\underline{a}) = \underline{x}^{\circ}(2;\underline{a}) + \underline{\Phi}(2,1;\underline{a})\underline{K}(1;\underline{a})\left[\underline{z}(1;\underline{a}_t) - \underline{H}(1)\underline{x}^{\circ}(1;\underline{a})\right]$$

$$\vdots$$

$$\bar{\underline{x}}(i;\underline{a}) = \underline{x}^{\circ}(i;\underline{a}) + \underline{\Phi}(i,i-1;\underline{a})\underline{K}(i-1;\underline{a})\left[\underline{z}(i-1;\underline{a}_t) - \underline{H}(i-1)\underline{x}^{\circ}(i-1;\underline{a})\right]$$
$$+ \left[\underline{\Phi}(i,i-1;\underline{a})\left[\underline{I}-\underline{K}(i-1;\underline{a})\underline{H}(i-1)\right]\right] \cdot$$
$$\cdot\left[\underline{\Phi}(i-1,i-2;\underline{a})\underline{K}(i-2;\underline{a})\left[\underline{z}(i-2;\underline{a}_t) - \underline{H}(i-2)\underline{x}^{\circ}(i-2;\underline{a})\right]\right] + \ldots$$

$$+\ldots+\left[\underline{\Phi}(i,i-1;\underline{a})\left[\underline{I}-\underline{K}(i-1;\underline{a})\underline{H}(i-1)\right]\right]\cdots\left[\underline{\Phi}(3,2;\underline{a})\left[\underline{I}-\underline{K}(2;\underline{a})\underline{H}(2)\right]\right]\cdot$$
$$\cdot\left[\underline{\Phi}(2,1;\underline{a})\underline{K}(1;\underline{a})\left[\underline{z}(1;\underline{a}_t)-\underline{H}(1)\underline{x}^\circ(1;\underline{a})\right]\right]$$

(4.2.24)

where the general term can be more conveniently expressed as

$$\overline{\underline{x}}(i;\underline{a}) = \underline{x}^\circ(i;\underline{a})+\underline{\Phi}(i,i-1;\underline{a})\underline{K}(i-1;\underline{a})\left[\underline{z}(i-1;\underline{a}_t)-\underline{H}(i-1)\underline{x}^\circ(i-1;\underline{a})\right]$$
$$+\sum_{k=0}^{i-3}\left[\prod_{j=0}^{k}\underline{\Phi}(i-j,i-j-1;\underline{a})\left[\underline{I}-\underline{K}(i-j-1;\underline{a})\underline{H}(i-j-1)\right]\right]\cdot$$
$$\cdot\left[\underline{\Phi}(i-k-1,i-k-2;\underline{a})\underline{K}(i-k-2)\cdot\right.$$
$$\left.\cdot\left[\underline{z}(i-k-2;\underline{a}_t)-\underline{H}(i-k-2)\underline{x}^\circ(i-k-2)\right]\right]$$

(4.2.25)

By means of this expression, $\underline{\mathcal{F}}(i,1;\underline{a})$ can be partitioned into i (m x m) submatrices:

$$\underline{\mathcal{F}}(i,1;\underline{a}) = \left[\underline{\mathcal{F}}_1(i,1;\underline{a}) \vdots \underline{\mathcal{F}}_2(i,1;\underline{a}) \vdots \cdots \vdots \underline{\mathcal{F}}_i(i,1;\underline{a})\right]$$

(4.2.26)

and these submatrices can be generated sequentially through:

$$\underline{\mathcal{F}}_1(1,1;\underline{a}) = \underline{H}(2)\underline{\Gamma}_{11} \qquad \underline{\Gamma}_{11} = \underline{\Phi}(2,1;\underline{a})\underline{K}(1;\underline{a})$$

(4.2.27)

$$\underline{\mathcal{F}}_1(2,1;\underline{a}) = \underline{H}(3)\underline{\Gamma}_{21} \qquad \underline{\Gamma}_{21} = \underline{\Phi}(3,2;\underline{a})\left[\underline{I}-\underline{K}(2;\underline{a})\underline{H}(2)\right]\underline{\Gamma}_{11}$$
$$\underline{\mathcal{F}}_2(2,1;\underline{a}) = \underline{H}(3)\underline{\Gamma}_{22} \qquad \underline{\Gamma}_{22} = \underline{\Phi}(3,2;\underline{a})\underline{K}(2;\underline{a})$$

(4.2.28)

$$\underline{\mathcal{F}}_1(3,1;\underline{a}) = \underline{H}(4)\underline{\Gamma}_{31} \qquad \underline{\Gamma}_{31} = \underline{\Phi}(4,3;\underline{a})[\underline{I}-\underline{K}(3;\underline{a})\underline{H}(3)]\,\underline{\Gamma}_{21}$$

$$\underline{\mathcal{F}}_2(3,1;\underline{a}) = \underline{H}(4)\underline{\Gamma}_{32} \qquad \underline{\Gamma}_{32} = \underline{\Phi}(4,3;\underline{a})[\underline{I}-\underline{K}(3;\underline{a})\underline{H}(3)]\,\underline{\Gamma}_{22}$$

$$\underline{\mathcal{F}}_3(3,1;\underline{a}) = \underline{H}(4)\underline{\Gamma}_{33} \qquad \underline{\Gamma}_{33} = \underline{\Phi}(4,3;\underline{a})\underline{K}(3;\underline{a})$$

(4.2.29)

$$\underline{\mathcal{F}}_1(i,1;\underline{a}) = \underline{H}(i+1)\underline{\Gamma}_{i1} \qquad \underline{\Gamma}_{i1} = \underline{\Phi}(i+1,i;\underline{a})[\underline{I}-\underline{K}(i;\underline{a})\underline{H}(i)]\,\underline{\Gamma}_{(i-1)1}$$

$$\underline{\mathcal{F}}_{i-1}(i,1;\underline{a}) = \underline{H}(i+1)\underline{\Gamma}_{i(i-1)} \qquad \underline{\Gamma}_{i(i-1)} = \underline{\Phi}(i+1,i;\underline{a})[\underline{I}-\underline{K}(i;\underline{a})\underline{H}(i)]\,\underline{\Gamma}_{(i-1)(i-1)}$$

$$\underline{\mathcal{F}}_i(i,1;\underline{a}) = \underline{H}(i+1)\underline{\Gamma}_{ii} \qquad \underline{\Gamma}_{ii} = \underline{\Phi}(i+1,i;\underline{a})\underline{K}(i;\underline{a})$$

(4.2.30)

These matrices are then used to evaluate the ambiguity function, which, from its definition and equation (4.2.11), can be written as

$$\psi_i\,[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = -\frac{im}{2}\,\ell n\,(2\pi) - \tfrac{1}{2}\sum_{j=1}^{i}\ell n\left[\left|\underline{A}(j;\underline{a})\right|\right] - \tfrac{1}{2}\sum_{j=1}^{i}\operatorname{tr}\left\{\underline{A}^{-1}(j;\underline{a})\left[\underline{d}(j)\underline{d}^{T}(j) + E\left\{\tilde{\underline{r}}(j)\tilde{\underline{r}}^{T}(j)\,\middle|\,\underline{a}_t\right\}\right]\right\}$$

(4.2.31)

Although the preceding expression can be evaluated in a straightforward manner, it requires a considerable amount of computation. It will now be shown that a different but equivalent expression can be developed that allows a sizable reduction in required calculations.

The last term in equation (4.2.31) can be rewritten as a summation of:

$$\text{tr}\left[\underline{A}^{-1}(j;\underline{a})E\left\{[\underline{z}(j;\underline{a}_t)-\underline{H}(j)\overline{\underline{x}}(j;\underline{a})][\underline{z}(j;\underline{a}_t)-\underline{H}(j)\overline{\underline{x}}(j;\underline{a})]^T \mid \underline{a}_t\right\}\right]$$

which, for a given value of the true parameters, $\underline{a}_t$, is a function of $\underline{a}$. To evaluate the expectation, express it as

$$E\left\{[\underline{z}(j;\underline{a}_t)-\underline{H}(j)\overline{\underline{x}}(j;\underline{a})][\underline{z}(j;\underline{a}_t)-\underline{H}(j)\overline{\underline{x}}(j;\underline{a})]^T \mid \underline{a}_t\right\} =$$

$$= \underline{H}(j)E\left\{[\underline{x}(j;\underline{a}_t)-\overline{\underline{x}}(j;\underline{a})][\underline{x}(j;\underline{a}_t)-\overline{\underline{x}}(j;\underline{a})]^T \mid \underline{a}_t\right\}\underline{H}^T(j) + \underline{R}(j) \tag{4.2.32}$$

For specified values of $\underline{a}_t$ and $\underline{a}$, this can be evaluated by means of the error sensitivity analysis of a Kalman filter based on parameter values $\underline{a}$, but which is operating on a system which actually has parameter values specified by $\underline{a}_t$. That is, the system is given by

$$\underline{x}(i+1;\underline{a}_t) = \underline{\Phi}(i+1,i;\underline{a}_t)\underline{x}(i;\underline{a}_t) + \underline{B}(i;\underline{a}_t)\underline{u}(i) + \underline{G}(i)\underline{w}(i) \tag{4.2.33}$$

$$\underline{z}(i;\underline{a}_t) = \underline{H}(i)\underline{x}(i;\underline{a}_t) + \underline{v}(i) \tag{4.2.34}$$

and the filter is given by

$$\overline{\underline{x}}(i+1;\underline{a}) = \underline{\Phi}(i+1,i;\underline{a})\hat{\underline{x}}(i;\underline{a}) + \underline{B}(i;\underline{a})\underline{u}(i) \tag{4.2.35}$$

$$\hat{\underline{x}}(i;\underline{a}) = \overline{\underline{x}}(i;\underline{a}) + \underline{M}(i;\underline{a})\underline{H}^T(i)\underline{A}^{-1}(i;\underline{a})[\underline{z}(i;\underline{a}_t) - \underline{H}(i)\overline{\underline{x}}(i;\underline{a})] \tag{4.2.36}$$

where

$$\underline{A}(i;\underline{a}) = \underline{H}(i)\underline{M}(i;\underline{a})\underline{H}^T(i) + \underline{R}(i) \tag{4.2.37}$$

and where the estimated covariances propagate as:

$$\underline{M}(i+1;\underline{a}) = \underline{\Phi}(i+1,i;\underline{a})\underline{P}(i;\underline{a})\underline{\Phi}^{T}(i+1,i;\underline{a}) + \underline{G}(i)\underline{Q}(i)\underline{G}^{T}(i) \tag{4.2.38}$$

$$\underline{P}(i;\underline{a}) = \underline{M}(i;\underline{a}) - \underline{M}(i;\underline{a})\underline{H}^{T}(i)\underline{A}^{-1}(i;\underline{a})\underline{H}(i)\underline{M}(i;\underline{a}) \tag{4.2.39a}$$

$$= [\underline{I}-\underline{K}(i;\underline{a})\underline{H}(i)]\,\underline{M}(i;\underline{a})\,[\underline{I}-\underline{K}(i;\underline{a})\underline{H}(i)]^{T}+\underline{K}(i;\underline{a})\,\underline{R}(i)\,\underline{K}^{T}(i;\underline{a}) \tag{4.2.39b}$$

where

$$\underline{K}(i;\underline{a}) = \underline{M}(i;\underline{a})\underline{H}^{T}(i)\underline{A}^{-1}(i;\underline{a}) \tag{4.2.40}$$

This is exactly the problem treated in section 2.2 on sensitivity analysis of state estimation. Upon examination of equation (4.2.32), it can be seen that the matrix $\underline{M}_t(i;\underline{a}_t,\underline{a})$, as determined recursively by equations (2.2.20) to (2.2.30), will completely specify the ambiguity function. Thus, the value of $\psi_i[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t]$ previously given by equation (4.2.31) can be written equivalently as

$$\psi_i[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = -\frac{im}{2}\ln(2\pi) - \tfrac{1}{2}\sum_{j=1}^{i}\ln\left[\left|\underline{A}(j;\underline{a})\right|\right] - \tfrac{1}{2}\sum_{j=1}^{i}\operatorname{tr}\left[\underline{A}^{-1}(j;\underline{a})[\underline{H}(j)\underline{M}_t(j;\underline{a}_t,\underline{a})\underline{H}^{T}(j) + \underline{R}(j)]\right] \tag{4.2.41}$$

Similarly, for a log-likelihood function $\ln f(\underline{Z}_N(i)|\underline{Z}(i-N,\underline{a}))$, the N-step ambiguity function is

$$\psi_{i,N}[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = -\frac{Nm}{2}\ln(2\pi) - \tfrac{1}{2}\sum_{j=i-N+1}^{i}\ln\left[\left|\underline{A}(j;\underline{a})\right|\right] - \tfrac{1}{2}\sum_{j=i-N+1}^{i}\operatorname{tr}\left[\underline{A}^{-1}(j;\underline{a})[\underline{H}(j)\underline{M}_t(j;\underline{a}_t,\underline{a})\underline{H}^{T}(j) + \underline{R}(j)]\right] \tag{4.2.42a}$$

for $i \geq N$; and, for $i \leq N$,

$$\psi_{i,N}[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = \psi_i[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] \qquad (i \leq N) \qquad (4.2.42b)$$

An analogous development can be made for the log-likelihood function $\ell n\ f(\underline{Z}_N(i)|\,\underline{a})$. The values of $\underline{\bar{x}}$, $\hat{\underline{x}}$, $\underline{M}$, $\underline{P}$, and $\underline{A}$ are replaced with $\tilde{\underline{x}}^-$, $\tilde{\underline{x}}^+$, $\tilde{\underline{M}}$, $\tilde{\underline{P}}$, and $\tilde{\underline{A}}$ respectively. The evaluation of the $\tilde{\underline{M}}_t$ and $\tilde{\underline{P}}_t$ matrices can be accomplished by an N-step iteration each time: propagate up to the beginning of the N-step arc as though no measurements were taken ($\tilde{\underline{P}}_t(i-N) = \tilde{\underline{M}}_t(i-N)$ = one step propagation from preceding initial condition), and then use equations (2.2.20) and (2.2.21) to generate the values within the arc. A conservative approximation that requires less computation can be effected as described in the filter derivations. With the appropriate definitions, the results for this case are:

$$\tilde{\psi}_{i,N}[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = -\frac{Nm}{2}\,\ell n\,(2\pi) - \tfrac{1}{2}\sum_{j=i-N+1}^{i}\ell n\left[\left|\tilde{\underline{A}}(j;\underline{a})\right|\right]$$
$$- \tfrac{1}{2}\sum_{j=i-N+1}^{i}\mathrm{tr}\left[\tilde{\underline{A}}^{-1}(j;\underline{a})\left[\underline{H}(j)\tilde{\underline{M}}_t(j;\underline{a}_t,\underline{a})\underline{H}^T(j) + \underline{R}(j)\right]\right] \qquad (4.2.43a)$$

when $i \geq N$; and when $i \leq N$,

$$\tilde{\psi}_{i,N}[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = \psi_i[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] \qquad (i \leq N) \qquad (4.2.43b)$$

Now consider the contribution to the ambiguity function due to $\ell n\ f(\underline{x}(i)|\,\underline{Z}(i),\underline{a})$ or $\ell n\ f(\underline{x}(i)|\,\underline{Z}_N(i),\underline{a})$. Through the same reasoning as used previously, the function corresponding to $\ell n\ f(\underline{x}(i)|\,\underline{Z}(i),\underline{a})$ can readily be written as

$$\psi_i'[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = -\frac{n}{2}\,\ell n\,(2\pi) - \tfrac{1}{2}\,\ell n\left[\left|\underline{P}(i;\underline{a})\right|\right]$$
$$- \tfrac{1}{2}\,\mathrm{tr}\left[\underline{P}^{-1}(i;\underline{a})\underline{P}_t(i;\underline{a}_t,\underline{a})\right] \qquad (4.2.44)$$

and, for $\ell n\ f(\underline{x}(i)|\underline{Z}_N(i),\underline{a})$,

$$\tilde{\psi}_i'[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t] = -\frac{n}{2}\ell n\ (2\pi) - \tfrac{1}{2}\ell n\left[|\tilde{\underline{P}}(i;\underline{a})|\right]$$

$$-\tfrac{1}{2}\,\mathrm{tr}\left[\tilde{\underline{P}}^{-1}(i;\underline{a})\tilde{\underline{P}}_t(i;\underline{a}_t,\underline{a})\right] \qquad (4.2.45)$$

The effect of these contributions upon the Cramér-Rao lower bound on the parameter estimate error covariance warrants special attention. Although they serve to decrease the bound, the result is an unrealizable lower bound. The last term in (4.2.44) and (4.2.45) correspond to the likelihood equation terms that contain the vector $[\underline{x}(i) - \hat{\underline{x}}(i)]$ and $[\underline{x}(i) - \tilde{\underline{x}}^{+}(i)]$ respectively, all of which vanish. Consequently, these terms are to be ignored in establishing the lower bound from the ambiguity function.

Finally, consider the effect of apriori information, $\ell n\ f(\underline{a})$. If a normal distribution is assumed, with mean $\underline{\mu}$ and covariance $\underline{\Pi}$, the log-likelihood function becomes

$$\ell n\ f(\underline{a}) = -\frac{p}{2}\ell n\ (2\pi) - \tfrac{1}{2}\ell n\,|\underline{\Pi}| - \tfrac{1}{2}[\underline{a}-\underline{\mu}]^T\underline{\Pi}^{-1}[\underline{a}-\underline{\mu}] \qquad (4.2.46)$$

where p is the dimension of the parameter vector. It can be seen that the contribution to the ambiguity function is just $\ell n\ f(\underline{a})$ itself, regardless of $\underline{a}_t$. This is, of course, true for other assumed densities as well.

Equations (4.2.41) to (4.2.46) provide the means to construct ambiguity functions for the various forms of maximum likelihood estimators. Table 4.2.1 delineates the results.

These results are especially convenient. One would normally perform an error sensitivity analysis of a standard Kalman filter in order to determine which parameters might be the most critical to estimate. By so doing, he would simultaneously provide much of the data required to predict the global performance characteristics of the various forms of

Table 4.2.1

Ambiguity Functions

For an estimator based upon a log-likelihood function of:	The ambiguity function, $\Psi[\underline{x},\underline{a};\underline{x}_t,\underline{a}_t]$, is
$\ell n\ f(\underline{Z}(i)\mid \underline{a})$	ψ_i
$\ell n\ f(\underline{Z}_N(i)\mid \underline{a})$	$\tilde{\psi}_{i,N}$
$\ell n\ f(\underline{Z}_N(i)\mid \underline{Z}(i-N),\underline{a})$	$\psi_{i,N}$
$\ell n\ f(\underline{x}(i)\mid \underline{Z}(i),\underline{a})$	ψ_i'
$\ell n\ f(\underline{x}(i)\mid \underline{Z}_N(i),\underline{a})$	$\tilde{\psi}_i'$
$\ell n\ f(\underline{x}(i),\underline{Z}(i)\mid \underline{a})$	$\psi_i + \psi_i'$
$\ell n\ f(\underline{x}(i),\underline{Z}_N(i)\mid \underline{a})$	$\tilde{\psi}_{i,N} + \tilde{\psi}_i'$
$\ell n\ f(\underline{x}(i),\underline{Z}_N(i)\mid \underline{Z}(i-N),\underline{a})$	$\psi_{i,N} + \psi_i'$
$\ell n\ f(\underline{x}(i),\underline{a}\mid \underline{Z}(i))$	$\psi_i + \psi_i' + \ell n\ f(\underline{a})$
$\ell n\ f(\underline{x}(i),\underline{a}\mid \underline{Z}_N(i))$	$\tilde{\psi}_{i,N} + \tilde{\psi}_i' + \ell n\ f(\underline{a})$

maximum likelihood parameter estimators that are available. In practice, one would choose an $\underline{a}_t$ and an $\underline{a}$ and generate the sensitivity analysis and the factors ψ_i, $\psi_{i,N}$ and $\tilde{\psi}_{i,N}$ (for selected values of N), ψ_i', $\tilde{\psi}_i'$, and $\ln f(\underline{a})$. This would be repeated for a range of $\underline{a}$ values, thereby evaluating the various ambiguity functions for specified $\underline{a}_t$ as a function of $\underline{a}$. The entire process could be repeated for ranges of $\underline{a}_t$, $\underline{R}$, $\underline{Q}$, or $\underline{P}_o$ to discern the effect of these factors on the expected performance. For instance, by computing the Cramér-Rao lower bound as a function of a_t, the relative sensitivity of the parameter estimate over the entire range of admissible parameter values can be discerned.

Under certain circumstances, as perfect measurements or no dynamic driving noise, the expressions for the ambiguity function become vastly simplified. Schweppe (1970) indicates the procedure for a single-state time-invariant system, and this can readily be generalized to the case of an n-dimensional state with parameters modelled as invariant over N steps. Even if noise is known to enter into the dynamics or measurements, a designer might well investigate these special cases in his preliminary analyses to gain insights into performance without expending substantial time or effort.

CHAPTER V

METHODS OF ATTAINING ON-LINE APPLICABILITY

The full-scale solution algorithm presented in section 3.4 requires a substantial amount of computation after each new measurement is incorporated. In the most complete form of the algorithm, local iterations are made until the estimate has converged suitably, each iteration itself entailing an N-step propagation of equations (3.4.14) through (3.4.37). Even when local iterations are not employed, the computational load is excessive. In order to attain an estimation technique with on-line applicability, i.e. one that can provide a solution in real time, it is imperative to minimize the calculations while retaining adequate estimation accuracy. It is similarly desirable, though not as critical, to make the algorithm efficient for off-line use as well. Therefore, this chapter will examine various means of reducing the number of calculations and the effects such methods have upon the estimator performance.

Because of the time constraint inherent in on-line applications, certain general considerations will prevail. First, the parameter estimation should be restricted to the least number of parameters that provide acceptable system performance. Although it is mathematically optimal to estimate all uncertain values, the penalty in computation time might well prevent practical implementation, nullifying all benefits of parameter estimation. Secondly, the problem characterization of parameters varying significantly more slowly than the state variables allows a number of simplifications in the solution procedure, such as estimating the parameters less frequently than the states. In many applications, it is important to provide an updated state estimate extremely rapidly after incorporation of a measurement;

especially when feedback control is computed, the time lag introduced by state estimation must be kept as small as possible. On the other hand, the parameter estimate, whose purpose is to improve the state estimation (and possibly to evaluate controller gains), is not needed as quickly, especially because of its slowly varying nature. Consequently, a system mechanization that calculates the state estimate immediately upon receiving a measurement, using a previous parameter estimate value, and then performs the parameter calculations in the interim of the sample period, is superior to one which requires an iteration of the parameter calculations before it makes its (mathematically superior) state estimate available. Finally, it is advantageous to extract the essence of the maximum likelihood estimation method rather than incorporate the full-scale solution. Certain terms in the likelihood equations provide substantially less sensitivity to the parameter values than others, and these can be neglected without serious performance deterioration. Other quantities can be precomputed, either on the basis of a nominal set of parameter values or as a function of the parameter values, to be evaluated with the most recent parameter estimate, and the tradeoff of computation time and system performance can be employed to determine the appropriate approximations to use. Other aspects of on-line applications, such as required wordlength in the computer, will also be discussed.

To appreciate the magnitude of the computations required by each of the methods presented in this chapter, a representative problem will be posed, and the number of multiplications, additions, subtractions, and inversions counted for each case. Appendix F presents the general expressions for these values, in terms of the dimensions of the various vectors and matrices involved. The hypothetical problem involves the estimation of a five-dimensional state, two

parameters influencing the state transition matrix $\underline{\Phi}(i,i+1)$, and one parameter affecting the control input matrix $\underline{B}(i)$. The control vector is two-dimensional, computed in feedback fashion using the current state estimate (the controller gains are assumed to be precomputed). In addition, the state is driven by a scalar dynamic noise, and two-dimensional measurements are made at each sample time. Finally, it has been found that the parameters can be modelled adequately as constants over an interval of ten sample times. This problem will serve as a basis of comparison for the various techniques discussed in this chapter, indicating the effect each technique has upon the feasibility of on-line implementation. For the purposes of this comparison, the estimation based upon $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})$ will be used throughout the chapter.

In order to put the resulting numbers into perspective, the capabilities of current on-line digital hardware should be ascertained. The Apollo Guidance Computer has a memory cycle time of 11.7 microseconds; additions require two such cycle times (23.4 μsec.), multiplications four (46.9 μsec.), divisions seven (82.0 μsec.) and subtractions two or three (23.4 or 36.1 μsec.), depending upon the order in which it is implemented with respect to other operations. Other operations of interest are the branch and transfer of storage, each of which require two memory cycle times (23.4 μsec.), transfer of control which needs one cycle time (11.7 μsec.), and double precision addition requiring four (46.9 μsec.). Its memory is composed of 580,000 read-only bits and 32,000 erasable bits for data. (See Alonso and Randa (1967) and Apollo Guidance Program Symbolic Listing (1967).) In comparison, the IBM 4-Pi System, developed four years later and announced in 1967, has a 2.5 microsecond memory cycle time and 295,000 bits (expandable to 1.18 million). Further improvements can be projected in the future, and these two

general purpose computer systems are indicative of the rate of change in the state of the art.

Section 5.1 will discuss certain simplifications to the solution procedure proposed in section 3.4. However, the result will still be a full-scale estimation method, not suitable for on-line implementation. Therefore, section 5.2 will present two conceptualizations that exploit the slowly varying nature of the parameters to produce an estimation technique feasible for on-line use. Further efficiency is attainable by approximating the likelihood equations themselves; section 5.3 shows that a profitable, and generally adequate, approximation is that of using only weighted least squares type terms that appear in the likelihood equations. To reduce the number of computations required on-line, precomputation of certain variables is possible, as seen in section 5.4. There are also a number of advantageous state space representations, as phase variables or canonical variables, that reduce the amount of computations through matrix sparcity; these topics will be explored in section 5.5. The final sections will discuss means of exploiting symmetry and modifying measurement processing that further expedite an on-line implementation.

5.1 Approximations to the Solution Algorithm

Section 3.4 developed the recursions necessary to propagate the state and parameter estimates by the full-scale solution method. From these relations,it is possible to assess the number of multiplications, additions, subtractions, and matrix inversions to be performed in a single sample period in order to achieve these estimates. Table 5.1 portrays these values for the hypothetical problem posed at the beginning of this chapter. The totals are separated into evaluations necessary to propagate the state $\underline{x}$, score $\underline{s}$, and conditional information matrix $\underline{J}$, and finally the

Table 5.1

Full-Scale Solution of Hypothetical Problem

Term	Multiplications	Additions	Subtractions	Inversions
$\underline{x}$	8800	6638	270	10 (2x2)
$\underline{s}$	25742	20618	60	1 (5x5)
$\underline{J}$	120437	89765	0	0
$\underline{a}^*$	9	9	0	1 (3x3)

additional computations to achieve a new parameter estimate. For the case of local iterations, the numbers correspond to the number of computations in each local iteration.

It can be seen from this table that even for the moderate state dimension of the hypothetical problem, the computational load is great. The general expressions in Appendix F demonstrates that the number of additions and multiplications grows in proportion to the cube of the state dimension, and the burden on the computer thus becomes overwhelming as the number of state variables is increased.

The inordinately large number of calculations required for the conditional information matrix motivates search for a means of approximating its value in a simple fashion. One possibility is off-line precomputation, but discussion of this method will be deferred until section 5.4. Another technique would be to employ the approximation

$$E\left\{ \frac{\partial \underline{\bar{x}}(j)}{\partial a_k} \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell} \,\middle|\, \underline{a}_*(i) \right\} \cong \frac{\partial \underline{\bar{x}}(j)}{\partial a_k} \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell} \qquad (5.1.1)$$

This states that the expectation of the matrix $(\partial \underline{\bar{x}}(j)/\partial a_k) \cdot (\partial \underline{\bar{x}}(j)^T/\partial a_\ell)$ over all possible noise sequences can be adequately represented by the value it would attain due to the particular sequence which is assumed to have generated the measurement data. The difference between the two matrices can be expected to average out over an N-step propagation, especially for large N. Thus, the approximation is related to an assumption of ergodicity, that ensemble averages can be equated to time averages performed on one representative member of the ensemble. A further heuristic justification of equation (5.1.1) is that the scoring approximation to the Newton-Raphson method removed the dependence of the Jacobian matrix upon the particular sequence of data taken, and this dependence is now regained

in a different, but related, manner. From a practical standpoint, this should reduce sensitivity of the estimates to incorrectly assumed values for the noise statistics.

By incorporating equation (5.1.1) into (3.4.27) and (3.4.39), the component terms of the conditional information matrix become

$$
\begin{aligned}
E\left\{ s^1{}_k[\underline{Z}(j),\underline{a}_*(i)]\ s^1{}_\ell[\underline{Z}(j),\underline{a}_*(i)] \,\middle|\, \underline{a}_*(i)\right\} &\cong \\
&\cong \tfrac{1}{2}\ \mathrm{tr}\left\{ \underline{A}^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_k}\ \underline{A}^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_\ell} \right. \\
&\qquad \left. + 2\ \underline{A}^{-1}(j)\underline{H}(j)\ \frac{\partial \underline{\bar{x}}(j)}{\partial a_k}\ \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell}\ \underline{H}^T(j) \right\} \\
&= \tfrac{1}{2}\ \mathrm{tr}\left\{ \underline{A}^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_k}\ \underline{A}^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_\ell} \right\} \\
&\qquad + \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_k}\ \underline{H}^T(j)\underline{A}^{-1}(j)\underline{H}(j)\ \frac{\partial \underline{\bar{x}}(j)}{\partial a_\ell}
\end{aligned}
\qquad (5.1.2)
$$

When this approximation is used, the single step conditional information matrix will be denoted by $\underline{J}^1[\underline{Z}(j),\underline{a}_*(i)]$ instead of $\underline{J}^1[j,\underline{a}_*(i)]$, to show its explicit dependence upon the actual measurements taken. Similarly, $J_{k\ell}[\underline{Z}(i),\underline{a}_*(i)]$ is the sum of the $J^1{}_{k\ell}[\underline{Z}(j),\underline{a}_*(i)]$ values and the term

$$
\begin{aligned}
E\left\{ \gamma_k[\underline{Z}(i),\underline{a}_*(i)]\ \gamma_\ell[\underline{Z}(i),\underline{a}_*(i)] \,\middle|\, \underline{a}_*(i)\right\} &\cong \\
&\cong \tfrac{1}{2}\ \mathrm{tr}\left\{ \underline{P}^{-1}(i)\ \frac{\partial \underline{P}(i)}{\partial a_k}\ \underline{P}^{-1}(i)\ \frac{\partial \underline{P}(i)}{\partial a_\ell} \right\} \\
&\qquad + \frac{\partial \underline{\hat{x}}^T(i)}{\partial a_k}\ \underline{P}^{-1}(i)\ \frac{\partial \underline{\hat{x}}(i)}{\partial a_\ell}
\end{aligned}
\qquad (5.1.3)
$$

From these relations can be seen the tremendous benefit of using the proposed approximation: the conditional information matrix can be evaluated using only the propagations necessary for the state and score computations. The matrix recursions (3.4.24) to (3.4.26) and (3.4.35) to (3.4.37), and

the associated vector propagations (3.4.28) and (3.4.29) are no longer required. For the hypothetical problem posed previously, 1598 multiplications and 1117 additions are required to evaluate $\underline{J}[\underline{Z}(i), \underline{a}_*(i)]$. Of these, only 790 multiplications and 464 additions are required for the ten-term sum of components given by equation (5.1.2), the remainder being for the final term given by (5.1.3): this is caused by the difference between the state and measurement dimensions. These numbers are vast reductions from those appearing in Table 5.1, and yet investigations have shown the corresponding accuracy and rate of convergence to be exceptionally high. Chapter VI will delineate some of these results.

One problem that may arise in practice is the accentuation of the small magnitude of the conditional information matrix near the initial time. Even in the full-scale evaluation, $\underline{J}[i, \underline{a}_*(i)]$ has a small magnitude for the first few sample times, before many terms enter the summations; the realized residual sequence may be such that the problem is more severe in the approximate case. In either event, the small magnitude of $\underline{J}$ results in a large $\underline{J}^{-1}$, and for a number of samples this value may not correspond at all to the second derivative matrix it is meant to represent. This potential stability problem can be alleviated simply in one of three ways. Simulation studies can be used to determine a minimum value for $\underline{J}$ (and corresponding maximum value of $\underline{J}^{-1}$) that produces acceptable results; the magnitude of the computed $\underline{J}$ can then be compared to this, and if it is too small, the precomputed value of $\underline{J}^{-1}$ can be used for the parameter estimation. More simply, a precomputed $\underline{J}^{-1}$ value could be employed for a prespecified number of initial stages of the estimation process. Finally, the first parameter estimate update can be deferred until a certain minimum number of samples, again determined by simulation, have been taken.

The slowly varying (or constant) nature of the system

parameters can be exploited to introduce other simplifications into the solution algorithm as well, especially once the parameter estimate has converged to a stable value (and is therefore probably tracking the true value with some accuracy). When recomputing the state estimates over the N-step interval with a new parameter estimate, it is possible to use a first order update rather than that specified in section 3.4. This would entail using the newly evaluated $\underline{\Phi}$ and $\underline{B}$ matrices to propagate the vector quantities $\underline{\bar{x}}(j)$ and $\partial\underline{\bar{x}}(j)/\partial a_k$, but retaining the same second order statistics, $\underline{M}(j)$ and $\partial\underline{M}(j)/\partial a_k$, as used previously unless the parameter estimate has changed by more than some prespecified amount. (If these second order quantities have not attained steady state values by the end of the N-step interval, a single step propagation is necessary to generate their values for the last stage in the interval.) Thus, the state estimate would incorporate the newest parameter estimate, except that the gains $\underline{K}(j)$ would be generally based on an older parameter estimate; only at points of total relinearization would the first and second order statistics depend on the same parameter estimate. Similarly, a "first order" parameter estimate can be used: the score is recomputed each time an estimate is made, but the inverse conditional information matrix is re-evaluated only periodically. Except for very small interval sizes, the N-step $\underline{J}^{-1}[\,i\,,\,\underline{a}_*(i)\,]$ matrix does not vary significantly with a small change in $\underline{a}_*(i)$, so it is not essential to recompute it every sample time.

If increased accuracy is desired in the iterations, as near the initial time before a stable parameter estimate is achieved, there is an alternative to local iterations. At the end of a single iteration of the full-scale method, estimates of $\underline{\hat{x}}[j;\underline{a}^*(i-1)]$ for $j = 1, \ldots, i$ (or starting at $(i-N+1)$ if i is greater than N) and $\underline{a}^*(i)$ have been made. It is then possible to process just the state equations again, using $\underline{a}^*(i)$, to obtain the values $\underline{\hat{x}}[j;\underline{a}^*(i)]$, yielding a current state estimate of $\underline{\hat{x}}[i;\underline{a}^*(i)]$. Near the initial

time, this may differ substantially from $\hat{\underline{x}}[i;\underline{a}^*(i-1)]$. This method still suffers from delaying the availability of the state estimate, though less so than local iterations. If a stable parameter estimate has not been attained by the time $i = N$, then it would be advantageous at least to perform a single stage of this state recalculation, to obtain $\hat{\underline{x}}[i-N+1;\underline{a}^*(i)]$ and $\underline{P}[i-N+1;\underline{a}^*(i)]$ as the initial conditions for the estimates at the next sample time. In the practical experience gained to date, however, state recalculation has not been necessary to insure adequate performance.

5.2 On-Line Conceptualizations

Although the previous section introduced certain simplifications which can reduce the number of required calculations quite substantially, the resulting formulation does not possess on-line applicability. Every sample period, it requires the regeneration of the state estimate and component terms of the parameter estimate over an N-step interval: at sample time i, $\hat{\underline{x}}(i-N)$ and $\underline{P}(i-N)$ are used to initialize the propagations of the state and the variables comprising the score and conditional information matrix, based upon the most recent parameter estimate, and at the end of the propagation, a new parameter estimate is calculated. On-line applications would essentially preclude the re-evaluation of all N residuals, $\overline{\underline{x}}(j)$ values, score and $\underline{J}$ component values within a sample period. Consequently, algorithms are sought to approximate the full-scale solution method, providing comparable performance while not necessitating an N-step propagation whenever an estimate is made. Such algorithms are conceivable because of the slowly varying character of the parameters.

Since the parameters are assumed to be essentially constant over N periods, once a good estimate is made, it should not vary significantly in N steps. Therefore, for certain applications, it is adequate to estimate the parameters only

every N sample periods. The implementation for this case is especially simple, requiring only the maintenance within specified computer registers of running sums for the score and conditional information matrix. To describe the procedure, let a parameter estimate, $\underline{a}^*(j)$, be made, and the registers for the score and conditional information matrix set to zero. For the next N sample periods, this parameter estimate is used to propagate the state, score, and conditional information matrix relations given in section 3.4, one step at a time. (The $\underline{J}$ approximation of section 5.1 would be incorporated.) At each sample time, the running sum registers, whose contents are defined as $\tilde{\underline{s}}(i)$ and $\tilde{\underline{J}}(i)$, are updated according to

$$\tilde{\underline{s}}(i) = \tilde{\underline{s}}(i-1) + \underline{s}^1[\underline{Z}(i),\underline{a}^*(j)] \tag{5.2.1}$$

$$\tilde{\underline{J}}(i) = \tilde{\underline{J}}(i-1) + \underline{J}^1[\underline{Z}(i),\underline{a}^*(j)] \tag{5.2.2}$$

where $\underline{s}^1[\underline{Z}(i),\underline{a}^*(j)]$ is the single step score defined by equation (3.4.22) and $\underline{J}^1[\underline{Z}(i),\underline{a}^*(j)]$ is the single step conditional information matrix defined by (5.1.2). After N samples of measurement data have been accumulated since the last parameter estimate, i.e. when i = j+N, the final components of the score and conditional information matrix are added to the current contents of the registers:

$$\underline{s}[\underline{Z}(j+N),\underline{a}^*(j)] = \tilde{\underline{s}}(j+N) + \underline{s}[\underline{Z}(j+N),\underline{a}^*(j)] \tag{5.2.3}$$

$$\underline{J}[\underline{Z}(j+N),\underline{a}^*(j)] = \tilde{\underline{J}}(j+N) + E\{\underline{s}(j+N)\ \underline{s}^T(j+N)\,|\,\underline{a}^*(j)\} \tag{5.2.4}$$

where these terms are defined by (3.4.38) and (5.1.3) respectively. Using these values, a new parameter estimate is computed as:

$$\underline{a}^*(j+N) = \underline{a}^*(j) + \underline{J}[\underline{Z}(j+N), \underline{a}^*(j)]^{-1} \underline{s}[\underline{Z}(j+N), \underline{a}^*(j)] \tag{5.2.5}$$

The advantage to this method is that the calculations necessary for updating the running sums can be performed between sample times, and only that contribution due to the most recent measurement residual has to be processed within a sample period. The state estimate is available almost instantaneously after the measurement is taken at each time, which is especially desirable if the state estimate is to be used for determining the control to be applied for the succeeding sample period. Even when feedback control is not to be used, it is usually very advantageous to provide a state estimate rapidly after incorporating a measurement; it is not usually critical to process the parameter estimate so quickly, especially since the parameters vary slowly with respect to the state. Besides the rapid availability of the state estimate and the relatively small computational load, this method does not require the saving of individual measurements or measurement residuals, but only the updating of running sum registers which accumulate the information over N steps and are then erased. Thus, the memory requirement is not excessive. It is noteworthy, however, that this procedure produces a new parameter estimate only every N samples, yielding slower initial convergence to a good estimate and an inherent lag in estimating a parameter change. Moreover, this method creates estimates based upon nonoverlapping batches of data, so there may be oscillations in the successive parameter estimates, as due to the effects of noises, even though the actual parameter values are known to be essentially constant over N sample periods. Solutions to these problems will be discussed later in this section.

If it is desired to evaluate a parameter estimate more

frequently than every N sample periods, a different technique must be utilized. It will replace the full-scale regeneration of N-step sums for the score vector and conditional information matrix with the addition of a term to the end of the interval and subtracting a corresponding term from the beginning. Once a single-step component term of the score or conditional information matrix has been evaluated, it will be considered invariant and will not be recalculated, but instead stored for such time that knowledge of its value is required. In other words, to propagate from one sample time to the next, only the values of $\underline{s}^1$ and $\underline{J}^1$ pertaining to the current sample time will be evaluated, rather than recomputing the previous (N-1) such values as well. To provide storage of the N most recent sets of component terms, an N-block read-write memory, as perhaps a drum (inexpensive storage) would be maintained. Running sum registers would also be provided to retain the latest evaluation of the conditional information matrix and score vector.

The procedure can be performed in the following manner. The single measurement values, $\underline{s}^1$ and $\underline{J}^1$, are evaluated every sample instant, whether or not a new parameter estimate is to be made. If N or fewer measurements have been made, these values are read into the read-write memory storage and also added to the running sums. For this case, the running sums are updated as

$$\tilde{\underline{s}}(i) = \tilde{\underline{s}}(i-1) + \underline{s}^1[\underline{Z}(i), \underline{a}^*(j)] \tag{5.2.6}$$

$$\tilde{\underline{J}}(i) = \tilde{\underline{J}}(i-1) + \underline{J}^1[\underline{Z}(i), \underline{a}^*(j)] \tag{5.2.7}$$

where j is the most recent time that the parameter estimate was made. For more than N measurements already processed, the read-write memory will contain blocks of component values from times (i-N), (i-N+1), ... (i-1) just before the i-th measurement. When the measurement is taken, the $\underline{s}^1$ and $\underline{J}^1$

values pertaining to time (i-N) are read out of the memory to evaluate the running sums as

$$\tilde{\underline{s}}(i) = \tilde{\underline{s}}(i-1) + \underline{s}^1[\underline{Z}(i),\underline{a}^*(j)] - \underline{s}^1[\underline{Z}(i-N),\underline{a}^*(k)] \tag{5.2.8}$$

$$\tilde{\underline{J}}(i) = \tilde{\underline{J}}(i-1) + \underline{J}^1[\underline{Z}(i),\underline{a}^*(j)] - \underline{J}^1[\underline{Z}(i-N),\underline{a}^*(k)] \tag{5.2.9}$$

The values $\underline{s}^1[\underline{Z}(i),\underline{a}^*(j)]$ and $\underline{J}^1[\underline{Z}(i),\underline{a}^*(j)]$ are then read into the locations previously assumed by $\underline{s}^1$ and $\underline{J}^1$ from time (i-N). In practice, it may be advisable to maintain an (N+1)-block memory of $\underline{s}^1$ and $\underline{J}^1$ values from (i-N) through i at all times, thus providing a simple computer restart capability to protect against loss of data in case of power failure or other mishap. At any time that a new parameter estimate is required, the total score and conditional information matrix are evaluated as

$$\underline{s}[\underline{Z}(i),\underline{a}^*(j)] = \tilde{\underline{s}}(i) + \underline{\mathscr{s}}[\underline{Z}(i),\underline{a}^*(j)] \tag{5.2.10}$$

$$\underline{J}[\underline{Z}(i),\underline{a}^*(j)] = \tilde{\underline{J}}(i) + E\{\underline{\mathscr{s}}(i)\underline{\mathscr{s}}^T(i)\,|\,\underline{a}^*(j)\} \tag{5.2.11}$$

with which the parameter estimate is calculated as

$$\underline{a}^*(i) = \underline{a}^*(j) + \underline{J}[\underline{Z}(i),\underline{a}^*(j)]^{-1}\,\underline{s}[\underline{Z}(i),\underline{a}^*(j)] \tag{5.2.12}$$

These equations are approximations since they do not entail recalculating all quantities in the N-step interval with each new parameter estimate, but merely adding a component to the end and subtracting a factor from the beginning of the summation terms. Considering the residuals that appear in these terms, the full-scale method would use (for the case of estimating the parameter every sample instant):

$$\underline{r}[i-N+1;\underline{a}^*(i-1)]\ ,\ \underline{r}[i-N+2;\underline{a}^*(i-1)]\ ,\ \ldots,\ \underline{r}[i;\underline{a}^*(i-1)]$$

whereas the approximate procedure would employ

$$\underline{r}[i-N+1;\underline{a}^*(i-N)]\ ,\ \underline{r}[i-N+2;\underline{a}^*(i-N+1)]\ ,\ \ldots,\ \underline{r}[i;\underline{a}^*(i-1)]$$

If the assumption that the parameters are slowly varying is correct, then these sequences should not differ significantly. However, because of the lack of recalculation, a term evaluated with a bad parameter estimate (as especially near the initial time before the estimator starts to track the true parameter values with accuracy) will be retained for the number of steps in the interval, at which time it will be subtracted (the subtraction of a large term can affect stability and other behavior very adversely). This may well cause unacceptable transient response in the parameter estimate. The problem can be eliminated in practice by decreasing the size of the interval, thereby retaining "old" data for a shorter period of time. This effectively reduces the lag in responding to information that the parameter value is different than previously estimated. Then, to incorporate the expectation that the parameters will be approximately invariant over N steps, the parameter estimates made over the most recent N-step interval can be averaged. Investigations have shown this to be a valid and advantageous procedure. Chapter VI will demonstrate its effect on parameter estimate tracking ability and overall performance in three different applications.

Table 5.2 reveals the reduction in computations afforded by the latter of the two conceptualizations. These numbers pertain to a sample period in which an estimate is made - the circumstance in which the most calculations must be performed within a sample period. Note that the approximation for $\underline{J}$ given in section 5.1 is employed.

For the instance of not updating the parameter estimate, the required calculations are as in Table 5.3. The differences

Table 5.2

On-Line Conceptualization for Hypothetical Problem

Term	Multiplications	Additions	Subtractions	Inversions
$\underline{x}$	889	671	27	1 (2x2)
$\underline{s}$	2801	2249	9	1 (5x5)
$\underline{J}$	887	694	6	0
$\underline{a}^*$	9	9	0	1 (3x3)

Table 5.3

On-Line Conceptualization - No Parameter Update

Term	Multiplications	Additions	Subtractions	Inversions
$\underline{x}$	889	671	27	1 (2x2)
$\underline{s}$	2549	2041	9	0
$\underline{J}$	79	47	6	0
$\underline{a}^*$	0	0	0	0

between this and the previous table are due mostly to the removal of the final components of the score and conditional information matrix given by equations (3.4.38) and (5.1.3), the magnitude of which depend upon the state dimension. The first conceptualization differs from these results in that it requires three fewer subtractions for $\underline{s}$ and six fewer for $\underline{J}$.

Thus far, the procedures of this section have pertained to a log-likelihood function of $\ell n\ f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$ and the related $\ell n\ f(\underline{x}(i), \underline{Z}(i) | \underline{a})$ which provides the startup for the first N measurements. For this case, the propagation of $\overline{\underline{x}}$ and $\hat{\underline{x}}$, $\underline{M}$ and $\underline{P}$, $\partial\overline{\underline{x}}/\partial a_\ell$ and $\partial\hat{\underline{x}}/\partial a_\ell$, and $\partial\underline{M}/\partial a_\ell$ and $\partial\underline{P}/\partial a_\ell$ are effected by simple one-step recursions, as

$$\underline{M}(i+1) = \underline{\Phi}(i+1,i)\underline{P}(i)\underline{\Phi}^T(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^T(i) \tag{5.2.13}$$

$$\underline{P}^{-1}(i) = \underline{M}^{-1}(i) + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i) \tag{5.2.14}$$

or such equations transformed by the matrix inversion lemma. However, if the estimator based upon $\ell n\ f(\underline{x}(i), \underline{Z}_N(i) | \underline{a})$ is contemplated for on-line use, the required recursions are considerably more complex. From equation (4.2.24), it can be seen that the contribution to the estimate of $\underline{x}(i)$ due to $\underline{z}(i-N)$ is

$$\left[\left[\underline{I}-\underline{K}(i)\underline{H}(i)\right]\underline{\Phi}(i,i-1)\right] \cdots \left[\left[\underline{I}-\underline{K}(i-N+1)\underline{H}(i-N+1)\right]\underline{\Phi}(i-N+1,i-N)\right] \cdot$$

$$\cdot\ \underline{K}(i-N)\left[\underline{z}(i-N) - \underline{H}(i-N)\underline{x}^\circ(i-N)\right]$$

When passing from instant (i-1) to i, this term should be subtracted from the state estimate. Even approximations, as those suggested in Appendix C, require substantially more calculations:

$$\tilde{\underline{M}}(i+1) = \underline{\Phi}(i+1,i)\tilde{\underline{P}}(i)\underline{\Phi}^{T}(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^{T}(i) \tag{5.2.15}$$

$$\tilde{\underline{P}}^{-1}(i) = \underline{M}^{-1}(i) + \underline{H}(i)\underline{R}^{-1}(i)\underline{H}(i) - \underline{\Phi}^{T}(i-N,i)\underline{H}^{T}(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i) \tag{5.2.16}$$

It would be necessary to justify this additional cost in computer loading by improvement in performance before this implementation could be considered as a practical alternative to the previous forms. It is possible to evaluate the performance of an estimator based upon $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{a})$ by means of the ambiguity function concept or by simulations of the full-scale solution using this log-likelihood function. For the examples investigated, no appreciable difference in estimator performance was apparent. Thus, it is expected that there will not generally be sufficient benefit from the additional calculations to warrant the use of this estimator for on-line applications.

A similar consideration concerns the propagation of the matrices $\partial\underline{M}/\partial a_\ell$ and vectors $\partial\underline{\bar{x}}/\partial a_\ell$ for the on-line conceptualization that creates parameter estimates more frequently than every N samples. In the full-scale solution, these N-step propagations are initialized by $\partial\underline{P}(i-N)/\partial a_\ell = \underline{0}$ and $\partial\underline{\hat{x}}(i-N)/\partial a_\ell = \underline{0}$, but the on-line method must reflect these reinitializations without actually performing an N-step propagation. This does not present as formidable a problem as does the subtraction of terms in the implementation of $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{a})$ since these initializations are known to be zero. In practical applications, these terms can be set to zero periodically, as every 2N sample periods, to attain adequate behavior. Furthermore, the matrices tend to steady state values that are a function of the most recent estimate, and in many applications will attain these values (or within a small magnitude difference of them) within N propagation

steps; thus, it is often possible to neglect the reinitializations of these matrices. Note that the on-line implementation that estimates the parameters only every N measurement times simply sets these vectors and matrices to zero immediately after every parameter estimate.

At this point it would be useful to list the factors that favor as large a choice of N as possible and those considerations which support a small interval length. The reasons for making N large are:

1) A parameter estimator based upon large N is less susceptible to errors due to poor quality measurement data than are systems that would incorporate an independent parameter estimate at each sample (N=1) or systems with small N. By reflecting the physical expectation that the parameter values are rather consistent from sample to sample, the parameter estimate high frequency oscillation can be reduced.
2) The asymptotic properties of the estimator, as being asymptotically unbiased, normally distributed, and efficient, pertain to the case of $N \rightarrow \infty$. Thus the larger the size of the interval, the more confident the user can be that the estimator behavior will not differ significantly from these theoretically derived properties.
3) By means of the ambiguity function development, the Cramér-Rao lower bound on the covariance of the error in the parameter estimate can be shown to be inversely related to N. As such, estimator performance is penalized heavily for very small interval sizes, but the rate of reduction of this lower bound quickly decreases for larger values of N. Chapter VI will portray typical shapes of the functional relationship between the lower bound and interval size. In effect, this consideration establishes a lower bound on N that can possibly provide adequate parameter estimation accuracy.
4) Certain approximations, as especially:

$$\frac{\partial^2 L}{\partial \underline{\xi}^2}\left[\underline{\xi}, \underline{Z}_N(i)\right] \Bigg|_{\underline{\xi} \to \underline{\xi}^*(i)} \cong -\underline{J}\left[i, \underline{\xi}^*(i)\right] \qquad (5.2.17)$$

and

$$E\left\{\frac{\partial \underline{\bar{x}}(j)}{\partial a_k} \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell} \,\middle|\, \underline{a}_*(i)\right\} \cong \frac{\partial \underline{\bar{x}}(j)}{\partial a_k} \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell} \qquad (5.2.18)$$

become more valid with increasing N. They are accurate for large samples of data, but their accuracy becomes less certain as the sample size decreases.

The reasons for seeking a small interval size are:

1) There is the upper limit on N imposed by the validity of the model that the system parameters are essentially constant over the entire interval. The definition of "essentially constant" is a subjective one.
2) Another limit is imposed by the capacity of the computer to accomplish the required calculations in an allotted portion of a sample period. Here, the idea of "portion" is stressed because many potential applications would employ a general-purpose on-line computer, which has to execute many other functions as well.
3) Memory storage requirements similarly restrict the interval size, as well as the amount of data saved to perform the estimates.
4) For the on-line applications described, a small interval is necessary to maintain adequate tracking capability. A large amount of "old" data retained can seriously affect the initial convergence to the neighborhood of the true parameter values, and also create a significant lag in the response to changes in these true values.

A tradeoff must be analyzed, and the value of N that is most compatible with these diverse factors chosen as the actual interval size to implement on-line.

5.3 Approximations to the Likelihood Equations

This section will discuss certain approximations that can be made to the likelihood equations themselves. The objective of an on-line estimation technique is to preserve the essence of a full-scale method with a minimum of computational effort. By separating the effects of the individual terms comprising the ambiguity functions, the sensitivity to parameter values of the corresponding terms in the likelihood equations can be assessed. These relative sensitivities can also be substantiated by simulations in which particular terms of the likelihood equations have been neglected. When significant discrepancies exist in the various terms' capability or efficiency in providing accurate parameter estimates, the "higher order" terms can be removed without notable loss in performance. Experience has shown that, of the possible implementations of this form, the most successful is the inclusion of only weighted-least-squares type of terms. Not only does this remove a considerable amount of computation, but it provides estimates of the same quality as attained with the more complex algorithm. In fact, certain terms removed in this manner can be shown to contribute a bias to the parameter estimate, further motivating their removal, provided that the remaining terms can effect an adequate estimate.

Section 5.3.1 describes the on-line algorithms that are produced by including only the weighted least squares type of terms in the estimator based on $\ell n\ f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$. In section 5.3.2 further approximations are introduced to yield a desirable algorithmic form. Appendix G employs the weighted least squares approximation idea to derive a closed form estimator (as opposed to an iterative form) of the uncertain elements in $\underline{\Phi}$ or $\underline{B}$; this development is separated into an appendix because it differs substantially from the trend of thought in this chapter.

5.3.1 Inclusion of Only Weighted Least Squares Terms

Maximum likelihood estimation techniques seek the mode of a specified conditional probability density to effect an estimate of the desired variables. If Gaussian statistics are involved, the maximum likelihood estimate of a variable $\underline{\xi}$ is in general that value which maximizes the log-likelihood function

$$L = \ell n \left[(2\pi)^{-\frac{n}{2}} |\underline{Y}|^{-\frac{1}{2}} \exp\left\{ - \tfrac{1}{2} \underline{\xi}^T \underline{Y}^{-1} \underline{\xi} \right\} \right] \tag{5.3.1}$$

as a function of $\underline{\xi}$. If $\underline{Y}$ is not a function of $\underline{\xi}$, this is equivalent to minimizing the term

$$\epsilon = \tfrac{1}{2} \underline{\xi}^T \underline{Y}^{-1} \underline{\xi} \tag{5.3.2}$$

as a function of $\underline{\xi}$. In comparison, the method of weighted least squares would obtain an estimate as the value of $\underline{\xi}$ which minimizes the function

$$\nu = \tfrac{1}{2} \underline{\xi}^T \underline{W} \underline{\xi} \tag{5.3.3}$$

with respect to $\underline{\xi}$, where $\underline{W}$ is an arbitrarily, but conveniently, chosen weighting matrix. Therefore, for the case of Gaussian statistics in which the covariance $\underline{Y}$ is not a function of the uncertain parameters $\underline{\xi}$, the method of weighted least squares is equivalent to the maximum likelihood technique if the weighting matrix $\underline{W}$ is set equal to the appropriate inverse covariance, $\underline{Y}^{-1}$. There is nothing inherent in the weighted least squares technique, however, that specifies the appropriate choice of this weighting matrix. If ξ_ℓ is the ℓ-th component of the p-dimensional vector $\underline{\xi}$, then the optimal estimate using either technique becomes the solution to the p simultaneous equations:

$$\left.\frac{\partial \underline{\xi}^T}{\partial \xi_\ell} \underline{Y}^{-1} \underline{\xi}\right|_{\underline{\xi} \to \underline{\xi}^*} = 0 \qquad (\ell = 1, 2, \ldots p) \tag{5.3.4}$$

However, if the value of $\underline{Y}$ is itself a function of $\underline{\xi}$, then the maximum likelihood estimate becomes the solution to the p simultaneous equations

$$\left.\frac{\partial \underline{\xi}^T}{\partial \xi_\ell} \underline{Y}^{-1} \underline{\xi} + \tfrac{1}{2} \operatorname{tr}\left\{ \underline{Y}^{-1} \frac{\partial \underline{Y}}{\partial \xi_\ell} \right\} - \tfrac{1}{2} \underline{\xi}^T \underline{Y}^{-1} \frac{\partial \underline{Y}}{\partial \xi_\ell} \underline{Y}^{-1} \underline{\xi}\right|_{\underline{\xi} \to \underline{\xi}^*} = 0 \qquad (\ell = 1, 2, \ldots p) \tag{5.3.5}$$

For this case, there are two plausible weighted least squares estimates. First, a $\underline{W}$ matrix can be chosen, and the value of $\tfrac{1}{2} \underline{\xi}^T \underline{W} \underline{\xi}$ minimized with respect to $\underline{\xi}$ by evaluating

$$\left.\frac{\partial \underline{\xi}^T}{\partial \xi_\ell} \underline{W} \underline{\xi} + \tfrac{1}{2} \underline{\xi}^T \frac{\partial \underline{W}}{\partial \xi_\ell} \underline{\xi}\right|_{\underline{\xi} \to \underline{\xi}^*} = 0 \qquad (\ell = 1, 2, \ldots p) \tag{5.3.6}$$

Allowing $\underline{W}$ to be set equal to $\underline{Y}^{-1}$ yields the same equations as (5.3.5), minus the trace term. Another means of obtaining a weighted least squares estimate would entail minimizing $\tfrac{1}{2} \underline{\xi}^T \underline{W} \underline{\xi}$ with respect to $\underline{\xi}$ for an arbitrarily chosen matrix $\underline{W}$, treating $\underline{W}$ only as matrix of weighting parameters and not as a function of $\underline{\xi}$. Then, in order to evaluate the estimate completely, a convenient value for $\underline{W}$ is used in the calculations. This procedure leads to the equations:

$$\left.\frac{\partial \underline{\xi}^T}{\partial \xi_\ell}\,\underline{W}\,\underline{\xi}\right|_{\underline{\xi}\to\underline{\xi}^*} = \left.\frac{\partial \underline{\xi}^T}{\partial \xi_\ell}\,\underline{Y}^{-1}\underline{\xi}\right|_{\underline{\xi}\to\underline{\xi}^*}$$

$$= 0 \qquad (\ell = 1,2,\ldots p) \qquad (5.3.7)$$

These relations suggest that a practical on-line implementation of a maximum likelihood estimator may be achieved by retaining only those terms of the likelihood equations that are of the weighted least squares form. Whereas the full-scale estimator would use equation (5.3.5), it may well be adequate to retain only the first and third terms, or just the first term, of this relation.

Neglecting the trace term in (5.3.5) is essentially the same as removing the factor $\ell n\left[(2\pi)^{-\frac{n}{2}}|\underline{P}(i)|^{-\frac{1}{2}}\right]$ and $\ell n\left[(2\pi)^{-\frac{m}{2}}|\underline{A}(j)|^{-\frac{1}{2}}\right]$ from the likelihood functions $\ell n\ f(\underline{x}(i)|\underline{Z}(i),\underline{a})$ and $\ell n\ f(\underline{z}(j)|\underline{Z}(j-1),\underline{a})$ respectively. It can be seen from the propagation relations in section 3.4 that these factors are not functionally dependent upon the actual measurements taken. Since these measurements are the single source of information for adjusting the parameter estimates, these factors cannot make a valid contribution to the estimates. Figure 5.1 is a representative portrayal of the two component terms of an ambiguity function plotted as a function of the parameter value a, for a particular value of the true parameter value a_t. With regard to equation (5.3.1), ψ_1 is the component due to $\ell n\left[(2\pi)^{-\frac{n}{2}}|\underline{Y}|^{-\frac{1}{2}}\right]$, and ψ_2 is due to $\left[-\tfrac{1}{2}\underline{\xi}^T\underline{Y}^{-1}\underline{\xi}\right]$. The latter is considerably more sensitive to parameter variations. Moreover, the peak of ψ_1 will not generally occur at a_t, but at some fixed value of a no matter what the true value, a_t, might be. This will be shown in Chapter VI for some particular examples. Thus, the factors comprising ψ_1 generally contribute an insignificant bias and no additional valid parameter sensitivity, so they can be neglected.

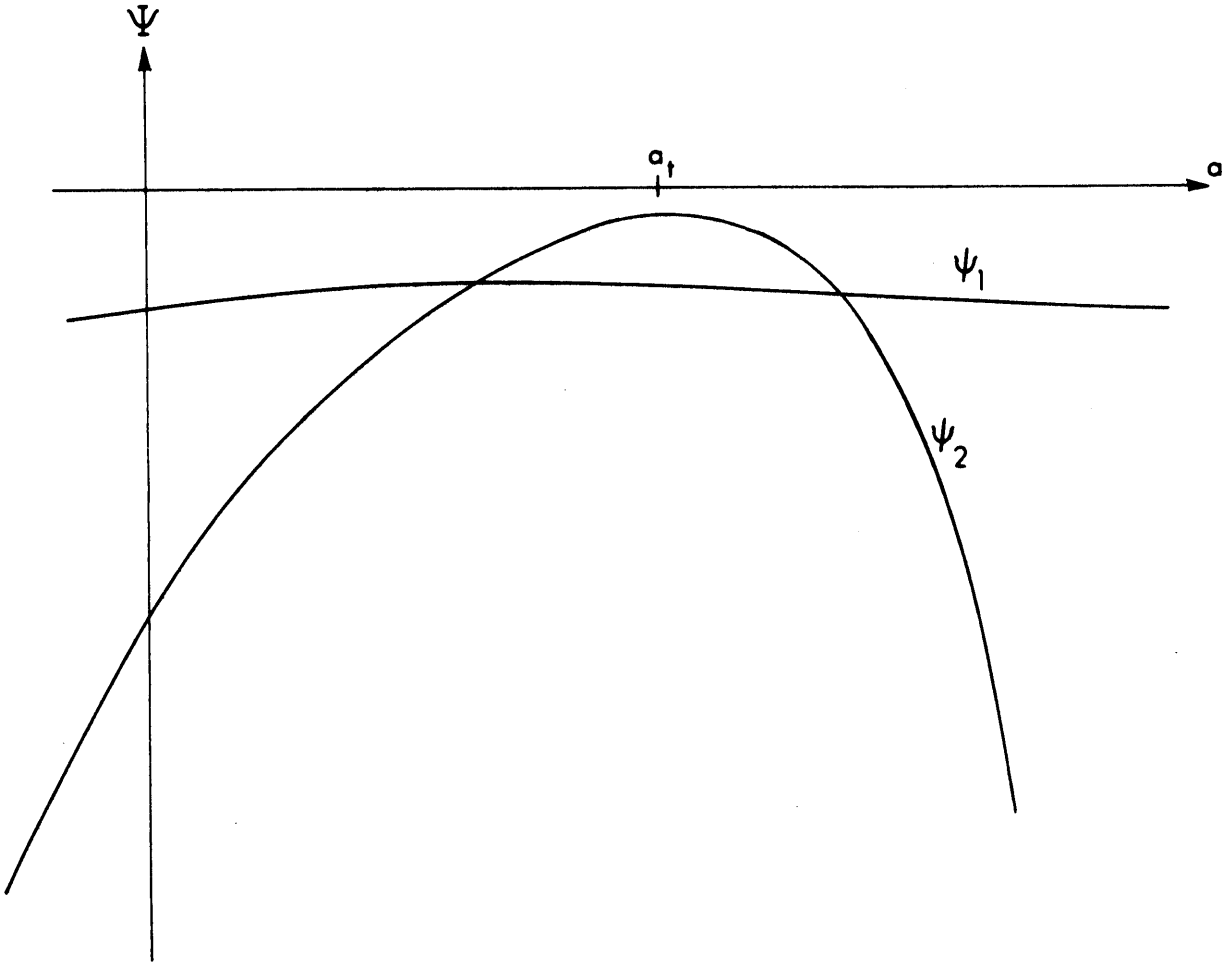


Figure 5.1 Ambiguity Function Component Terms

If, in addition, the last term in equation (5.3.5) can be ignored with respect to the first term, and adequate estimator performance is maintained, then further simplifications are possible. This would imply that $\partial \underline{M}(j)/\partial a_\ell$ can be neglected compared to the effects of $\partial \underline{\bar{x}}(j)/\partial a_\ell$, and so the score equations of section 3.4 could be altered by the removal of the trace term in (3.4.22), the term that differentiates (3.4.34) from (3.4.34'), and the entire equation (3.4.38). If the approximation of section 5.1 is used, then the trace terms in (5.1.2) and (5.1.3) are similarly removed. Since the score effects of $\underline{\mathcal{F}}[\underline{Z}(i), \underline{a}_*(i)]$ have already been removed, it is in fact appropriate to neglect all of equation (5.1.3), thereby removing the requirement to invert $\underline{P}(i)$ in the computations. (Neglecting these effects of the terms due to $\ell n\ f(\underline{x}(i)|\underline{Z}(i), \underline{a})$ is further substantiated by the fact that their removal does not affect the asymptotic behavior of the estimator, as shown in section 4.1.7.) Under these circumstances, there is no longer any need for propagating equations (3.4.18), (3.4.19), (3.4.20), and (3.4.33). In other words, an n-by-n matrix $\partial \underline{M}(j)/\partial a_\ell$ need not be propagated for each uncertain parameter, the vector propagation of $\partial \underline{\bar{x}}(j)/\partial a_\ell$ sufficing for estimation purposes. Since this is a substantial benefit to the computational load, the adequacy of such methods warrants investigation.

Table 5.4 depicts the required computations for the hypothetical problem described earlier, using the second of the two on-line conceptualizations in section 5.2. These results are directly comparable to Table 5.2, based on using all likelihood equation terms. If no parameter estimate is made, only the last row of the table is affected. As before, the first on-line conceptualization would differ only in requiring three fewer subtractions for $\underline{s}$ and six fewer for $\underline{J}$. The reductions afforded by using only the weighted least squares type terms is very apparent when the two tables are compared.

Table 5.4

On-Line Method Using Weighted Least Squares Terms Only

Term	Multiplications	Additions	Subtractions	Inversions
$\underline{x}$	889	671	27	1 (2x2)
$\underline{s}$	275	189	3	0
$\underline{J}$	36	24	6	0
$\underline{a}^*$	9	9	0	1 (3x3)

The implementation of the propagations corresponding to equations (3.4.14) through (3.4.39) would be as follows. To propagate between sample times, the state equations are the same as equations (3.4.14) to (3.4.17). The score equations become

$$\frac{\partial \bar{\underline{x}}(i)}{\partial a_\ell} = \underline{\Phi}(i,i-1)\,\frac{\partial \hat{\underline{x}}(i-1)}{\partial a_\ell} + \frac{\partial \underline{\Phi}(i,i-1)}{\partial a_\ell}\,\hat{\underline{x}}(i-1) + \frac{\partial \underline{B}(i-1)}{\partial a_\ell}\,\underline{u}(i-1) \tag{5.3.8}$$

$$s^1{}_\ell[\underline{Z}(i),\underline{a}_*(i)] = \frac{\partial \bar{\underline{x}}^T(i)}{\partial a_\ell}\,\underline{H}^T(i)\underline{\eta}(i) \tag{5.3.9}$$

and the conditional information matrix relations are

$$J^1_{k\ell}[\underline{Z}(i),\underline{a}_*(i)] = \frac{\partial \bar{\underline{x}}^T(i)}{\partial a_k}\,\underline{H}^T(i)\underline{A}^{-1}(i)\underline{H}(i)\,\frac{\partial \bar{\underline{x}}(i)}{\partial a_\ell} \tag{5.3.10}$$

At the time of a measurement, the state relationships remain unchanged from (3.4.30) to (3.4.32). However, the score equations become

$$\frac{\partial \hat{\underline{x}}(i)}{\partial a_\ell} = [\underline{I}-\underline{K}(i)\underline{H}(i)]\,\frac{\partial \bar{\underline{x}}(i)}{\partial a_\ell} \tag{5.3.11}$$

The contribution to the conditional information matrix corresponding to equation (3.4.39) would become $[(\partial \hat{\underline{x}}(i)/\partial a_k)^T \underline{P}^{-1}(i) \cdot (\partial \hat{\underline{x}}(i)/\partial a_\ell)]$, but since the related score contribution goes to zero, this term need not be added to $J_{k\ell}[\underline{Z}(i),\underline{a}_*(i)]$, as previously mentioned.

5.3.2 Towards a Desirable Form

If it is desired to make an estimate of the system parameters based upon the most recent N measurement residuals, then a very convenient form for the estimation algorithm would be

$$\hat{\underline{a}}(i) = \hat{\underline{a}}(i-1) + \underline{\mathcal{Y}}(i)\,[\underline{z}(i)-\underline{H}(i)\overline{\underline{x}}(i)] - \underline{\mathcal{Y}}(i-N)\,[\underline{z}(i-N)-\underline{H}(i-N)\overline{\underline{x}}(i-N)] \tag{5.3.12}$$

where $\underline{\mathcal{Y}}(i)$ and $\underline{\mathcal{Y}}(i-N)$ are appropriate weighting matrices. It will now be shown that the estimator of the previous section can be expressed in a similar form, but that the approximation needed to attain the precise form of (5.3.12) significantly alters the structure of the estimator.

Assume that a parameter estimate is to be made every sample period by the second on-line conceptualization of section 5.2, using only weighted least squares terms. Then two successive parameter estimates would be calculated as

$$\underline{a}^*(i) = \underline{a}^*(i-1) + \underline{J}[\underline{Z}(i),\underline{a}^*(i-1)]^{-1}\,\underline{s}[\underline{Z}(i),\underline{a}^*(i-1)] \tag{5.3.13}$$

$$\underline{a}^*(i-1) = \underline{a}^*(i-2) + \underline{J}[\underline{Z}(i-1),\underline{a}^*(i-2)]^{-1}\,\underline{s}[\underline{Z}(i-1),\underline{a}^*(i-2)] \tag{5.3.14}$$

where the N-step score and conditional information matrix are defined, using the notation $\partial\overline{\underline{x}}/\partial\underline{a} = [\partial\overline{\underline{x}}/\partial a_1 \;\ldots\; \partial\overline{\underline{x}}/\partial a_p]$, as:

$$\begin{aligned}\underline{s}[\underline{Z}(i),\underline{a}^*(i-1)] &= \sum_{j=i-N+1}^{i} \underline{s}^1[\underline{Z}(j),\underline{a}^*(j-1)] \\ &= \sum_{j=i-N+1}^{i} \left(\frac{\partial\overline{\underline{x}}(j)}{\partial\underline{a}}\right)^T \underline{H}^T(j)\underline{A}^{-1}(j)[\underline{z}(j)-\underline{H}(j)\overline{\underline{x}}(j)]\end{aligned} \tag{5.3.15}$$

$$\underline{J}[\underline{Z}(i),\underline{a}^*(i-1)] = \sum_{j=i-N+1}^{i} \underline{J}^1[\underline{Z}(j),\underline{a}^*(j-1)]$$

$$= \sum_{j=i-N+1}^{i} \left(\frac{\partial \underline{\bar{x}}(j)}{\partial \underline{a}}\right)^T \underline{H}^T(j)\underline{A}^{-1}(j)\underline{H}(j) \frac{\partial \underline{\bar{x}}(j)}{\partial \underline{a}} \tag{5.3.16}$$

Now assume that the value of $\sum_{j=i-N+1}^{i-1} \underline{J}^1[\underline{Z}(j),\underline{a}^*(j-1)]$ dominates the difference $\{\underline{J}^1[\underline{Z}(i),\underline{a}^*(i-1)] - \underline{J}^1[\underline{Z}(i-N),\underline{a}^*(i-N-1)]\}$, so that

$$\underline{J}[\underline{Z}(i),\underline{a}^*(i-1)]^{-1} =$$

$$= \left[\underline{J}^1[\underline{Z}(i),\underline{a}^*(i-1)] + \sum_{j=i-N+1}^{i-1} \underline{J}^1[\underline{Z}(j),\underline{a}^*(j-1)]\right]^{-1}$$

$$\cong \left[\underline{J}^1[\underline{Z}(i-N),\underline{a}^*(i-N-1)] + \sum_{j=i-N+1}^{i-1} \underline{J}^1[\underline{Z}(j),\underline{a}^*(j-1)]\right]^{-1}$$

$$= \underline{J}[\underline{Z}(i-1),\underline{a}^*(i-2)]^{-1} \tag{5.3.17}$$

Using this approximation, the difference between equations (5.3.13) and (5.3.14) becomes

$$\underline{a}^*(i) - \underline{a}^*(i-1) \cong \underline{a}^*(i-1) - \underline{a}^*(i-2) + \underline{J}[\underline{Z}(i),\underline{a}^*(i-1)]^{-1} \cdot$$

$$\cdot \left[\underline{s}[\underline{Z}(i),\underline{a}^*(i-1)] - \underline{s}[\underline{Z}(i-1),\underline{a}^*(i-2)]\right]$$

$$= \underline{a}^*(i-1) - \underline{a}^*(i-2) + \underline{J}[\underline{Z}(i),\underline{a}^*(i-1)]^{-1} \cdot$$

$$\cdot \left[\underline{s}^1[\underline{Z}(i),\underline{a}^*(i-1)] - \underline{s}^1[\underline{Z}(i-N),\underline{a}^*(i-N-1)]\right] \tag{5.3.18}$$

In other words, the on-line parameter estimate of the previous section can be expressed approximately as:

$$\underline{a}^*(i) = \underline{a}^*(i-1) + [\underline{a}^*(i-1) - \underline{a}^*(i-2)]$$
$$+ \underline{J}[\underline{Z}(i),\underline{a}^*(i-1)]^{-1} \left(\frac{\partial \underline{\bar{x}}(i)}{\partial \underline{a}}\right)^T \underline{H}^T(i)\underline{A}^{-1}(i)\underline{r}(i)$$
$$- \underline{J}[\underline{Z}(i),\underline{a}^*(i-1)]^{-1} \left(\frac{\partial \underline{\bar{x}}(i-N)}{\partial \underline{a}}\right)^T \underline{H}^T(i-N)\underline{A}^{-1}(i-N)\underline{r}(i-N)$$

(5.3.19)

where $\underline{r}(i)$ and $\underline{r}(i-N)$ are the residuals at times i and (i-N) respectively. This would be exactly the desired form if it were not for the term $[\underline{a}^*(i-1) - \underline{a}^*(i-2)]$. Although this is approximately equal to zero, neglecting it changes the estimator dynamics in the following manner. Assume that a single parameter is being estimated, so that (5.3.19) becomes

$$a^*(i) - 2\,a^*(i-1) + a^*(i-2) = \frac{1}{J}[s^1(i) - s^1(i-N)]$$

(5.3.20)

so that the Z-transform transfer function from the single measurement score to the parameter estimate is

$$\frac{a^*}{s^1} = \frac{1}{J}\,\frac{1-z^{-N}}{1-2z^{-1}+z^{-2}}$$
$$= \frac{1}{J}\,\frac{z^N-1}{(z-1)^2\,z^{(N-2)}} \qquad (5.3.21)$$

In the z plane, this is comprised of (N-2) poles at the origin, a single pole at the point (+1) on the real axis, the other pole there being cancelled by a zero, and the rest of the zeroes spaced at intervals of $(2\pi)/N$ radians around the unit circle (starting from the cancelled zero at +1). Neglecting the term $[\underline{a}^*(i-1) - \underline{a}^*(i-2)]$ would remove the remaining pole at +1 on the real axis, thereby changing the dynamics of the single measurement score incorporation significantly. Investigations have shown that this pole

removal can seriously deteriorate the estimator performance, whereas equation (5.3.19) yields suitable performance.

By considering the vector difference $[\underline{a}^*(j) - \underline{a}^*(j-1)]$ as a variable $\Delta\underline{a}^*(j)$, equation (5.3.19) can be put into the form

$$\Delta\underline{a}^*(i) = \Delta\underline{a}^*(i-1) + \underline{\mathcal{G}}(i)\underline{r}(i) - \underline{\mathcal{G}}(i-N)\underline{r}(i-N) \qquad (5.3.22)$$

$$\underline{a}^*(i) = \underline{a}^*(i-1) + \Delta\underline{a}^*(i) \qquad (5.3.23)$$

which is a convenient algorithm that provides the same performance as equation (5.3.19).

5.4 Feasible Precomputations

In order to improve the efficiency of the estimator algorithms, it would be advantageous to precompute various terms and store them for on-line use, rather than calculate all variables in real time. This task is more complex for the combined state and parameter estimator than for the linear state estimator, in which the covariances and gains can be precomputed exactly. Here, the precomputations must be determined parametrically in terms of the components of $\underline{a}$. Two modes of implementation are then possible: (1) use the most recent estimate of $\underline{a}$ to evaluate the required values functionally, or (2) through simulations determine a nominal value of $\underline{a}$ that provides adequate performance over the range of possible parameter values, and use precomputed values based on this nominal $\underline{a}$.

First of all, the N-step conditional information matrix, when evaluated as in section 3.4, typically reaches a steady state value very quickly, and need only be recomputed infrequently, if at all, after it has converged to a certain value. In fact, there is considerable advantage to using such a precomputed value before such time as an on-line computed

$\underline{J}^{-1}$ would converge. Before N measurements are taken, $\underline{J}[\underline{Z}(i),\underline{a}_*(i)]$ would be of lower magnitude than the steady state value, especially for the first sample instants. Therefore, $\underline{J}^{-1}[\underline{Z}(i),\underline{a}_*(i)]$ would be extremely large, adversely affecting the convergence of the parameter estimate to the true parameter values. By using a "steady state" value of $\underline{J}^{-1}[\,i\,,\underline{a}_*(i)\,]$, the estimator would in fact be employing a form of gradient iteration, instead of scoring, to obtain $\underline{a}^*(i)$. Consequently, convergence from a bad initial parameter estimate would be more assured in the first stages than if $\underline{J}^{-1}[\underline{Z}(i),\underline{a}_*(i)]$ were calculated on-line.

Matrix propagations that can be precomputed parametricaly would be the $\underline{M}(i)$, $\underline{P}(i)$ sequence and related $\underline{A}(i)$, $\underline{A}^{-1}(i)$, and $\underline{K}(i)$; the $\partial\underline{M}(i)/\partial a_\ell$, $\partial\underline{P}(i)/\partial a_\ell$ sequences and associated $\partial\underline{A}(i)/\partial a_\ell$; and finally the controller gains $\underline{C}(i)$ if feedback control is used. These sequences could be evaluated over the interval of interest using a number of fixed values of the parameters. Simple, but adequately accurate, approximating functions would be curve-fitted to express these sequences as a function of $\underline{a}$.

Precomputed estimator and controller gains raise the question of validity. Typical state estimator gain time histories for two values of $\underline{a}$ appear in figure 5.2. Especially for time invariant systems, a rapid (with respect to total time of interest) transient is followed by a steady-state (or slowly varying) value. Assume that the estimate of a at 0 and t_1 is a_1, but at t_2 a new estimate is determined as a_2. If all gain calculations were performed on-line (using only forward processing of data, without backward smoothing or recomputation of gains from the initial time), the gain variation would be as in the second diagram in figure 5.2. If precomputed gains were used, the gain would be evaluated as though its entire history were recomputed from the initial time using a_2, as seen in the third plot of figure 5.2. The forward processing would be appropriate

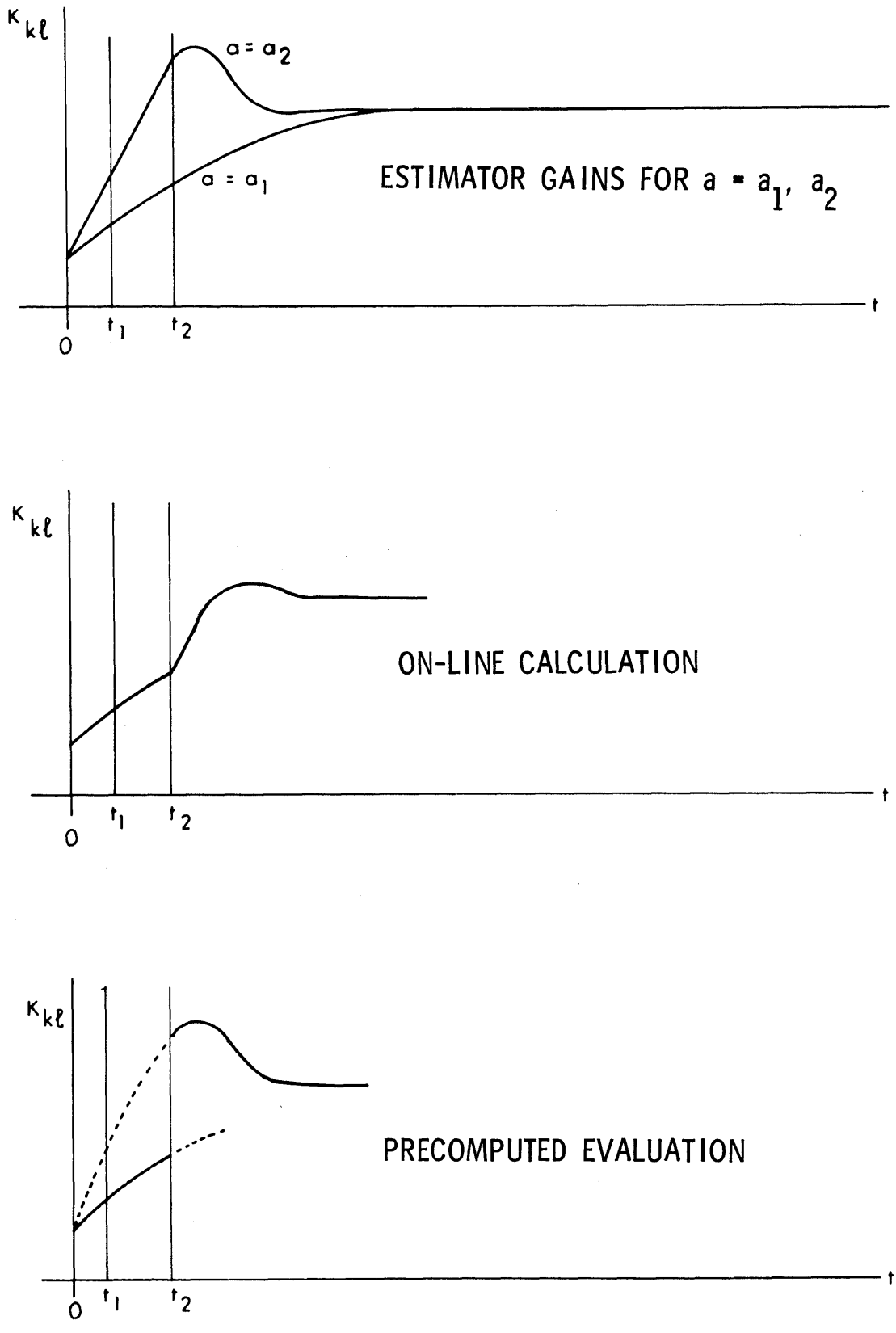


Figure 5.2 Precomputed Estimator Gains

if the difference between $a^*(t_1)$ and $a^*(t_2)$ were caused by variation in the actual parameter values, but if this difference were caused by $a^*(t_2)$ being a better estimate than $a^*(t_1)$, the optimal procedure would be to recompute the gains (and state estimate) from the initial time using a_2. However, such total recomputation is not feasible for on-line applications. Since the actual parameters are assumed to be slowly varying, and since the parameter estimator is in the process of converging to a good estimate near the initial time, precomputed gains may well yield as accurate a state estimate as provided by the forward processing approach.

Controller gains would be computed backward from the terminal time as a function of $\underline{a}$. Evaluating these functions with $\underline{a}^*(i)$ would be the philosophy of evaluating the optimal control as though the parameters would stay constant until the terminal time, and using such control until a future parameter estimate contradicts the assumption. This conceptual approach is often termed Open Loop Feedback Optimal (O.L.F.O.) Control.

Steady-state estimator and controller gains as a function of $\underline{a}$ suffice in many applications, as opposed to implementing the total gain time histories. These require particularly simple functional evaluations, and should be exploited if they yield adequate performance.

When the real system is continuous-time in nature, the functional dependence of the $\underline{\Phi}$ and $\underline{B}$ entries upon the value of uncertain parameters in the continuous model would also be precomputed. Curve-fitting of linear or piecewise linear functions to these relations would be most suitable for subsequent on-line usage, as discussed in the introduction to Chapter III.

Table 5.5 presents the calculations required for the hypothetical problem, using the on-line implementation of section 5.2, but with precomputed $\underline{J}^{-1}$ and matrix propagations as described in this section. These results are

Table 5.5

On-Line Method; Precomputed Matrices

Term	Multiplications	Additions	Subtractions	Inversions
$\underline{x}$	69	56	2	0
$\underline{s}$	309	229	9	0
$\underline{J}$	0	0	0	0
$\underline{a}^*$	9	9	0	0

Table 5.6

On-Line Method with WLS Terms Only; Precomputed Matrices

Term	Multiplications	Additions	Subtractions	Inversions
$\underline{x}$	69	56	2	0
$\underline{s}$	275	189	9	0
$\underline{J}$	0	0	0	0
$\underline{a}^*$	9	9	0	0

directly comparable to Table 5.2. Similarly, Table 5.6 presents the corresponding results for the case of including only weighted least squares terms; this table can be compared to Table 5.4 in section 5.3.1 to reveal the marked improvement afforded by precomputation.

5.5 Advantageous State Space Representations

A single system input-output relation can be represented by an infinite variety of state space representations, the different forms being related by similarity transformations. That is to say, if a system can be modelled as

$$\underline{x}(i+1) = \underline{\Phi}(i+1,i)\underline{x}(i) + \underline{B}(i)\underline{u}(i) + \underline{G}(i)\underline{w}(i) \qquad (5.5.1)$$

$$\underline{z}(i) = \underline{H}(i)\underline{x}(i) + \underline{v}(i) \qquad (5.5.2)$$

then it is possible to define a new state vector through use of an invertible transformation matrix $\underline{T}(i)$ as

$$\underline{x}(i) = \underline{T}(i)\underline{x}'(i) \qquad (5.5.3)$$

$$\underline{x}'(i) = \underline{T}^{-1}(i)\underline{x}(i) \qquad (5.5.4)$$

and obtain a model with an input-output relation identical to that of (5.5.1) and (5.5.2) in the form of

$$\underline{x}'(i+1) = \underline{\Phi}'(i+1,i)\underline{x}'(i) + \underline{B}'(i)\underline{u}(i) + \underline{G}'(i)\underline{w}(i) \qquad (5.5.5)$$

$$\underline{z}(i) = \underline{H}'(i)\underline{x}'(i) + \underline{v}(i) \qquad (5.5.6)$$

where the system matrices are defined as

$$\underline{\Phi}'(i+1,i) = \underline{T}^{-1}(i+1)\underline{\Phi}(i+1,i)\underline{T}(i) \tag{5.5.7}$$

$$\underline{B}'(i) = \underline{T}^{-1}(i+1)\underline{B}(i) \tag{5.5.8}$$

$$\underline{G}'(i) = \underline{T}^{-1}(i+1)\underline{G}(i) \tag{5.5.9}$$

$$\underline{H}'(i) = \underline{H}(i)\underline{T}(i) \tag{5.5.10}$$

Although such time varying system models can be developed, the parameters are assumed to be slowly varying, so the system will be modelled as time-invariant over N sample instants. Moreover, since on-line implementations would essentially require precomputation of the system matrices as a function of the parameter values (practical only for constant matrices), the time invariant transformations will be emphasized.

Since it is desirable to minimize the computational and storage requirements on the computer, certain state space descriptions may be preferable to that originally chosen to represent the system dynamics. If explicit estimates of the original state variables are required in a particular application, equation (5.5.3) can be used to transform back to these variables.

Of the possible state space representations, two are particularly convenient, the standard observable phase variable form and the modified Jordan canonical form. The first of these reduces the amount of mathematical computations because of the sparse form of its state transition matrix, and the second separates the different system modes. Sections 5.5.1 and 5.5.2 present these two forms for the case of scalar measurements, discussing the merits and disadvantages of each. Section 5.5.3 then discusses the case of vector measurements, which differs from the previous results in that uncertainties must generally be allowed to exist in the $\underline{H}'(i)$ matrix for non-scalar measurements.

5.5.1 Standard Observable Phase Variables; Scalar Measurements

In the case of scalar measurements, $\underline{H}(i)$ becomes a 1 by n matrix, or a vector transpose, so that equation (5.5.2) can be written as

$$z(i) = \underline{h}^T(i)\underline{x}(i) + v(i) \tag{5.5.11}$$

If the system is assumed to be observable over each successive set of n measurements (in the time-invariant case, this reduces to the assumption of system observability), then an invertible transformation matrix can be defined, using (5.5.1) and (5.5.11), as

$$\underline{T}^{-1}(i) = \begin{bmatrix} \underline{h}^T(i) \\ \underline{h}^T(i+1)\underline{\Phi}(i+1,i) \\ \vdots \\ \underline{h}^T(i+n-1)\underline{\Phi}(i+n-1,i+n-2) \cdots \underline{\Phi}(i+1,i) \end{bmatrix} \tag{5.5.12}$$

that will generate a system state transition matrix as

$$\underline{\Phi}'(i+1,i) = \begin{bmatrix} \underline{0} & \vdots & \underline{I} \\ \cdots & \vdots & \cdots\cdots \\ \underline{\phi}'^T(i+1,i) & & \end{bmatrix} \tag{5.5.13}$$

where $\underline{0}$ is an (n-1) by 1 zero vector, $\underline{I}$ is an (n-1) by (n-1) identity matrix, and $\underline{\phi}'^T(i+1,i)$ is defined by

$$\underline{\phi}'^T(i+1,i) = \underline{h}^T(i+n)\underline{\Phi}(i+n,i+n-1) \cdots \underline{\Phi}(i+1,i)\underline{T}(i) \tag{5.5.14}$$

and the measurement "matrix," $\underline{h}'^T(i)$, becomes

$$\underline{h}'^T(i) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \tag{5.5.15}$$

From this expression it can be seen that there are no uncertain parameters in $\underline{h}'^{T}(i)$.

The validity of these equations can be proven by a simple extension of similar proofs by Lee (1964) for time invariant systems. Widnall (1968) provides a somewhat different proof, along with a recursion for generating the transformation matrix.

For the time invariant case, equations (5.5.12) and (5.5.14) reduce to

$$\underline{T}^{-1} = \begin{bmatrix} \underline{h}^{T} \\ \underline{h}^{T}\underline{\Phi} \\ \vdots \\ \underline{h}^{T}\underline{\Phi}^{n-1} \end{bmatrix} \tag{5.5.16}$$

$$\underline{\phi}'^{T} = \underline{h}^{T}\underline{\Phi}^{n}\underline{T} \tag{5.5.17}$$

This state representation does substantially reduce the amount of mathematical computations because of the simple form of $\underline{\Phi}'(i+1,i)$. However, additional advantages are gained from the modified Jordan canonical form, described in the next section.

5.5.2 Modified Jordan Canonical Form; Scalar Measurements

It can be shown (Brockett (1969)) that if $\underline{\Phi}$ is a real (not complex) matrix, then there exists a real nonsingular transformation matrix $\underline{T}_1$ such that

$$\underline{\Phi}' = \underline{T}_1^{-1}\underline{\Phi}\,\underline{T}_1 \tag{5.5.18}$$

is of modified Jordan canonical form, expressed as the block diagonal matrix:

$$\underline{\Phi}' = \begin{bmatrix} \underline{\Phi}_1' & & & & 0 \\ & \ddots & & & \\ & & \underline{\Phi}_\alpha' & & \\ & & & \underline{\Phi}_{\alpha+1}' & \\ & & & & \ddots \\ 0 & & & & \underline{\Phi}_\beta' \end{bmatrix} \tag{5.5.19}$$

where each $\underline{\Phi}_k'$, for $1 \leq k \leq \alpha$, represents a second order mode of multiplicity μ_k, with eigenvalues $(\sigma_k \pm j\omega_k)$, expressed as the $2\mu_k$-by-$2\mu_k$ matrix

$$\underline{\Phi}_k' = \begin{bmatrix} \underline{S}_k & \underline{I} & & & 0 \\ & \underline{S}_k & \underline{I} & & \\ & & \ddots & \ddots & \\ & & & \underline{S}_k & \underline{I} \\ 0 & & & & \underline{S}_k \end{bmatrix} \quad \text{for } 1 \leq k \leq \alpha \tag{5.5.20}$$

where

$$\underline{S}_k = \begin{bmatrix} \sigma_k & \omega_k \\ -\omega_k & \sigma_k \end{bmatrix} \tag{5.5.21}$$

and the $\underline{\Phi}_k'$ for $(\alpha+1) \leq k \leq \beta$ represent a first order mode of multiplicity μ_k, with eigenvalue λ_k, expressed as the μ_k-by-μ_k matrix

$$\underline{\Phi}_k' = \begin{bmatrix} \lambda_k & 1 & & 0 \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{bmatrix} \quad \text{for } (\alpha+1) \leq k \leq \beta \tag{5.5.22}$$

Each $\underline{\Phi}_k'$ along the diagonal in equation (5.5.19) corresponds to a single real eigenvalue or a single complex conjugate pair. The most general Jordan canonical form of a real square matrix allows more than one $\underline{\Phi}_k'$ to correspond to a single eigenvalue or complex pair, so that, as an example, for a first order repeated root of multiplicity 3, there would be an ambiguity of proper Jordan form among

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad \left[\begin{array}{cc:c} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ \hdashline 0 & 0 & \lambda \end{array}\right], \quad \left[\begin{array}{c:c:c} \lambda & 0 & 0 \\ \hdashline 0 & \lambda & 0 \\ \hdashline 0 & 0 & \lambda \end{array}\right]$$

depending on whether one, two, or three independent eigenvectors (respectively) could be determined. However, the second and third of these are invalid for state space models since there is no $\underline{h}'^T$ that could generate an observable model with these $\underline{\Phi}'$'s. In general, it can be shown that to produce an observable system model, each $\underline{\Phi}_k'$ must correspond to a distinct real eigenvalue or complex pair. For unrepeated roots, these $\underline{\Phi}_k'$'s reduce to

$$\underline{\Phi}_k' = \underline{S}_k \qquad \text{for} \quad 1 \leq k \leq \alpha \tag{5.5.20'}$$

$$\underline{\Phi}_k' = \lambda_k \qquad \text{for} \quad (\alpha+1) \leq k \leq \beta \tag{5.5.22'}$$

The transformation matrix is not unique. It is known that a $\underline{T}_1$ as in equation (5.5.18) does exist, and for any $\underline{T}_2$ such that

$$\underline{T}_2^{-1}\underline{\Phi}'\underline{T}_2 = \underline{\Phi}' \tag{5.5.23}$$

the composite transformation $(\underline{T}_1\underline{T}_2)$ also transforms $\underline{\Phi}$ into $\underline{\Phi}'$, since

$$\underline{T}_2^{-1}\underline{T}_1^{-1}\underline{\Phi}\,\underline{T}_1\underline{T}_2 = \underline{T}_2^{-1}\underline{\Phi}'\underline{T}_2 = \underline{\Phi}' \tag{5.5.24}$$

The general form of the transformation $\underline{T}_2$ that satisfies equation (5.5.23), with $\underline{\Phi}'$ defined by (5.5.19), can be shown to be

$$\underline{T}_2 = \begin{bmatrix} \underline{T}_{21} & & & 0 \\ & \ddots & & \\ & & \underline{T}_{2\alpha} & \\ & & & \underline{T}_{2(\alpha+1)} \\ 0 & & & & \ddots \\ & & & & & \underline{T}_{2\beta} \end{bmatrix} \tag{5.5.25}$$

where, for $1 \leq k \leq \alpha$,

$$\underline{T}_{2k} = \begin{bmatrix} \underline{T}_{2k1} & \underline{T}_{2k2} & \cdots\cdots & \underline{T}_{2k\mu_k} \\ & \underline{T}_{2k1} & \underline{T}_{2k2} & \vdots \\ & & \ddots & \underline{T}_{2k2} \\ 0 & & & \underline{T}_{2k1} \end{bmatrix} \tag{5.5.26}$$

where μ_k is the multiplicity of the k-th second order root, and

$$\underline{T}_{2k\ell} = \begin{bmatrix} t_{2k\ell} & s_{2k\ell} \\ -s_{2k\ell} & t_{2k\ell} \end{bmatrix} \tag{5.5.27}$$

and, for $(\alpha+1) \leq k \leq \beta$,

$$\underline{T}_{2k} = \begin{bmatrix} t_{2k1} & t_{2k2} & \cdots\cdots & t_{2k\mu_k} \\ & t_{2k1} & t_{2k2} & \vdots \\ & & \ddots & t_{2k2} \\ 0 & & & t_{2k1} \end{bmatrix} \tag{5.5.28}$$

Investigation of these equations reveals that there are n independent variables in the $\underline{T}_2$ matrix. By specifying these n variables, a unique transformation from $\underline{\Phi}$ to $\underline{\Phi}'$ can be achieved. Since it is desired to make the measurement matrix devoid of uncertain parameters if possible (because of the simplified estimator relations), the n entries of $\underline{h}'^T$ can be specified in the case of scalar measurements, and thereby produce the unique transformation. A convenient choice of $\underline{h}'^T$ that can be used to represent any observable system is

$$\underline{h}'^T = \left[\underline{h}_1'^T \cdots \underline{h}_\alpha'^T \; \underline{h}_{\alpha+1}'^T \cdots \underline{h}_\beta'^T\right] \tag{5.5.29}$$

$$\underline{h}_k'^T = \left[1 \quad 1 \quad 0 \cdots 0\right] \quad \text{for } 1 \leq k \leq \alpha \tag{5.5.30}$$

$$\underline{h}_k'^T = \left[1 \quad 0 \cdots 0\right] \quad \text{for } (\alpha+1) \leq k \leq \beta \tag{5.5.31}$$

The procedure would be to take the original $\underline{\Phi}$ and $\underline{h}^T$, calculate the eigenvalues of $\underline{\Phi}$, and with these eigenvalues write down the $\underline{\Phi}'$ defined by (5.5.19) to (5.5.22) and the $\underline{h}'^T$ defined by (5.5.29) to (5.5.31). Then the unique transformation $\underline{T}$ is the matrix that satisfies the equations

$$\underline{T}\,\underline{\Phi}' = \underline{\Phi}\,\underline{T} \tag{5.5.32}$$

$$\underline{h}'^T = \underline{h}^T\underline{T} \tag{5.5.33}$$

These two equations generate a set of n^2+n linear equations in the elements of $\underline{T}$ which are solved to determine $\underline{T}$. If the system is observable, then only n^2 of these equations are linearly independent, so their solution does in fact yield a unique transformation matrix $\underline{T}$. Once $\underline{T}$ is determined, $\underline{B}'$ is calculated as

$$\underline{B}' = \underline{T}^{-1}\underline{B} \tag{5.5.34}$$

and similarly for $\underline{G}'$. This procedure would be repeated for a number of values of the uncertain parameters, so as to determine the (approximate, curve-fitted) functional dependence of the elements of $\underline{\Phi}'$, $\underline{B}'$, and $\underline{G}'$ upon the value of $\underline{a}$. These precomputed relations are then implemented for the on-line application.

Note that this procedure does not require an eigenvector evaluation; rather, a set of linear equations is solved by standard row reduction methods.

The primary advantage of the modified Jordan canonical form is that it separates the system modes. Typically, the effect of a particular system parameter will enter into only one or a few system modes, rather than all of them. (When more than one mode is affected, there is also the possibility of neglecting the less sensitive modes if the time element is critical.) For each particular parameter a_ℓ, the system model matrices can be partitioned into modes affected by a_ℓ, denoted by "a," and those not affected by a_ℓ, denoted as "b." (Here the state variables have been reordered to locate the "a" modes in the upper portions of $\underline{\Phi}'$ and $\underline{B}'$ for simplicity of description; this is not necessary in practice.) Thus,

$$\underline{\Phi}' = \begin{bmatrix} \underline{\Phi}'_a & \vdots & \underline{0} \\ \cdots\cdots & \vdots & \cdots\cdots \\ \underline{0} & \vdots & \underline{\Phi}'_b \end{bmatrix} \tag{5.5.35}$$

$$\underline{B}' = \begin{bmatrix} \underline{B}'_a \\ \cdots \\ \underline{B}'_b \end{bmatrix} \tag{5.5.36}$$

where $\underline{\Phi}'_a$ and $\underline{\Phi}'_b$ are themselves block diagonal matrices, and the dimension of $\underline{\Phi}'_a$ is typically much smaller than that of $\underline{\Phi}'$. As a result of such partitioning, the matrices $\partial\underline{\Phi}'/\partial a_\ell$ and $\partial\underline{B}'/\partial a_\ell$ can be written as:

$$\frac{\partial \underline{\Phi}'}{\partial a_\ell} = \begin{bmatrix} \partial \underline{\Phi}_a' / \partial a_\ell & \vdots & \underline{0} \\ \cdots\cdots & \vdots & \cdots\cdots \\ \underline{0} & \vdots & \underline{0} \end{bmatrix} \tag{5.5.37}$$

$$\frac{\partial \underline{B}'}{\partial a_\ell} = \begin{bmatrix} \partial \underline{B}_a' / \partial a_\ell \\ \cdots\cdots \\ \underline{0} \end{bmatrix} \tag{5.5.38}$$

The nonzero partition of these matrices is usually very small; if a_ℓ affects a single first order mode, then $\partial \underline{\Phi}'/\partial a_\ell$ has one nonzero entry, and $\partial \underline{B}'/\partial a_\ell$ has at most r (the dimension of the $\underline{u}$ vector).

Thus, the modified Jordan canonical form affords reduced computational and memory requirements in a number of ways. First, because of the block diagonal $\underline{\Phi}'$, it generally requires fewer state estimate mathematical computations than the equivalent system expressed in physical variables, and about the same as that expressed in phase variables. Secondly, in practical applications, the entries of $\underline{\Phi}$, $\underline{B}$, $\partial \underline{\Phi}/\partial a_\ell$, and $\partial \underline{B}/\partial a_\ell$ are evaluated through curve-fitted functions and the most recent parameter estimate, and this form yields the least number of entries that require such evaluation. Finally, the propagations required for the parameter estimate are simplified considerably because of the sparse form of $\partial \underline{\Phi}'/\partial a_\ell$ and $\partial \underline{B}'/\partial a_\ell$:

$$\frac{\partial \underline{\bar{x}}'(i+1)}{\partial a_\ell} = \begin{bmatrix} \partial \underline{\Phi}_a'/\partial a_\ell \; \underline{\hat{x}}_a'(i) + \partial \underline{B}_a'/\partial a_\ell \; \underline{u}(i) + \underline{\Phi}_a' \; \partial \underline{\hat{x}}_a'(i)/\partial a_\ell \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \underline{\Phi}_b' \; \partial \underline{\hat{x}}_b'(i)/\partial a_\ell \end{bmatrix} \tag{5.5.39}$$

$$\frac{\partial \underline{M}'(i+1)}{\partial a_\ell} = \begin{bmatrix} \partial \underline{M}_{aa}'(i+1)/\partial a_\ell & \vdots & \partial \underline{M}_{ab}'(i+1)/\partial a_\ell \\ \cdots\cdots\cdots\cdots\cdots\cdots & \vdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ \partial \underline{M}_{ab}'^{T}(i+1)/\partial a_\ell & \vdots & \partial \underline{M}_{bb}'(i+1)/\partial a_\ell \end{bmatrix} \tag{5.5.40}$$

where

$$\frac{\partial \underline{M}'_{aa}(i+1)}{\partial a_\ell} = \frac{\partial \underline{\Phi}'_a}{\partial a_\ell} \underline{P}'_{aa}(i)\underline{\Phi}'^{T}_a + \underline{\Phi}'_a \underline{P}'_{aa}(i) \frac{\partial \underline{\Phi}'^{T}_a}{\partial a_\ell} + \underline{\Phi}'_a \frac{\partial \underline{P}'_{aa}(i)}{\partial a_\ell} \underline{\Phi}'^{T}_a$$

$$\frac{\partial \underline{M}'_{ab}(i+1)}{\partial a_\ell} = \frac{\partial \underline{\Phi}'_a}{\partial a_\ell} \underline{P}'_{ab}(i)\underline{\Phi}'^{T}_b + \underline{\Phi}'_a \frac{\partial \underline{P}'_{ab}(i)}{\partial a_\ell} \underline{\Phi}'^{T}_b$$

$$\frac{\partial \underline{M}'_{bb}(i+1)}{\partial a_\ell} = \underline{\Phi}'_b \frac{\partial \underline{P}'_{bb}(i)}{\partial a_\ell} \underline{\Phi}'^{T}_b \qquad (5.5.41)$$

Moreover, the modified Jordan form allows a shorter computer wordlength to achieve the same precision in system representation as that attained by other state variable forms. For example, if the Z-transform transfer function of a system were

$$G(z) = \frac{1}{(z+.4)\ (z+.7)} = \frac{1}{z^2+2.2\ z+2.8} \qquad (5.5.42)$$

then the Jordan form of $\underline{\Phi}'$ would be

$$\underline{\Phi}' = \begin{bmatrix} -.4 & 0 \\ 0 & -.7 \end{bmatrix} \qquad (5.5.43)$$

whereas the phase variable form $\underline{\Phi}'$ would be

$$\underline{\Phi}' = \begin{bmatrix} 0 & 1 \\ -.28 & -2.2 \end{bmatrix} \qquad (5.5.44)$$

For this case, the phase variable form needs twice the significant figures as the Jordan form to provide an equally accurate model. Since the wordlength is a critical

factor to minimize for on-line use, the Jordan form is preferable.

Even in applications that require estimates in physical variable coordinates, the (approximate) value of $\underline{T}(i)$ can be maintained, and equation (5.5.3) used to provide the desired estimate. The advantages of the Jordan form will often warrant such a procedure, rather than performing all calculations with the physical variables.

5.5.3 Vector Measurements

If vector measurements are to be incorporated, the $\underline{H}'(i)$ of either phase variable or modified Jordan canonical forms can no longer be made void of uncertainties in general. There are two means of treating this case, either to allow uncertainties in $\underline{H}'(i)$ or to perform an equivalent update using m scalar measurements instead of a single m-dimensional vector update.

Appendix A presents the estimator formulation that allows uncertainties in the measurement matrix. To use this in conjunction with either phase or canonical variables, a row of the measurement matrix, $\underline{H}(i)$, with respect to which the system is completely observable, is denoted as $\underline{h}^T(i)$, and the procedure of either of the two previous sections is followed to obtain the corresponding $\underline{h}'^T(i)$, $\underline{\Phi}'(i+1,i)$, $\underline{B}'(i)$, and $\underline{G}'(i)$ as a function of $\underline{a}$. The remaining portion of the measurement matrix is obtained from

$$\underline{H}'(i) = \underline{H}(i)\underline{T}(i) \tag{5.5.45}$$

Note that if $\underline{h}^T(i)$ is the r-th row of $\underline{H}(i)$, then the r-th row of $\underline{H}'(i)$ is $\underline{h}'^T(i)$, so that the r-th row of $\partial\underline{H}'(i)/\partial a_\ell$ will always be $\underline{0}^T$.

If no such row of $\underline{H}(i)$ can be found, the procedure must be modified. In the case of canonical variables, a row $\underline{h}^T(i)$ is chosen and the $\underline{h}_k'^T$ vectors corresponding to states that

are unobservable with respect to $\underline{h}^T$ are set to $\underline{0}^T$. Another row is chosen, and an $\underline{h}_k'^T$ of the form given by equations (5.5.30) or (5.5.31) is inserted into the corresponding row of $\underline{H}'(i)$ for each k for which a mode was previously unobservable, but is observable with respect to the current row. This is continued until every column of $\underline{H}'(i)$ contains one element corresponding to an $\underline{h}_k'^T$ as prescribed by (5.5.30) or (5.5.31).

When there is no single row of $\underline{H}$ with respect to which all states are observable, the specification of the off-diagonal elements of the $\underline{\Phi}'$ matrix for the case of repeated eigenvalues is somewhat more complicated than previously. A simple example would be a system composed of two independent, identical first order lags, each with its own measuring device: the appropriate $\underline{\Phi}'$ would be a diagonal matrix, without an off-diagonal "1" as in equation (5.5.22). For each row of $\underline{H}$, the maximal system partition (or partitions) which is (are) completely observable with respect to that row would be treated as in section 5.5.2; if no such partition accounts for all of the repetitions of a root, then more than one $\underline{\Phi}_k'$ will correspond to that single eigenvalue or complex pair (i.e., more than one independent eigenvector or eigenvector pair can be determined). For time invariant systems, the multiplicity of the roots in the denominator of the transfer functions between each input-measurement pair specifies the placement of off-diagonal ones or identity matrices: if a real eigenvalue λ were repeated three times but the highest order pole corresponding to it were second order, the appropriate $\underline{\Phi}'$ would contain two $\underline{\Phi}_k'$ blocks corresponding to λ. Often the block diagram or original state space representation of a system will readily reveal whether repeated roots are cascaded in a single channel and thus correspond to a single $\underline{\Phi}_k'$ block, or occur in parallel channels and correspond to separate $\underline{\Phi}_k'$ blocks.

It is also possible to employ m scalar updates instead of a single vector update, and this will be described further

in section 5.7. For this technique, the $\underline{H}(i)$ matrix is partitioned into m rows, denoted as $\underline{h}_j^T(i)$ for j=1,2,...m. Again, the procedures of section 5.5.1 or 5.5.2 could generate the corresponding $\underline{h}_j'^T(i)$, $\underline{\Phi}'(i+1,i)$, $\underline{B}_j'(i)$, and $\underline{G}_j'(i)$ as a function of the parameter values; note the absence of subscript j on $\underline{\Phi}'(i+1,i)$, since it is the same for all j. However, this technique is somewhat unattractive for use with phase or canonical variables because it requires m coordinate rotations on-line every sample period, and thus the previous method would usually be more practical for implementation.

Table 5.7 presents the required number of calculations for the hypothetical problem, using canonical variables and the on-line estimator formulation including $\underline{H}'(i)$ uncertainties. For this table, all terms of the likelihood function are employed. Table 5.8 relates the corresponding values when only weighted least squares terms are used. Finally, Table 5.9 adds the precomputations of section 5.4 to reduce the computations still further. For all of these tables, it is assumed that there are three first order modes and one second order; one uncertain parameter affects the second order mode, another a first order mode, and the last is an uncertain parameter in $\underline{B}$ that affects a first order mode. The numbers cited do not include the computation of the required functional evaluations for the system matrices, which would further motivate use of canonical variables. The numbers in parentheses are the portion of the totals due to allowing uncertain parameters in $\underline{H}'$: for scalar measurements, there would be no such contribution.

5.6 Exploiting Symmetry

Various matrices in the estimator equations are known to be symmetric, and the numerical algorithms should ensure this symmetry. Although all of the off-diagonal terms can

Table 5.7

On-Line Method; Canonical Form

Term	Multiplications	Additions	Subtractions	Inversions
x	606	473	27	1 (2x2)
s	1394 (115)	1136 (95)	69 (60)	1 (5x5)
J	887	698 (4)	6	0
a*	9	9	0	1 (3x3)

Table 5.8

On-Line Method with WLS Terms Only; Canonical Form

Term	Multiplications	Additions	Subtractions	Inversions
x	606	473	27	1 (2x2)
s	162 (25)	137 (11)	13 (10)	0
J	36	28 (4)	6	0
a*	9	9	0	1 (3x3)

Table 5.9

On-Line Method; WLS Terms; Precomputed Matrices; Canonical Form

Term	Multiplications	Additions	Subtractions	Inversions
x	46	43	2	0
s	162 (25)	137 (11)	13 (10)	0
J	0	0	0	0
a*	9	9	0	0

be computed and then the i-j-th and j-i-th elements set equal to the average of their computed values, a more expedient method would be to calculate only lower triangular (or upper triangular) forms and then set the corresponding off-diagonal terms equal. For an n-by-n matrix, only $\frac{1}{2}\, n\,(n+1)$ terms require evaluation, instead of n^2.

Symmetry can be exploited further by using a square root covariance formulation rather than the conventional filtering techniques. This algebraically equivalent approach was first suggested by Potter in Battin (1964), and further developed by Potter (1964), Bellantoni and Dodge (1967), Dyer and McReynolds (1969), Speyer and Desai (1970), and Potter and Deckert in the Fourth Quarterly Progress Report for Contract NAS 9-10386 (1971) by the Draper Laboratory. Essentially, it entails the propagation of the triangular form square root error covariance matrix along with the state estimate itself. Survey papers by Schmidt (1970) and Kaminski, Bryson, and Schmidt (1971) demonstrate that the square root approach can double the effective precision of the conventional estimator, especially in ill-conditioned problems, so that shorter wordlengths can be employed in the computer (or, fewer double precision calculations). The numerical conditioning of the square root form is generally better than that of the conventional form, and the product $\sqrt{\underline{P}}\ \sqrt{\underline{P}}^{T} = \underline{P}$ can never be indefinite, even in the face of roundoff errors that often cause indefiniteness of the conventional $\underline{P}$. Even if the covariances are calculated on-line, the computational cost of the later square root implementations is comparable to that of the conventional, greater than that of $\underline{P} = \underline{M} - \underline{K}\,\underline{H}\,\underline{M}$, but less than that of $\underline{P} = (\underline{I} - \underline{K}\,\underline{H})\,\underline{M}\,(\underline{I} - \underline{K}\,\underline{H})^{T} + \underline{K}\,\underline{R}\,\underline{K}^{T}$. Furthermore, Kaminski, Bryson and Schmidt (1971) have shown that even the second of these is still subject to certain ill-conditioned roundoff problems that are avoided by the square root formulation. Thus, the square root

covariance estimator warrants investigation for many proposed applications.

5.7 Modified Measurement Incorporation

As originally derived in section 2.1, the maximum likelihood estimate of the state just after the i-th measurement is incorporated, $\underline{\hat{x}}(i)$, can be expressed in terms of the estimate just before the i-th measurement is incorporated, $\underline{\bar{x}}(i)$, as

$$\underline{\hat{x}}(i) = \left[\underline{M}^{-1}(i)+\underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{H}(i)\right]^{-1} \cdot \cdot \left[\underline{M}^{-1}(i)\underline{\bar{x}}(i)+\underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{z}(i)\right] \qquad (5.7.1)$$

However, evaluation of this relation required the inversion of n-by-n matrices, and therefore was undesirable for on-line implementation. As a result, the matrix inversion lemma was invoked to produce an equivalent update expression as

$$\underline{\hat{x}}(i) = \underline{\bar{x}}(i) + \underline{K}(i)\left[\underline{z}(i)-\underline{H}(i)\underline{\bar{x}}(i)\right] \qquad (5.7.2)$$

$$\underline{K}(i) = \underline{M}(i)\underline{H}^{T}(i)\left[\underline{H}(i)\underline{M}(i)\underline{H}^{T}(i)+\underline{R}(i)\right]^{-1} \qquad (5.7.3)$$

This is computationally advantageous because the n-by-n matrix inversions are replaced with m-by-m matrix inversions, and the dimension of the measurement vector is typically much smaller than the number of state variables. Although a substantial improvement, this does not constitute a trivial solution form if m is greater than one or two. One is thus led to considering the possibility of incorporating a large vector measurement as a succession of scalar or two-dimensional vectors. A state estimate update of this form will now be developed and shown to be equivalent to the two previous $\underline{\hat{x}}(i)$ relations, and its potential benefits will then be discussed.

For the sake of simplicity, the method will be demonstrated for the case of partitioning the measurement vector into two subvectors, but the result is easily generalized. Assume that the measurement $\underline{z}(i)$ has been taken:

$$\underline{z}(i) = \underline{H}(i)\underline{x}(i) + \underline{v}(i) \tag{5.7.4}$$

Partition this expression as

$$\begin{bmatrix} \underline{z}_1(i) \\ \cdots \\ \underline{z}_2(i) \end{bmatrix} = \begin{bmatrix} \underline{H}_1(i) \\ \cdots \\ \underline{H}_2(i) \end{bmatrix} \underline{x}(i) + \begin{bmatrix} \underline{v}_1(i) \\ \cdots \\ \underline{v}_2(i) \end{bmatrix} \tag{5.7.5}$$

and assume that $\underline{v}_1(i)$ and $\underline{v}_2(i)$ are not correlated, so that the measurement noise covariance $\underline{R}(i)$ can be written as

$$\underline{R}(i) = \begin{bmatrix} \underline{R}_1(i) & \vdots & \underline{0} \\ \cdots & \vdots & \cdots \\ \underline{0} & \vdots & \underline{R}_2(i) \end{bmatrix} \tag{5.7.6}$$

It will now be demonstrated that the state estimate $\hat{\underline{x}}(i)$ achieved by either equation (5.7.1) or (5.7.2) and (5.7.3) can be obtained by means of

$$\hat{\underline{x}}_1(i) = \overline{\underline{x}}(i) + \underline{K}_1(i)\left[\underline{z}_1(i) - \underline{H}_1(i)\overline{\underline{x}}(i)\right] \tag{5.7.7}$$

$$\hat{\underline{x}}(i) = \hat{\underline{x}}_1(i) + \underline{K}_2(i)\left[\underline{z}_2(i) - \underline{H}_2(i)\hat{\underline{x}}_1(i)\right] \tag{5.7.8}$$

where the gain matrices $\underline{K}_1(i)$ and $\underline{K}_2(i)$ are defined by

$$\underline{K}_1(i) = \underline{M}(i)\underline{H}_1^T(i)\left[\underline{H}_1(i)\underline{M}(i)\underline{H}_1^T(i) + \underline{R}_1(i)\right]^{-1} \tag{5.7.9}$$

$$\underline{K}_2(i) = \underline{M}_1(i)\underline{H}_2^T(i)\left[\underline{H}_2(i)\underline{M}_1(i)\underline{H}_2^T(i) + \underline{R}_2(i)\right]^{-1} \tag{5.7.10}$$

where $\underline{M}_1(i)$ is obtained from

$$\underline{M}_1(i) = \left[\underline{M}^{-1}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{H}_1(i)\right]^{-1} \qquad (5.7.11a)$$

$$= [\underline{I}-\underline{K}_1(i)\underline{H}_1(i)]\underline{M}(i)[\underline{I}-\underline{K}_1(i)\underline{H}_1(i)]^T + \underline{K}_1(i)\underline{R}_1(i)\underline{K}_1^T(i) \qquad (5.7.11b)$$

or an equivalent relation.

To prove this, substitute the partitioned equations, (5.7.5) and (5.7.6), into equation (5.7.1) to obtain

$$\hat{\underline{x}}(i) = \left[\underline{M}^{-1}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{H}_1(i)+\underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{H}_2(i)\right]^{-1} \cdot \left[\underline{M}^{-1}(i)\bar{\underline{x}}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{z}_1(i)+\underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{z}_2(i)\right] \qquad (5.7.12)$$

Now generate the result due to an update based upon (5.7.7) to (5.7.11). By reversing the procedure that yielded (5.7.2) from (5.7.1), the intermediate variable $\hat{\underline{x}}_1$ can be expressed as

$$\hat{\underline{x}}_1(i) = \left[\underline{M}^{-1}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{H}_1(i)\right]^{-1} \cdot \left[\underline{M}^{-1}(i)\bar{\underline{x}}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{z}_1(i)\right] = \underline{M}_1(i)\left[\underline{M}^{-1}(i)\bar{\underline{x}}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{z}_1(i)\right] \qquad (5.7.13)$$

Similarly, equation (5.7.8) becomes:

$$\hat{\underline{x}}(i) = \left[\underline{M}_1^{-1}(i)+\underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{H}_2(i)\right]^{-1} \cdot \left[\underline{M}_1^{-1}(i)\hat{\underline{x}}_1(i)+\underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{z}_2(i)\right] \qquad (5.7.14)$$

Now use equations (5.7.11a) and (5.7.13) to replace $\underline{M}_1^{-1}(i)$ in the first brackets and $\hat{\underline{x}}_1(i)$ in the second:

$$\hat{\underline{x}}(i) = \left[\underline{M}^{-1}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{H}_1(i)+\underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{H}_2(i)\right]^{-1} \cdot$$

$$\cdot\left[\underline{M}_1^{-1}(i)\underline{M}_1(i)\left\{\underline{M}^{-1}(i)\bar{\underline{x}}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{z}_1(i)\right\}\right.$$

$$\left. + \underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{z}_2(i)\right]$$

$$= \left[\underline{M}^{-1}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{H}_1(i)+\underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{H}_2(i)\right]^{-1} \cdot$$

$$\cdot\left[\underline{M}^{-1}(i)+\underline{H}_1^T(i)\underline{R}_1^{-1}(i)\underline{z}_1(i)+\underline{H}_2^T(i)\underline{R}_2^{-1}(i)\underline{z}_2(i)\right]$$

(5.7.15)

which is the same as equation (5.7.12), obtained by treating $\underline{z}(i)$ as a single vector update.

The same arguments can be applied for any number of partitions of the measurement vector, as long as the corresponding partitions of $\underline{R}(i)$ yield a block diagonal matrix. For instance, to transform an m-dimensional vector update to m separate scalar updates ($\underline{R}(i)$ must be diagonal), m equations of the form

$$\hat{\underline{x}}_j(i) = \hat{\underline{x}}_{j-1}(i) + \underline{k}_j(i)\left[z_j(i)-\underline{h}_j^T(i)\hat{\underline{x}}_{j-1}(i)\right] \tag{5.7.16}$$

$$\underline{k}_j(i) = \frac{\underline{M}_{j-1}(i)\underline{h}_j(i)}{\underline{h}_j^T(i)\underline{M}_{j-1}(i)\underline{h}_j(i) + R_j(i)} \tag{5.7.17}$$

$$\underline{M}_j(i) = \left[\underline{M}_{j-1}^{-1}(i) + \underline{h}_j(i)\frac{1}{R_j(i)}\underline{h}_j^T(i)\right]^{-1} \tag{5.7.18}$$

$$= \left[\underline{I} - \underline{k}_j(i)\underline{h}_j^T(i)\right] \underline{M}_{j-1}(i) \qquad (5.7.18')$$

would be employed, where it is understood that $\underline{M}_0(i) = \underline{M}(i)$, $\hat{\underline{x}}_0(i) = \overline{\underline{x}}(i)$, $\underline{M}_m(i) = \underline{P}(i)$, and $\hat{\underline{x}}_m(i) = \hat{\underline{x}}(i)$. Other equivalent forms of (5.7.18) are also possible, as discussed previously.

Since the two evaluations of $\hat{\underline{x}}(i)$ are identical for all values of the system parameters, $\underline{a}$, the partial derivative of $\hat{\underline{x}}(i)$ with respect to $\underline{a}$ must also be equivalent when evaluated by the two methods. Thus, the computation of $\partial\hat{\underline{x}}(i)/\partial a_\ell$ and $\partial\underline{P}(i)/\partial a_\ell$ as accomplished through equations (3.4.33) and (3.4.34) is equivalent (assuming $\underline{R}(i)$ is diagonal) to m iterations of

$$\frac{\partial\hat{\underline{x}}_j(i)}{\partial a_\ell} = \left[\underline{I}-\underline{k}_j(i)\underline{h}_j^T(i)\right]\left[\frac{\partial\hat{\underline{x}}_{j-1}(i)}{\partial a_\ell} + \frac{\partial\underline{M}_{j-1}(i)}{\partial a_\ell}\underline{h}_j(i)\frac{1}{\underline{h}_j^T(i)\underline{M}_{j-1}(i)\underline{h}_j(i)+R_j(i)} \cdot \left[\underline{z}_j(i)- \underline{h}_j^T(i)\hat{\underline{x}}_{j-1}(i)\right]\right] \qquad (5.7.19)$$

$$\frac{\partial\underline{M}_j(i)}{\partial a_\ell} = \left[\underline{I} - \underline{k}_j(i)\underline{h}_j^T(i)\right]\frac{\partial\underline{M}_{j-1}(i)}{\partial a_\ell}\left[\underline{I} - \underline{k}_j(i)\underline{h}_j^T(i)\right]^T \qquad (5.7.20)$$

If $\underline{R}(i)$ is not diagonal, but is block diagonal, the appropriate partitions can be made to effect a similar procedure.

Moreover, the single-step score, $s_\ell^1[\underline{Z}(i), \underline{a}_*(i)]$, as evaluated by equation (3.7.22), can be shown to be equal to an m-term summation of similar terms, each due to a single component of $\underline{z}(i)$:

$$s_\ell^1 [\underline{Z}(i), \underline{a}_*(i)] = \sum_{j=1}^{m} \left[\frac{\partial \underline{\hat{x}}_j^T(i)}{\partial a_\ell} \underline{h}_j(i) [z_j(i) - \underline{h}_j^T(i) \underline{\hat{x}}_{j-1}(i)] - \tfrac{1}{2} \Omega_j(i) \frac{\partial A_j(i)}{\partial a_\ell} \right] \qquad (5.7.21)$$

To show this, $\ell n\ f(\underline{z}(i) | \underline{Z}(i-1), \underline{a})$ can be expanded, using Bayes' Rule, as

$$\begin{aligned} \ell n\ f(\underline{z}(i) | \underline{Z}(i-1), \underline{a}) &= \ell n\ f(z_1(i) | \underline{Z}(i-1), \underline{a}) \\ &+ \ell n\ f(z_2(i) | z_1(i), \underline{Z}(i-1), \underline{a}) + \ldots \\ &+ \ell n\ f(z_m(i) | z_1(i), \ldots, z_{m-1}(i), \underline{Z}(i-1), \underline{a}) \end{aligned} \qquad (5.7.22)$$

which, upon differentiation with respect to a_ℓ, yields equation (5.7.21). Similarly, the single-step conditional information matrix can be expressed as an m-term summation.

The computational benefits of this technique become substantial for large m because of the difficulty of inverting an m-by-m matrix for large m; the computational load increases as a function of mn^2, instead of m^3 as in the case of a direct vector update, so this method is especially advantageous for m significantly greater than n. For moderate measurement dimension (especially if numerous parameters are being estimated), the additional iterations require as much computer time as the inversion. However, even if there is no reduction in the computational load, the iterative method might be preferable. If a central processor is employed, high priority jobs might interrupt the updating, and the iterative scheme allows more points of easy interruption. Moreover, if the computer is too overloaded to process an m-vector measurement, it would be better to incorporate as many of the separate partitions as possible than to lose all of the information contained

in $\underline{z}(i)$. In the situation that the iterative scheme does reduce the computer load, it should definitely be used since it is an exact, not approximate, technique.

CHAPTER VI

COMPUTATIONAL RESULTS

This chapter applies the simultaneous estimation of states and parameters to three practical problems. The simulation results will demonstrate the performance achievable from the estimation concept, as well as the effects of the various approximations suggested to achieve on-line applicability. Furthermore, these results will be correlated with the performance predicted by the methods of Chapter IV.

The first problem entails the estimation of the states and uncertain parameter(s) of $\underline{\Phi}(i+1,i)$ in a discrete-time second order system with no control inputs. In the second example, the system is again second order, but the uncertain parameter is in the control input matrix $\underline{B}(i)$, and the control input is completely predetermined. Finally, the last problem is a higher dimensioned, continuous-time system employing feedback control, with an uncertain parameter in the continuous-time dynamics model.

In these examples, the actual system is simulated by means of a set of linear equations driven by white Gaussian noise. This noise, and that corrupting the measurements, are produced by a random number generator whose output is weighted to yield samples distributed according to a normal density. The state and parameter estimates are then subtracted from the corresponding true system values to generate the error time histories presented in this chapter.

6.1 Second Order System with Uncertain Parameters in $\underline{\Phi}$

The first example to be considered is a second order discrete system, driven by a scalar white Gaussian noise sequence, and upon which noise-corrupted scalar measurements are taken. For convenience, the state space model will be

expressed in phase variable coordinates, as

$$\begin{bmatrix} x_1(i+1) \\ x_2(i+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} \begin{bmatrix} x_1(i) \\ x_2(i) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(i) \tag{6.1.1}$$

$$z(i) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(i) \\ x_2(i) \end{bmatrix} + v(i) \tag{6.1.2}$$

$$= y(i) + v(i) \tag{6.1.2'}$$

where $y(i)$ is the actual system output and $z(i)$ is the noisy measurement of its value. Figure 6.1.1 presents the block diagram of the system; the notation $\Longrightarrow$ denotes vector quantities. Also depicted is the equivalent Z-transform model for this sampled data system, for which the transfer function from $w(i)$ to $y(i+1)$ is

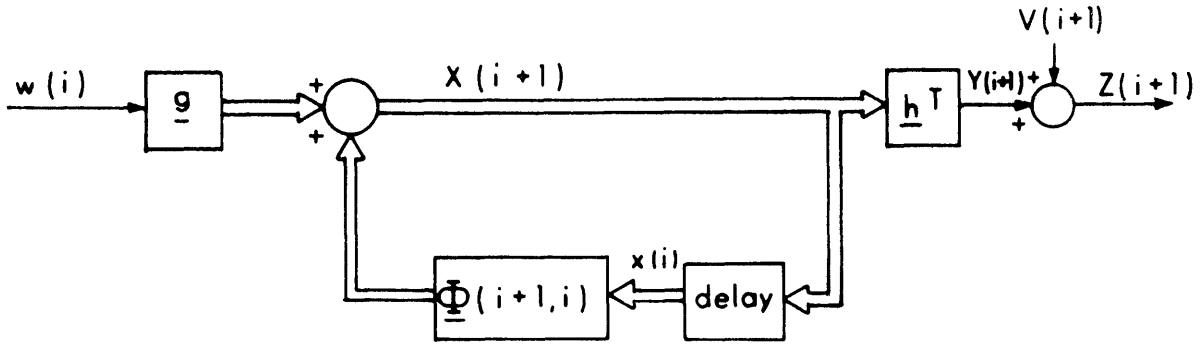
$$\frac{y(i+1)}{w(i)} = \frac{z}{z^2 + a_1 z + a_0} \tag{6.1.3}$$

The values of the parameters a_0 and a_1 are known only with some uncertainty. Their true values determine the z-plane pole locations of the actual system: since the eigenvalues of the system are:

$$\lambda = -\frac{a_1}{2} \pm \tfrac{1}{2} \sqrt{a_1^2 - 4a_0} \tag{6.1.4}$$

then the poles will be on the real axis (two first order system modes) when $a_1^2 - 4a_0 \geq 0$, whereas for $a_1^2 - 4a_0 < 0$, they will be a conjugate pair (a single second order mode). If these true values are in fact constants, then stability is

Time Domain



$$\underline{g} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \underline{h}^T = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$\underline{\Phi}(i+1,i) = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}$$

Equivalent Z-Transform Representation

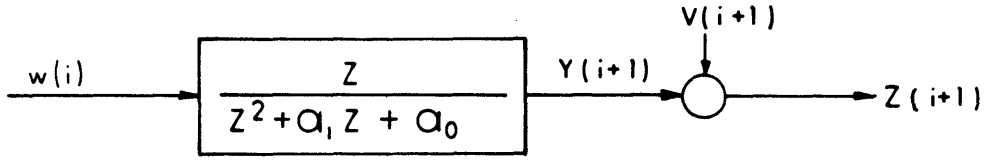


Figure 6.1.1 System Representations

insured by the poles lying inside the unit circle, or $|\lambda| < 1$, and neutral stability resulting from their being on the unit circle, or $|\lambda| = 1$. Thus, the stability conditions for the two previously-mentioned cases are

$$|a_1| + \left|\sqrt{a_1^2 - 4a_0}\right| \leq 2 \quad \text{when } a_1^2 - 4a_0 \geq 0 \tag{6.1.5}$$

$$a_1^2 - 2a_0 \leq 2 \quad \text{when } a_1^2 - 4a_0 < 0 \tag{6.1.6}$$

For the purpose of simulation, it will be assumed that the true parameter values are $a_1 = -1$ and $a_0 = 0.8$, so that a damped second order mode results. Unless specifically stated otherwise, the value of a_1 will be estimated along with the state variables, the initial estimate being $\hat{a}_1 = -0.5$. The assumed apriori state statistics are

$$\underline{\hat{x}}_\circ = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{6.1.7}$$

$$\underline{P}_\circ = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix} \tag{6.1.8}$$

In order to observe the transient behavior of the state estimator, the actual state initial conditions will be set at the 1σ values with the same sign:

$$\underline{x}(0) = \begin{bmatrix} 10 \\ 10 \end{bmatrix} \tag{6.1.9}$$

The driving noise covariance will be assumed time invariant, with magnitude 3 1/3, and similarly the measurement noise

covariance is 0.1 for all time. The effect of altering these defining values will be investigated, but those just described will serve as the basis of comparison. The sample period will be 0.1 sec.

6.1.1 Prediction of Performance

For the system described, a typical state trajectory for $x_1(i)$ is portrayed in figure 6.1.2, which plots $x_1(i)$ versus i for 200 sample times. The $x_2(i)$ trajectory is identical, except that it is displaced one time step ahead. Since the measurement is a rather accurate determination of $x_1(i)$, the errors in the estimate $\hat{x}_1(i)$ will be rather small regardless of the parameter estimate, whereas the errors in $\hat{x}_2(i)$ (a prediction of the next value of x_1) will be more dependent upon the accuracy of the system model. This is borne out in figure 6.1.3: plots a and b depict the errors in the estimates $\hat{x}_1(i)$ and $\hat{x}_2(i)$ of a Kalman filter with a correctly evaluated system model, i.e., $a_1 = -1$, whereas c and d plot the same errors for a Kalman filter with $a_1 = -0.5$. (The generally larger errors in $\hat{x}_2(i)$ are due to its being a prediction of a quantity driven by a relatively strong white noise.) Thus, plots a and b of 6.1.3 represent the best state estimate performance attainable from an estimator of the state and parameters combined, a performance level approached as the parameter estimate converges to the true value.

Figure 6.1.4 presents the ambiguity function plotted as a function of the parameter $-a_1$ (with sign reversed to correspond to the entry in the transition matrix), for $i = 50$ and $N = 30$. Ψ is the total ambiguity function, whereas ψ_1 and ψ_2 are the component terms described in section 5.3.1; these components will be investigated further in section 6.1.6. From this figure it is seen that the ambiguity function is unimodal, with sufficient curvature

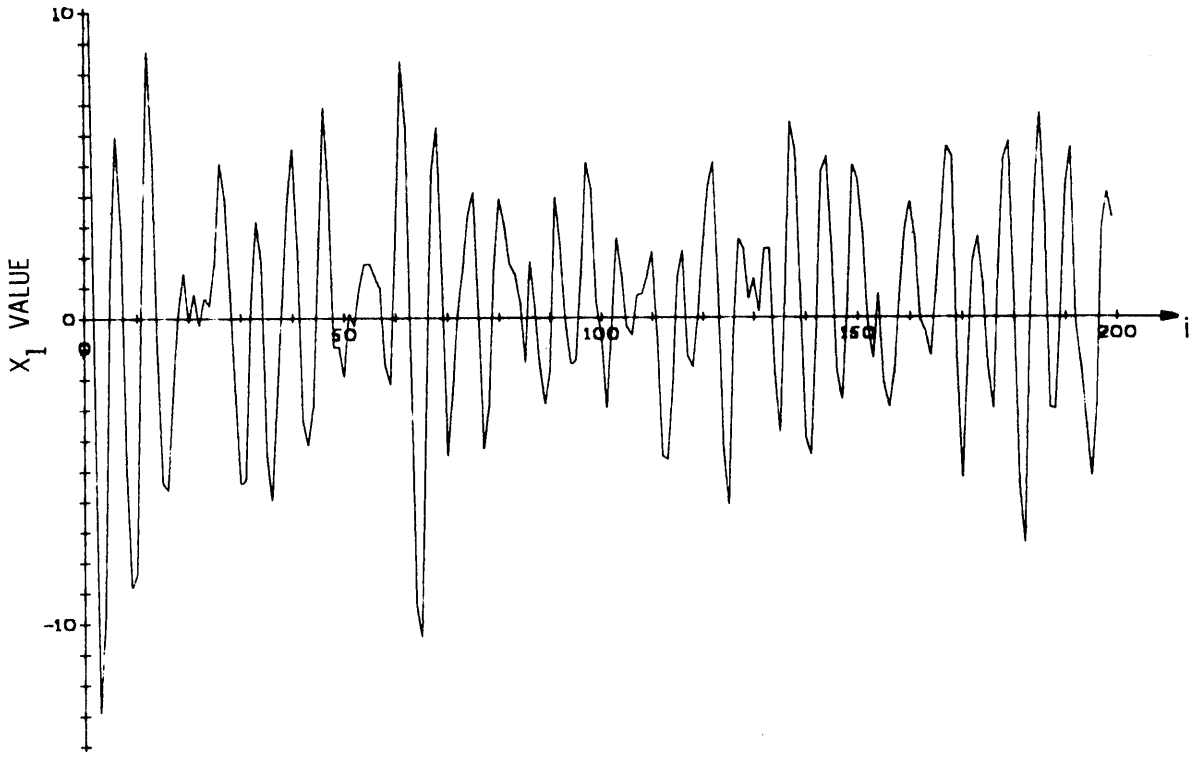


Figure 6.1.2 Typical x_1 State Trajectory

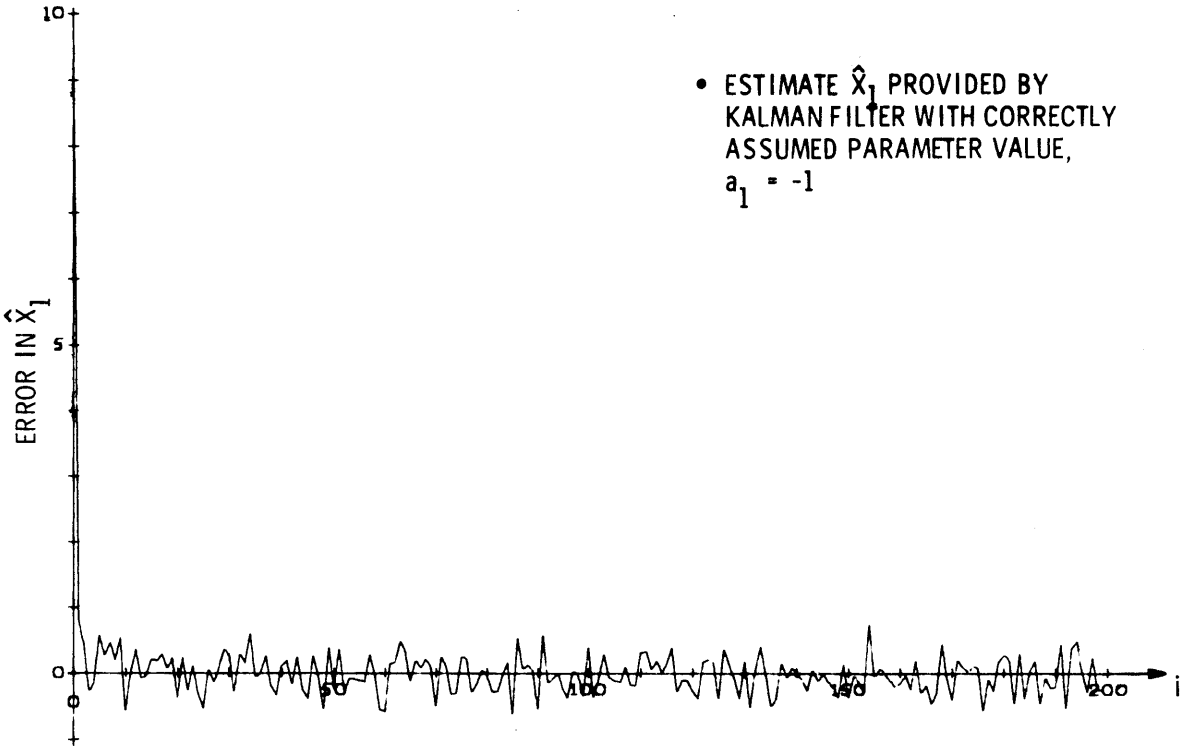


Figure 6.1.3a Error in $\hat{x}_1$ for Assumed Parameter Value of -1

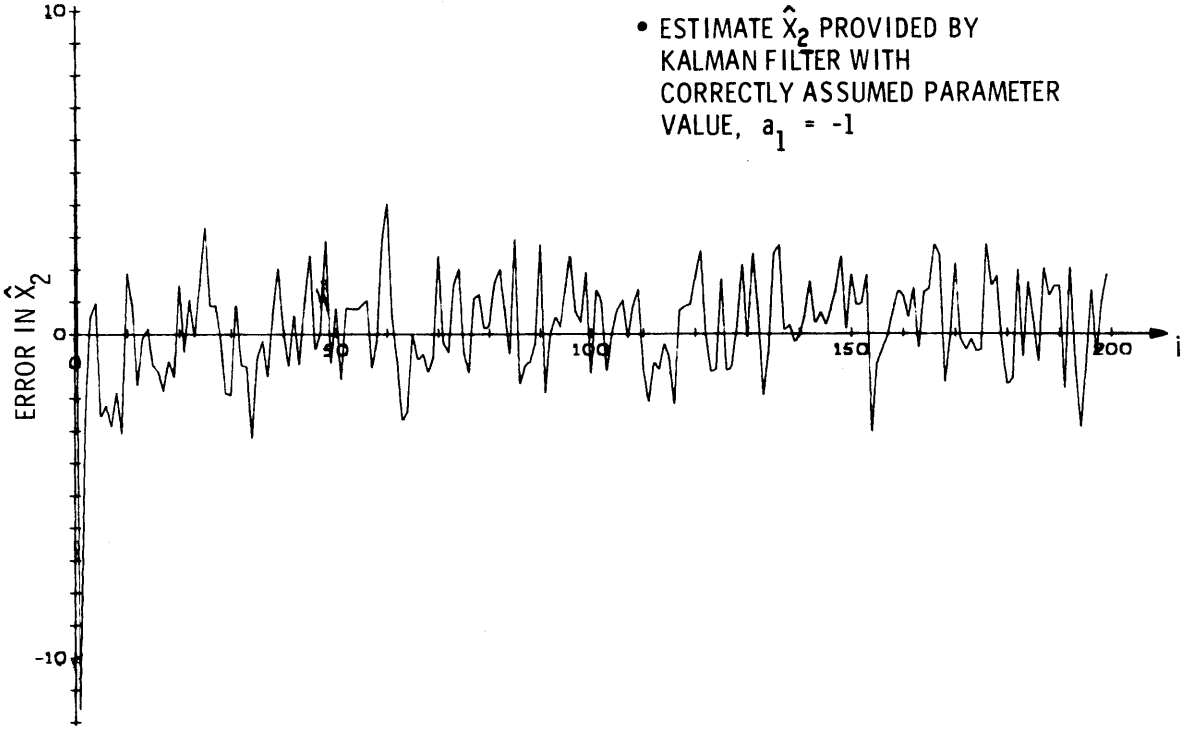


Figure 6.1.3b Error in $\hat{x}_2$ for Assumed Parameter Value of -1

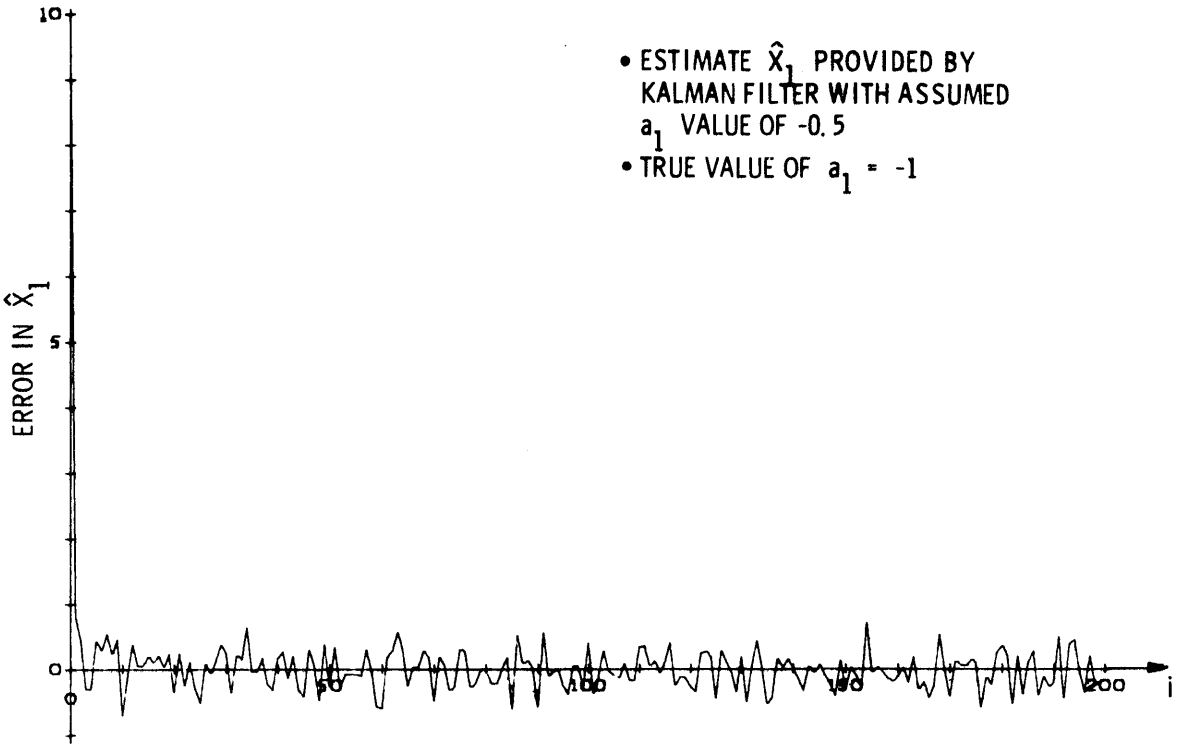


Figure 6.1.3c Error in $\hat{x}_1$ for Assumed Parameter Value of -0.5

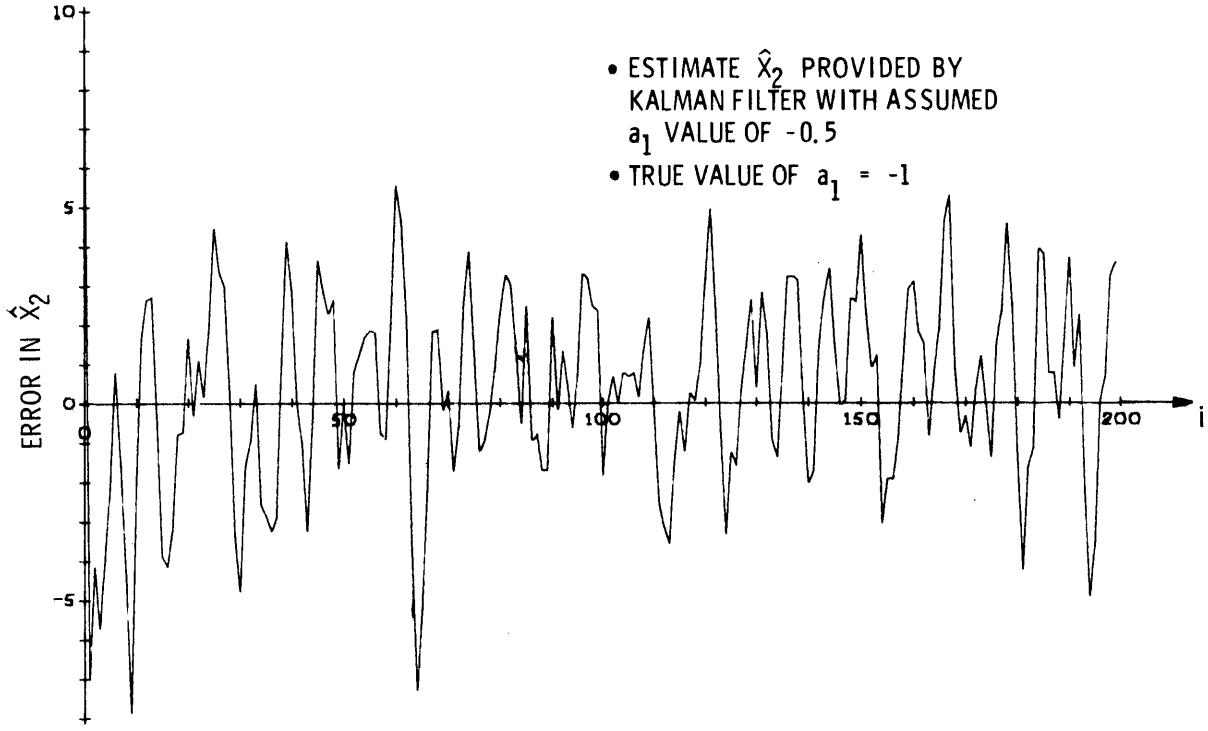


Figure 6.1.3d Error in $\hat{x}_2$ for Assumed Parameter Value of -0.5

at the peak value to predict successful parameter estimation. By calculating the ambiguity function in similar fashion for other values of N, equation (4.3.10) can be used to establish the Cramér-Rao lower bound on an unbiased parameter estimate error covariance as a function of N, the size of the fixed length interval of measurement data that generates the estimate. This is presented in figure 6.1.5. Since this graph represents a lower bound on the achievable performance, it specifies a lower bound on the N required to yield a desired accuracy in parameter estimation for particular values of $\underline{R}$, $\underline{Q}$, and $\underline{P}_o$. Moreover, it relates how practical it would be to attempt performance improvement by increasing N: for small N, enlarging the data interval can substantially improve the accuracy, but for N above approximately 30, the interval must be enlarged appreciably to effect a moderate reduction in error variance. This would be invaluable knowledge to aid determination of the N to be implemented on-line.

6.1.2 Full-Scale Estimator

To analyze the performance attainable through a practical full-scale estimator, the equations of section 3.4 were implemented for this example using the approximation of equations (5.1.2) and (5.1.3). The state and parameter estimates were then compared to the corresponding values in the actual system to yield time histories of the errors in the estimates. In figure 6.1.6 are presented typical parameter error trajectories, plot a pertaining to an estimator with N = 10, and b differing only by letting N = 30. As predicted by the ambiguity function analysis, the larger interval is very effective in reducing the magnitude of the error fluctuations. From figure 6.1.5, the lower bounds on the 1σ values settle out to approximately .16 for N = 10

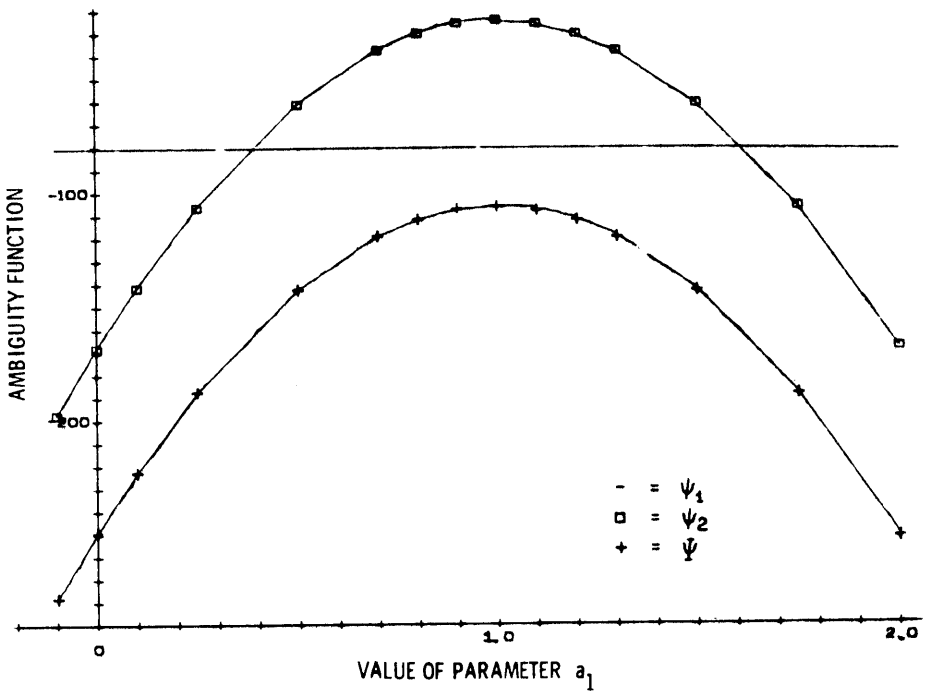


Figure 6.1.4 Ambiguity Function for i = 50 and N = 30

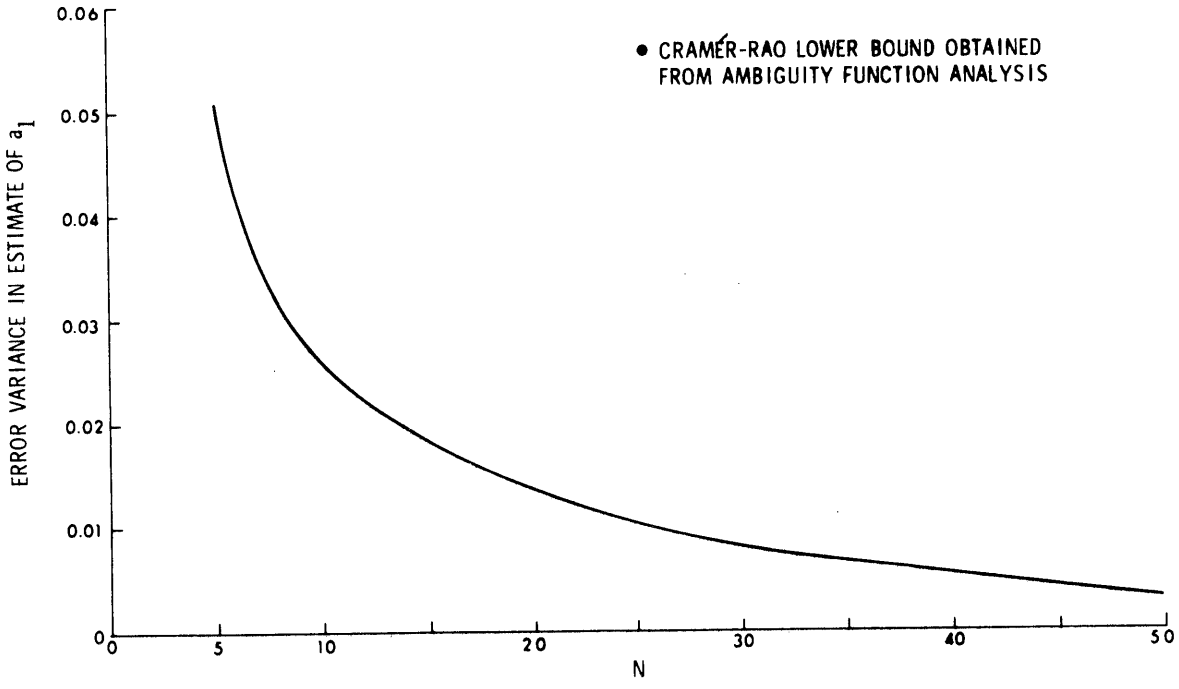


Figure 6.1.5 Lower Bound on Parameter Estimate Error Variance vs. N at i = 50

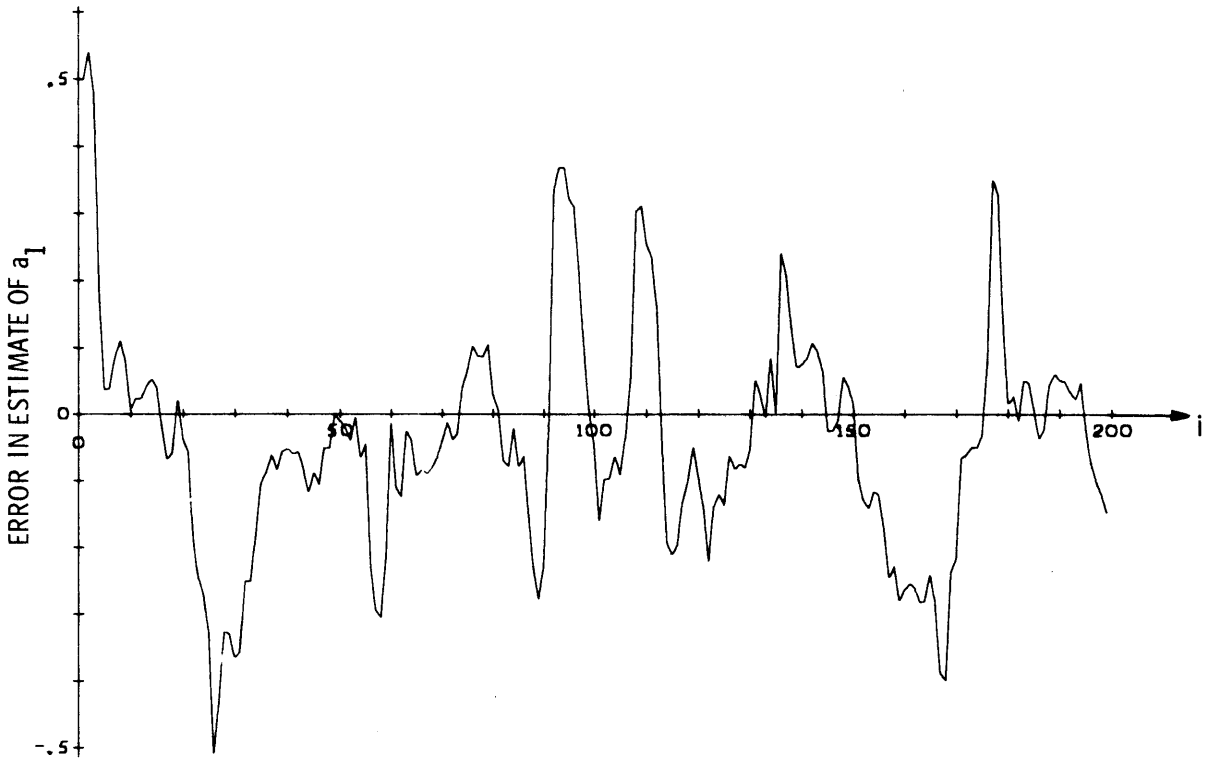


Figure 6.1.6a Parameter Estimate Error Trajectory; Full-Scale Estimator with N = 10

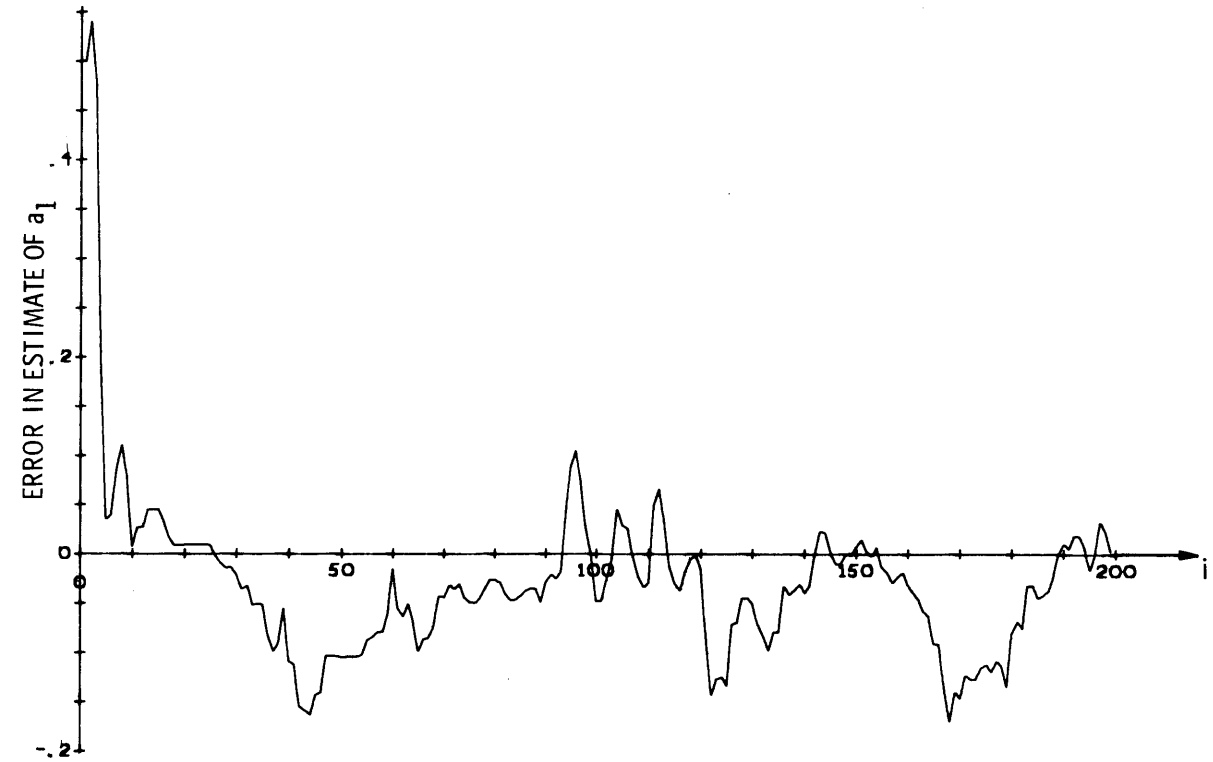


Figure 6.1.6b Parameter Estimate Error Trajectory; Full-Scale Estimator with N = 30

and .09 for N = 30, almost a halving of the expected fluctuations. This is verified by the trajectories in figure 6.1.6: a data interval of 10 samples results in rather large error fluctuations, whereas an interval of 30 samples approximately doubles the average accuracy.

Note that the estimate is somewhat biased after the initial transient. Although an ensemble average over various simulations may yield an average error that is not consistently positive or negative (see figure 6.1.7), each particular simulation typically resulted in a slightly biased parameter estimate.

The ambiguity function analysis can be further substantiated by a Monte Carlo evaluation. For this purpose, five individual simulations were run, differing only in the particular w(j) and v(j) sequences used, and their results averaged at each time instant. Figure 6.1.7 is a plot of the resulting mean and 1σ values of the error in the parameter estimate versus the sample period number (or, time in tenths of a second). It can be seen that the 1σ values have converged to the close vicinity of the lower bound provided by the ambiguity function analysis. The initial transient is undesirable in that the average parameter error grows rapidly for a few samples before converging to a small "steady state" value; the error grows because of the particular initial conditions (not changed for the Monte Carlo runs), and its rate of change is large partially due to the large magnitude of J^{-1} for the first few samples. Practical experience has shown that this undesirable behavior can be removed entirely by delaying the first parameter estimate by the amount of time the average transient requires to settle out, in this case until i = 5.

For all of the cases discussed, the state error trajectories do in fact converge to the behavior of the Kalman filter with correctly evaluated parameters as the parameter estimate converges to the true parameter values. This is

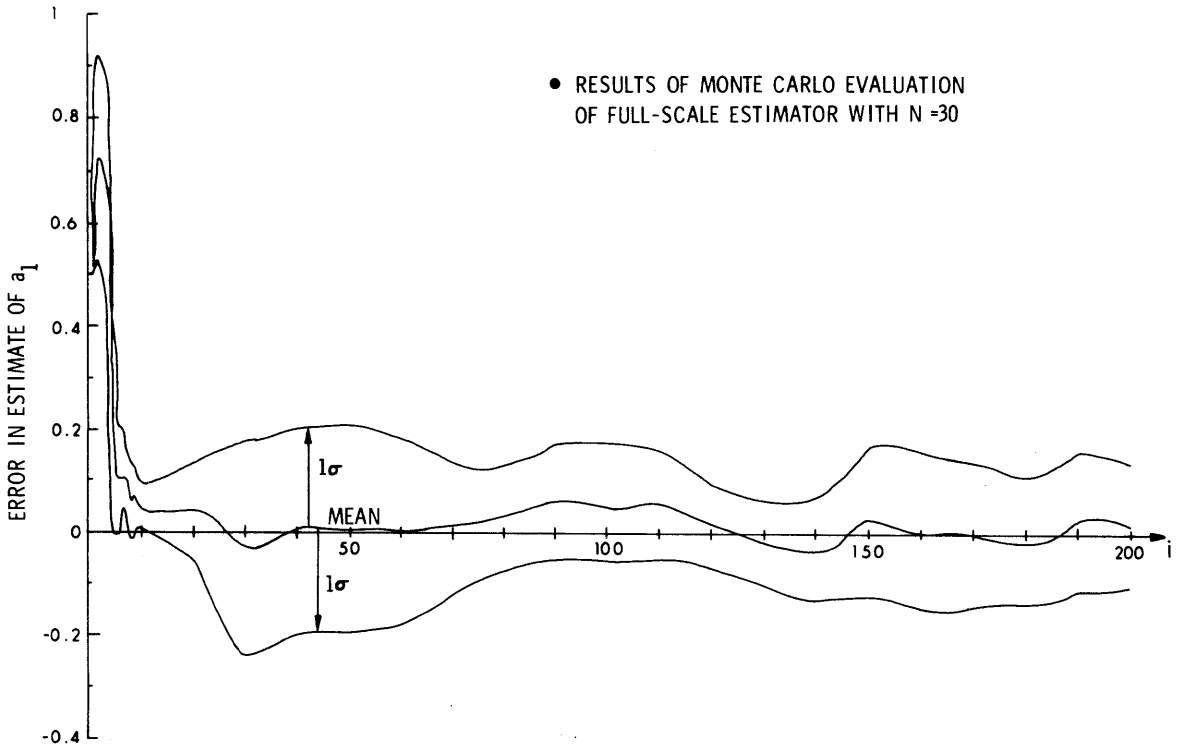


Figure 6.1.7 Mean ± 1σ Values of Full-Scale Parameter Estimate Errors

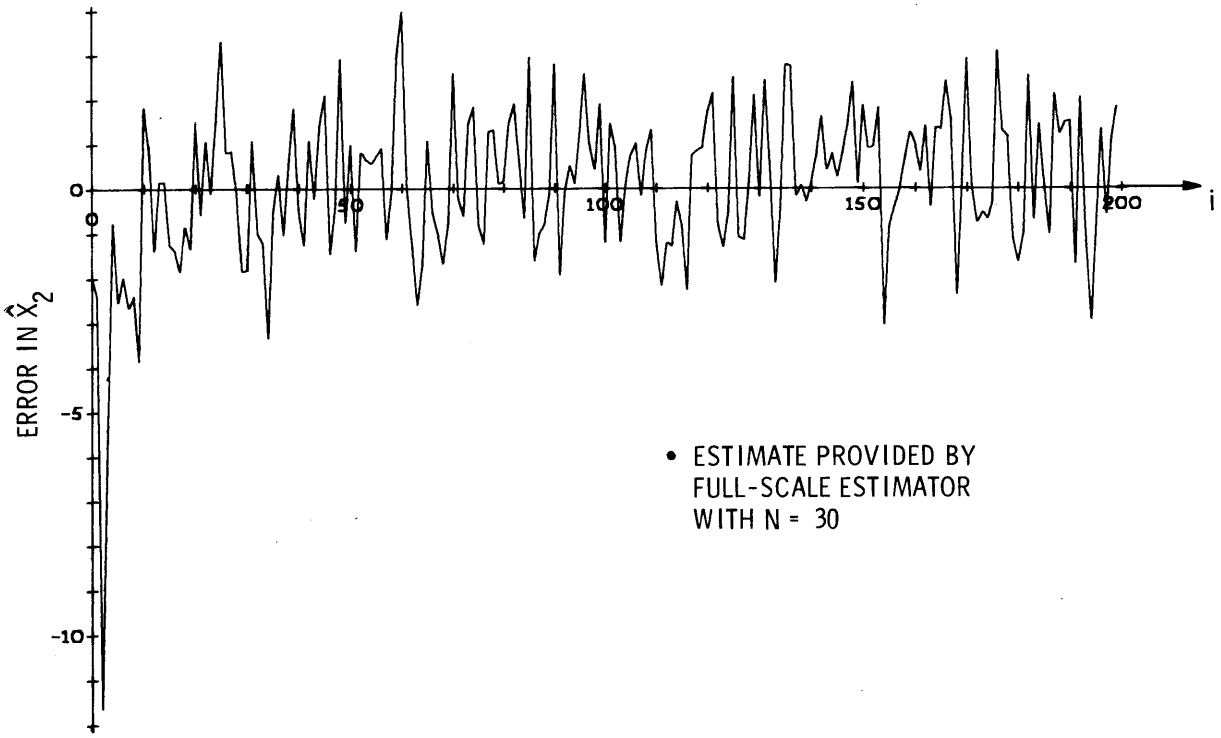


Figure 6.1.8 Error in $\hat{x}_2$ when Parameter is Estimated Simultaneously

readily seen by comparing figure 6.1.8, a plot of the error in $\hat{x}_2$ generated by the estimator that yielded figure 6.1.6b, to figure 6.1.3b.

The previous results were obtained with the estimator based upon a log-likelihood function of $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})$. If instead the estimator were generated from $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{a})$, the computations would be considerably more complex for on-line implementations, so it would be important to know if it is potentially superior enough to warrant further investigation. Comparisons between the two estimators were made over a range of noise levels (Q and R) and parameter estimate initial conditions. Figure 6.1.9 is typical: plot a corresponds to $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{Z}(i-N),\underline{a})$ and b to $\ell n\ f(\underline{x}(i),\underline{Z}_N(i)|\underline{a})$ for $Q = 3\ 1/3$, $R = 0.05$, $\hat{a}(0) = 0$, $N = 30$. The trajectories are almost identical, so the first of the two forms is preferable because of its smaller computational requirement.

Figure 6.1.9a also illustrates two other aspects of the parameter estimator's behavior. First of all, the ability to acquire a good estimate from an arbitrary initial condition is not strongly dependent upon the magnitude of the error in the initial condition. Over a large range of $\hat{a}(0)$, the estimator was able to converge to the immediate neighborhood of a_t within five sample periods; the same behavior was exhibited in the other examples as well. In fact, the choice of an initial error of 0.5 in the parameter estimate for most of the plots in this section was motivated by the desire to keep the "steady state" behavior readily discernable relative to the scale of the graphs. Thus, for practical applications, if the problem is not pathological (as, multiple-peaked ambiguity and likelihood functions), the validity of such approximations as linear curve-fitted function evaluations would restrict the range of allowable parameter values more than the ability of the initial transient to converge.

Secondly, when figure 6.1.9a is compared to 6.1.6b, the effect of reducing the measurement noise variance R(j)

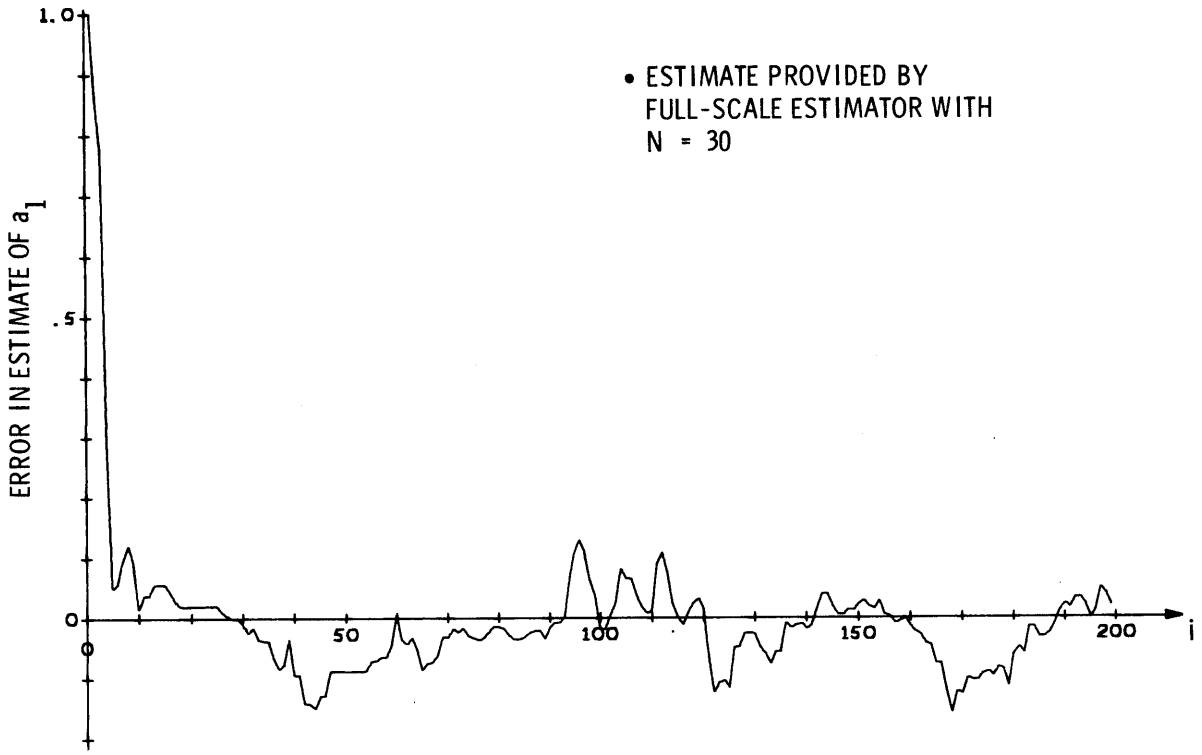


Figure 6.1.9a Parameter Estimate Error; Estimator Based on $\ell n\ f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$

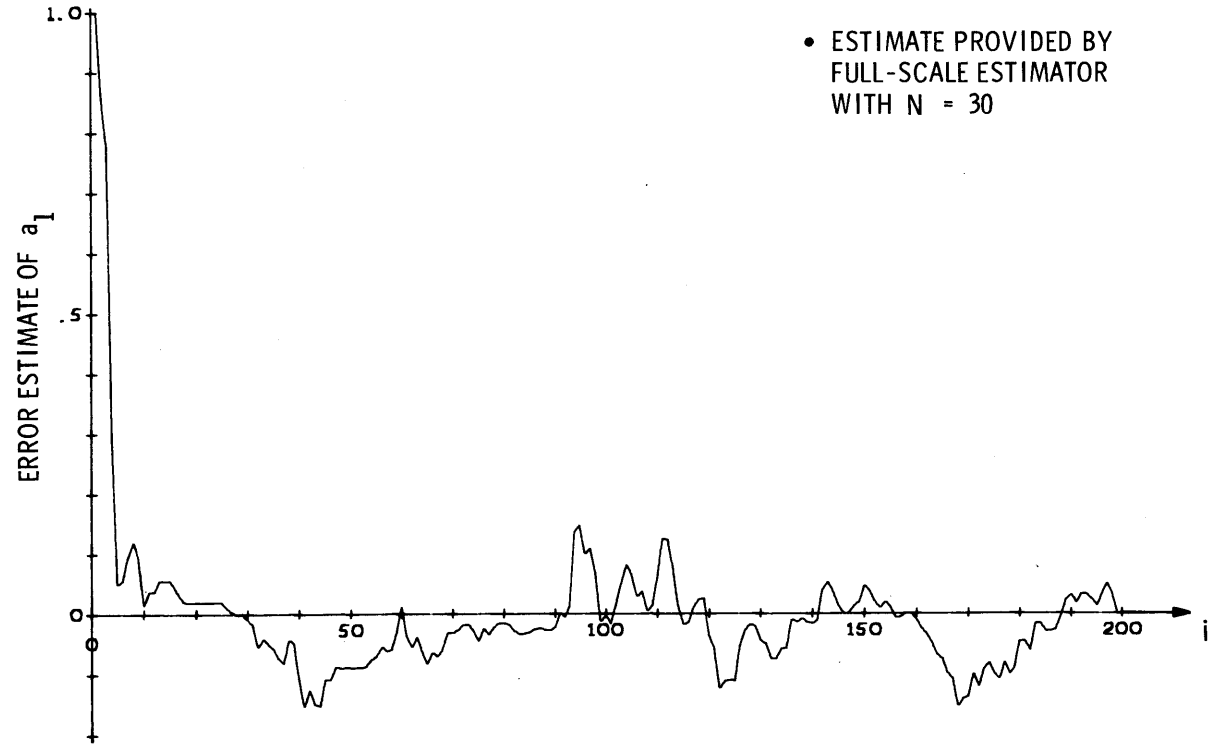


Figure 6.1.9b Parameter Estimate Error; Estimator Based on $\ell n\ f(\underline{x}(i), \underline{Z}_N(i) | \underline{a})$

from 0.1 to 0.05 is not very apparent in the post-transient, or "tracking", phase. This and other runs confirm the trends discernable from figure 6.1.10, which is a plot of the Cramér-Rao lower bound as a function of N, at i = 50, for various values of R. For R less than or equal to 0.1, increasing or decreasing the accuracy of the measurements does not perceptibly affect the accuracy of the parameter estimate, whereas above this variance level, additional measurement precision can significantly improve performance, especially for smaller values of the data interval size, N.

Simulations and ambiguity function analyses similarly display a low sensitivity to driving noise variance Q over a large range of its value. Since there are no deterministic inputs, a certain amount of driving noise is necessary to maintain enough system activity to observe its dynamics and generate an accurate parameter estimate. Above a Q value of 1, additional driving noise essentially does not affect the lower bound, and simulations demonstrated that a threefold increase of Q to 10 increased the absolute value of the parameter error by at most 25 percent and on the average by approximately 2 percent.

Figure 6.1.11 plots the typical parameter and state estimate errors incurred when both a_1 and a_0 are estimated; plot a is the error in the a_1 estimate, b pertains to a_0, and c to $\hat{x}_2$. The only difference between this simulation and that portrayed by figure 6.1.6b is the addition of the a_0 estimate and the delay of the first parameter estimate to reduce the magnitude of initial transients: i.e., the identical random numbers simulated the noise inputs, and initial conditions were the same except that $a_0(0)$ was erroneously assumed to be 0.3 (again note that the plots are actually errors in $-a_1$ and $-a_0$). Therefore, figure 6.1.11a is directly comparable to figure 6.1.6b, and the ability to perform an accurate estimate is seen to be maintained in the face of additional parameter uncertainty.

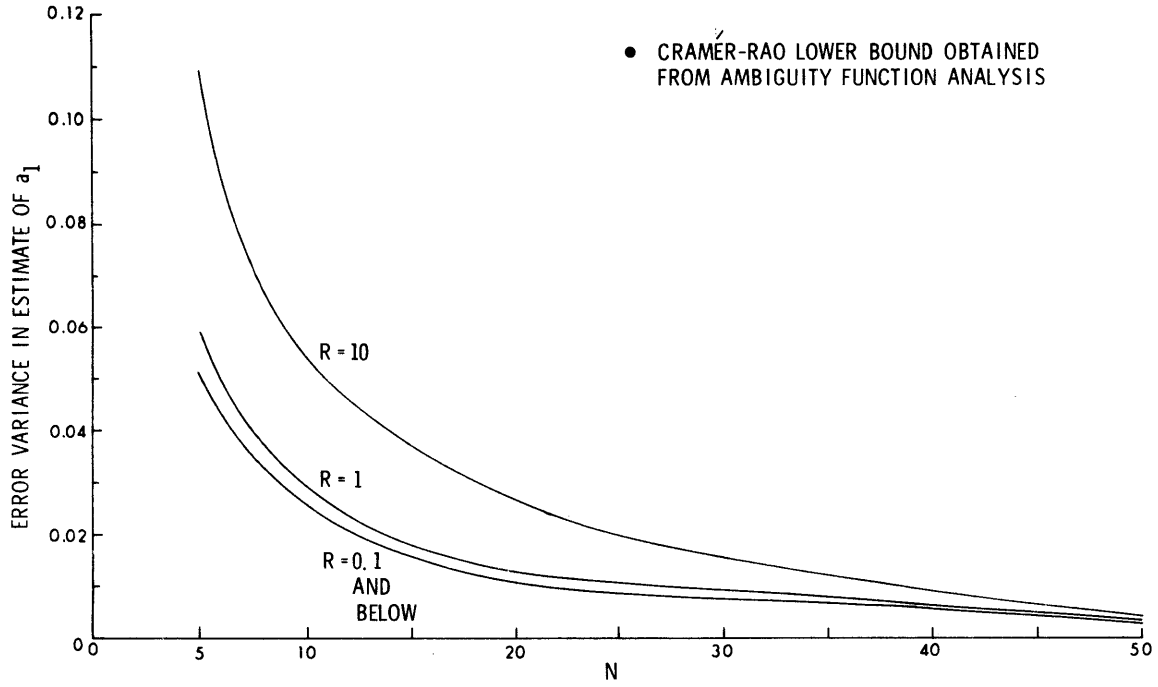


Figure 6.1.10 Lower Bound on Parameter Estimate Error Variance vs. N at i = 50

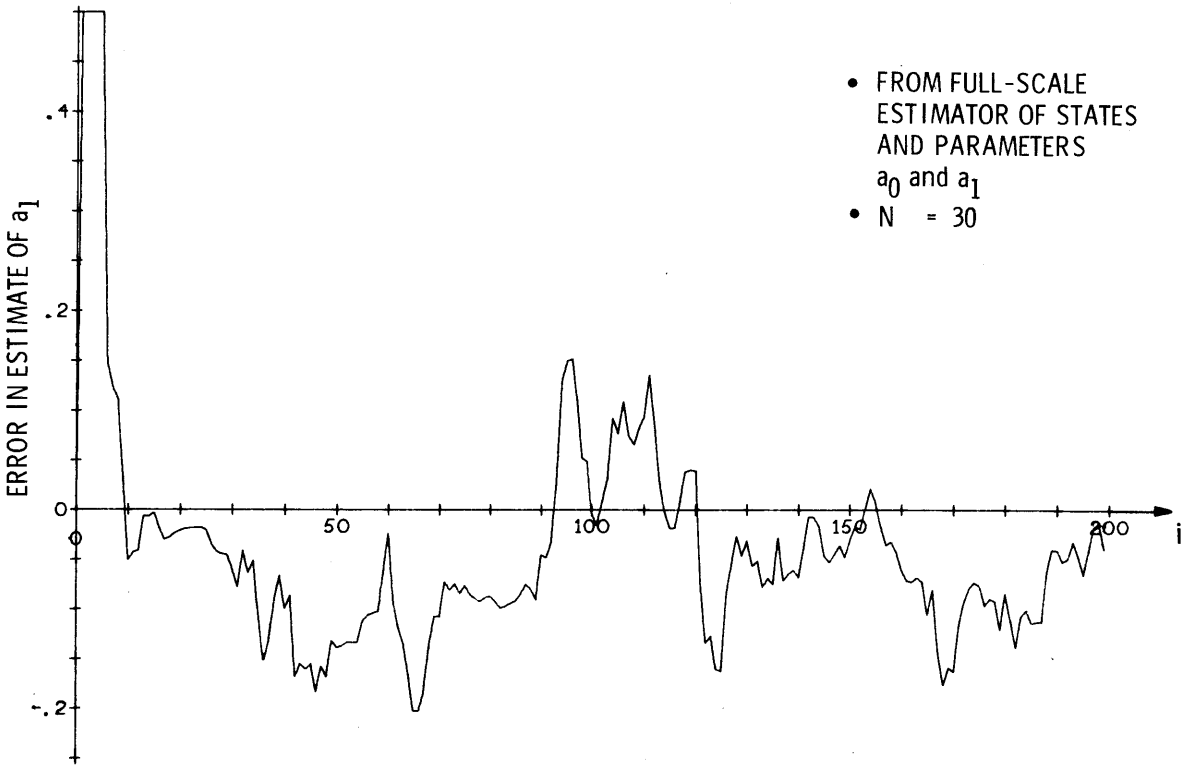


Figure 6.1.11a Error in Estimate of the Parameter a_1

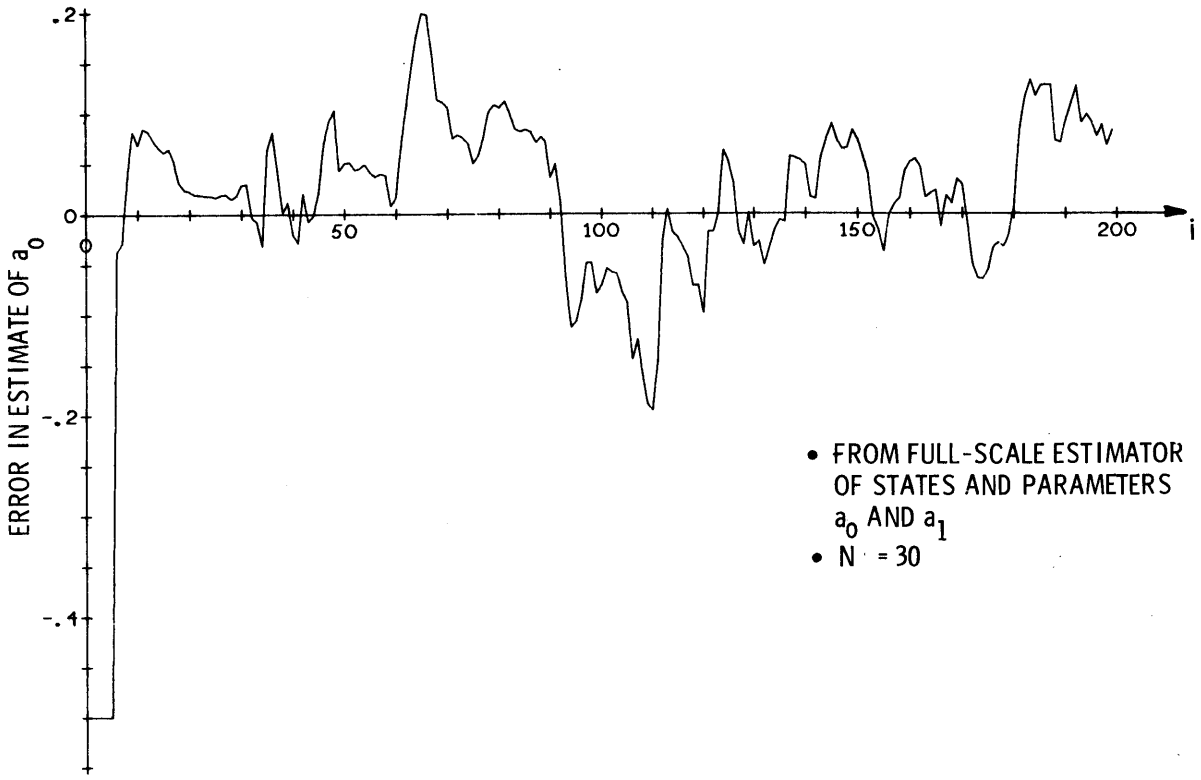


Figure 6.1.11b Error in Estimate of the Parameter a_0

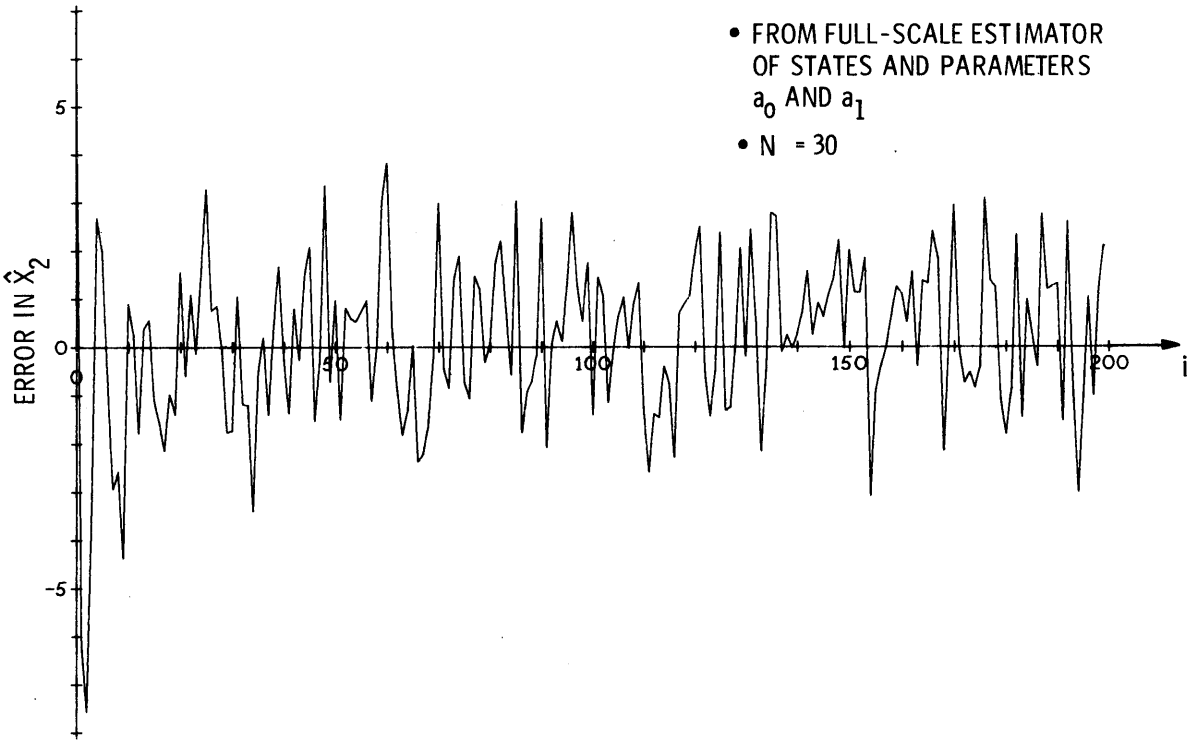


Figure 6.1.11c Error in Estimate of the State x_2

Direct comparison of the first two plots of figure 6.1.11 reveals that the increased error magnitude for a_1 corresponds directly to an error of opposite sign in estimating a_0, which is to be expected since the parameter estimator is basically fitting the system model to the acquired data. As a result, the errors in the state estimates for the two cases are virtually indistinguishable. This is demonstrated by relating figure 6.1.11c to figure 6.1.8 (the improved transient behavior of figure 6.1.11c is due to the delay in performing the first parameter estimate).

6.1.3 Effects of Apriori Statistical Information

Section 3.2.2 presented arguments against the incorporation of apriori parameter statistical information into the likelihood function unless this knowledge is rather precise. The influence of such information upon estimator behavior will now be investigated.

Figure 6.1.12a presents the parameter estimate error trajectory for a simulation identical to that producing figure 6.1.6b except that an apriori normal $f(a_1)$ is assumed, centered at the true value and with a variance of one. The initial change in the parameter estimate is in the direction of the mean of this density, but by the fifth sample period, the effect of $f(a_1)$ upon the parameter estimate has essentially been removed. Similar behavior is exhibited in figure 6.1.12b, which is the parameter estimate error caused by using an erroneous apriori normal density with zero mean and unity variance: initial convergence towards the mean of $f(a_1)$ followed by a rapid decay of its influence upon the estimate. Although densities centered very near the true parameter values with considerably smaller variances can substantially improve the estimator performance, such accurate apriori data is seldom available for a parameter that requires estimation.

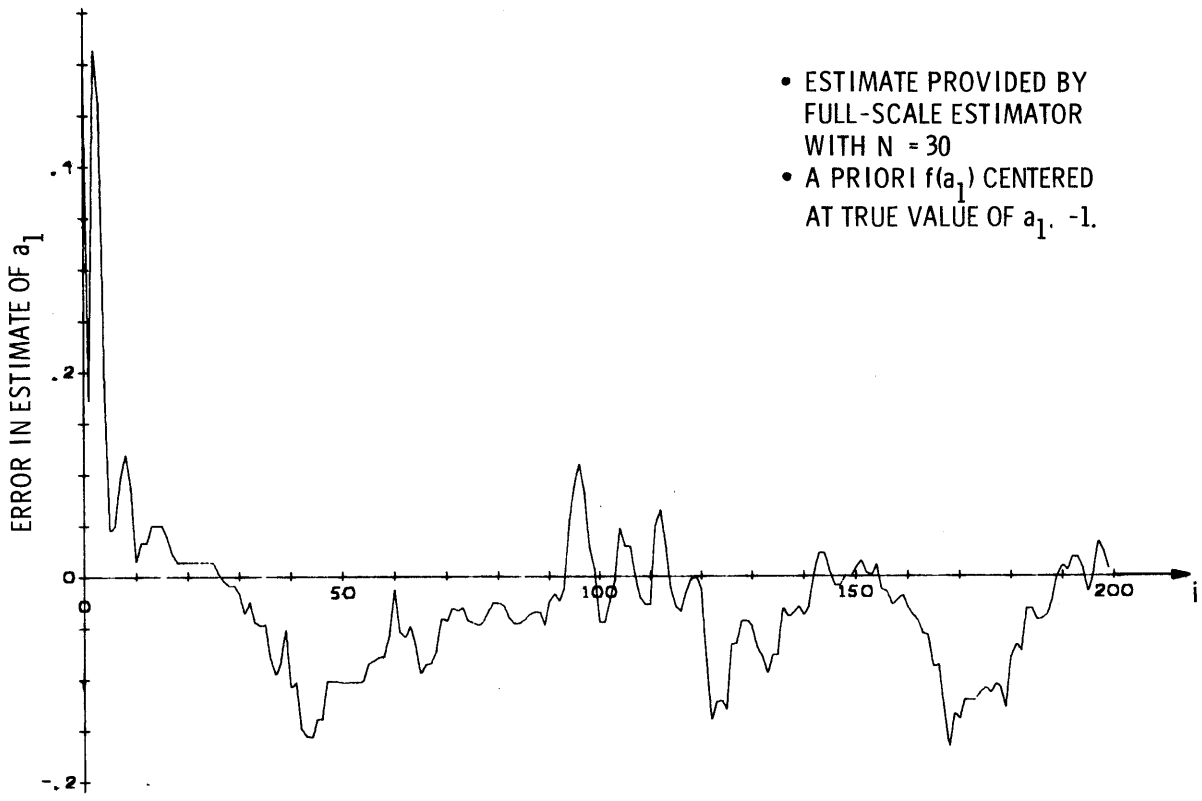


Figure 6.1.12a Parameter Estimate Error; Apriori $f(a_1) = N(-1,1)$

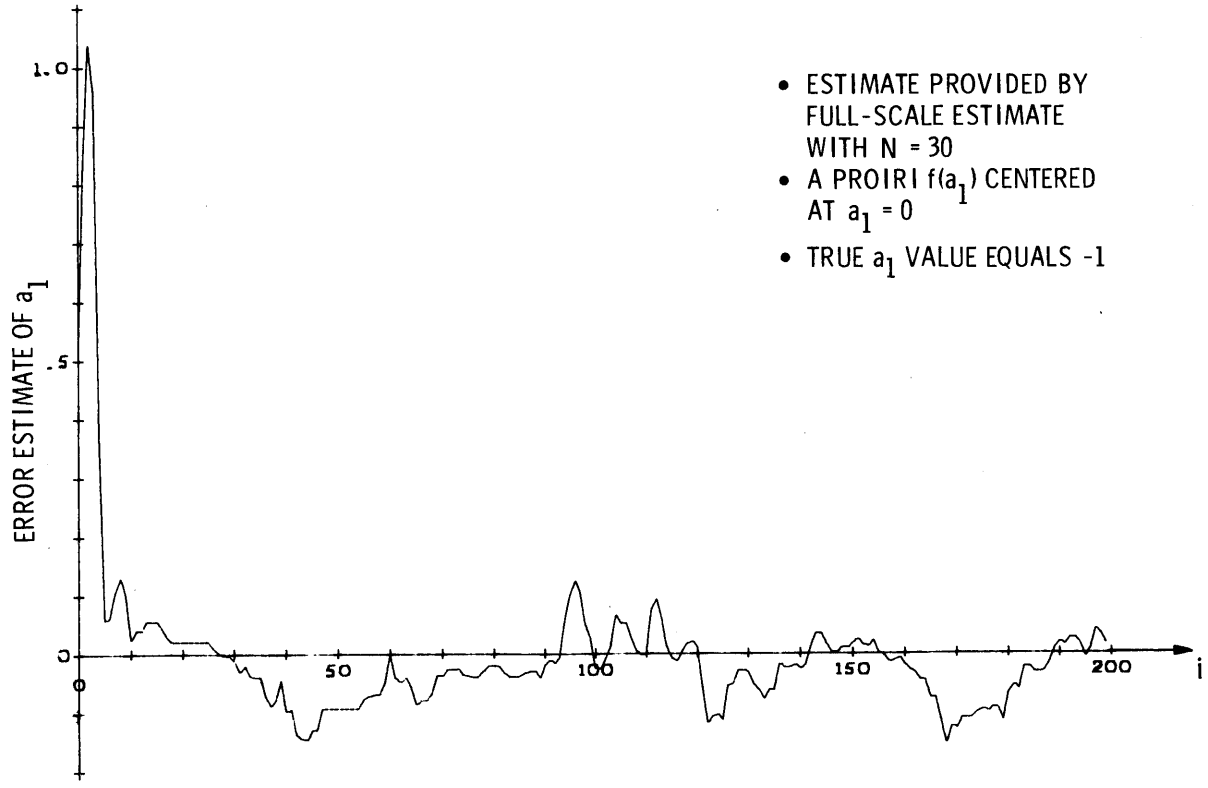


Figure 6.1.12b Parameter Estimate Error; Apriori $f(a_1) = N(0,1)$

Typically, the information contained in $f(\underline{a})$ is insignificant relative to that available from N samples of data, so that $f(\underline{a})$ bears an effect only upon the initial transients of the parameter estimate. Since the mean of any available $f(\underline{a})$ would serve as the first parameter estimate, even this effect is reduced. Especially if the first parameter estimate is delayed for a number of sample instants, the incorporation of $\ell n\ f(\underline{a})$ into the likelihood equation is often superficial. This tends to support the contention that the estimator based upon $\ell n\ f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$ will usually be the most useful in practice.

Apriori state statistical information is usually more complete and is thus included in the model. However, its effect also is substantially reduced after an initial transient. Figure 6.1.13 presents the time progression of the parameter estimate error variance over the first 30 sample periods for $\underline{P}_o$ equal to $\underline{I}$ and 100 $\underline{I}$, assuming $\hat{\underline{x}}_o$ is unbiased. Whereas the initial bounds differ widely, their separation quickly decreases to being almost undiscernable at 30 samples.

6.1.4 On-Line Technique; Parameter Estimate Every N Samples

This section investigates the on-line conceptualization that performs a parameter estimate every N samples, as presented in section 5.2. In order to provide a direct basis of comparison, the representative graphs that appear in this section correspond to an identical sequence of w(j) and v(j) as that which generated the full-scale results of figure 6.1.6.

Figure 6.1.14a portrays the parameter estimate error committed by an on-line system that estimates the parameter every ten samples (with N = 10). For the first ten sample periods, a parameter estimate is made every period in order to improve the acquisition characteristics from a large initial error. Without this modification, the first change

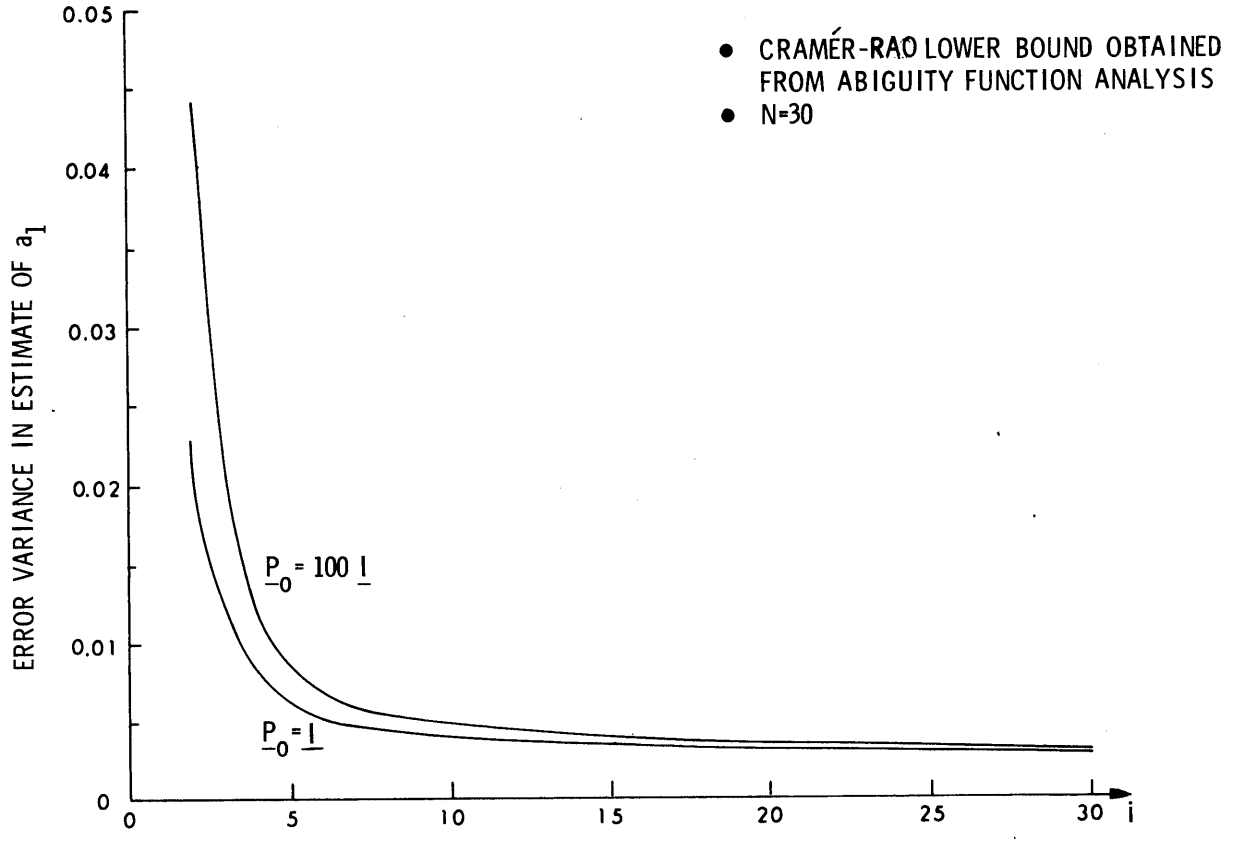


Figure 6.1.13 Time Progression of Lower Bound on Parameter Estimate Error Variance

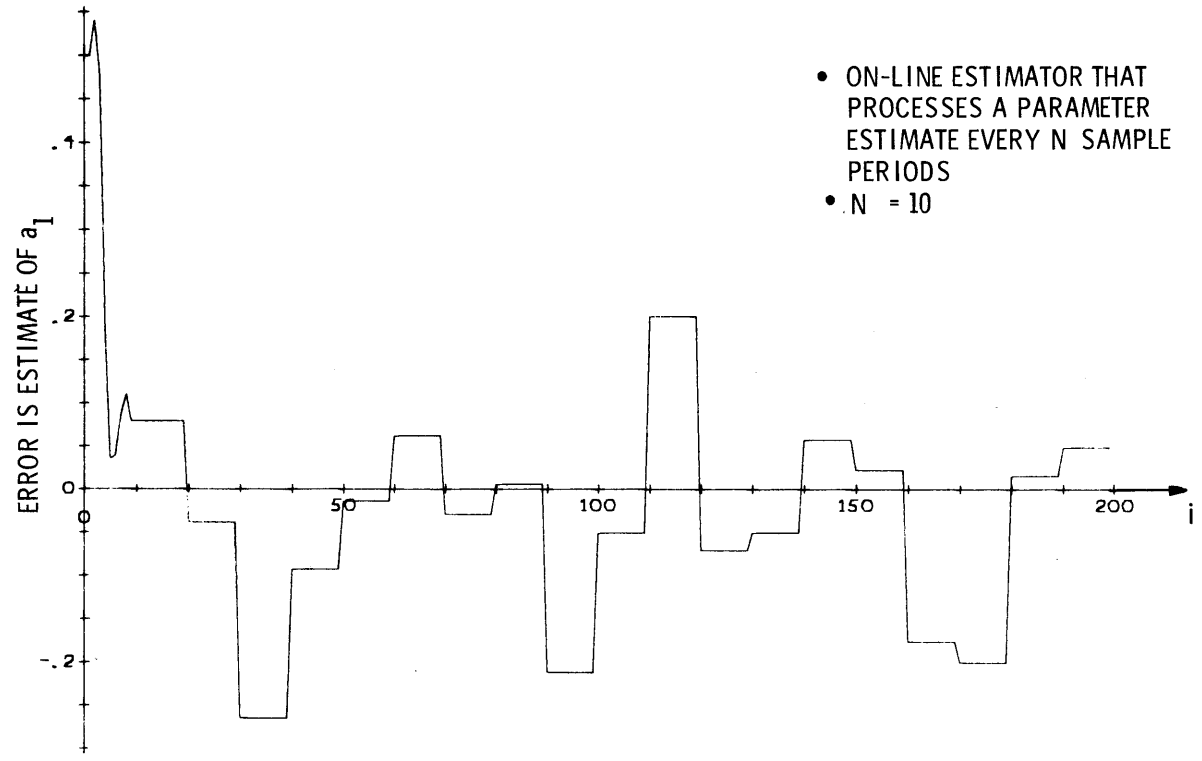


Figure 6.1.14a Parameter Estimate Error; On-Line Estimator with N = 10

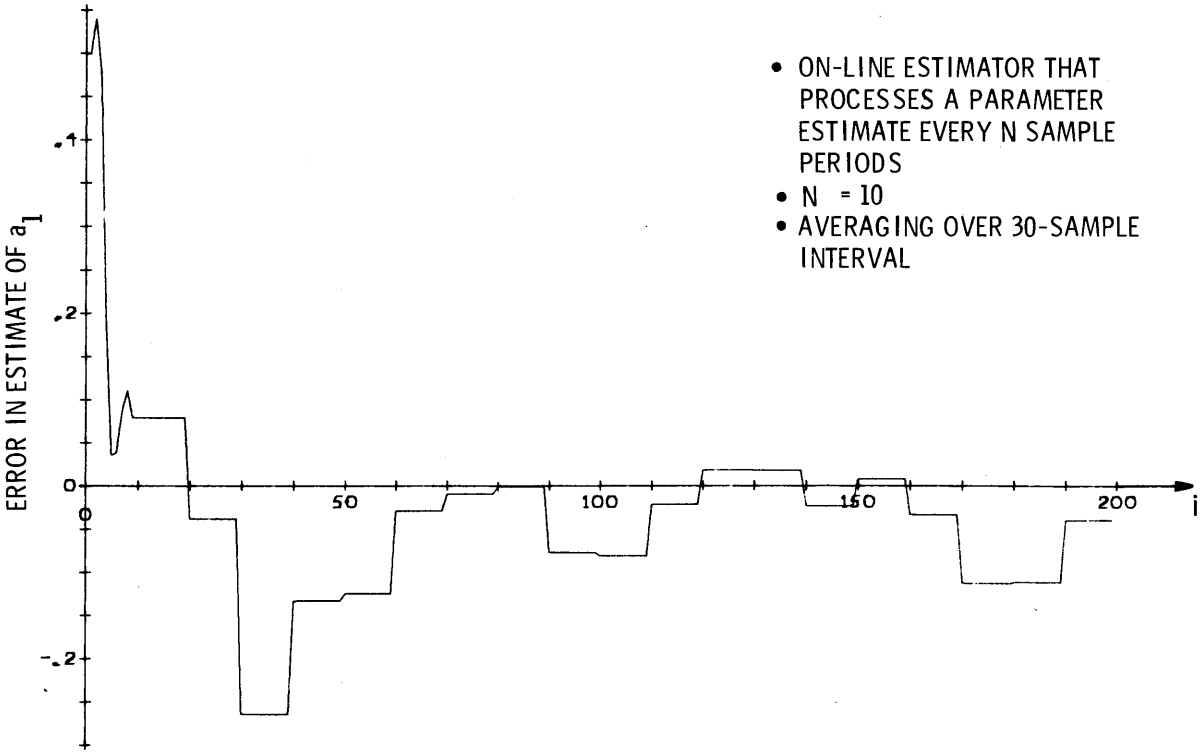


Figure 6.1.14b Parameter Estimate Error; On-Line Estimator with N = 10 plus Averaging

in the parameter estimate is generated by an N-step propagation of data using the original guess of the parameter values. Since the estimate requires a few iterations to converge, this number of iterations multiplied by N yields the number of single step score and conditional information matrix values that will be generated with potentially poor quality parameter values. Although the more rapid parameter estimation in the early stages does not incorporate recalculations of any $\underline{s}^1$ or $\underline{J}^1$ values, by allowing N iterations upon the parameter estimate in the first N samples, instead of one, it generally produces greater convergence to a good estimate. Furthermore, as the estimate begins to converge, succeeding values of $\underline{s}^1$ and $\underline{J}^1$ are generated with more valid parameter values, thereby improving the convergence. Tests conducted without the initially accelerated estimate frequency required two to three parameter estimates (20 to 30 sample instants) for the transient to settle out: on the average, five times as long as required for the decay of the transient in figure 6.1.14a. Since the additional computations are minimal, the benefits of this method advocate its implementation in general.

By observing figures 6.1.6a and 6.1.14a, the error magnitudes of the full-scale and on-line techniques are seen to be comparable. This is true, despite the fact that the full-scale parameter estimator required 12.95 seconds of IBM 360/75 computation time for the 200 sample periods (with no precomputations, the associated Kalman state estimator required 0.55 seconds; simulation and program setup times are subtracted from all times cited), whereas the on-line conceptualization needed only 1.10 seconds.

Since an appropriate value of N has been determined as equal to 30, this on-line technique can simply be implemented so as to make a parameter estimate every 30 periods (after an initial 10 samples of more rapid estimation.

Without this increased estimation frequency, the problem of slow convergence and $\underline{s}^1$ and $\underline{J}^1$ values generated with inaccurate parameter values is accentuated by the larger N: the transient remains in effect for about two estimates, i.e., 60 sample periods, which comprises an unacceptably large amount of time.) However, the expectation that the parameters will be approximately invariant over 30 time steps can be incorporated in an alternative fashion. The estimator with N = 10 can generate parameter estimates every ten periods, each based upon the most recent ten samples of data by means of $\ell n\ f(\underline{x}(i), \underline{Z}_{10}(i) | \underline{Z}(i-10), \underline{a})$. Once a new parameter estimate is made, it is averaged with the previous two estimates, generated through $\ell n\ f(\underline{x}(i-10), \underline{Z}_{10}(i-10) | \underline{Z}(i-20), \underline{a})$ and $\ell n\ f(\underline{x}(i-20), \underline{Z}_{10}(i-20) | \underline{Z}(i-30), \underline{a})$, to obtain the parameter value that best fits the data taken over the three most recent 10-step intervals of data. The original (unaveraged) estimate is stored for future averaging, while the averaged value is used as the actual parameter estimate. Figure 6.1.14b presents the results of this procedure. Here the averaging was started at i = 40 to yield a graph scale identical to that of figure 6.1.14a, thereby facilitating a comparison of the post-transient error magnitudes; averaging two estimates at i = 20 and three at i = 30 further reduces the error oscillations and would be recommended for actual use. The accuracy of the estimates following i = 40 strongly resembles that attained by the full-scale estimator with N = 30, as seen in figure 6.1.6b. For this case, the reduction in computation time is even more pronounced, from 38.98 to 1.10 seconds on an IBM 360/75. This method has an additional advantage over simply setting N to 30; it decreases the inherent lag time in responding to parameter variations. If a parameter value does slowly vary after a good estimate is achieved at time instant i, the latter technique will not begin to respond until time (i+30), whereas the former would begin

to respond at (i+10). Moreover, the former would make three iterations by time (i+30), with more valid $\underline{s}^1$ and $\underline{J}^1$ for instants (i+11) to (i+30), and thus would probably converge to a better estimate at time (i+30) than the estimator that uses an N of 30.

Figure 6.1.15 presents the mean and 1σ values obtained from a Monte Carlo simulation using N = 10 and averaging over the three most recent estimates. The same sets of noise sequences were used to generate this as the full-scale simulations yielding figure 6.1.7, and comparison of these two figures verifies the fact that the on-line technique attains almost identical accuracy as the more complex version.

The effects of varying initial conditions, driving noise variance Q, and measurement noise variance R duplicate the trends established by the full-scale estimator in the two previous sections.

6.1.5 On-Line Technique; More Frequent Parameter Estimation

In section 5.2 a second on-line conceptualization, suitable for estimating the parameters more frequently than every N samples, was described. Rather than regenerate all terms in an N-step interval when the parameter estimate was updated, it simply added a term to the end and subtracted one from the beginning of the summations as the system progressed forward by a sample period. This section examines this form of estimator for the case of making a parameter estimate every sample time.

Although an appropriate N is 30, the lack of term regeneration requires a smaller interval to provide acceptable transient and tracking behavior, as discussed in the previous section. The estimates over the preceding 30-step interval are then averaged to incorporate the constant-over-N-step model into the parameter estimate and remove high frequency fluctuations. (Without the averaging, high frequency oscillations with unacceptably large amplitude may be exhibited

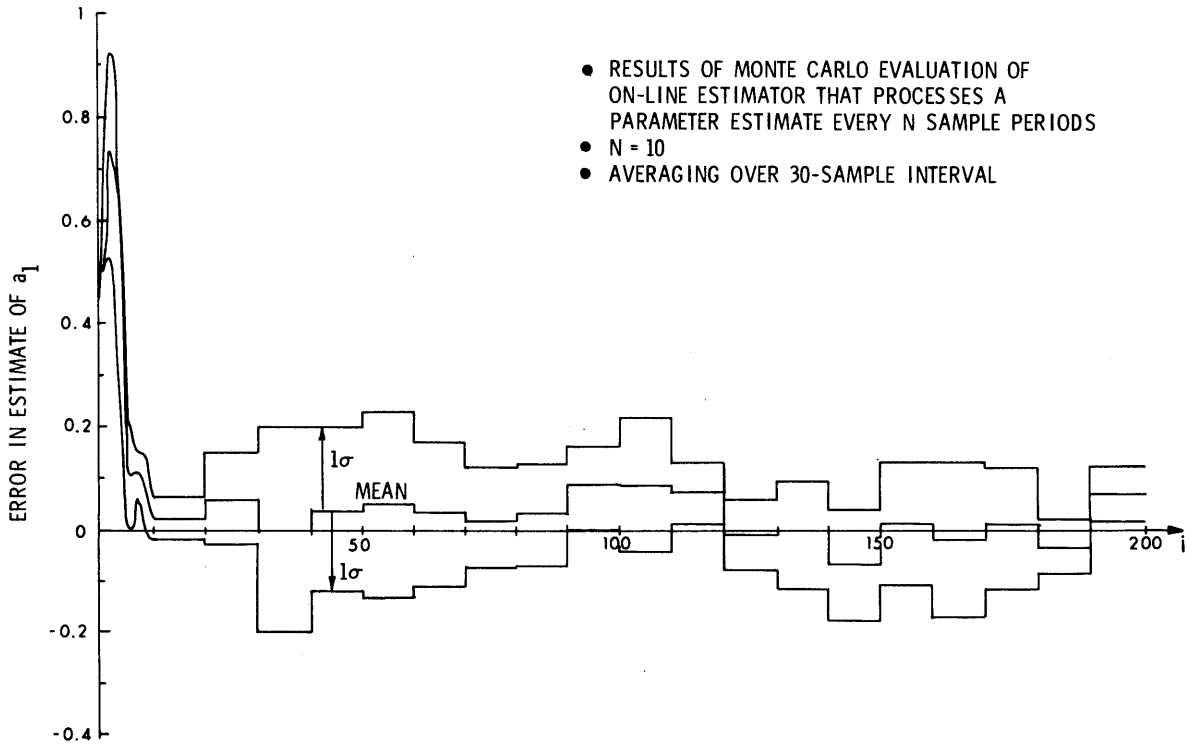


Figure 6.1.15 Mean ± 1σ Values of On-Line Parameter Estimate Errors

by the parameter estimate.) In figure 6.1.16a is plotted the parameter estimate error from such an estimator using an interval of 5 sample periods, produced by the same noise sequences as generated figures 6.1.9 and 6.1.14. Averaging was started at the initial time, causing the initial transient to settle out somewhat more slowly; this aspect is removed by starting the averaging later. As seen from the figure, the tracking behavior of this on-line method is comparable to that of the full-scale estimator. Very similar performance is achievable, with less memory load, by including only every second or third estimate in the averaging or by making a parameter estimate only every second or third sample time. The undesirable initial deviation can be removed by delaying the time of the first estimate by two or three sample periods. If this particular application only required a parameter estimate every five sample periods, then this estimator is equivalent to that of the previous section with N = 5; since that estimator lended itself to a simpler implementation, it would then be employed.

Figure 6.1.16b presents the results of the same Monte Carlo simulation as performed on the full-scale and previous on-line estimators. Although this graph further substantiates the fact that this on-line method provides the same estimation precision as the full-scale estimator, it also demonstrates that the typical estimate is biased. The time-average of the bias is approximately 0.065 for the runs in the Monte Carlo analysis.

6.1.6 Inclusion of Weighted Least Squares Terms Only

Section 5.3.1 proposed including only weighted least squares type terms in the likelihood equations. This suggestion can be justified in part by looking at figure 6.1.4, which displays the ambiguity function components ψ_1 and ψ_2 described in section 5.3.1. As seen in the

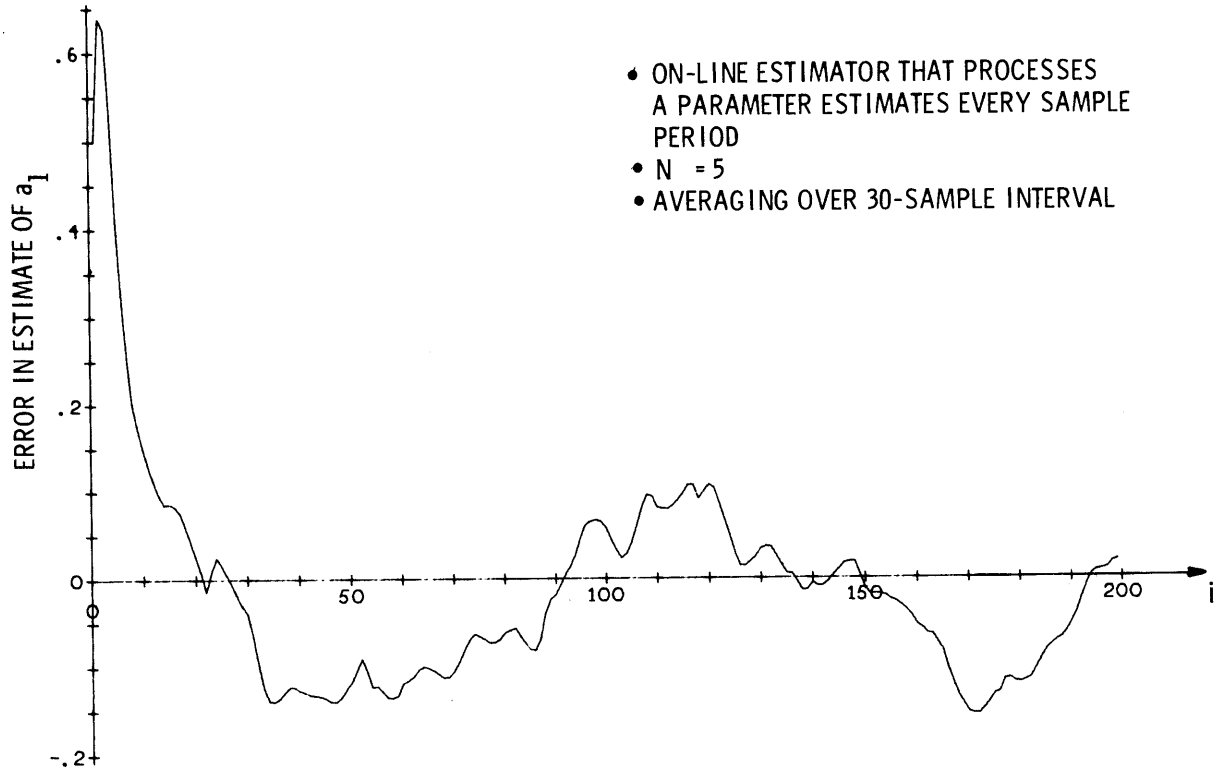


Figure 6.1.16a Parameter Estimate Error; On-Line Estimator with N = 5 plus Averaging

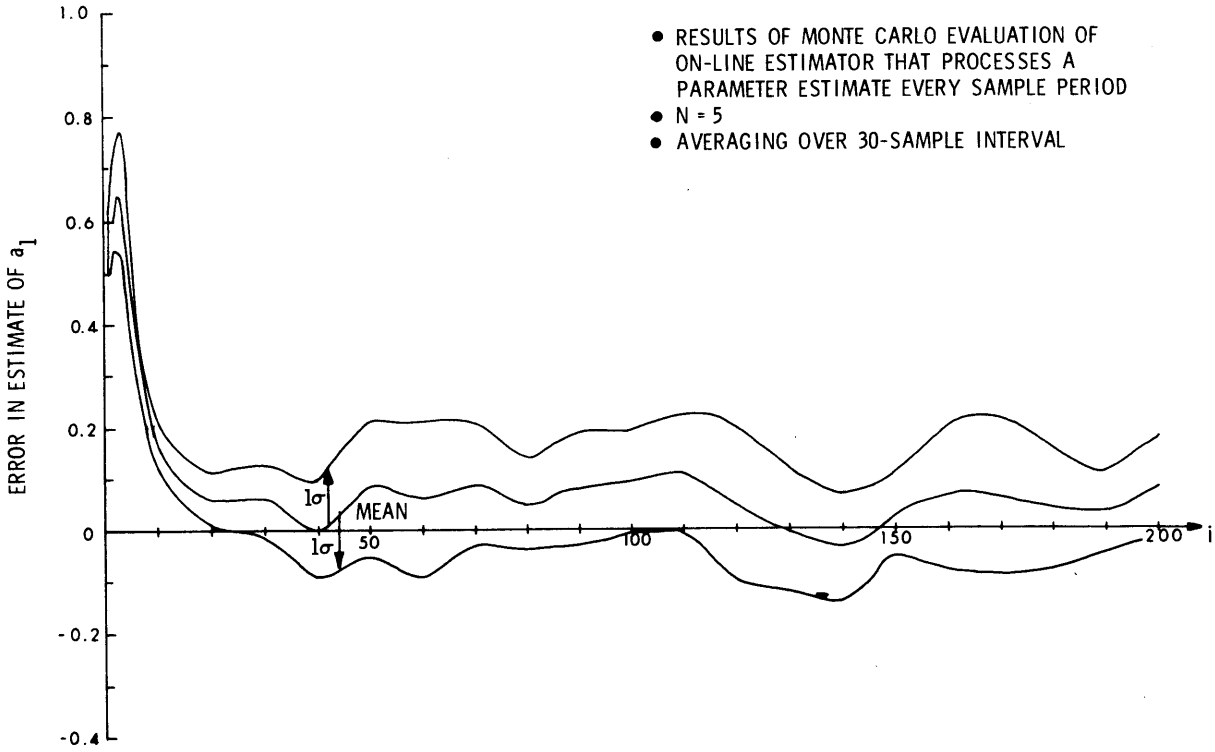


Figure 6.1.16b Mean ± 1σ Values of On-Line Parameter Estimate Errors

figure, ψ_1 is not as strong a function of the parameter a_1 as is ψ_2. Furthermore, when plotted on an expanded ordinate scale, ψ_1 reaches a maximum value at $a_1 = 0$, regardless of the true value of the parameter. This result has been verified by programming an estimator that included only terms corresponding to ψ_1, and the estimate converged to zero for all cases. Since the maximization of terms $(\underline{h}^T \underline{M}\, \underline{h} + R)^{-\frac{1}{2}}$ and $|\underline{P}|^{-\frac{1}{2}}$ does not depend upon the actual sequence of measurements, they cannot provide a valid means of setting model parameters to best agree with output from a real system. Thus, these terms contribute nothing at best, and a bias or destabilization at worst.

Implementations corresponding to both equations (5.3.6) and (5.3.7) have been examined and found to be successful. Because of its greater potential in reducing the computational load, the estimator of the latter form, employing the algorithm of equations (5.3.8) to (5.3.11), will be explicitly described. The on-line technique making a parameter estimate every period will serve as a basis of comparison; conclusions for this version are verified by the other on-line method and the full-scale estimator as well.

Figure 6.1.17 is directly comparable to the previous figure: plot a is the parameter estimate error for the same representative run, and plot b presents the results of a Monte Carlo simulation. Two recurring characteristics appear through a comparison of the graphs. First, the 1σ bands are wider in 6.1.17b for the initial stages, but become narrower than those in 6.1.16b as time progresses: there is greater variation in the transitory behavior, but less in the "steady state", from one simulation to another when only weighted least squares terms are used. More importantly, though, the bias is seen to be decreased in figure 6.1.17b. This aspect is made more evident by performing the same Monte Carlo runs on the two estimators

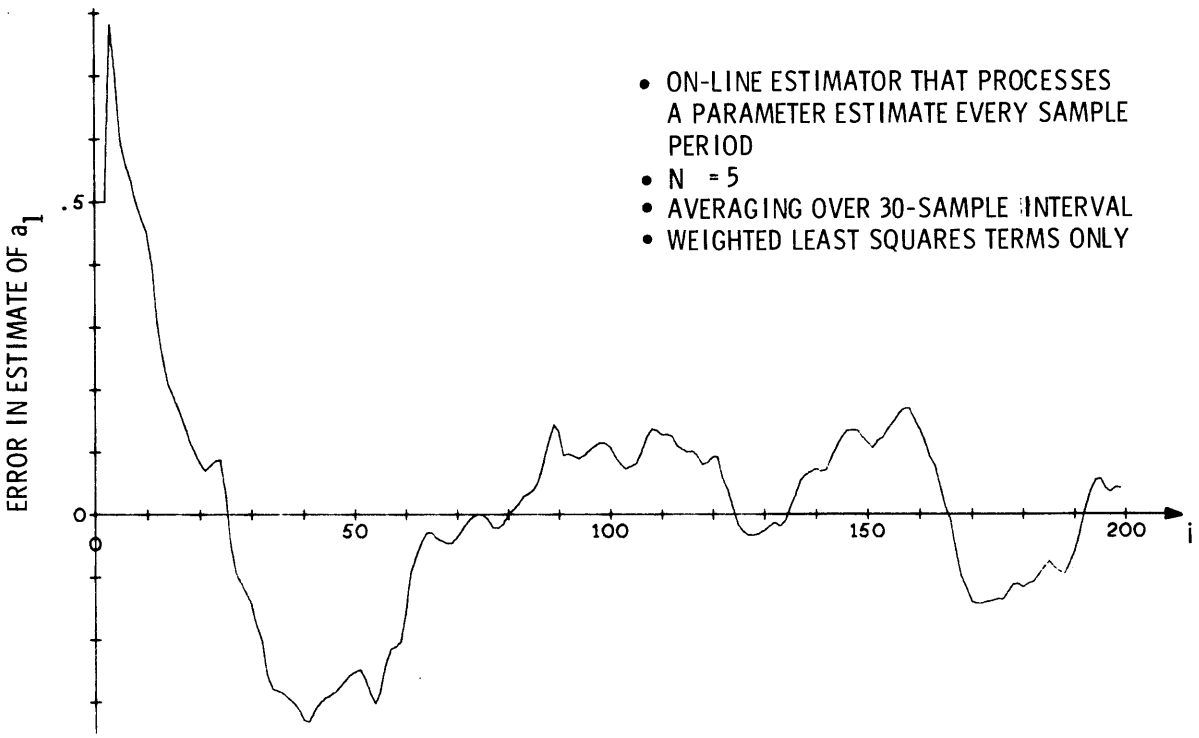


Figure 6.1.17a Parameter Estimate Error for On-Line Estimator Using Only WLS Terms

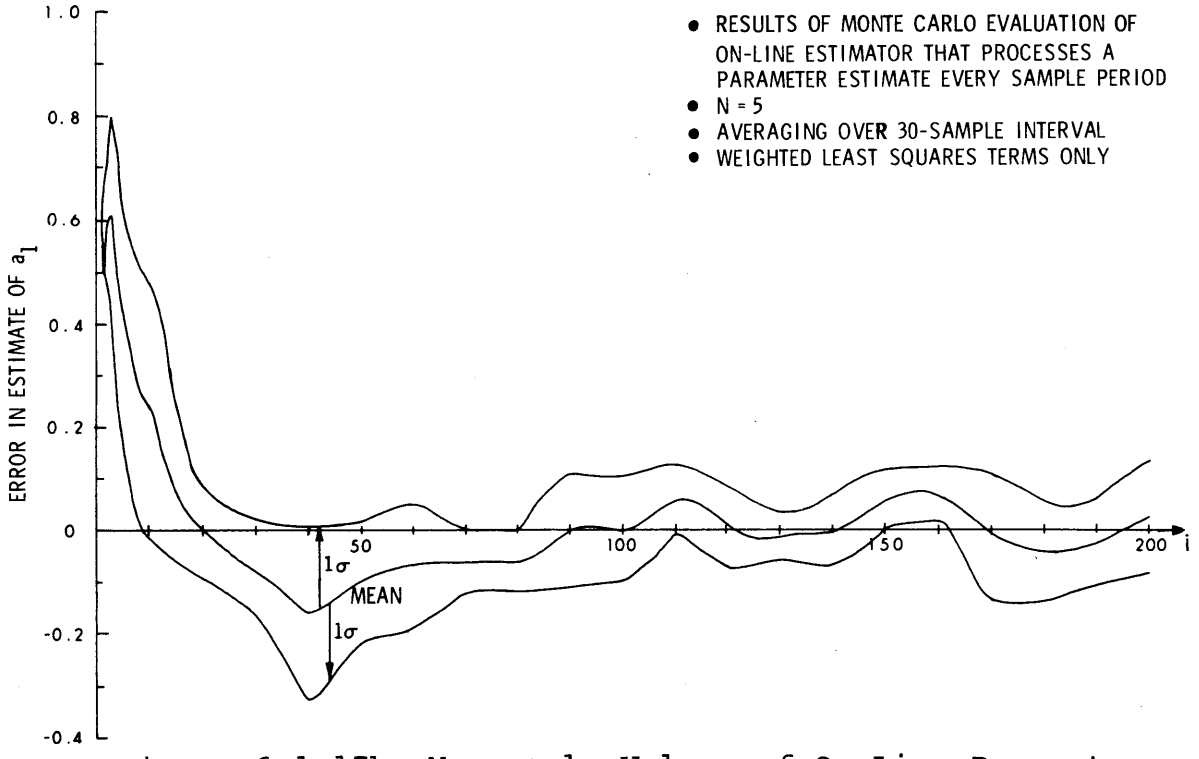


Figure 6.1.17b Mean ± 1σ Values of On-Line Parameter Estimate Errors; WLS Terms Only

(the former using all terms of the likelihood equation, and the latter including just the WLS terms), but averaging over the entire 200-step interval rather than the original 30-step interval; this is allowed since the parameter is actually time invariant. Figure 6.1.18 relates the results: whereas the first estimator maintains a biased estimate, the average bias steadily decreases in the second estimator.

As described in section 5.3, these benefits are gained along with a reduction of computation time. IBM 360/75 time for the parameter estimator using only weighted least squares terms was 2.68 seconds, as opposed to 6.40 for the estimator incorporating all terms, the major difference attributed to the removal of the matrix propagation of $\partial \underline{M}/\partial a$.

6.1.7 Precomputations

If the Monte Carlo simulations of the previous section are examined, the "inverse" of the conditional information matrix, 1/J, is seen to oscillate about a value of approximately 0.05 in the later stages, after the transient period. Therefore, the estimator was modified by equating 1/J to 0.05 for all time. The resulting error is plotted in figure 6.1.19. In relating this to figure 6.1.17a, the most notable change is the improved transient behavior: the initial rise in error is less; it decreases to 0.05 in approximately half the time, and yet overshoots zero by less. In certain other simulations that previously required a delay in making the first parameter estimate, the use of precomputed 1/J yielded adequate performance without the delay, since the initial values of 1/J are no longer disproportionately large. Substantiating results were obtained both with the other on-line implementation and full-scale version of the estimator.

Precomputed state estimator gains as a function of the parameter value also proved successful for this problem.

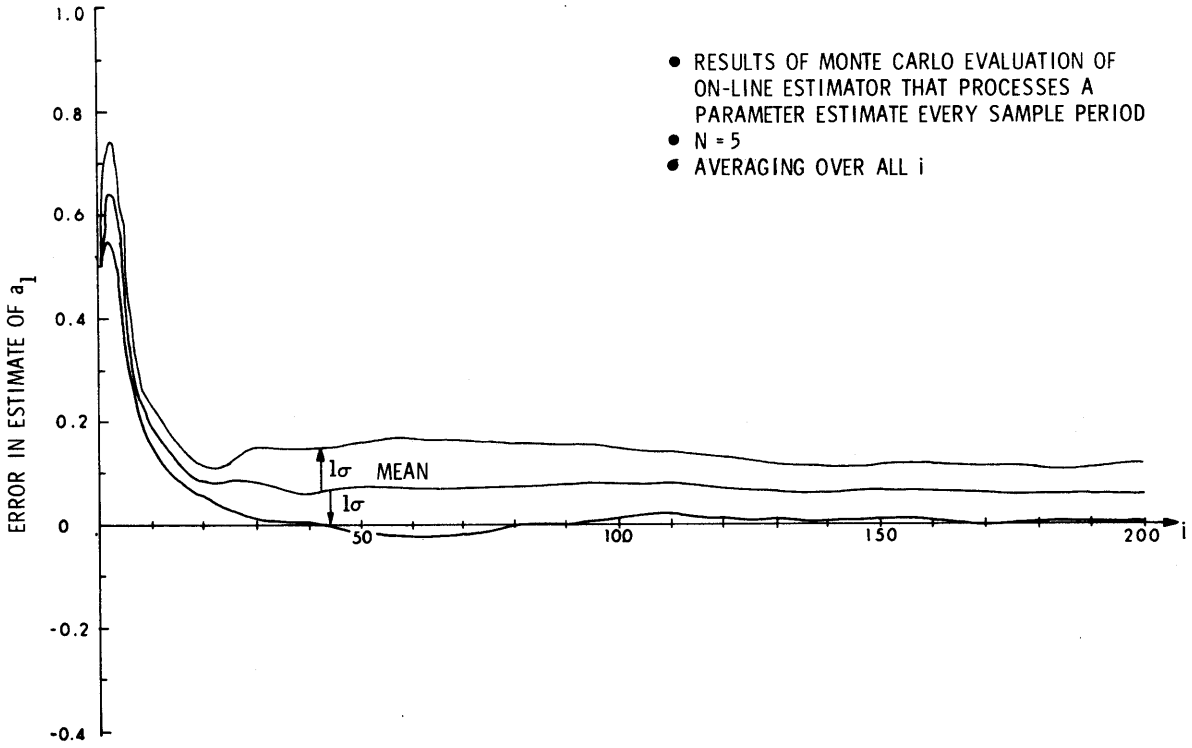


Figure 6.1.18a Mean ± 1σ Values of On-Line $\hat{a}$ Errors; Averaging over All i

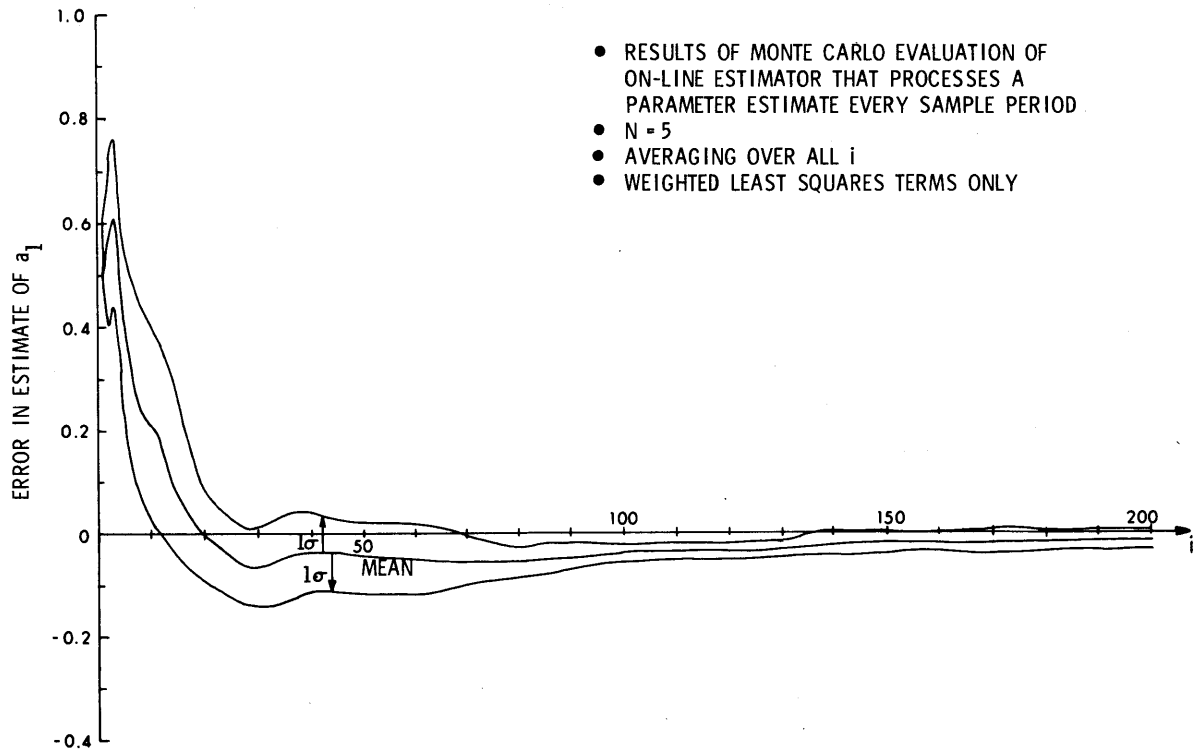


Figure 6.1.18b Mean ± 1σ Values of On-Line $\hat{a}$ Errors; WLS Terms Only; Averaging over All i

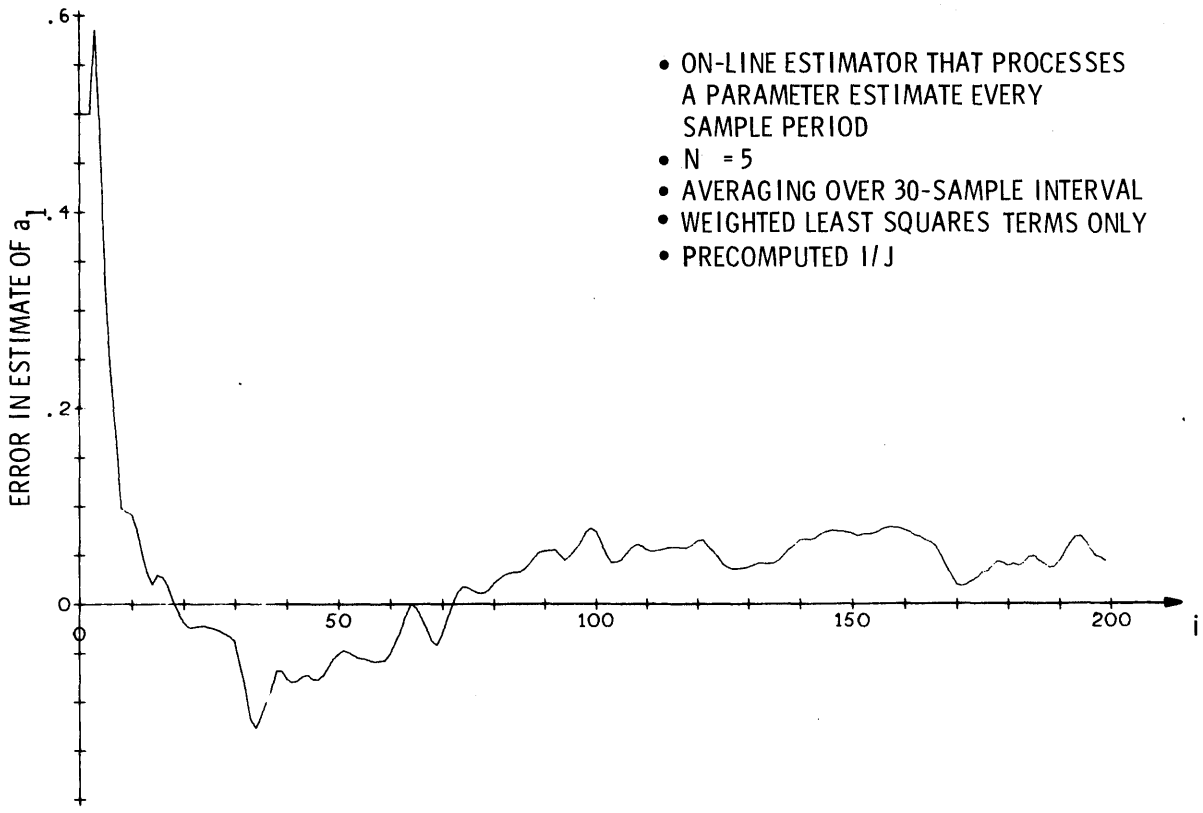


Figure 6.1.19 Parameter Estimate Error for Estimator Using Precomputed 1/J

The error analysis that generates the ambiguity functions also provides information about how fast these gains and associated covariance matrices reach steady state values. In this case, the steady state values were attained from the initial conditions to four significant figures in four sample periods over the entire range of parameter values. Consequently, only the steady state gains were computed as a function of the parameter. Least squares curve-fitting techniques yielded an approximation of

$$\left.\begin{aligned} k_1 &= 0.971392 - 0.000035a_1 + 0.000680a_1^2 \\ k_2 &= -0.9505a_1 \end{aligned}\right\} \quad (6.1.10)$$

that commits an error of, at worst, one part in 10^4 over a range of a_1 from -2 to 0. As expected from the rapid transients in $\underline{k}$ and the accuracy of the approximation, the estimation performance remained essentially the same. In fact, the simulation results agreed to two significant figures by the end of the second sample period, and were virtually indistinguishable thereafter.

6.1.8 Time Varying System Parameter

The ability of the estimators to track a time varying parameter value was examined by letting the true a_1 be a linear function of time. Figures 6.1.20a, b, and c portray representative errors produced by the full-scale estimator, the modified on-line estimator described in section 6.1.4, and the on-line technique described in section 6.1.5, respectively, for the case of a_1 doubling its magnitude in the 200 sample period interval. Figure 6.1.20b displays the slope of this growth, since the simulation computed the parameter estimate error every period while this estimate

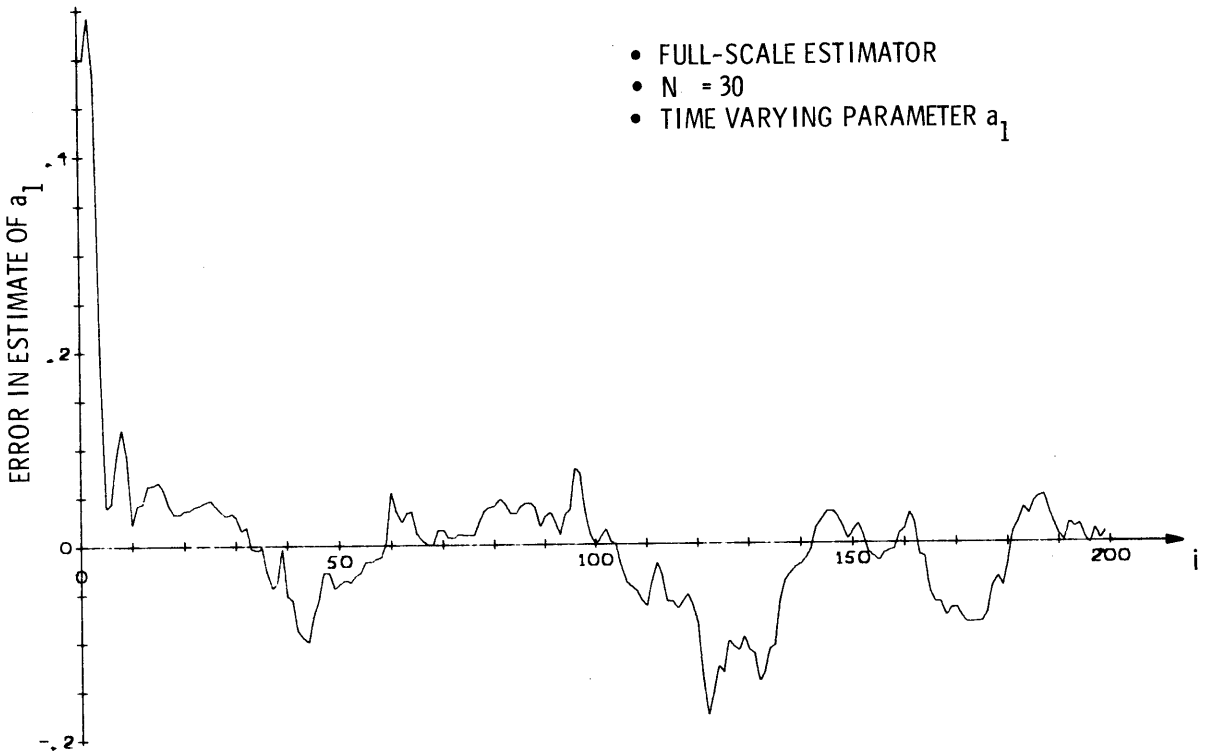


Figure 6.1.20a Full-Scale Parameter Estimate Error; Variable True Parameter

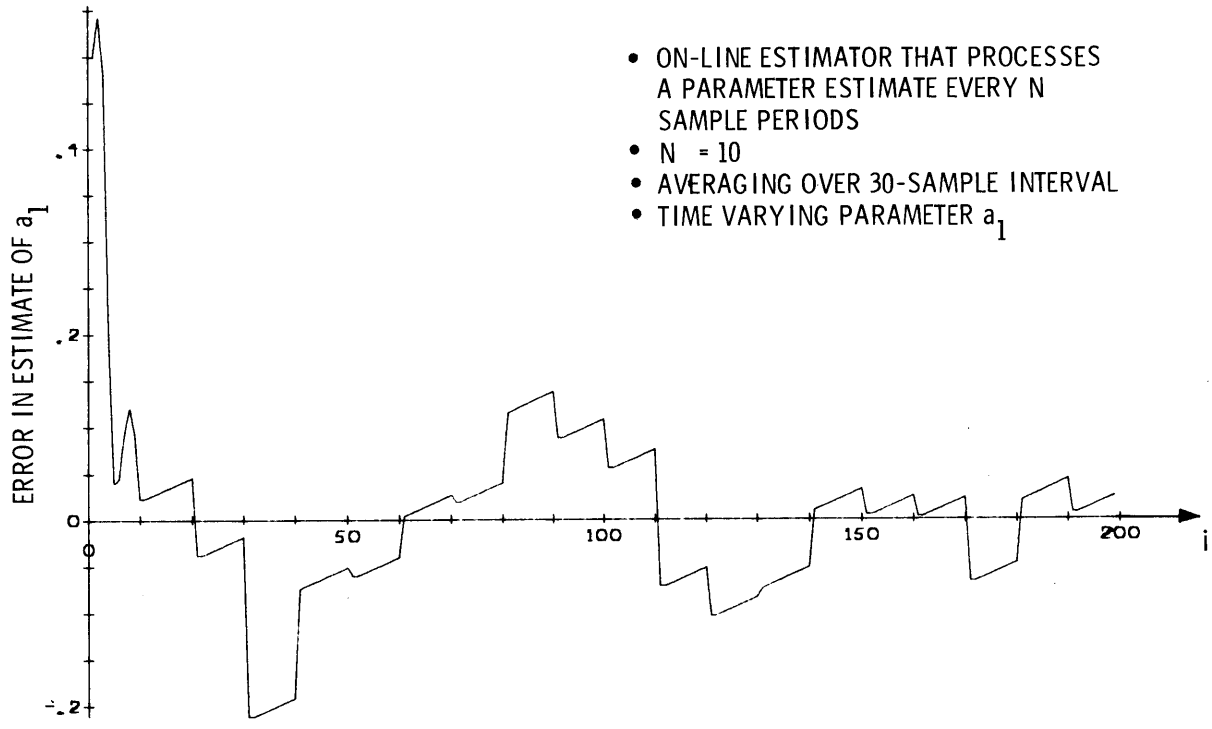


Figure 6.1.20b On-Line Parameter Estimate Error; Variable True Parameter

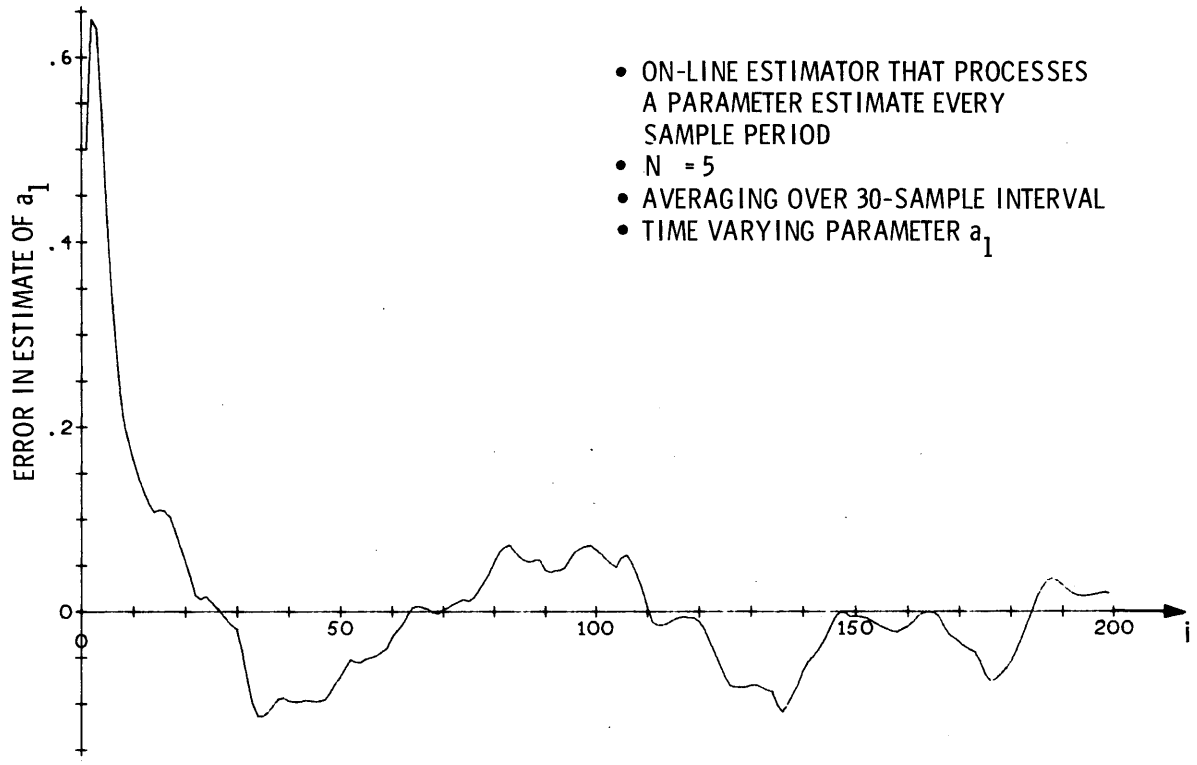


Figure 6.1.20c On-Line Parameter Estimate Error; Variable True Parameter

was only made every ten samples after i = 10. Thus, although the parameter actually changed by 0.075 every 30-step interval, i.e., 15% of its original value, the estimators were able to track its value very adequately with a model of the parameter as constant over 30 periods.

Tests were also performed with the parameter value purposely set to produce an unstable system. For these cases, the parameter estimate accuracy was greater than that obtained with a stable system. Physically, as the state variables grow unstably, the system modes become very visible, overwhelming the uncertainty introduced by the noises, and therefore the parameter estimation can yield more precise values.

6.2 Second Order System with Uncertain Parameters in $\underline{B}$

This example demonstrates the estimation of parameters in the $\underline{B}(i)$ matrix when the entire control history, u(j), is precomputed. Furthermore, it is based upon a continuous-time system model, thereby allowing the uncertain parameter to be other than a simple entry in the $\underline{B}(i)$ matrix. The physical problem that motivates this example, that of estimating an unknown acceleration, reveals that ingenuity in system modelling can extend the methods of this thesis beyond the scope of uncertain control input matrices or transition matrices in descriptions of real systems.

Suppose a radar operator desires to track an object subjected to a constant (or slowly varying) acceleration α, which he is trying to determine. The reference coordinates are set so that at $t = 0$, the body is approximately at position zero and velocity zero. Let the two states be x_1 = position and x_2 = velocity. In order to identify the value of α, it is assumed that an adequate model is

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha \end{bmatrix} 1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t) \tag{6.2.1}$$

Note that a constant input, u(i) = 1 for all i, has been introduced so that identification of α is reduced to the estimation of a parameter in the control input matrix (in this case, control input vector). A measurement of position is taken every τ seconds, and it is corrupted by the white Gaussian noise sequence v(i):

$$z(t_i) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} + v(t_i) \tag{6.2.2}$$

The equivalent discrete model of this system would be the following representation. It maintains the same state space coordinate directions, so $x_1(i)$ is the position at time t_i and $x_2(i)$ is the corresponding velocity:

$$\begin{bmatrix} x_1(i+1) \\ x_2(i+1) \end{bmatrix} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(i) \\ x_2(i) \end{bmatrix} + \begin{bmatrix} \frac{1}{2}\tau^2\alpha \\ \tau\alpha \end{bmatrix} 1 + \begin{bmatrix} \frac{1}{2}\tau^2 \\ \tau \end{bmatrix} w(i) \tag{6.2.3}$$

$$z(i) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(i) \\ x_2(i) \end{bmatrix} + v(i) \tag{6.2.4}$$

where now w(i) is modelled as a white Gaussian noise sequence. Let x_1 be measured in ft., x_2 in ft./sec., α and w in ft./sec.^{2}, and let u be unitless. Further assume that the sample period

τ is 0.2 seconds, causing equation (6.2.3) to become

$$\begin{bmatrix} x_1(i+1) \\ x_2(i+1) \end{bmatrix} = \begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(i) \\ x_2(i) \end{bmatrix} + \begin{bmatrix} 0.02\alpha \\ 0.2\alpha \end{bmatrix} 1 + \begin{bmatrix} 0.02 \\ 0.2 \end{bmatrix} w(i) \tag{6.2.5}$$

As mentioned previously, the operator attempts to set the reference coordinatesat zero for t = 0. Thus, the best initial state estimate would be:

$$\hat{\underline{x}}_o = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{6.2.6}$$

but since there is some uncertainty in this initialization, a reasonable value for $\underline{P}_o$ would be

$$\underline{P}_o = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \tag{6.2.7}$$

A reasonable initial guess for the value of α would be 32 ft./sec.^{2}, i.e., that the object is in free fall. Further assume that experience has shown that the types of bodies of interest are restricted to an 8-g dive and a 1.5-g climb, so that the bounding values on α are 250 and -50. Finally, the noise covariances are Q = 50 and R = 5.

For the simulations, a body with a time acceleration of 100 ft./sec.^{2} is to be tracked, starting from the initial conditions of $\underline{x}^T(0) = [1, 1]$.

Note that there are three possible "parameters" to estimate:

1) estimate $b_1 = 0.02\alpha$ and let $(\partial \underline{b}/\partial a)^T = [1, 10]$

2) estimate $b_2 = 0.2\alpha$ and let $(\partial \underline{b}/\partial a)^T = [0.1, 1]$

3) estimate the continuous-time parameter α itself, using
$(\partial \underline{b}/\partial a)^T = [0.02, 0.2]$

This serves to demonstrate the nonuniqueness of the particular choice of parameters to be estimated when they are functionally dependent.

6.2.1 Prediction of Performance

An ambiguity function analysis was performed to provide a measure of the best performance attainable from maximum likelihood techniques. Figure 6.2.1 is representative of the information acquired through such an analysis: it plots the Cramér-Rao minimum estimate error variance versus the data interval length at time instant 50 (10 seconds). The characteristic of strong dependence of the lower bound upon the data interval for small N, but substantially decreased dependence for larger N, is evident. From this figure, it would be advantageous to choose a value of N above the region of twenty, but computational requirements would dictate against a substantially larger choice because of the marginal benefit of the additional length.

6.2.2 Full-Scale Estimator

On the basis of the preceding analysis, the data interval length was established at 30 sample periods. Portrayed in figure 6.2.2 are the errors generated by the full-scale estimator, employing the approximations of equations (5.1.2) and (5.1.3), in a typical simulation. Plot a is the error in the position estimate, b the error in velocity, and c the error in estimating the second component of the $\underline{b}$ vector, $b_2 = 0.2\alpha$. The parameter error is very characteristic of the simulation results: an initial transient of about ten sample periods, followed by good tracking behavior but maintaining a small bias. Furthermore, comparisons demonstrate

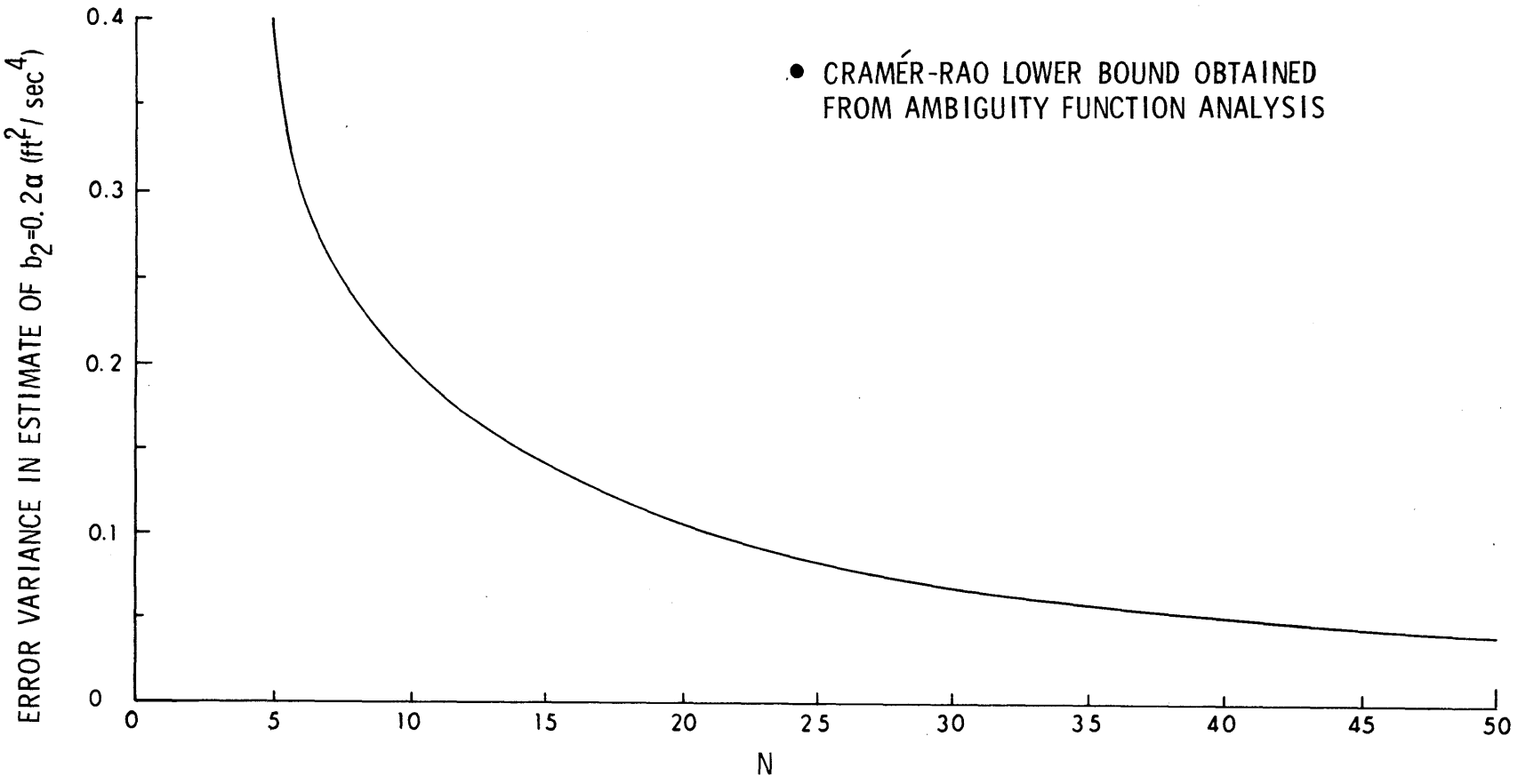


Figure 6.2.1 Lower Bound on Parameter Error Variance vs. N at i = 50

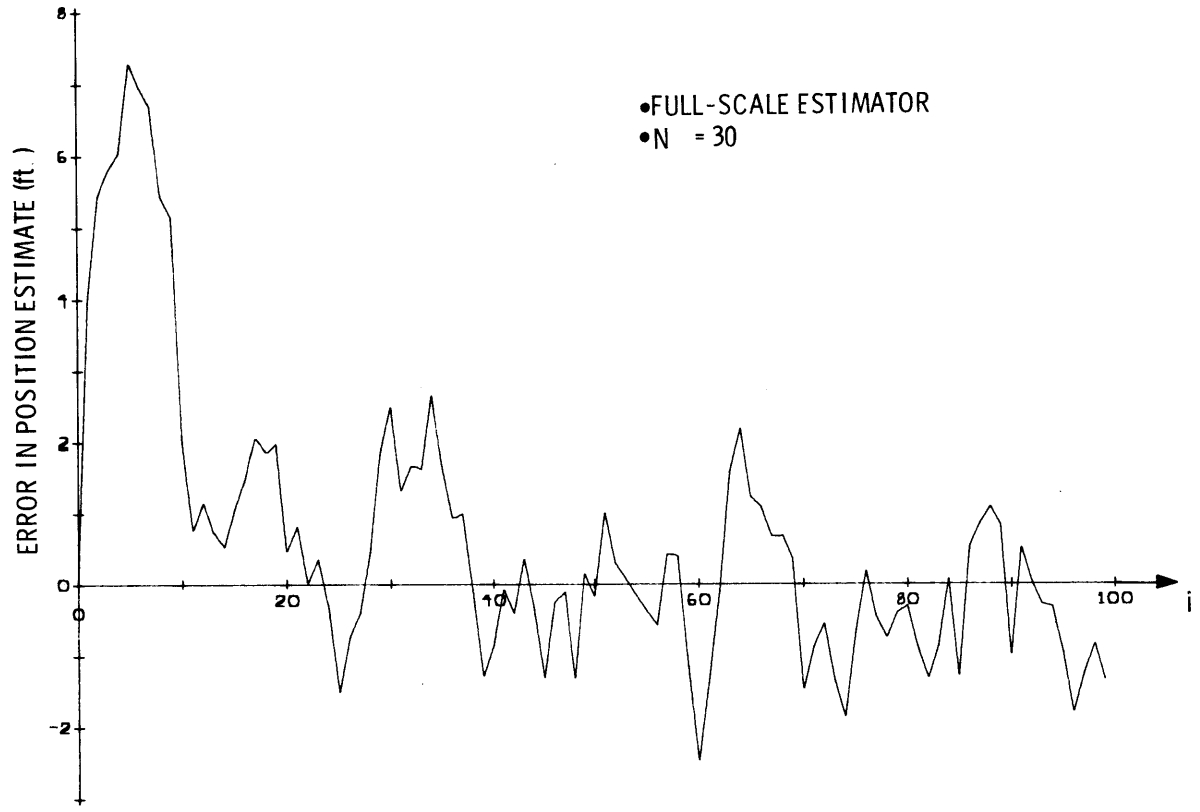


Figure 6.2.2a Error in Position Estimate; Full-Scale Estimator

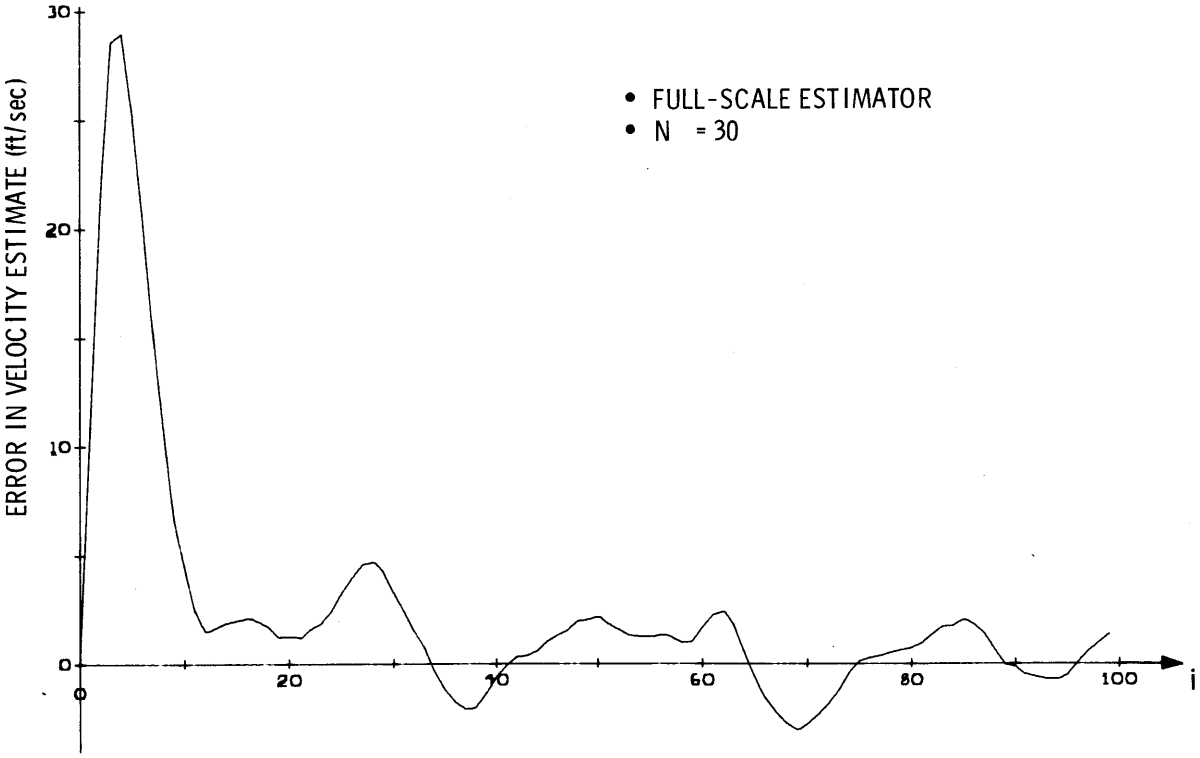


Figure 6.2.2b Error in Velocity Estimate; Full-Scale Estimator

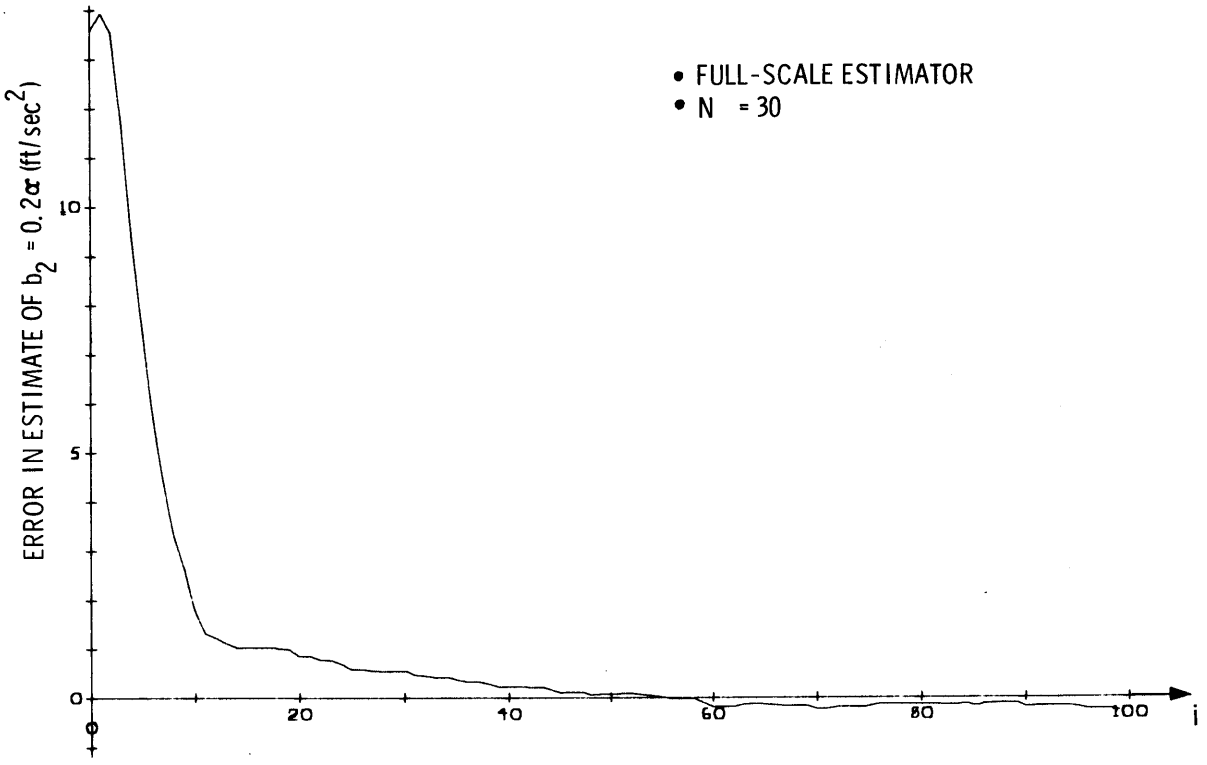


Figure 6.2.2c Error in Estimate of Parameter b_2; Full-Scale Estimator

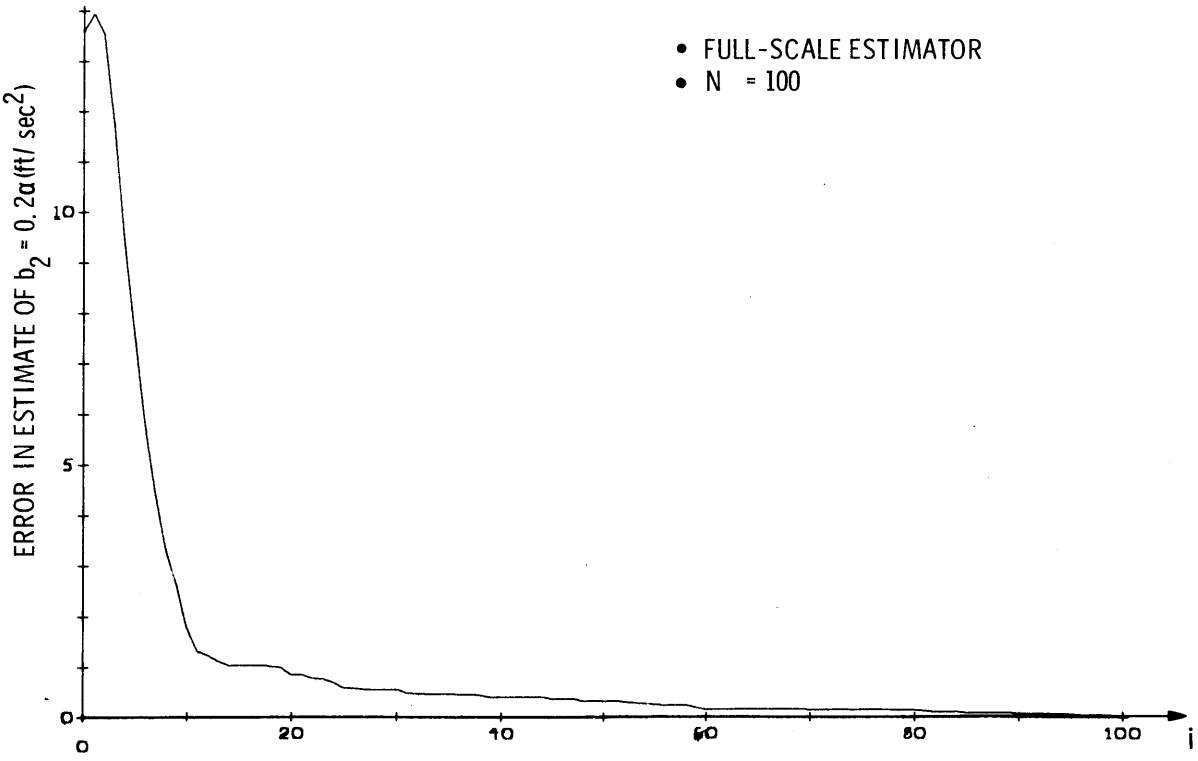


Figure 6.2.3 Parameter Estimate Error for N = 100

that as the parameter estimate converges to the true value, the state error trajectories duplicate the performance of a Kalman state estimator utilizing the true parameter value in its dynamic model.

The ambiguity function analysis predicts that increasing the data interval length above 30 will not appreciably affect the estimation accuracy. This is corroborated by figure 6.2.3, which depicts the parameter estimate error for the case of N = 100. Such a large N is not practical, but it serves to demonstrate a fact about asymptotic properties. Here the parameter value is actually constant for all time, so that an N equal to the entire time interval of interest is in fact a valid model. Under these circumstances, the bias in the parameter estimate tends to zero as the number of samples increases, whereas the bias remains for any fixed size of data interval.

Thus the ability of this estimation technique to handle uncertain parameters in $\underline{B}(i)$ has been demonstrated. Although the difference in the three "parameters" estimated in the simulations is trivial (a scale factor), the equivalent precision of the resulting estimators indicates that the parameters need not be entries in $\underline{\Phi}(i+1,i)$ or $\underline{B}(i)$. Section 6.3 will consider a case in which the estimator can be formulated more conveniently when the parameter is chosen to be the square of an uncertain frequency, rather than the frequency itself, or a particular entry of the associated $\underline{\Phi}$ matrix. Thus, there is some flexibility in the choice of uncertain parameters to use, and those that yield the least complex implementation (as, a set of independent parameters that yield the most nearly linear functional evaluations for such terms as $\partial\underline{\Phi}/\partial a_\ell$) are generally preferable.

Note that, since $\partial\underline{M}/\partial a_\ell$ is zero for a_ℓ equal to a parameter in $\underline{B}$, only weighted least squares type terms appear in the propagations that generate the parameter estimate.

6.2.3 On-Line Techniques

Figure 6.2.4 is a graph of the parameter estimate error produced by the modified on-line technique described in section 6.1.4, using the same noise sequences and initial conditions as the preceding full-scale estimator simulation. Here the estimator again uses an N of 10, combined with averaging over the three most recent estimates, as well as a more rapid estimation frequency for the first ten samples. Comparing this to figure 6.2.2c demonstrates the feasibility of this on-line conceptualization as a satisfactory replacement for the full-scale implementation. Further simulations have substantiated this conclusion.

For the same simulation noise sequences, the on-line technique described in section 6.1.5, using an N of 5 and an averaging length of 30 periods, yields a parameter error depicted in figure 6.2.5a. Again, the initial transient is slow because the early erroneous estimates are averaged into the later evaluations. This can effectively be alleviated by initiating the averaging at a later time: simulations can reveal the mean time for the unaveraged estimate transient to decay suitably; at this point, averaging would be started to increase the consistency of the estimated value from sample to sample. Figure 6.2.5b portrays the error that results from delaying the averaging until the fifth sample instant: the transient effects are removed considerably more rapidly and an accurate value is maintained thereafter. The slope discontinuities of the two plots are caused by the initial estimates being removed from the averaging interval for the first time.

Because of its lower computational requirements, the first of these on-line methods would be preferable in practice, unless more frequent parameter estimation is desired.

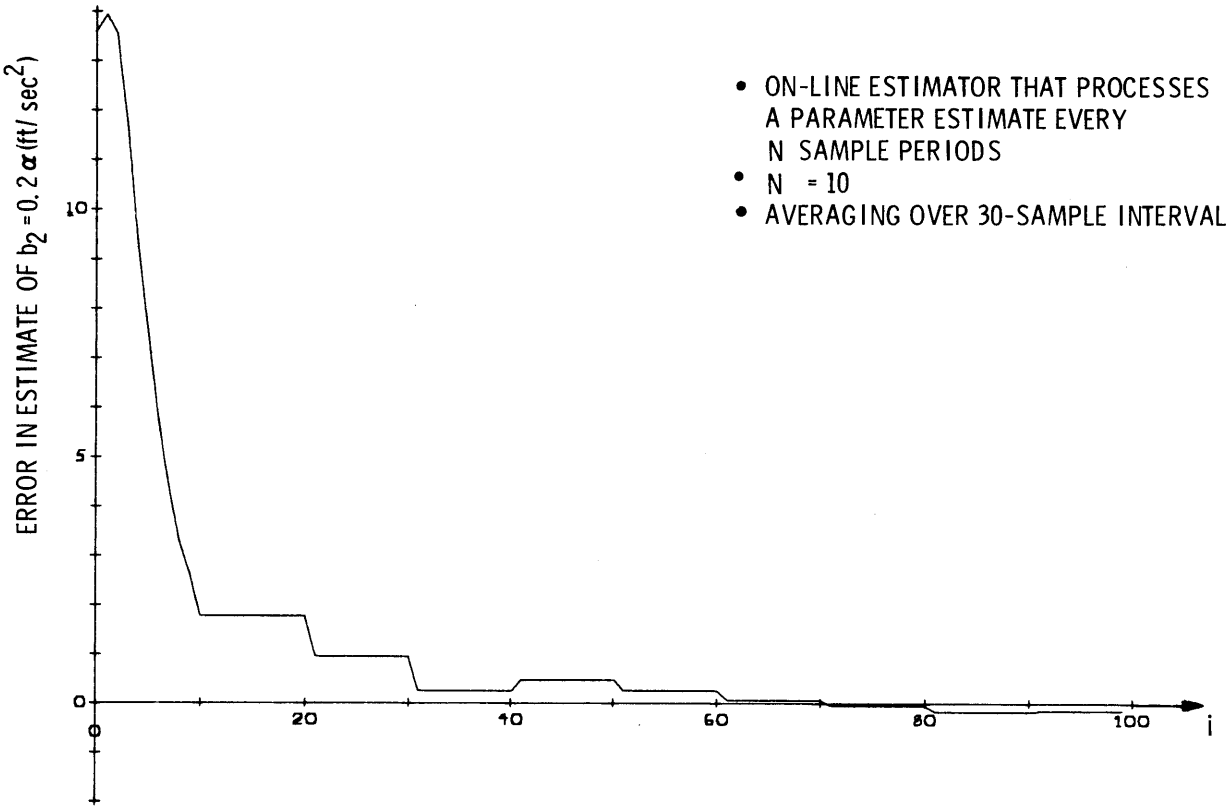


Figure 6.2.4 Parameter Estimate Error; On-Line Estimator

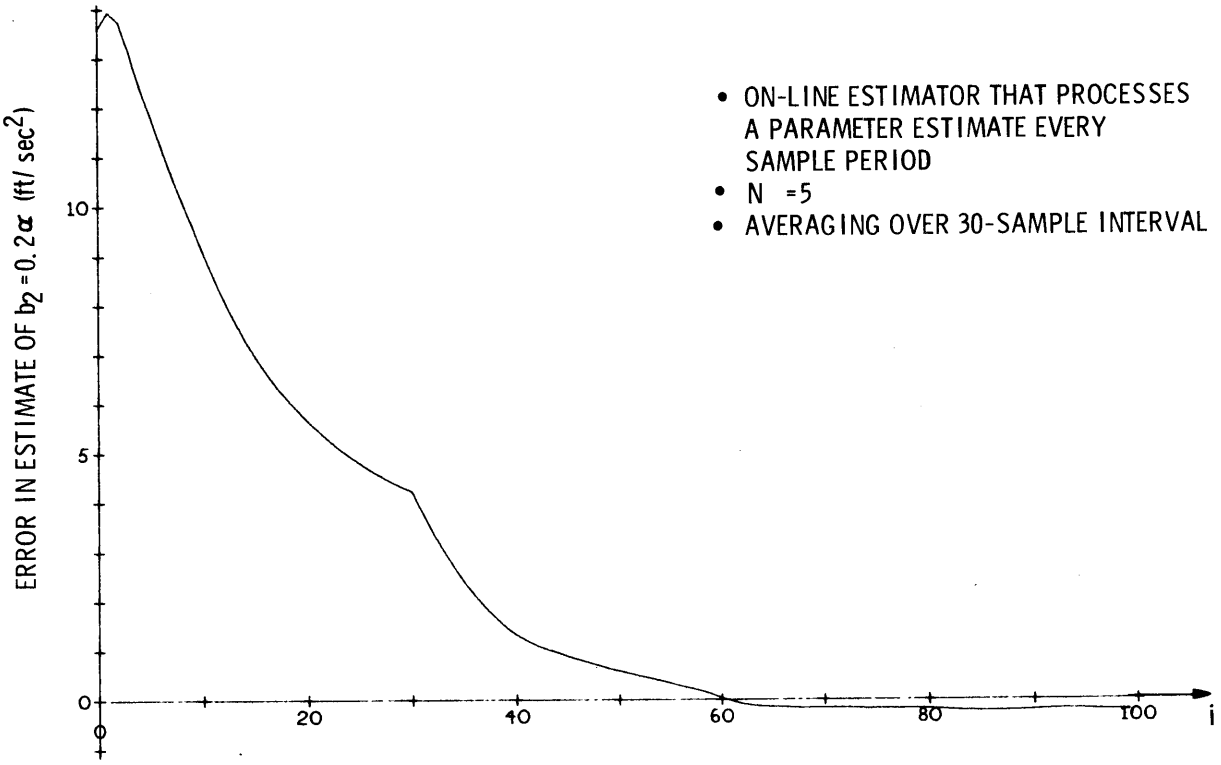


Figure 6.2.5a Parameter Estimate Error; On-Line Estimator

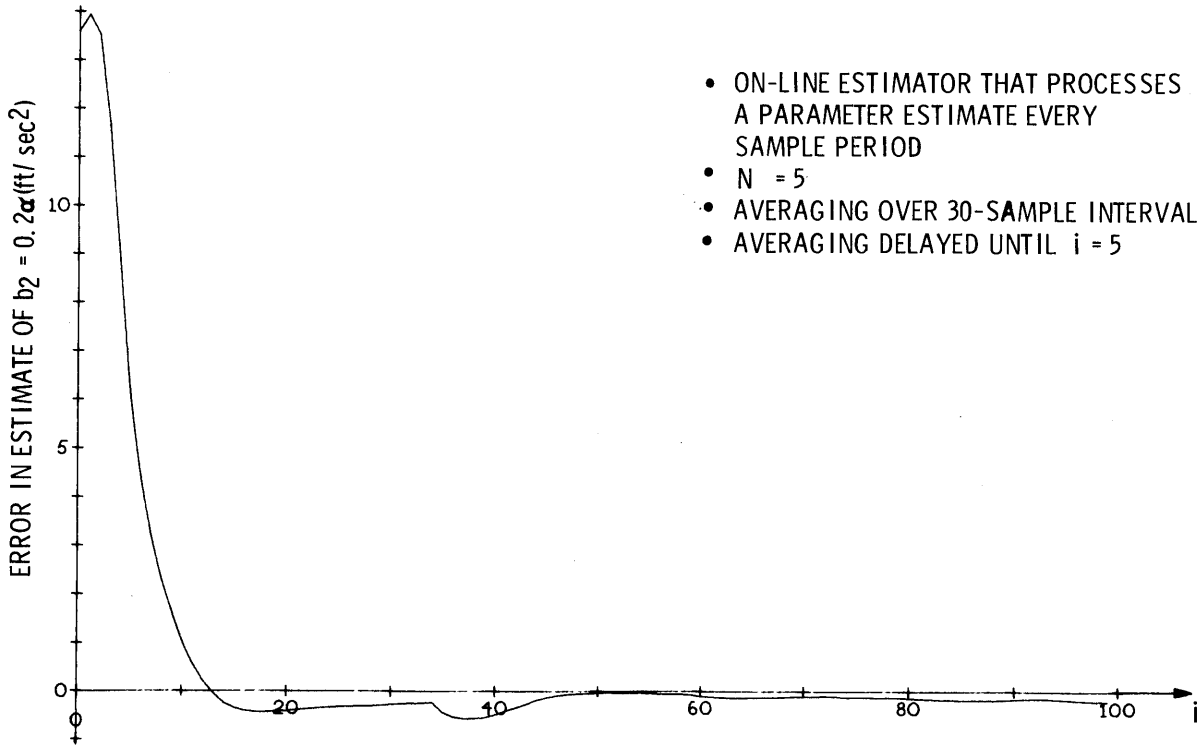


Figure 6.2.5b Parameter Estimate Error When Averaging is Delayed

6.3 Apollo Thrust Vector Controller

This section considers the problem of thrust vector control of the docked configuration of the Apollo CSM (Command and Service Module) and LM (Lunar Module). The particular choice of this vehicle was based upon availability of information, but the results are applicable in a wider context: vehicle bending frequencies will always be somewhat uncertain, and the ensuing problems will probably become more critical as the spacecraft become more complexly shaped, as Skylab, or elongated, as the proposed designs for nuclear vehicles, thereby causing bending mode effects to become more dominant.

During an Apollo mission, trajectory corrections are performed by igniting the main engine of the Service Module. Initial alignment is accomplished through use of small attitude control jets, but once the main engine is ignited, this desired attitude is maintained by gimballing the main engine nozzle via onboard computer-generated commands. Not only do these commands steer the vehicle to desired cutoff conditions, but also counteract any rigid body rotations or bending mode motions. This thrust vector controlling is necessary so as not to jeopardize the structural integrity of the docking tunnel joining the two vehicles.

Linear optimal control theory would propose cascading a linear optimal state estimator with a linear optimal controller to achieve an overall optimum stochastic controller design. Measurements from the IMU (Inertial Measurement Unit) would serve as inputs to this system, and the optimal commanded control would be the output. However, when analyzed in the frequency domain, it can be seen that this optimal controller utilizes cancellation compensation (see Widnall (1968)). This would be entirely suitable if all of the parameters in the state dynamics equations were known exactly. Unfortunately, the frequency of the bending mode can take on a range of different values, due to such factors as the amount of fuel in

the propellent tanks. Moreover, the engineering approximations used to determine the frequency result in some uncertainty in its assumed value; and in fact the actual frequencies obtained in the first space flights differed substantially from those anticipated by the various analyses. Not only does this uncertainty cause suboptimal control, but it can, with a frequency variation of only a few radians per second, cause instability.

This example will be developed for the pitch plane only, and it will concentrate upon the dynamics of the rigid body vehicle and the most significant bending mode at the docking tunnel connecting the CSM and LM. The coupling of fundamental yaw plane and torsional mode effects can be considered a perturbation upon this planar analysis; there are higher modes as well, but they are substantially higher in frequency and lower in energy content. In order to simplify the dynamic model, the effects of fuel slosh and center-of-gravity shift are assumed sufficiently small to be neglected. These assumptions are made so as to concentrate attention upon the most significant aspects of the dynamics, keeping the results from being obscured by a combination of smaller, less relevant, effects. Such assumptions might not be made in a complete system analysis intended for actual software design, since for example the stability of the fuel slosh effects must be assured, but they are employed to expedite this feasibility study.

Under the previous assumptions, the fundamental differential equations governing the state variables can be developed in the following manner. The rigid body motion can be described as a second order differential equation, or the two equivalent first order relations:

$$\dot{\omega}(t) = -\frac{T(L\sigma_T + d_T)}{I}\, q(t) + \frac{TL}{I}\,\left[\delta(t)+w(t)\right] \qquad (6.3.1)$$

$$\dot{\theta}(t) = \omega(t) \qquad (6.3.2)$$

where the state variables $\omega(t)$ and $\theta(t)$ are the rigid body motion angular velocity and attitude, relative to the inertial reference direction. The other variables are $q(t)$, the generalized bending coordinate; $\delta(t)$, the deterministic part of the main engine nozzle angle relative to the Service Module; and $w(t)$, a white noise superimposed upon the deterministic $\delta(t)$. The constants of these equations are:

- T = thrust of engine; assumed to be 22,000 lbs.
- I = pitch moment of inertia of the rigid vehicle; 370,000 slug-ft.^{2}
- L = distance between the vehicle center of mass and the engine; 19 ft.
- σ_T = the slope of the bending mode at the engine station (per unit displacement of the generalized coordinate); -0.13 radian/ft.
- d_T = the displacement of the bending mode at the engine station (per unit of displacement of the generalized coordinate); 1.1 ft./ft.

The bending mode dynamics are described by the state variables $v_b(t)$ and $q(t)$, the velocity and position of the generalized bending coordinate, which satisfy:

$$\dot{v}_b(t) = -\omega_b^2\, q(t) - a\, d_T\, [\delta(t) + w(t)] \qquad (6.3.3)$$

$$\dot{q}(t) = v_b(t) \qquad (6.3.4)$$

In equation (6.3.3), the constant a is the magnitude of the vehicle acceleration due to thrusting, assumed to be 10 ft./sec.^{2}. The quantity ω_b is the natural frequency of the bending mode, whose value is uncertain. Note that these equations imply zero damping, which is very nearly the case for the actual vehicle in the space environment.

Finally, the servo mechanism on the engine gimbal can be adequately modelled as a first order lag:

$$\dot{\delta}(t) = -\frac{1}{\tau}\,\delta(t) + \frac{1}{\tau}\,\delta_{com}(t) \qquad (6.3.5)$$

where δ_{com} is the commanded value of the gimbal angle (the output of the controller) and τ is the lag time constant with which the nozzle angle follows the command, with a value of 0.1 sec. In reality, the actuator drive system has a peak rate limit of 0.1 radian per second, but this is accounted for in the formulation of the cost function of the optimal controller so as to keep the system description linear.

The dynamic driving noise, w(t), is a vibrational disturbance at the bottom end of the Service Module, due to the rocket engine. It is assumed to enter the equations as a white noise random thrust vector angle. As a result, the true nozzle angle is divided into the completely deterministic $\delta(t)$ and the statistically random portion, w(t), with mean value of zero and low frequency spectral density of 0.0004 radians^2 per cycle-per-second. Such a disturbance can be expected to cause a lateral velocity of about 2 ft./sec. in 100 seconds of engine burn time.

Thus, for a given value of the bending frequency ω_b, the vehicle dynamics can be described by the five-dimensional vector equation

$$\frac{d}{dt}\begin{bmatrix}\omega(t)\\ \theta(t)\\ v_b(t)\\ q(t)\\ \delta(t)\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0.0815 & 1.13\\ 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -\omega_b^2 & -11\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & -10\end{bmatrix}\begin{bmatrix}\omega(t)\\ \theta(t)\\ v_b(t)\\ q(t)\\ \delta(t)\end{bmatrix} + \begin{bmatrix}0\\ 0\\ 0\\ 0\\ 10\end{bmatrix}\delta_{com}(t) + \begin{bmatrix}1.13\\ 0\\ -11\\ 0\\ 0\end{bmatrix} w(t) \qquad (6.3.6)$$

Before the main engine ignition, the controlled member of the Inertial Measurement Unit can be commanded to a convenient orientation, but during the burn, the commanded angular velocity is zero. Therefore, except for interfering torques, transmission of base motion through the gimbals, and gyro drift, the controlled member is actually fixed with respect to inertial space. A sampled readout from the IMU will then give the angular orientation of the outside case (fixed to the vehicle) with respect to the stable member. By processing the gimbal angle readouts through a suitable transformation, it is possible to provide a measurement of the angular orientation of the spacecraft, relative to inertial space, in the pitch plane of the vehicle. This is actually a measurement of the sum of three effects:

$$z(t_i) = \theta(t_i) + \sigma_a \, q(t_i) + v(t_i) \qquad (6.3.7)$$

where $\theta(t_i)$ is the rigid body angular attitude; $q(t_i)$ is the generalized bending coordinate and is multiplied by σ_a, the slope of the bending mode at the IMU station per unit displacement of the generalized coordinate (with value of -0.13 radian/ft.); and $v(t_i)$ is a random measurement noise, used to model the interfering effects described previously and also the quantization of the readout. If the quantization error is assumed to be uniformly distributed over the quantization interval of 0.0002 radians associated with the analog-to-digital conversion, then the measurement error statistics are a mean of zero and a variance of $1/12\ (0.0002)^2$ radians^2. Since this is the major contribution to measurement uncertainty, $v(t_i)$ is modelled as a white Gaussian sequence with the above mean and variance. (The variance could be increased to account for the interfering disturbances, the influence of higher order bending modes upon the measurements, and the like.) Thus, the measurement made available every 0.1 seconds is:

$$z(t_i) = \begin{bmatrix} 0 & 1 & 0 & -0.13 & 0 \end{bmatrix} \begin{bmatrix} \omega(t_i) \\ \theta(t_i) \\ v_b(t_i) \\ q(t_i) \\ \delta(t_i) \end{bmatrix} + v(t_i) \tag{6.3.8}$$

Equations (6.3.6) and (6.3.8) provide the state space description of the system under investigation. It is desired to develop a discrete-time estimator of the five state variables and the uncertain parameter, (ω_b^2). This would then be cascaded with a linear controller of the form

$$\delta_{com}(i) = - \underline{c}(i;t_f,\omega_b^2)^T \hat{\underline{x}}(i) \tag{6.3.9}$$

to provide a suboptimal stochastic controller, but with better performance than one which did not attempt estimation of the uncertain parameter. This value of $\delta_{com}(i)$ would then be held constant over the interval $t_i \leq t < t_{i+1}$. In equation (6.3.9) the notation is meant to display the fact that the controller gains are a function of the time the burn is terminated, t_f, and the value of ω_b^2. The gains can be computed using an assumed value of ω_b^2, or they can be expressed as a function of ω_b^2 and evaluated using the most recent estimate of this uncertain parameter.

Although this state space description is valid for developing the controller (the corresponding deterministic system is observable and controllable), there is a problem associated with the inclusion of δ in the estimator state space. From figure 6.3.1, it is evident that noise does not enter into this state variable. Thus, if the initial condition $\delta(t_o)$ is known exactly, then $\delta(t)$ will be known perfectly for all time, resulting in singular covariance matrices. Even if there were some small initial uncertainty (it would have to be very small since the angle could be

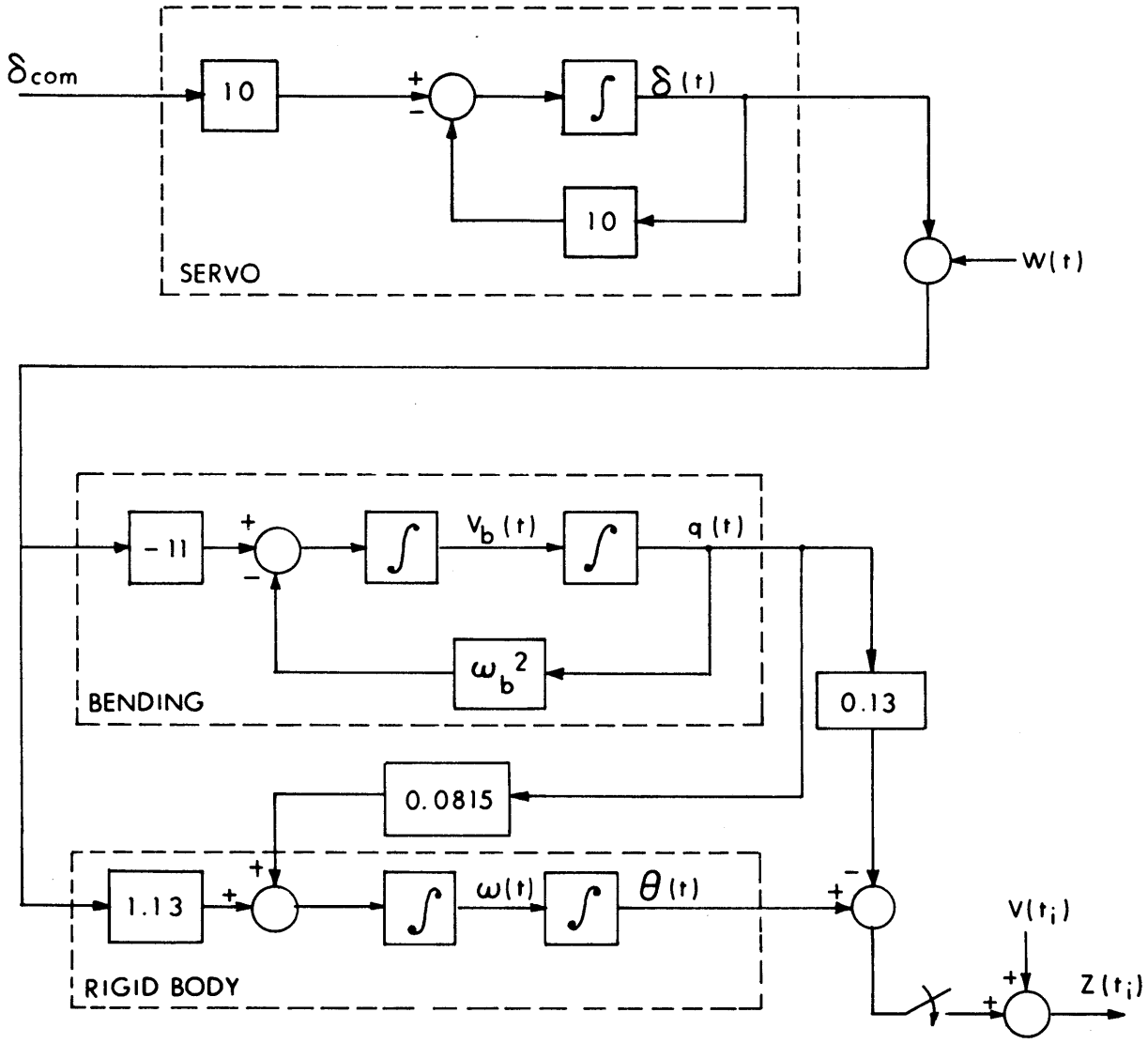


Figure 6.3.1 Thrust Vector Control Problem Dynamics

measured directly, and previous to the burn, the servo would be at rest), it would converge to zero because of the absence of noise to drive the stable first order lag which models the servo. From a mathematical point of view, the five dimensional state space representation is uncontrollable with respect to the points of entry of the dynamic driving noise w. The result is that the fifth row and column of the covariance matrices $\underline{M}(j)$ and $\underline{P}(j)$ would be (or converge to) zeroes for all time. This would preclude calculation of $\underline{P}^{-1}(i)$ as required in equations (3.4.38) and (3.4.39) of the estimator algorithm.

Thus, whereas the value of the five states must be maintained at all times, only the first four necessitate actual estimation (if $\delta(t_o)$ is known exactly), the fifth being calculated deterministically. In other words, the state estimator gain corresponding to δ would always be zero, and the incorporation of a measurement would not affect the "estimate" of δ. Moreover, the fifth element of $\partial\hat{\underline{x}}(j)/\partial a$ and $\partial\bar{\underline{x}}(j)/\partial a$, and the fifth rows and columns of $\underline{M}(j)$, $\underline{P}(j)$, $\partial\underline{M}(j)/\partial a$, and $\partial\underline{P}(j)/\partial a$ would be zero for all time, and need not be calculated. For actual implementation, the time propagation relation for $\bar{\underline{x}}$, equation (3.4.14), would be a five dimensional vector, but all other vectors and matrices in the estimator would be four dimensional.

It is desired to obtain a discrete-time system model in the form of

$$\underline{x}(i+1) = \underline{\Phi}(i+1,i)\underline{x}(i) + \underline{B}(i)\underline{u}(i) + \underline{G}(i)\underline{w}(i) \qquad (6.3.10)$$

which is equivalent to the continuous-time model, equation (6.3.6). Since the control function $\delta_{com}(t)$ is constant over a sample period, the methods of Appendix B can be applied conveniently to perform the desired conversion. Furthermore, the model is time invariant, so the required

matrices need only be calculated once for all time as a function of the uncertain parameter ω_b^2.

The initial conditions for the state estimator are determined as follows. Since positive and negative values of ω, θ, v_b, and q are equally likely, $\underline{\hat{x}}_o$ is assumed to be $\underline{0}$. To establish the covariance values, the rigid body and bending modes can be assumed to contain nearly equal energy. Combined with typical resolution capabilities and the fact that the bending coordinate q is to be kept below a value of 0.1 ft., this assumption leads to a diagonal $\underline{P}_o$ with elements

$$\left.\begin{aligned} P_{o11} &= (1 \text{ degree/sec.})^2 \\ P_{o22} &= (1 \text{ degree})^2 \\ P_{o33} &= (1 \text{ ft./sec.})^2 \\ P_{o44} &= (0.1 \text{ ft.})^2 \end{aligned}\right\} \quad (6.3.11)$$

The best apriori estimate of the parameter value is $a = \omega_b^2 = 100 \text{ radians}^2/\text{sec.}^2$.

For simulation purposes, the following initial conditions will be used for the actual system:

$$\left.\begin{aligned} \omega(0) &= 0.8 \text{ deg./sec.} \\ \theta(0) &= 0.8 \text{ deg.} \\ v_b(0) &= 0.7 \text{ ft./sec.} \\ q(0) &= 0.07 \text{ ft.} \end{aligned}\right\} \quad (6.3.12)$$

and the true value of ω_b^2 will be set at 150.

6.3.1 The Need for Parameter Estimation

The need for parameter estimation in this application was demonstrated by attempting to use a Kalman state estimator cascaded with a linear optimal controller to provide the desired thrust vector controlling. Figure 6.3.2 por-

trays typical state estimate error trajectories produced by an estimator and controller that assume $\omega_b^2 = 100$, while the true system is simulated with $\omega_b^2 = 150$. These are approximately ten times as large as the corresponding rigid body state estimate errors, and about fifteen times the bending mode errors, produced when the correct value of ω_b^2 is employed. Furthermore, these error magnitudes are very significant compared to the actual state RMS values, which are on the order of 0.006, 0.015, 0.7, and 0.06 respectively. The accuracy problem is magnified as the true ω_b^2 value increases: above approximately 250, the state estimate errors grow unstably. As a result, the controller applies inappropriate commands, causing the overall closed loop system to be unstable.

An ambiguity function analysis predicted that the value of ω_b^2 could be estimated simultaneously with the state variables. One result of this analysis is presented in figure 6.3.3: the Cramér-Rao lower bound on parameter estimate error variance as a function of the size of the data interval, N, evaluated at time instant 50. This graph suggests that N should be chosen to be greater than about 20, but that much larger values would not yield significantly better performance.

6.3.2 Full-Scale Estimator Performance

In order to implement the full-scale estimator, the elements of $\underline{\Phi}$, $\underline{B}$, $\underline{G}\underline{Q}\underline{G}^T$, and $\underline{C}$ (the controller gains) were precomputed as a function of the parameter ω_b^2, and then approximate linear functions were curve-fitted to them. Subsequent to each parameter estimate, these elements were reevaluated; in real applications, the reevaluations could be disregarded if the parameter estimate had not changed by more than a prespecified amount since the last reevaluation had been processed. Figure 6.3.4a presents a representative parameter estimate error trajectory as generated

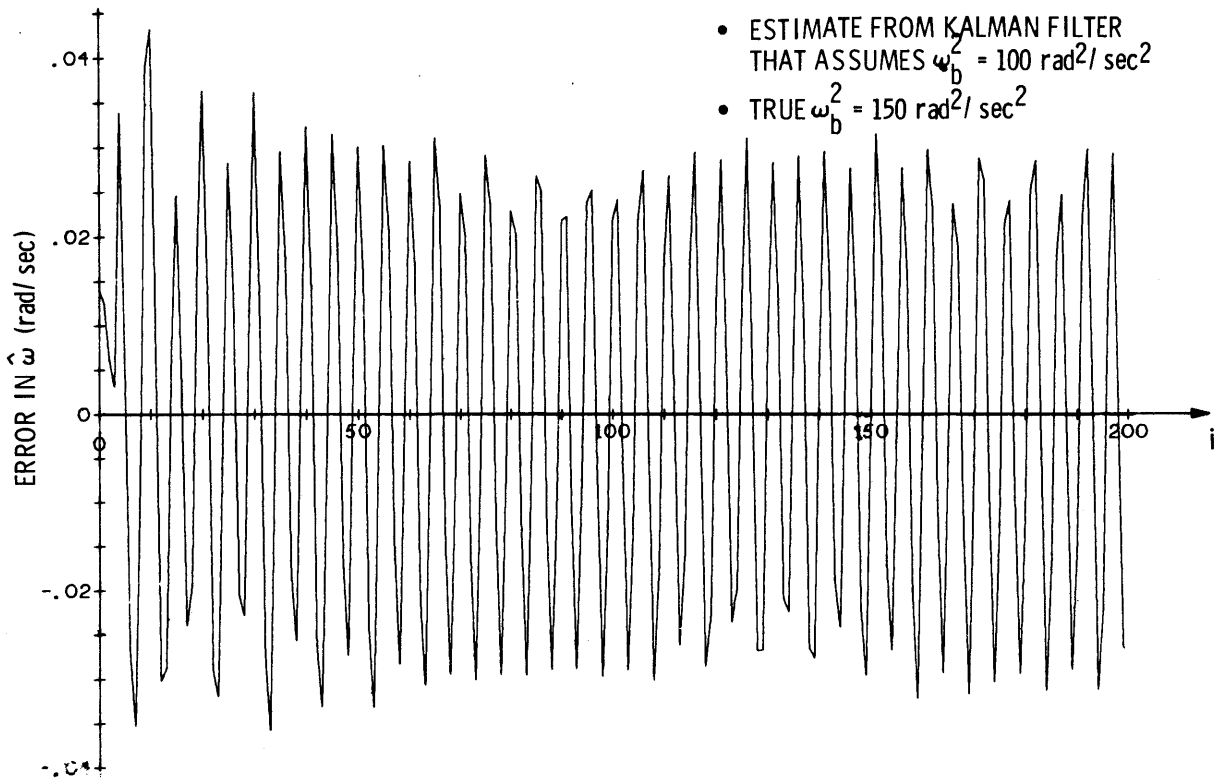


Figure 6.3.2a $\hat{\omega}$ Error When Parameter is Not Estimated

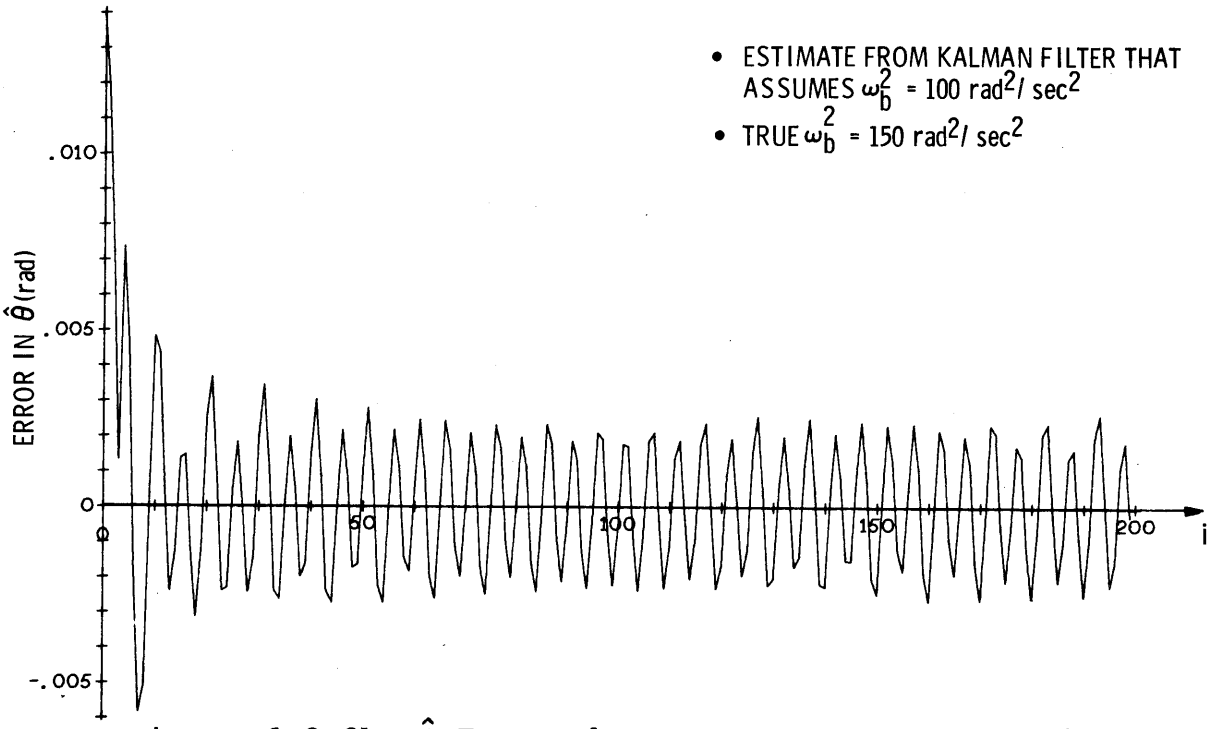


Figure 6.3.2b $\hat{\theta}$ Error When Parameter Is Not Estimated

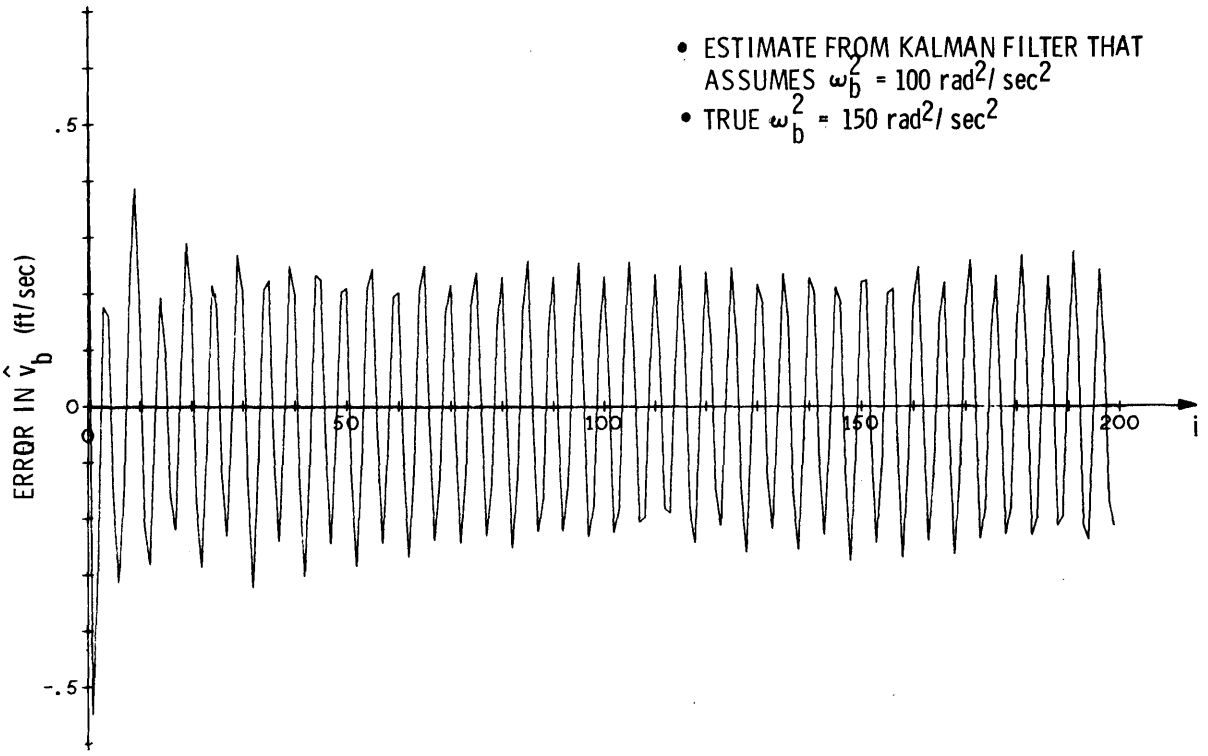


Figure 6.3.2c $\hat{v}_b$ Error When Parameter Is Not Estimated

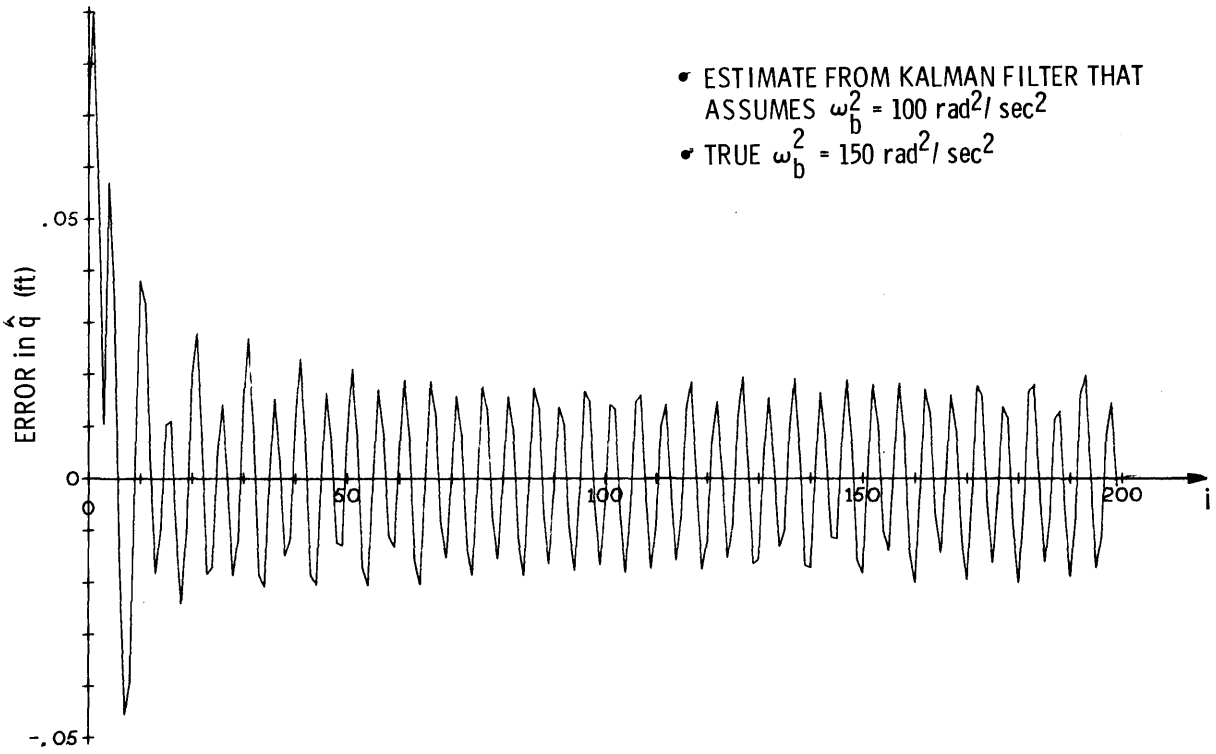


Figure 6.3.2d $\hat{q}$ Error When Parameter Is Not Estimated

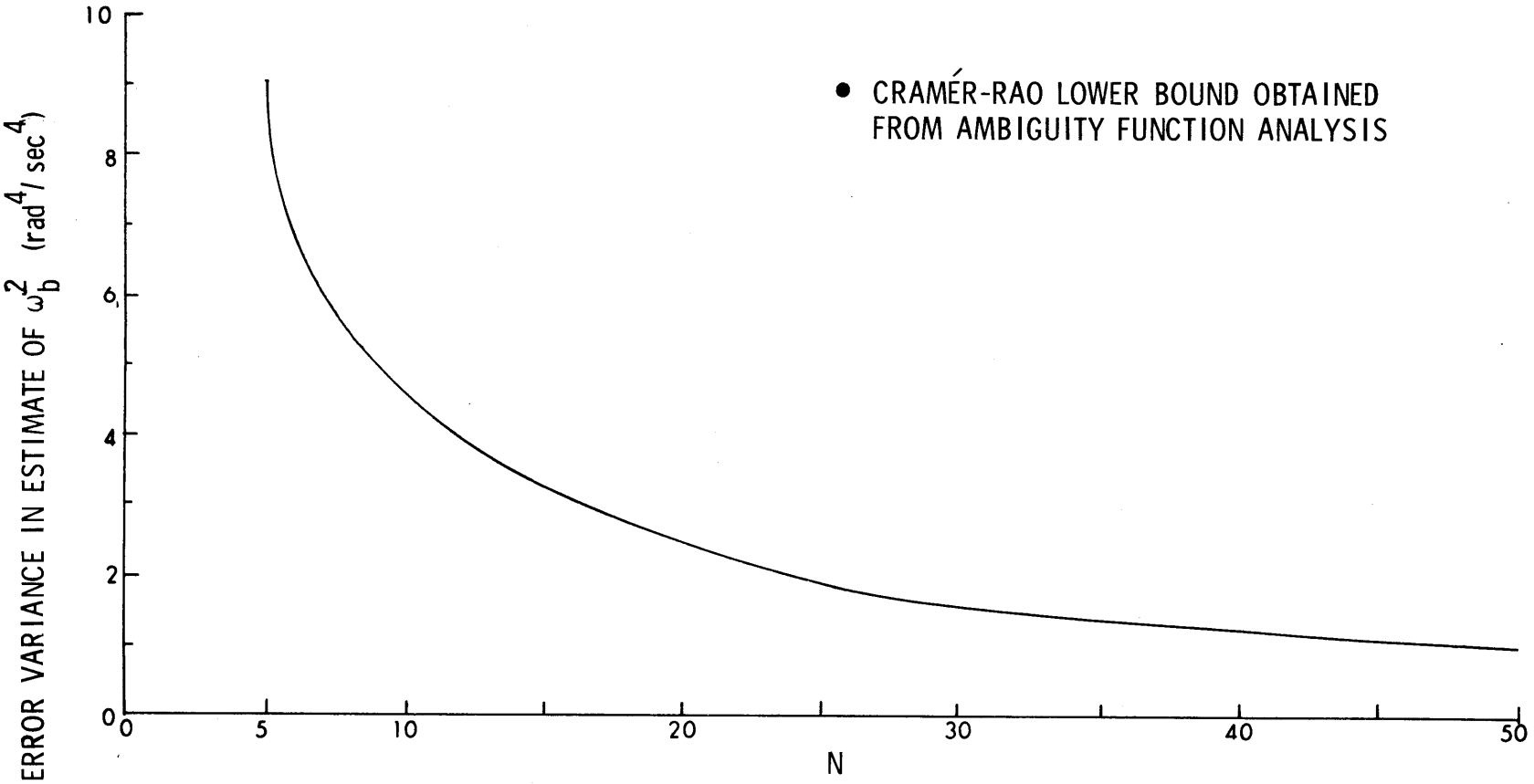


Figure 6.3.3 Lower Bound on Parameter Estimate Error Variance vs. N at i = 50

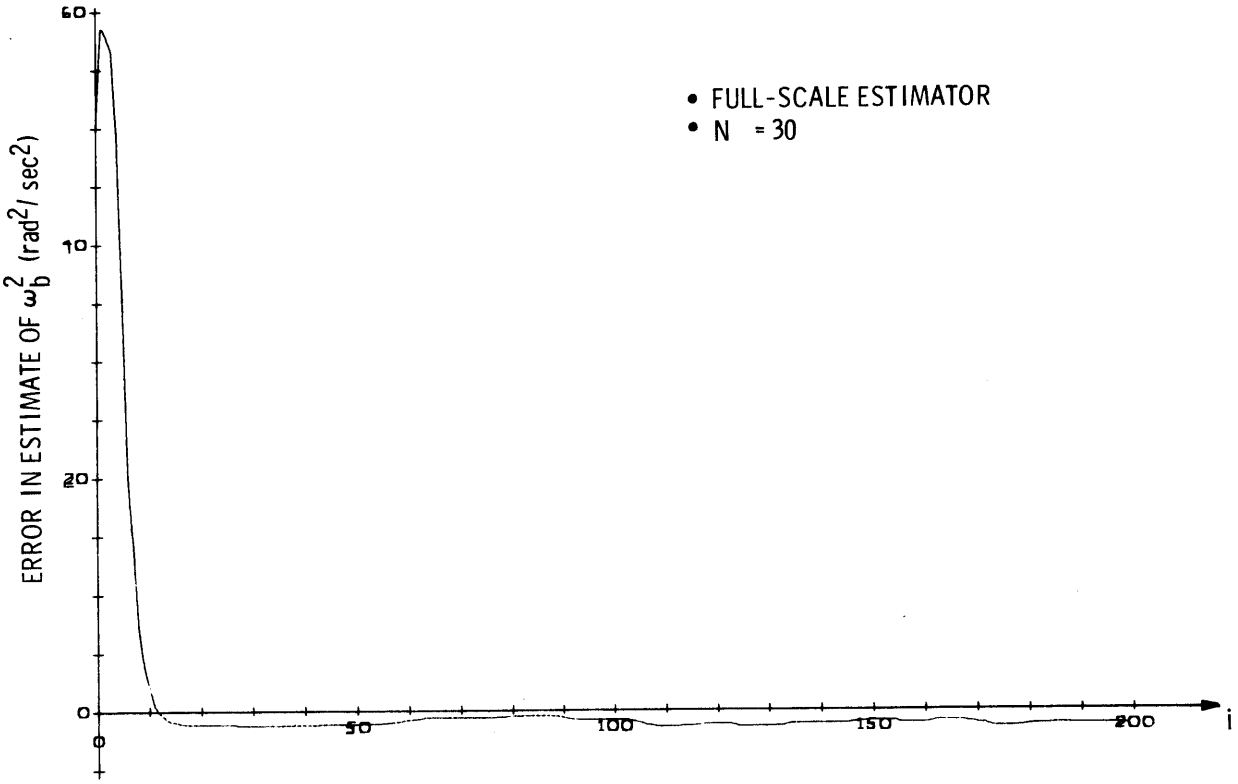


Figure 6.3.4a Full-Scale Parameter Estimate Error for N = 30

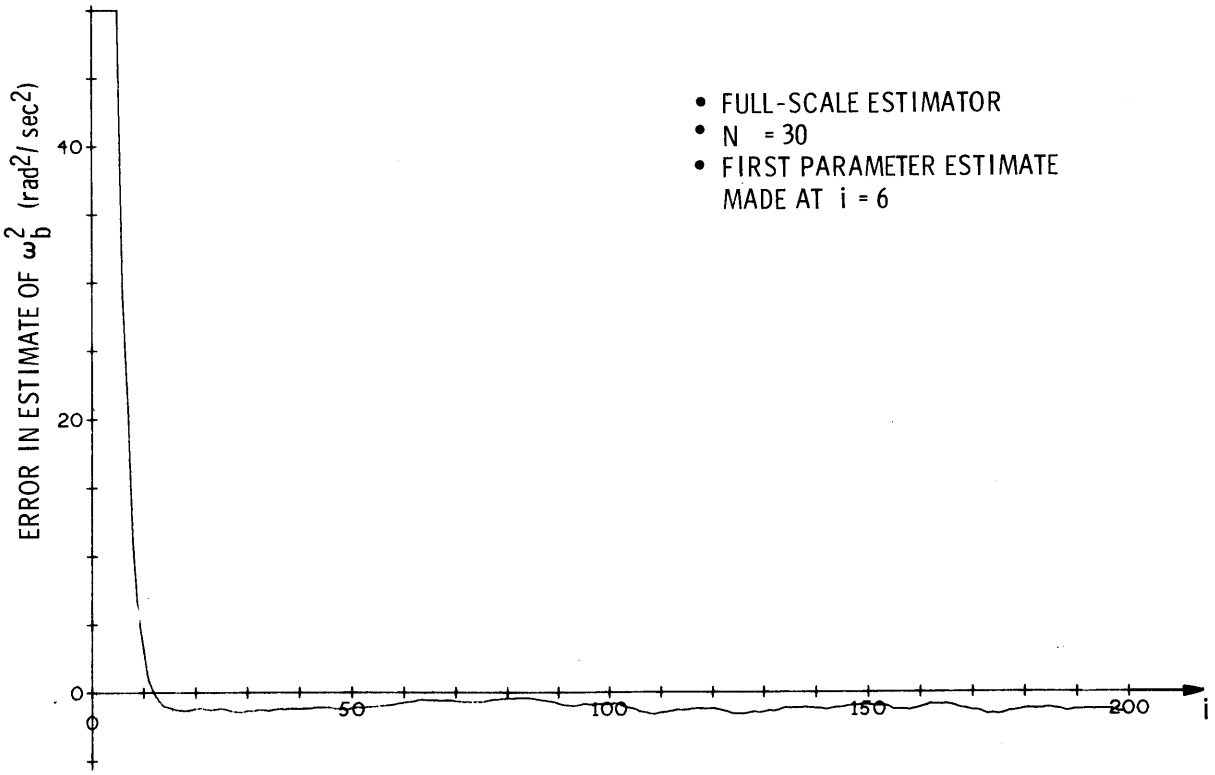


Figure 6.3.4b Full-Scale Parameter Estimate Error When First Estimate Is Delayed

by the full-scale estimator with N equal to 30. As in previous examples, the initial behavior characteristics are improved by delaying the time of the first estimate, thereby allowing the computed J^{-1} to attain a reasonable magnitude before using it in the estimation. By postponing the first parameter estimate until the sixth sample instant, the initial deviation is completely removed, as seen in figure 6.3.4b.

The corresponding state estimate error trajectories are portrayed in figure 6.3.5; comparing these to figure 6.3.2 reveals a vast improvement in estimation accuracy. In fact, because of the precision of the parameter estimate, these trajectories are essentially equivalent to the output of a Kalman filter tuned to the correct value of ω_b^2. Consequently, the performance of the estimator-controller combination very nearly duplicated that of the optimal stochastic linear controller based on the correct ω_b^2 value. Because it was more strongly dependent upon the parameter estimate, the accuracy of the state estimates was more instrumental in improving the feedback control than the correct evaluation of the controller gains, as evidenced by running a simulation in which these gains were not recomputed as a function of a*. For practical applications, this would suggest a simplified implementation that used either nominal controller gains or gains recomputed infrequently (as, only once, after the initial convergence of the parameter estimate, and thereafter only if the parameter estimate varied beyond a prespecified amount from this value).

The system maintained stability for all parameter values investigated; i.e., up through $\omega_b^2 = 700$, at which point the linear curve-fit approximations used to evaluate the system matrices became rather questionable. This distinctly contrasts with the linear estimator cascaded with a linear controller, which yielded unacceptable performance even at $\omega_b^2 = 150$.

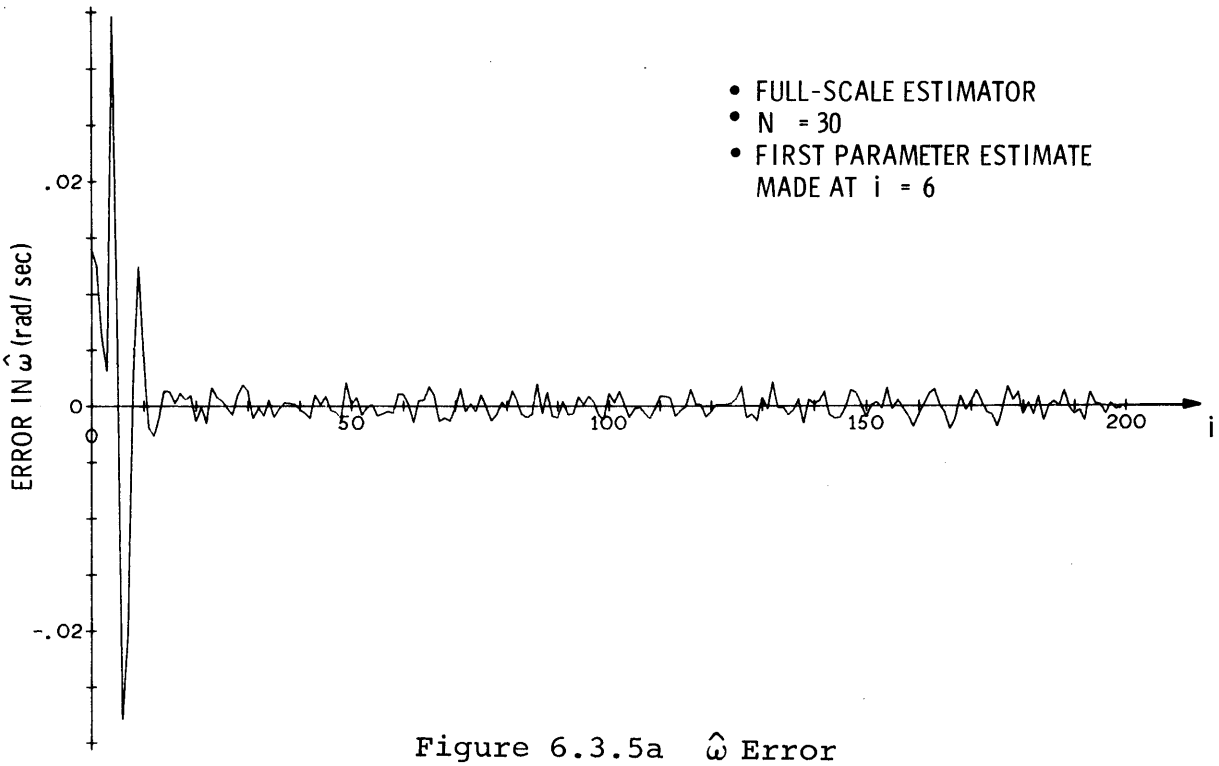


Figure 6.3.5a $\hat{\omega}$ Error

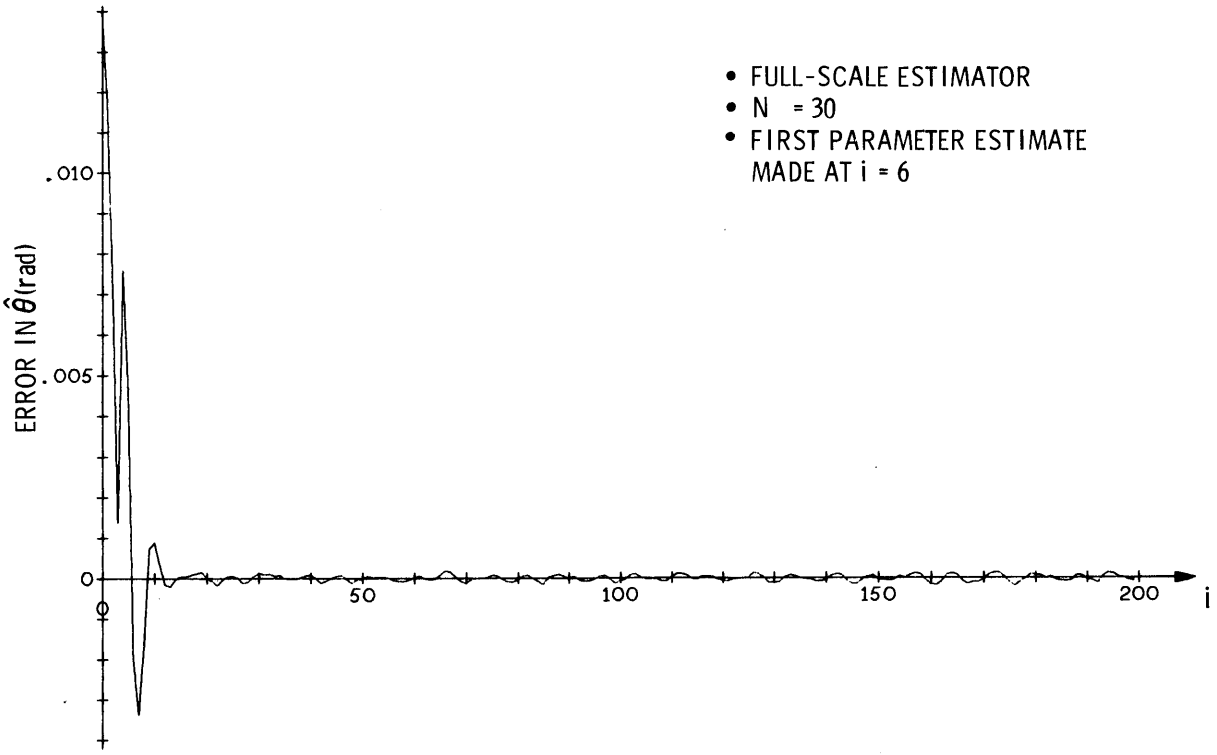


Figure 6.3.5b $\hat{\theta}$ Error

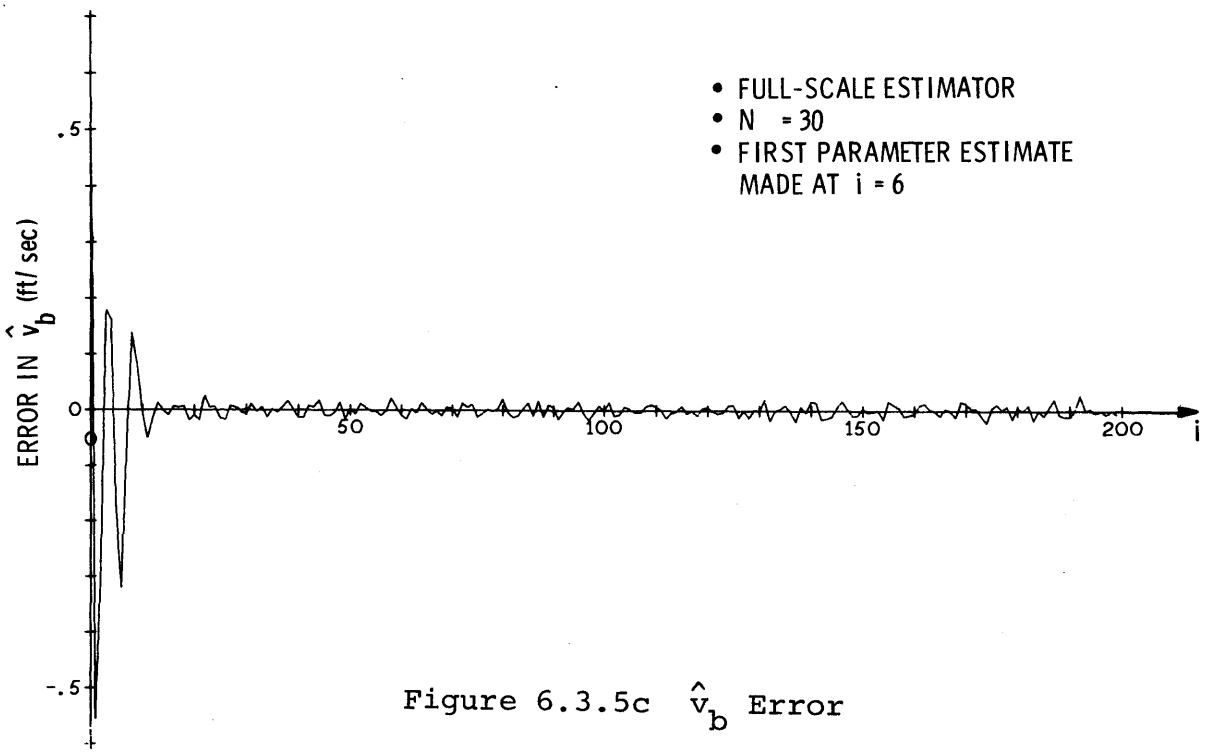


Figure 6.3.5c $\hat{v}_b$ Error

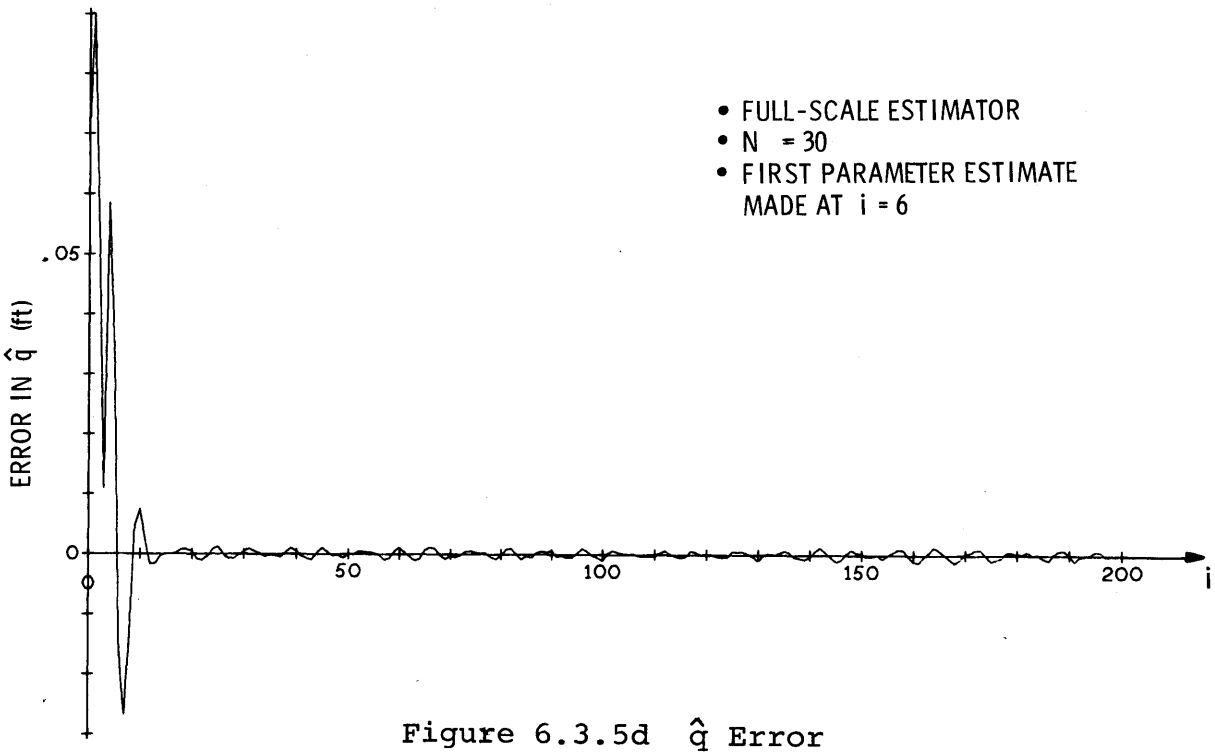


Figure 6.3.5d $\hat{q}$ Error

6.3.3 On-Line Estimation

The on-line estimator that employs an N of ten combined with averaging over the three most recent estimates was applied to this problem, and figure 6.3.6a is typical of the resulting parameter estimate error trajectories. As in the previous applications, by performing a parameter estimate more often for the first few instants, convergence to a good value was very rapid, subsequently followed by estimation accuracy that equals that of the full-scale estimator.

The error trajectory of figure 6.3.6b was generated by the same sequence of noise inputs, but the estimator itself differed in two respects. First, it incorporated only weighted least squares type terms, and the influence of dropping the other terms is minimal. Secondly, the initial parameter estimation frequency was not increased, and the resulting slow transient response is characteristic: although not as desirable as the behavior exhibited in figure 6.3.6a, it may well be satisfactory, and thereby be preferable because of the reduced computational load and implementation complexity.

In figure 6.3.7 are presented typical parameter estimate error trajectories resulting from the second on-line procedure, which produces a new parameter estimate every period. Plot a corresponds to an estimate with N = 5 and averaging over an interval of 30 periods. Because the averaging is initiated at the first sample time, the initial transient is slower than in the previous case or full-scale estimator. This is rectified by starting the averaging later, or more simply, performing the first parameter estimate later, as seen from plot b, in which the first estimate is made at i = 6. This avoids the problem of early deviation of the estimate, and subsequent averaging of erroneous values. Note that both cases incorporated only weighted least squares terms, resulting in a computationally efficient algorithm.

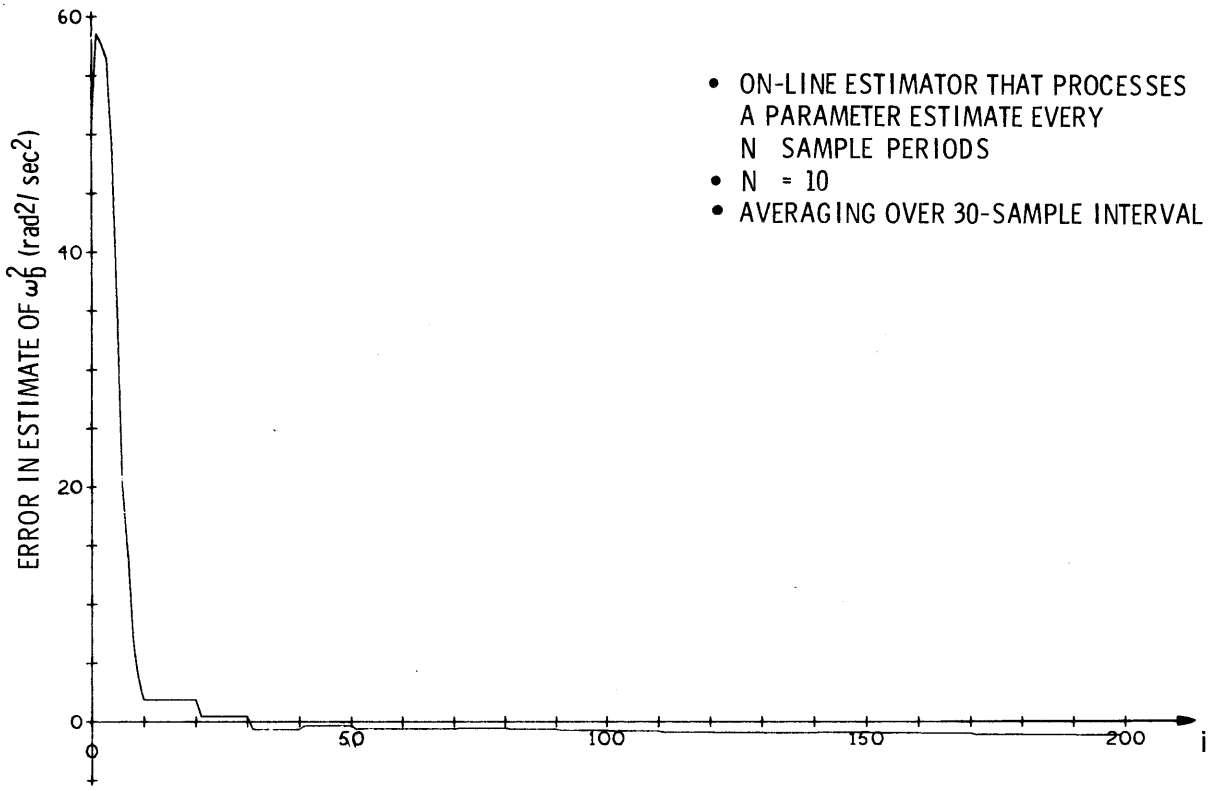


Figure 6.3.6a On-Line Parameter Estimate Error

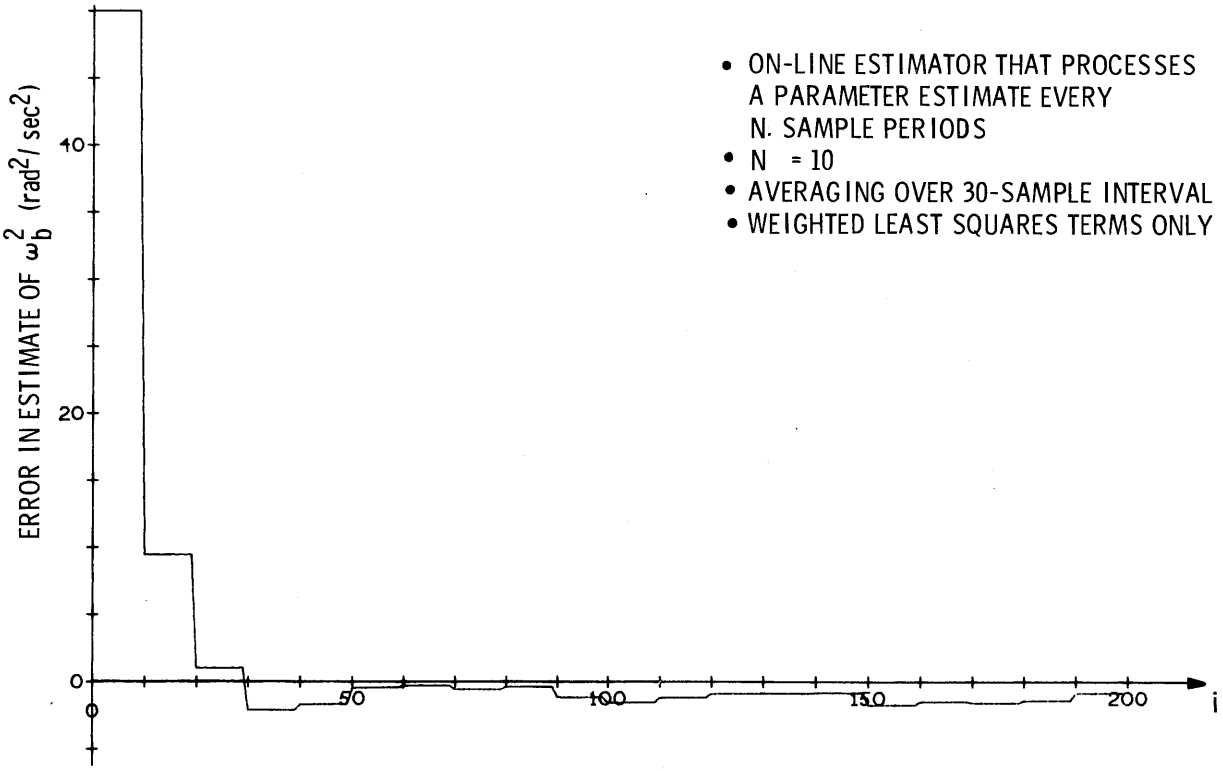


Figure 6.3.6b On-Line Parameter Estimate Errors; WLS Terms Only

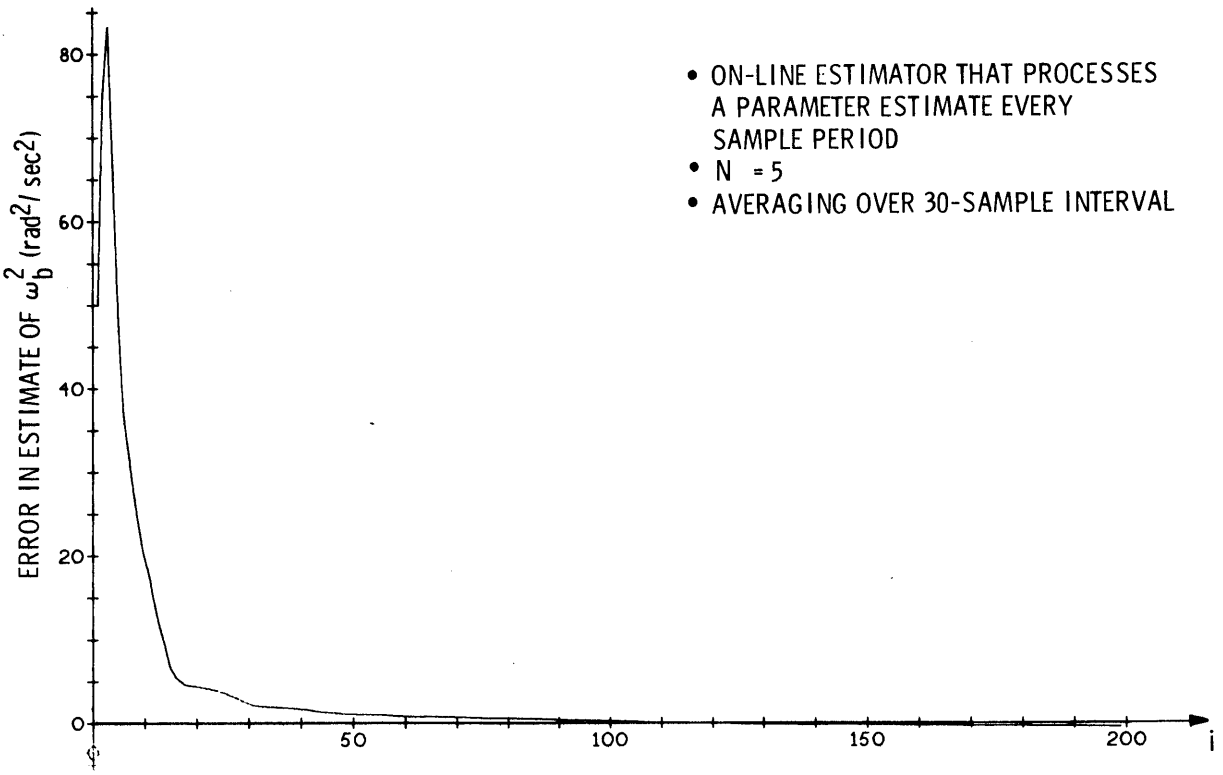


Figure 6.3.7a On-Line Parameter Estimate Error

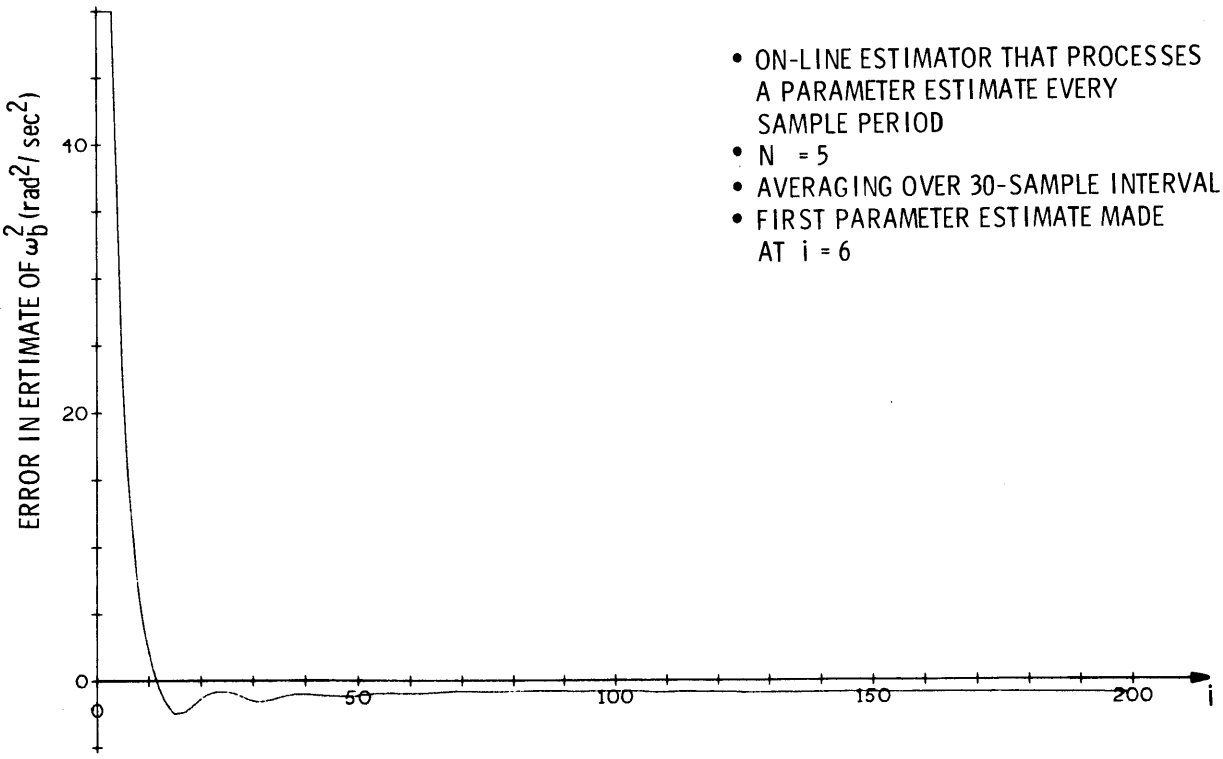


Figure 6.3.7b On-Line Parameter Estimate Error When First Estimate Is Delayed

As before, it can be concluded that the on-line conceptualizations provide very adequate parameter estimation, nearly duplicating that of the full-scale technique, but with considerably less computational expense. Because of this similarity, the state estimation accuracy is essentially identical to that of figure 6.3.5.

6.3.4 Canonical State Space Form

Section 5.5.2 presented a procedure to convert a given state space representation into the computationally more advantageous modified Jordan canonical form. Applying this procedure to the current example yields the system matrices

$$\underline{\Phi}' = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \sigma & \omega & 0 \\ 0 & 0 & -\omega & \sigma & 0 \\ 0 & 0 & 0 & 0 & 0.368 \end{bmatrix} \tag{6.3.13}$$

$$\underline{b}' = \begin{bmatrix} 0.041541 \\ 0.001492 \\ b_3' \\ b_4' \\ 0.63212 \end{bmatrix} \tag{6.3.14}$$

$$\underline{h}'^{T} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{6.3.15}$$

where σ and ω are the real and imaginary parts, respectively, of the upper bending mode root in figure 6.3.8a, a z-plane portrait of the system poles, and b_3' and b_4' are as plotted in figures 6.3.8b and 6.3.8c. Piecewise linear functions have been fitted to b_3' and b_4'; in actual use, the linear elements valid for $50 \leq \omega_b^2 \leq 200$ suffice, and these are

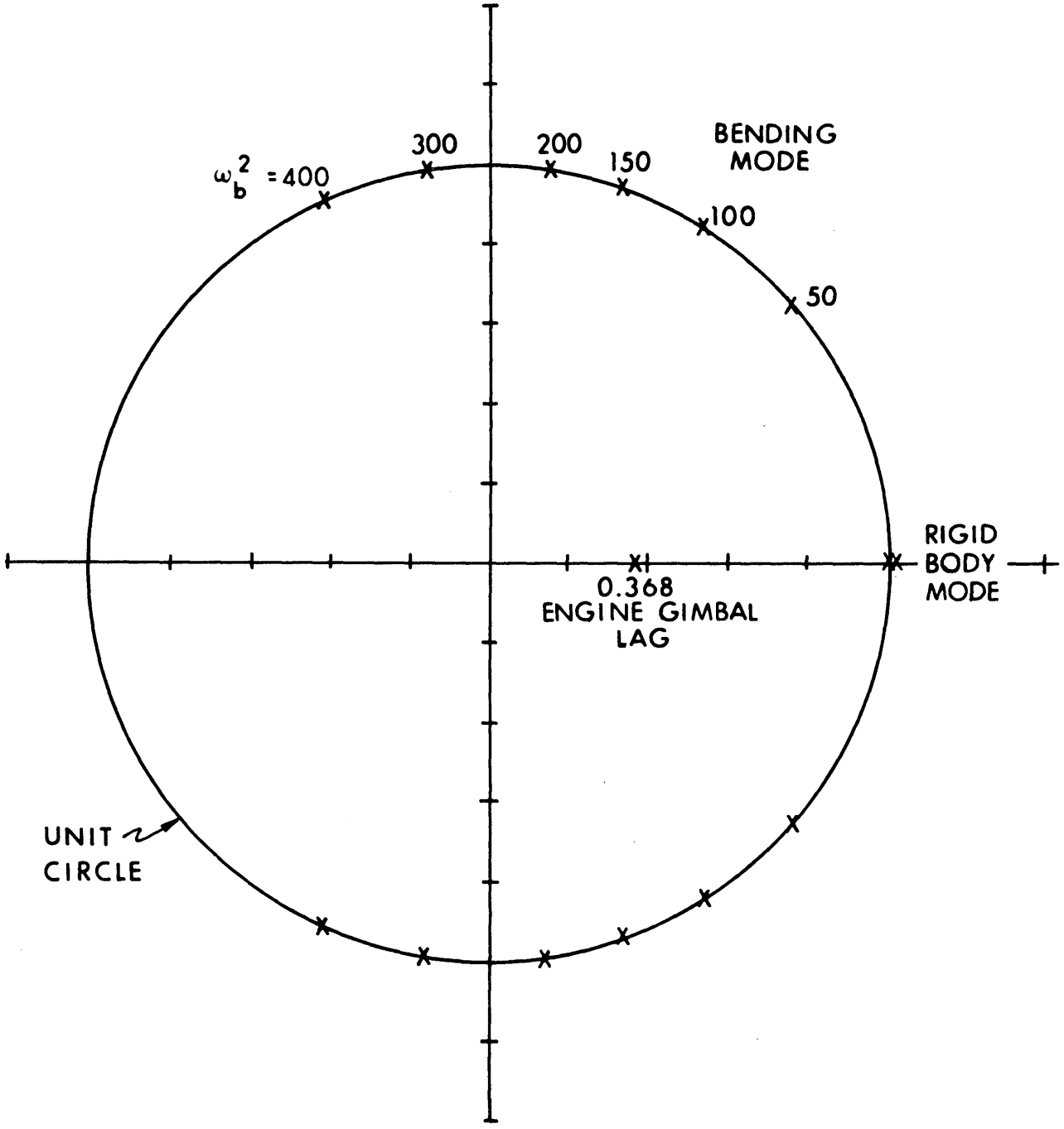


Figure 6.3.8a Z-Plane Portrait of System Poles

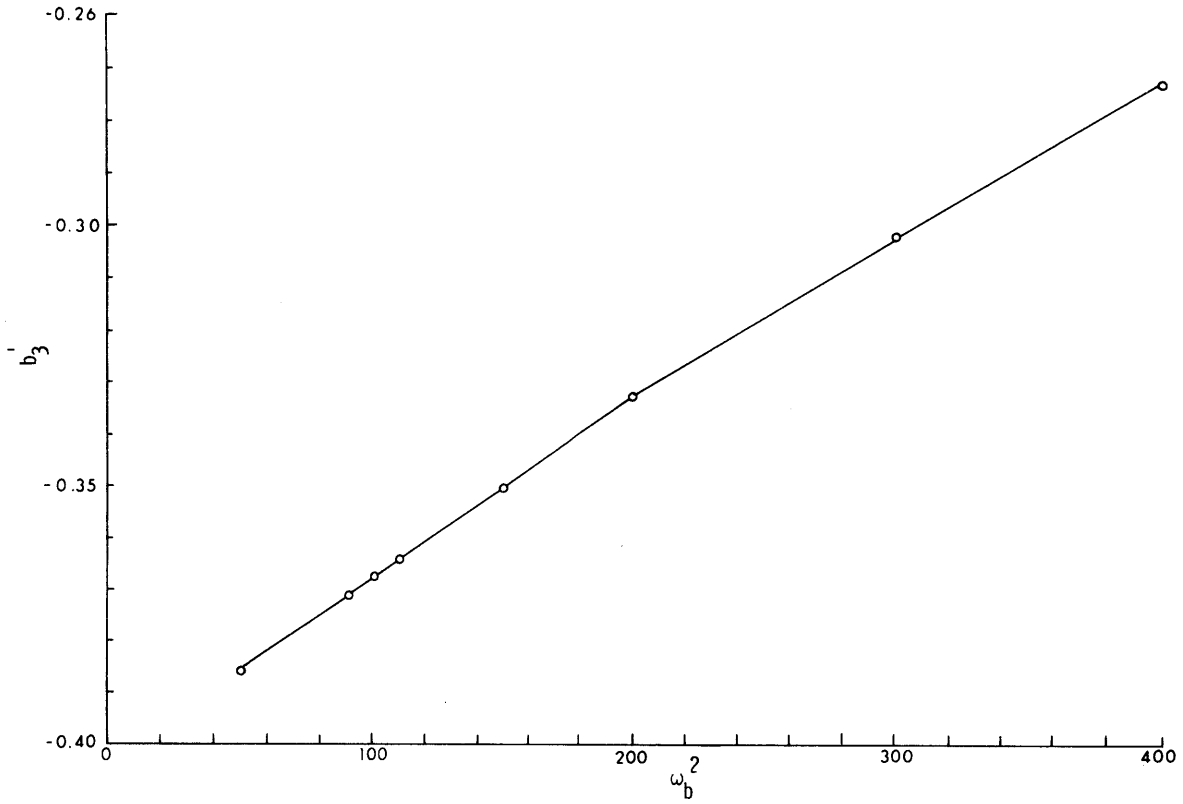


Figure 6.3.8b b_3' vs. Parameter Value

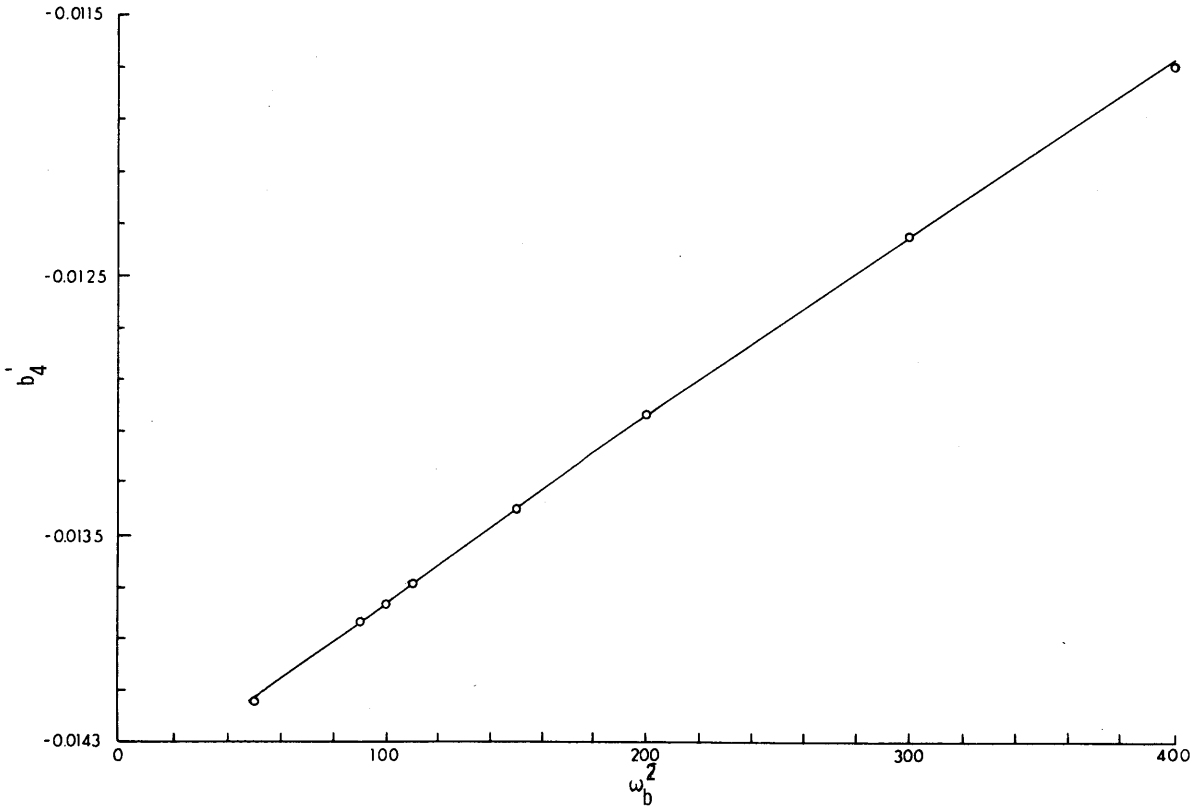


Figure 6.3.8c b_4' vs. Parameter Value

(note that $a = \omega_b^2$)

$$b_3' = -0.403 + 0.00035\ a \tag{6.3.16}$$

$$b_4' = -0.0145 + 0.0000073\ a \tag{6.3.17}$$

If the state estimates are to be expressed in terms of the original state variables, then the elements of $\underline{T}$ can likewise be curve-fitted, and then equation (5.5.3) used to generate the desired values. However, this is not essential since the purpose of the system is to provide control of vehicle oscillations, rather than to maintain estimates of particular variables. Thus, the alternative would be to calculate and curve fit the elements of the vector transpose ($\underline{c}^T\underline{T}$) as a function of the parameter, this being the controller gain matrix in the canonical state space:

$$u = -\underline{c}^T\hat{\underline{x}} = -\underline{c}^T\underline{T}\,\hat{\underline{x}}' \tag{6.3.18}$$

The benefits of the canonical form are apparent from the system matrices $\underline{\Phi}'$, $\underline{b}'$ and $\underline{h}'^T$. In $\underline{\Phi}'$ only four elements need be computed as a function of the uncertain parameter (and only two separate functional evaluations are required to determine their values), whereas ten elements of the original transition matrix were functionally dependent upon ω_b^2. Furthermore, a product such as $\underline{\Phi}\,\hat{\underline{x}}$ required fourteen multiplications and eleven additions, while $\underline{\Phi}'\hat{\underline{x}}'$ is accomplished with five multiplications and three additions. As mentioned in section 5.5.2, the propagations needed for the score are also simplified considerably, and a shorter computer wordlength can be used to achieve the same modelling precision as in the original state space. Since no approximation is involved in the state space transformation, these benefits are attained with no loss in performance.

6.3.5 Comparison to Extended Kalman Filter

This example can also be treated as a nonlinear estimation problem, and solved by means of an extended Kalman filter. For such a formulation, the uncertain parameter is considered as an additional state variable, thereby increasing the dimension of the filter state space. This approach was investigated to provide a basis of comparison for the performance achieved with the techniques of this thesis.

The extended Kalman filter solution is developed in the following way for a continuous-time system and discrete-time measurements. If the system state is assumed to satisfy

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}(t),\underline{u}(t),t) + \underline{G}(t)\underline{w}(t) \tag{6.3.19}$$

then the perturbation of the state from an assumed reference trajectory, $\delta\underline{x}(t) = \underline{x}(t) - \underline{x}_{\eta}(t)$, satisfies, to first order,

$$\delta\dot{\underline{x}}(t) = \underline{F}(\underline{x}_{\eta}(t),\underline{u}(t),t)\ \delta\underline{x}(t) + \underline{G}(t)\underline{w}(t) \tag{6.3.20}$$

where $\underline{x}_{\eta}(t)$ is the nominal trajectory generated by

$$\dot{\underline{x}}_{\eta}(t) = \underline{f}(\underline{x}_{\eta}(t),\underline{u}(t),t) \tag{6.3.21}$$

and $\underline{F}(\underline{x}_{\eta}(t),\underline{u}(t),t)$ is the matrix of partial derivatives of $\underline{f}$ with respect to $\underline{x}$, evaluated along the nominal trajectory:

$$\underline{F}(\underline{x}_{\eta}(t),\underline{u}(t),t) = \left.\frac{\partial \underline{f}(\underline{x}(t),\underline{u}(t),t)}{\partial \underline{x}}\right|_{\underline{x}(t) = \underline{x}_{\eta}(t)} \tag{6.3.22}$$

Let $\underline{\Phi}(i+1,i;\underline{x}_{\eta}(t_i),\underline{u}(t_i))$ be the state transition matrix corresponding to $\underline{F}$ for propagating from sample time t_i to t_{i+1}.

The basis of the extended Kalman filter is to relinearize about each estimate $\underline{\hat{x}}(t_i)$ once it has been computed. As soon as a new state estimate is made, a new and better reference trajectory is incorporated into the estimation, thereby maintaining the validity of the assumption that the deviations from the reference trajectory are small enough to allow linear perturbation equations. Thus, the estimate is propagated between measurements by means of

$$\underline{\bar{x}}(t_{i+1}) = \underline{\hat{x}}(t_i) + \int_{t_i}^{t_{i+1}} \underline{f}(\underline{\hat{x}}(t),\underline{u}(t),t)\, dt \qquad (6.3.23)$$

$$\underline{M}(i+1) = \underline{\Phi}(i+1,i)\underline{P}(i)\underline{\Phi}^T(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^T(i) \qquad (6.3.24)$$

where $\underline{\Phi}(i+1,i)$ is evaluated as $\underline{\Phi}(i+1,i;\underline{\hat{x}}(t_i),\underline{u}(t_i))$ and the term $\underline{G}(i)\underline{Q}(i)\underline{G}^T(i)$ is evaluated by means of the technique described in Appendix B. Although equation (6.3.24) has the same form as similar equations presented in section 3.4, the state space is now of higher dimension. Similarly, the relations for incorporating the measurement data have the same form as the $\underline{\hat{x}}(i)$ and $\underline{P}(i)$ equations presented previously, but the actual realization of the equations differ because of the augmented state space.

Equations (6.3.23) and (6.3.24) are identical to the result of integrating

$$\dot{\underline{\bar{x}}}(t) = \underline{f}(\underline{\bar{x}}(t),\underline{u}(t),t) \qquad (6.3.25)$$

$$\dot{\underline{M}}(t) = \underline{F}(\underline{\bar{x}}(t),\underline{u}(t),t)\underline{M}(t) + \underline{M}(t)\underline{F}^T(\underline{\bar{x}}(t),\underline{u}(t),t) + \underline{G}(t)\underline{Q}(t)\underline{G}^T(t) \qquad (6.3.26)$$

from time t_i to t_{i+1}, using the initial conditions:

$$\underline{\bar{x}}(t_i) = \underline{\hat{x}}(i) \tag{6.3.27}$$

$$\underline{M}(t_i) = \underline{P}(i) \tag{6.3.28}$$

For the thrust vector control problem, the estimator state vector would be $(\omega, \Theta, v_b, q, a)^T$, where a is the parameter ω_b^2; $\delta(t)$ is updated deterministically. Thus, the matrix $\underline{F}(\underline{\bar{x}}(t),\underline{u}(t),t)$ is

$$\underline{F}(\underline{\bar{x}}(t),\underline{u}(t),t) = \left[\begin{array}{ccccc} 0 & 0 & 0 & 0.0815 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a & -q \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]\Bigg|_{\substack{a = \bar{a}(t) \\ q = \bar{q}(t)}} \tag{6.3.29}$$

To evaluate equation (6.3.26) precisely, the time history of $\bar{q}(t)$ must be evaluated for $t_i \leq t < t_{i+1}$; $\bar{a}(t)$ would be equal to $\hat{a}(t_i)$ over the entire interval. As a result, $\underline{\Phi}(i+1,i;\underline{\hat{x}}(t_i),\underline{u}(t_i))$ would be $\underline{\Phi}(i+1,i;\hat{v}_b(i),\hat{q}(i),\hat{a}(i))$: if precomputation of $\underline{\Phi}$, $\underline{M}$, etc. is considered for on-line applications, the functional evaluations are dependent upon three parameters instead of one as in the previous sections. Even if $\underline{F}$ were instead evaluated with $a = \hat{a}(t_i)$ and $q = \hat{q}(t_i)$, two-parameter families of functions would be required; incorporating this simplification into the solution of equation (6.3.26) caused little change in the estimation because the sample rate was fast compared to the system dynamics. However, the relinearization using at least $\hat{q}(t_i)$ was required to achieve parameter tracking.

Figure 6.3.9 presents a representative error trajectory for the estimate of ω_b^2, using empirically determined "good" values for the assumed variance of $\hat{a}(0)$ and the assumed strength of the white noise driving a(t) in the system

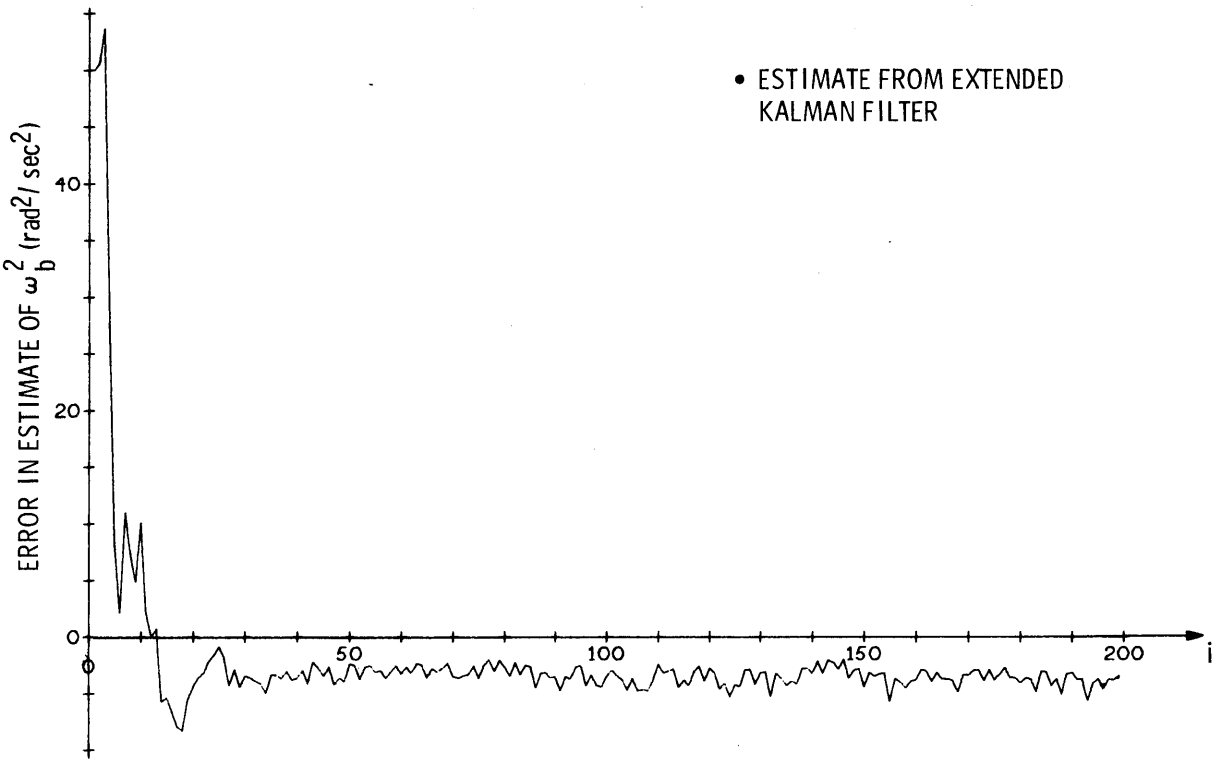


Figure 6.3.9 Parameter Estimate Error Produced by Extended Kalman Filter

model (1000 rad.^{4}/sec.^{4} and 200 rad.^{4}/sec.^{5} respectively). The corresponding error trajectories for the other four states are portrayed in figure 6.3.10. A comparison of these plots to those of figure 6.3.5 reveals a similar, but somewhat inferior, accuracy attained by the extended Kalman filter.

This estimator behavior is, however, strongly dependent upon the two statistical parameters used to model a(t). Higher values than those used to generate the above results will yield greater oscillations in the post-transient period, while lower values cause a slower transient response. Figure 6.3.11 is indicative of this fact: it is the error trajectory for $\hat{a}(t)$ caused by decreasing the two variances by a factor of 100. Unfortunately, the physics of a particular problem does not directly yield appropriate variance values, so they must be chosen in a rather ad hoc manner.

Thus, the methods of this thesis are seen to produce estimates of at least the precision attainable by an extended Kalman filter, without requiring the specification of the apriori statistics of $\underline{a}$ or the strength of a contrived white noise source in order to provide parameter estimation. Furthermore, the resulting estimator equations yield more easily to approximations and precomputations than do those of the extended Kalman filter, and they do not depend upon the validity of linearizing a nonlinear set of equations about a nominal trajectory. Moreover, they allow the parameter estimation portion to be turned on or off very easily, as needs and computational loadings might dictate - an attractive feature for on-line computations performed by a central processor. These factors tend to favor implementation of the algorithms of this thesis over those of the extended Kalman filter concept.

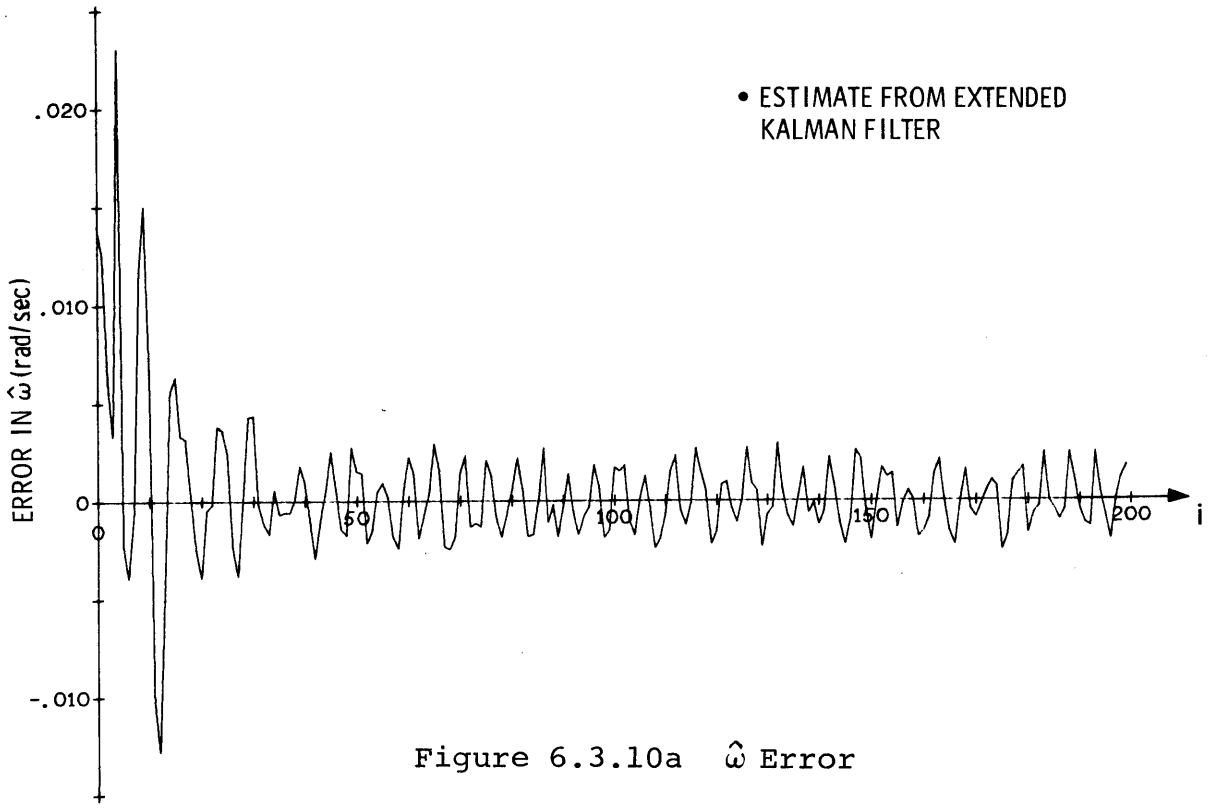


Figure 6.3.10a $\hat{\omega}$ Error

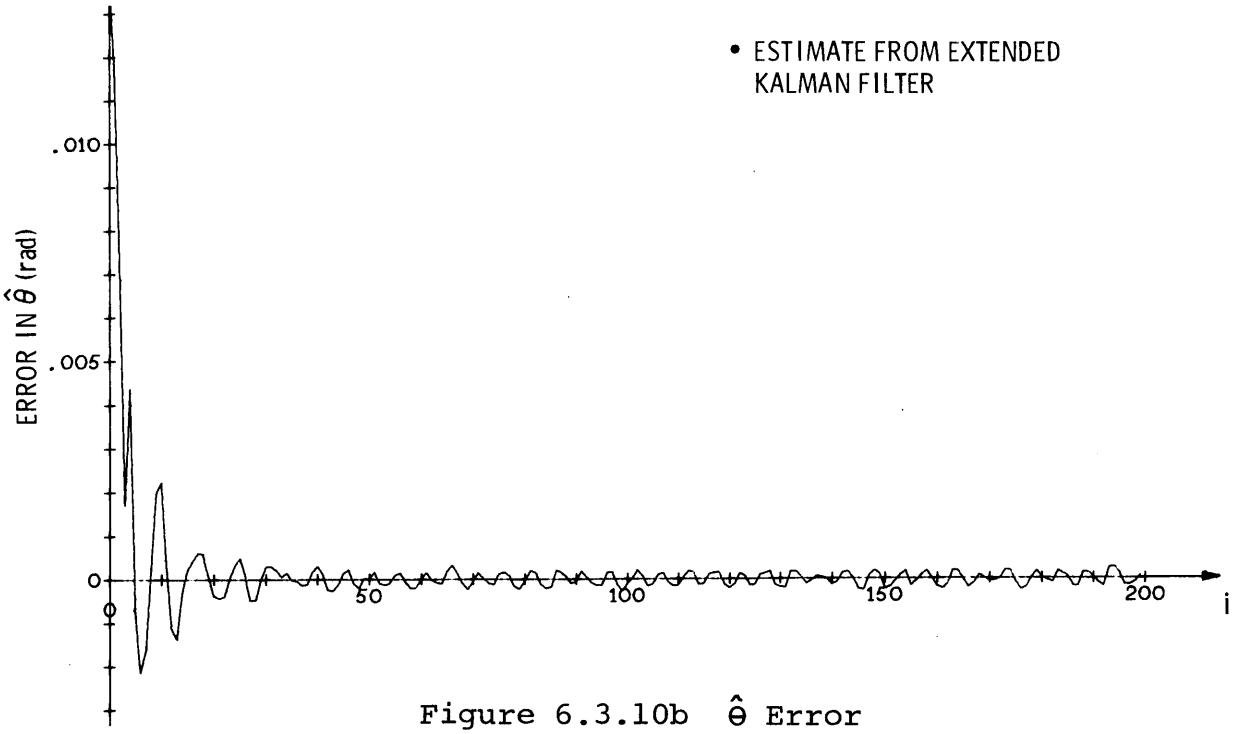


Figure 6.3.10b $\hat{\theta}$ Error

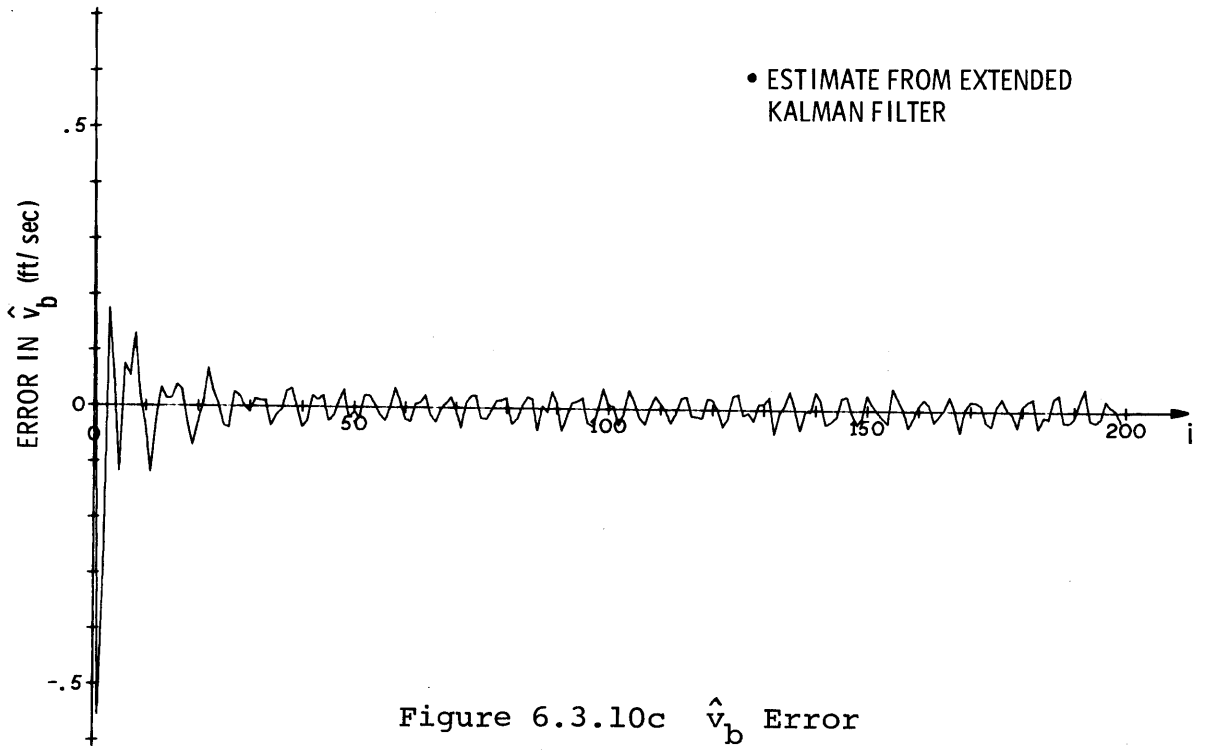


Figure 6.3.10c $\hat{v}_b$ Error

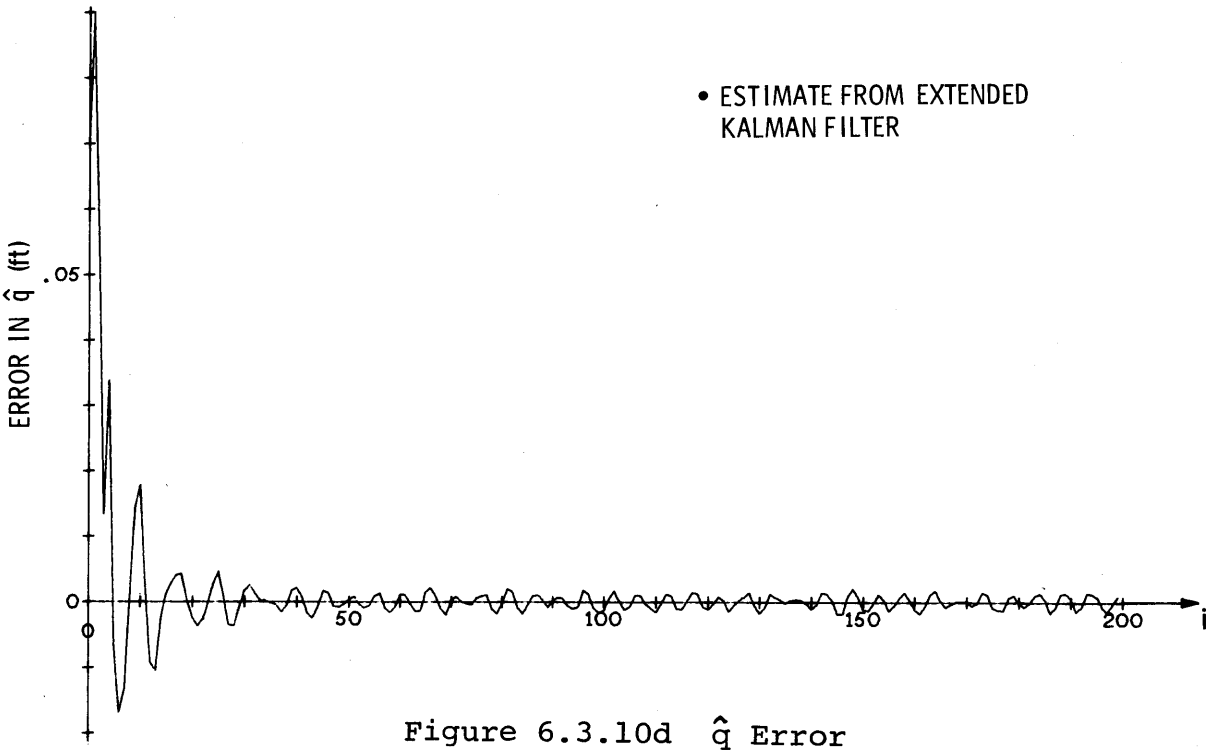


Figure 6.3.10d $\hat{q}$ Error

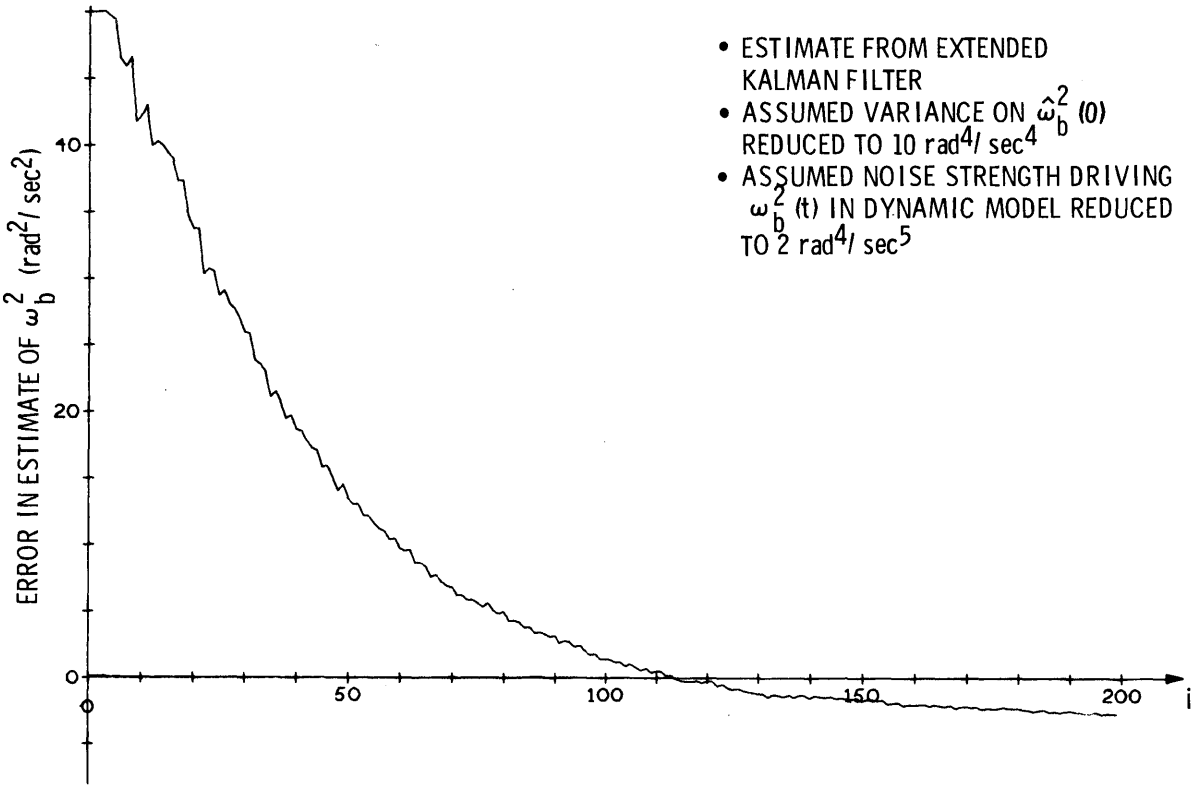


Figure 6.3.11 Parameter Estimate Error When Variances for $\hat{a}$ Model Are Reduced

CHAPTER VII

CONCLUSION

7.1 Summary of Results

The major contribution of this thesis has been to provide an effective and practical means of simultaneously estimating the states and parameters of a linear dynamic system model, using noisy measurements of the actual system outputs. The resulting estimator yields state estimation accuracy that can far surpass that of a filter that does not attempt to identify the uncertain parameter values. Especially since such uncertainty is usually present in prospective filtering applications, the state and parameter estimator has the potential not only of improving the performance of existing filter systems, but also of being successfully implemented in situations for which linear state estimators have proven totally inadequate.

Although other techniques have been suggested for parameter estimation, the method proposed in this thesis is unique in its ease of application. The engineer can systematically and directly use his knowledge that the parameters are more slowly varying than the states, rather than to transform this knowledge into the specification of a contrived noise source to drive a stochastic model for the parameters. Furthermore, he need not specify an apriori probability density for the parameters, as required by many other techniques, but this formulation is flexible enough to incorporate as much apriori parameter information as he can, in fact, provide. An original, improved formulation of the inverse covariance form of the state propagations additionally allows treatment of the cases of incomplete apriori state information or the apriori uncertainties differing by large orders of magnitude for different directions in state space.

Theoretical performance analyses, as substantiated by the empirical evidence of simulations, have demonstrated

the maximum likelihood technique to be a very suitable basis for state and parameter estimation. Under very general assumptions, the parameter estimator is consistent and asymptotically unbiased, efficient, and normally distributed with mean $\underline{a}_t$ and covariance $\underline{J}\,[\,i\,,\,\underline{a}_t\,]^{-1}$. Moreover, the behavior of the state estimator converges to that of the maximum likelihood state estimator based upon the true parameter values.

Besides providing a global performance analysis, ambiguity functions can serve as invaluable design tools. They are generated, in part, by an error sensitivity analysis of linear optimal state estimators, which also informs the user whether parameter estimation is required to attain desired performance, and if so, which are the critical parameters to estimate. Once these decisions have been made, ambiguity functions can compare the relative effectiveness of different likelihood functions, or even individual terms within these functions, in providing estimation accuracy. Also, they can be used to predict the sensitivity of this performance to the form of dynamics model, length of data interval N, and the types and precision of measurements taken: variables which the designer can control to some extent in order to ensure adequate performance.

The full-scale estimator derived from the likelihood function $\ell n\, f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$, developed in conjunction with the design tools provided above, is presented in section 3.4. To be of greatest use to the practicing engineer, these results and those of Chapter V have been derived from the point of view of developing the algorithms with which to implement the estimators. Since the full-scale estimator is suitable only for off-line applications, Chapter V is devoted to means of reducing the computational load, while maintaining adequate performance. The approximation of equation (5.1.1) and the on-line conceptualizations of section 5.2 have been shown to yield essentially identical performance to that of the full-scale estimator, yet reduce the required

number of calculations in typical applications by orders of magnitude. Including only the most significant terms in the likelihood equations, as suggested in section 5.3, similarly produced a vast reduction in required computer time, especially by eliminating certain matrix propagations in processing the parameter estimate, while retaining the same fundamental estimate behavior. Furthermore, precomputations proved feasible for many of the needed quantities, and when combined with the other approximations, it furnished an estimator with true on-line potential. This potential can be developed further by performing the computations in an equivalent, but more advantageous, canonical state space representation, and section 5.5.2 describes a particularly convenient means of performing the coordinate transformation. Here, the computational benefit is due not only to the high density of zeroes and ones in the $\underline{\Phi}$ and $\underline{H}$ matrices, but also to the separation of system modes, only a small subset of which are affected by each uncertain parameter; it also allows a shorter computer wordlength than other state space forms for equivalent modelling precision. Symmetry can be exploited by computing only the lower triangular form of symmetric matrices; an extension of this idea would be to implement the square root formulation of the state estimator, which is advantageous in that it also allows greater precision with a restricted wordlength. Finally, there may be significant utility in modifying a vector measurement update into a sequence of equivalent scalar or lower dimensional vector updates, as described in section 5.7. By employing these various techniques, an on-line state and parameter estimator can be developed that at least rivals other methods in performance and computational feasibility, and far surpasses them in ease of application.

7.2 Recommendations for Future Investigations

Although this research has produced a viable estimation

technique, there are certain areas that warrant further investigation:

1) The scope of this study was confined to the estimation of state variables and uncertain parameters, no attempt being made to solve the optimal control problem simultaneously. Although the third computational example utilized Open Loop Feedback Optimal linear control, with gain histories precomputed as a one-parameter family of curves evaluated with the most recent parameter estimate, this was performed as a logical approach to the control problem and no claim of overall system optimality was implied. This may in fact be the most practical logic to utilize, but further research might indicate a more beneficial means of formulating and solving the overall problem.
2) More definitive and practically verifiable conditions for the ability to perform the estimation would be useful. Although sections 3.1 and 4.2 provide considerable information about necessary conditions, it would be desirable to obtain sufficient conditions in a usable form. Unfortunately, their extraction has not been successful as yet.
3) The conditions and mechanisms of stability and convergence for both the full-scale and on-line implementations could be studied analytically for the case of a finite number of measurements being processed. In fact, the propagation of errors and their norms was investigated using a linearized perturbation model of the state and parameter estimator, but no generally applicable analytical results could be discerned from the analysis. (With regard to stability, only conditions that insure system instability could possibly be obtained from such analysis, since stability of the perturbation model yields no conclusive information about stability of the original total system.) As desirable as such results would be, the

behavior of maximum likelihood estimators for finite sample size has long eluded definitive theoretical description: this is a difficult, and possibly very unfruitful, area of research.

4) It would be valuable to apply this state and parameter estimator to situations in which the assumed form of the dynamics model does not correspond well to physical reality. Both theoretical and simulation results might be used to define the behavior of the estimator when the "true" system were a higher dimensional linear system model, or a nonlinear dynamics model. Hypothesis testing would be one means of investigating the adequacy of a certain model form for predicting the input-output behavior of the true system. This would further validate the estimation technique by explicitly accounting for the fact that any assumed dynamics model form is necessarily an approximation to physical reality.

Appendix A: Estimator Equations that Allow Uncertainties in $\underline{H}(i)$

Section 3.4 presented the recursions for the state, score, and conditional information matrix of the estimator based upon $\ell n\ f(\underline{x}(i), \underline{Z}_N(i) | \underline{Z}(i-N), \underline{a})$. As explained in the introduction to Chapter III, this formulation did not consider uncertainties in the measurement matrix $\underline{H}(i)$. This appendix relates the changes required to allow for such uncertainties; the equation numbers used here are the numbers of the corresponding equations that assume $\partial \underline{H}(i)/\partial a_\ell = \underline{0}$ for all ℓ. The time index will be omitted, since it will be apparent from the original equations.

The propagations between measurement times are as follows. The state equations are unaltered, but the score relations become

$$\frac{\partial \underline{A}}{\partial a_\ell} = \underline{H}\, \frac{\partial \underline{M}}{\partial a_\ell}\, \underline{H}^T + \frac{\partial \underline{H}}{\partial a_\ell}\, \underline{M}\, \underline{H}^T + \underline{H}\, \underline{M}\, \frac{\partial \underline{H}^T}{\partial a_\ell} \tag{3.4.20}$$

$$s^1_\ell = \frac{\partial \bar{\underline{x}}^T}{\partial a_\ell}\, \underline{H}^T\, \underline{\eta} - \tfrac{1}{2}\, \mathrm{tr}\left\{ \underline{\Omega}\, \frac{\partial \underline{A}}{\partial a_\ell} \right\} + \bar{\underline{x}}^T\, \frac{\partial \underline{H}^T}{\partial a_\ell}\, \underline{\eta} \tag{3.4.22}$$

and the changes to the conditional information matrix computations are

$$\begin{aligned} E\left\{ s^1_k\, s^1_\ell \,\middle|\, a_* \right\} = \tfrac{1}{2}\, \mathrm{tr}\Bigg[& \underline{A}^{-1}\, \frac{\partial \underline{A}}{\partial a_k}\, \underline{A}^{-1}\, \frac{\partial \underline{A}}{\partial a_\ell} \\ & + 2\, \underline{A}^{-1} \underline{H}\, E\left\{ \frac{\partial \bar{\underline{x}}}{\partial a_k}\, \frac{\partial \bar{\underline{x}}^T}{\partial a_\ell} \,\middle|\, \underline{a}_* \right\} \underline{H}^T + 2\, \underline{A}^{-1}\, \frac{\partial \underline{H}}{\partial a_k}\, E\left\{ \bar{\underline{x}}\, \frac{\partial \bar{\underline{x}}^T}{\partial a_\ell} \,\middle|\, \underline{a}_* \right\} \underline{H}^T \\ & + 2\, \underline{A}^{-1} \underline{H}\, E\left\{ \frac{\partial \bar{\underline{x}}}{\partial a_k}\, \bar{\underline{x}}^T \,\middle|\, \underline{a}_* \right\} \frac{\partial \underline{H}^T}{\partial a_\ell} + 2\, \underline{A}^{-1}\, \frac{\partial \underline{H}}{\partial a_k}\, E\left\{ \bar{\underline{x}}\, \bar{\underline{x}}^T \,\middle|\, \underline{a}_* \right\} \frac{\partial \underline{H}^T}{\partial a_\ell} \Bigg] \end{aligned} \tag{3.4.27}$$

$$E\left\{\frac{\partial \hat{\underline{x}}}{\partial a_\ell}\,\middle|\,\underline{a}_*\right\} = [\underline{I}-\underline{K}\underline{H}]\left[\underline{\Phi}\,E\left\{\frac{\partial \hat{\underline{x}}}{\partial a_\ell}\,\middle|\,\underline{a}_*\right\} + \frac{\partial \underline{\Phi}}{\partial a_\ell}\,E\left\{\hat{\underline{x}}\,\middle|\,\underline{a}_*\right\} + \frac{\partial \underline{B}}{\partial a_\ell}\,\underline{u}\right] - \underline{K}\,\frac{\partial \underline{H}}{\partial a_\ell}\,E\left\{\hat{\underline{x}}(j)\,\middle|\,\underline{a}_*\right\} \tag{3.4.29}$$

where the time index j is included in equation (3.4.29) to distinguish it from the term $E\left\{\hat{\underline{x}}\,\middle|\,\underline{a}_*\right\}$, which is understood to be evaluated at time (j-1).

After the measurement, the changes to the score relations are

$$\frac{\partial \underline{P}}{\partial a_\ell} = [\underline{I}-\underline{K}\underline{H}]\,\frac{\partial \underline{M}}{\partial a_\ell}\,[\underline{I}-\underline{K}\underline{H}]^T - \underline{P}\,\frac{\partial \underline{H}^T}{\partial a_\ell}\,\underline{K}^T - \underline{K}\,\frac{\partial \underline{H}}{\partial a_\ell}\,\underline{P} \tag{3.4.33}$$

$$\frac{\partial \hat{\underline{x}}}{\partial a_\ell} = [\underline{I}-\underline{K}\,\underline{H}]\left[\frac{\partial \overline{\underline{x}}}{\partial a_\ell} + \left(\frac{\partial \underline{M}}{\partial a_\ell}\,\underline{H}^T + \underline{M}\,\frac{\partial \underline{H}^T}{\partial a_\ell}\right)\underline{\eta}\right] - \underline{K}\,\frac{\partial \underline{H}}{\partial a_\ell}\,\hat{\underline{x}} \tag{3.4.34}$$

where $\underline{P}$ and $\hat{\underline{x}}$ are evaluated at time j. The conditional information matrix equations are changed by

$$\begin{aligned} E\left\{\frac{\partial \hat{\underline{x}}}{\partial a_k}\,\frac{\partial \hat{\underline{x}}^T}{\partial a_\ell}\,\middle|\,\underline{a}_*\right\} &= [\underline{I}-\underline{K}\,\underline{H}]\,E\left\{\frac{\partial \overline{\underline{x}}}{\partial a_k}\,\frac{\partial \overline{\underline{x}}^T}{\partial a_\ell}\,\middle|\,\underline{a}_*\right\}\,[\underline{I}-\underline{K}\,\underline{H}]^T \\ &+ \left[[\underline{I}-\underline{K}\,\underline{H}]\left(\frac{\partial \underline{M}}{\partial a_k}\,\underline{H}^T + \underline{M}\,\frac{\partial \underline{H}^T}{\partial a_k}\right) - \underline{K}\,\frac{\partial \underline{H}}{\partial a_k}\,\underline{M}\,\underline{H}^T\right]\underline{A}^{-1}\cdot \\ &\cdot\left[[\underline{I}-\underline{K}\underline{H}]\left(\frac{\partial \underline{M}}{\partial a_\ell}\,\underline{H}^T + \underline{M}\,\frac{\partial \underline{H}^T}{\partial a_\ell}\right) - \underline{K}\,\frac{\partial \underline{H}}{\partial a_\ell}\,\underline{M}\,\underline{H}^T\right]^T \\ &+ [\underline{I}-\underline{K}\,\underline{H}]\,E\left\{\frac{\partial \overline{\underline{x}}}{\partial a_k}\,\overline{\underline{x}}^T\,\middle|\,\underline{a}_*\right\}\,\frac{\partial \underline{H}^T}{\partial a_\ell}\,\underline{K}^T + \underline{K}\,\frac{\partial \underline{H}}{\partial a_k}\,E\left\{\overline{\underline{x}}\,\frac{\partial \overline{\underline{x}}^T}{\partial a_\ell}\,\middle|\,\underline{a}_*\right\}[\underline{I}-\underline{K}\underline{H}]^T \\ &+ \underline{K}\,\frac{\partial \underline{H}}{\partial a_k}\,E\left\{\overline{\underline{x}}\,\overline{\underline{x}}^T\,\middle|\,\underline{a}_*\right\}\,\frac{\partial \underline{H}^T}{\partial a_\ell}\,\underline{K}^T \end{aligned} \tag{3.4.35}$$

$$E\left\{\frac{\partial \hat{\underline{x}}}{\partial a_\ell}\,\hat{\underline{x}}^T \middle| \underline{a}_*\right\} = [\underline{I}-\underline{K}\,\underline{H}]\left[E\left\{\frac{\partial \overline{\underline{x}}}{\partial a_\ell}\,\overline{\underline{x}}^T \middle| \underline{a}_*\right\} + \left[\frac{\partial \underline{M}}{\partial a_\ell}\,\underline{H}^T + \underline{M}\,\frac{\partial \underline{H}^T}{\partial a_\ell}\right]\underline{K}^T\right]$$

$$- \underline{K}\,\frac{\partial \underline{H}}{\partial a_\ell}\left[E\{\overline{\underline{x}}\,\overline{\underline{x}}^T \middle| \underline{a}_*\} + \underline{M}\,\underline{H}^T\,\underline{K}^T\right] \qquad (3.4.37)$$

If the approximation of section 5.1 is utilized, then the resulting one-step conditional information matrix is computed as

$$J^1{}_{k\ell} = \tfrac{1}{2}\,\mathrm{tr}\left\{\underline{A}^{-1}\,\frac{\partial \underline{A}}{\partial a_k}\;\underline{A}^{-1}\,\frac{\partial \underline{A}}{\partial a_\ell}\right\}$$

$$+ \left[\frac{\partial \overline{\underline{x}}^T}{\partial a_k}\,\underline{H}^T + \overline{\underline{x}}^T\,\frac{\partial \underline{H}^T}{\partial a_k}\right]\underline{A}^{-1}\left[\underline{H}\,\frac{\partial \overline{\underline{x}}}{\partial a_\ell} + \frac{\partial \underline{H}}{\partial a_\ell}\,\overline{\underline{x}}\right]$$

(5.1.2)

Appendix B: Discrete Representation of Continuous Dynamic Systems

Let the continuous-time model of the system be given as the stochastic differential equation

$$d\underline{x}(t) = \underline{F}(t)\underline{x}(t)\ dt + \underline{B}(t)\underline{u}(t)\ dt + \underline{G}(t)d\underline{\beta}(t) \quad \text{(B-1)}$$

or the equivalent, though mathematically less precise, linear differential equation

$$\dot{\underline{x}}(t) = \underline{F}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) + \underline{G}(t)\underline{w}(t) \quad \text{(B-2)}$$

where $\underline{w}(t)$ is a Gaussian white noise with zero mean and covariance $\underline{Q}(t)\ \delta(t-\tau)$ with $\underline{Q}(t)$ chosen to duplicate the low frequency power spectral density of the actual noise entering the system:

$$E\left\{\underline{w}(t)\right\} = \underline{0} \quad \text{(B-3)}$$

$$E\left\{\underline{w}(t_1)\underline{w}^T(t_2)\right\} = \underline{Q}(t_1)\ \delta(t_1-t_2) \quad \text{(B-4)}$$

This appendix will develop the discrete system model that is equivalent to this continuous-time description, as seen in sampled data fashion. It is assumed that the measurements are taken at discrete instants, and are of the form

$$\underline{z}(t_i) = \underline{H}(t_i)\underline{x}(t_i) + \underline{v}(t_i) \quad \text{(B-5)}$$

Furthermore, it is assumed that a digital computer will provide the control input, so that $\underline{u}(t)$ will be piecewise constant: a measurement would be taken, the information processed, and a control input created and held constant until the following sample time. This development is based upon that originally presented by Widnall (1968).

The solution to equation (B-1) or (B-2) is

$$\underline{x}(t) = \underline{\Phi}(t,t_i)\underline{x}(t_i) + \int_{t_i}^{t} \underline{\Phi}(t,\tau)\underline{B}(\tau)\underline{u}(\tau)\, d\tau + \\ + \int_{t_i}^{t} \underline{\Phi}(t,\tau)\underline{G}(\tau)d\underline{w}(\tau) \qquad \text{(B-6)}$$

where the state transition matrix $\underline{\Phi}(t,t_i)$ satisfies the differential equation and initial condition

$$\frac{d}{dt}\underline{\Phi}(t,t_i) = \underline{F}(t)\underline{\Phi}(t,t_i) \qquad \text{(B-7)}$$

$$\underline{\Phi}(t_i,t_i) = \underline{I} \qquad \text{(B-8)}$$

Since the control is held constant over a given sample period, the solution within a single sample period is

$$\underline{x}(t) = \underline{\Phi}(t,t_i)\underline{x}(t_i) + \underline{D}(t,t_i)\underline{u}(t_i) + \int_{t_i}^{t} \underline{\Phi}(t,\tau)\underline{G}(\tau)d\underline{w}(\tau) \qquad \text{(B-9)}$$

where the matrix $\underline{D}(t,t_i)$ is defined as

$$\underline{D}(t,t_i) = \int_{t_i}^{t} \underline{\Phi}(t,\tau)\underline{B}(\tau)\, d\tau \qquad \text{(B-10)}$$

Differentiating this with respect to time yields the equivalent defining relations

$$\frac{d}{dt}\underline{D}(t,t_i) = \underline{B}(t) + \underline{F}(t)\underline{D}(t,t_i) \qquad \text{(B-11)}$$

$$\underline{D}(t_i,t_i) = \int_{t_i}^{t_i} \underline{\Phi}(t,\tau)\underline{B}(\tau)\, d\tau = \underline{0} \qquad \text{(B-12)}$$

Thus, if an equation of the form

$$\underline{x}(i+1) = \underline{\Phi}(i+1,i)\underline{x}(i) + \underline{B}(i)\underline{u}(i) + \underline{G}(i)\underline{w}(i) \qquad \text{(B-13)}$$

is to duplicate the state at the (i+1)-th sample time, $\underline{x}(t_{i+1})$, as given by equation (B-9), then it can be seen that

$$\underline{\Phi}(i+1,i) = \underline{\Phi}(t_{i+1},t_i) \tag{B-14}$$

$$\underline{B}(i) = \underline{D}(t_{i+1},t_i) \tag{B-15}$$

For filter applications, the last term in (B-13) need not be evaluated. Rather, an expression for the covariance of its contribution is required:

$$\underline{G}(i)\underline{Q}(i)\underline{G}^T(i) = \int_{t_i}^{t_{i+1}} \underline{\Phi}(t_{i+1},\tau)\underline{G}(\tau)\underline{Q}(\tau)\underline{G}^T(\tau)\underline{\Phi}^T(t_{i+1},\tau)\, d\tau \tag{B-16}$$

$$= \underline{N}(t_{i+1},t_i) \tag{B-17}$$

where (B-17) serves as a definition of $\underline{N}(t_{i+1},t_i)$. As previously, a more convenient form for evaluations would be obtained by differentiating (B-16) to yield

$$\frac{d}{dt}\underline{N}(t,t_i) = \underline{F}(t)\underline{N}(t,t_i) + \underline{N}(t,t_i)\underline{F}^T(t) + \underline{G}(t)\underline{Q}(t)\underline{G}^T(t) \tag{B-18}$$

$$\underline{N}(t_i,t_i) = \underline{0} \tag{B-19}$$

Thus, to obtain the equivalent discrete model for a continuous-time system, equations (B-7), (B-11), and (B-18) are integrated forward from the initial conditions (B-8), (B-12), and (B-19) to the time t_{i+1}. The required matrices $\underline{\Phi}(i+1,i)$, $\underline{B}(i)$, and $\underline{G}(i)\underline{Q}(i)\underline{G}^T(i)$ are then determined from equations (B-14), (B-15), and (B-17) respectively.

The influence of uncertain parameters in $\underline{F}(t)$ or $\underline{B}(t)$ upon the discrete model can be expressed analytically in simple cases. In more complex situations, the dependence

can be found by numerical integration of these equations for various values of the parameters, from which functional relationships between the elements of $\underline{\Phi}(i+1,i)$, $\underline{B}(i)$, and $\underline{G}(i)\underline{Q}(i)\underline{G}^T(i)$ and the parameter values can be established. Since the system is assumed time invariant (and noise inputs assumed stationary) over N samples, a single set of integrations suffice for all sample periods; slowly varying parameters that are known functions of time can be treated functionally in the same manner as the uncertain parameters, though in many practical applications the known parameters are in fact time-invariant. For on-line applications, the functional relations would be approximated by curve fitting of piecewise linear or low order polynomial functions to the actual results of the integrations.

For certain applications in which the system matrices are slowly varying or time invariant, and in which the sample period is short compared to the system's natural transients, a first order approximation to the solution of the differential equations may be adequate. These approximations are:

$$\underline{\Phi}(i+1,i) \cong \underline{I} + \underline{F}(t_i)\,[t_{i+1} - t_i] \tag{B-20}$$

$$\underline{B}(i) \cong \underline{B}(t_i)\,[t_{i+1} - t_i] \tag{B-21}$$

$$\underline{G}(i)\underline{Q}(i)\underline{G}^T(i) \cong \underline{G}(t_i)\underline{Q}(t_i)\underline{G}^T(t_i)\,[t_{i+1} - t_i] \tag{B-22}$$

Appendix C : Approximate One-Step Recursions for Fixed-Length Memory State Estimators

The concept of forming a state estimate from the most recent N samples of data was presented in sections 2.1 (equations (2.1.86) to (2.1.94)) and 3.2 (equations (3.2.25) to (3.2.38)). If such a form were to be contemplated for on-line use, a means of approximating an N-step propagation every sample time by a single step propagation, as developed in section 5.2 for the parameter estimator, would be required. However, in section 5.2 the algebraic form of the terms to be added to and subtracted from the score and conditional information matrix were readily computable. For the case of a fixed-length memory state estimator, the exact form is too complex to be used conveniently. An approximation is proposed and shown to yield a conservative estimator.

The covariance propagations for the recursive, growing-length memory state estimator are

$$\underline{M}(i+1) = \underline{\Phi}(i+1,i)\underline{P}(i)\underline{\Phi}^{T}(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^{T}(i) \tag{C-1}$$

$$\underline{P}^{-1}(i) = \underline{M}^{-1}(i) + \underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{H}(i) \tag{C-2}$$

When there is no dynamic driving noise, the second term in equation (C-1) becomes zero. For this case of $\underline{Q}(i) = \underline{0}$ for all i, equation (2.1.91) demonstrates that the corresponding fixed-length memory filter equations are

$$\tilde{\underline{M}}(i+1) = \underline{\Phi}(i+1,i)\tilde{\underline{P}}(i)\underline{\Phi}^{T}(i+1,i) \tag{C-3}$$

$$\begin{aligned}\tilde{\underline{P}}^{-1}(i) = {} & \tilde{\underline{M}}^{-1}(i) + \underline{H}^{T}(i)\underline{R}^{-1}(i)\underline{H}(i) \\ & - \underline{\Phi}^{T}(i-N,i)\underline{H}^{T}(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i)\end{aligned} \tag{C-4}$$

The exact equations for these propagations are very complicated when the system is subjected to driving noise. Noting the similarity of the relations for the growing-length and fixed-length memory filters suggests the following approximation for the case of $\underline{Q}(i)$ nonzero:

$$\tilde{\underline{M}}^*(i+1) = \underline{\Phi}(i+1,i)\tilde{\underline{P}}^*(i)\underline{\Phi}^T(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^T(i) \tag{C-5}$$

$$\tilde{\underline{P}}^{*-1}(i) = \tilde{\underline{M}}^{*-1}(i) + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i) - \underline{\Phi}^T(i-N,i)\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i) \tag{C-6}$$

It will now be shown that, in fact, the recursions are

$$\tilde{\underline{M}}(i+1) = \underline{\Phi}(i+1,i)\tilde{\underline{P}}(i)\underline{\Phi}^T(i+1,i) + \underline{G}(i)\underline{Q}(i)\underline{G}^T(i) \tag{C-7}$$

$$\tilde{\underline{P}}^{-1}(i) = \tilde{\underline{M}}^{-1}(i) + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i) + \underline{C}(i) - \underline{\Phi}^T(i-N,i)\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i) \tag{C-8}$$

where $\underline{C}(i)$ is positive semidefinite. In other words, the result of using equations (C-5) and (C-6) is an estimate of $\tilde{\underline{P}}^{-1}(i)$ that satisfies

$$\tilde{\underline{P}}^{*-1}(i) \leq \tilde{\underline{P}}^{-1}(i) \tag{C-9}$$

or equivalently,

$$\tilde{\underline{P}}^*(i) \geq \tilde{\underline{P}}(i) \tag{C-10}$$

Thus, this method of updating will provide a "conservative" estimator, in that it will never believe its state estimate

to be more reliable than it really is.

To derive this result, the N-step propagations from time (i-N) to (i-1) and from (i-N+1) to i will be computed and compared. These will be denoted as $\underline{P}_{i-1}^{-1}(j)$ and $\underline{P}_{i}^{-1}(j)$ respectively; the former assumes that $\underline{z}(i-N)$ is the first measurement and the latter assumes $\underline{z}(i-N+1)$ is the first:

$$\underline{P}_{i-1}^{-1}(i-N) = \underline{M}^{-1}(i-N) + \underline{H}^{T}(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N) \tag{C-11}$$

$$\underline{P}_{i}^{-1}(i-N) = \underline{M}^{-1}(i-N) \tag{C-12a}$$

$$= \underline{P}_{i-1}^{-1}(i-N) - \underline{H}^{T}(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N) \tag{C-12b}$$

Now equations (2.1.60) and (2.1.61) are used to propagate these variables to just before the measurement at time (i-N+1), and (C-2) used to incorporate this measurement. Here the matrices are implicitly assumed to be evaluated at time (i-N) unless specifically noted, $\underline{\Phi}$ equals $\underline{\Phi}(i-N+1,i-N)$, $\underline{\Phi}^{-1}$equals $\underline{\Phi}(i-N,i-N+1)$, and $\underline{\Phi}^{-T}$ equals $\underline{\Phi}^{T}(i-N,i-N+1)$:

$$\underline{P}_{i}^{-1}(i-N+1) = \underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1} + \underline{H}^{T}(i-N+1)\underline{R}^{-1}(i-N+1)\underline{H}(i-N+1)$$
$$- \underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\left[\underline{Q}^{-1} + \underline{G}^{T}\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1} \underline{G}^{T}\underline{\Phi}^{-T}\underline{M}\underline{\Phi}^{-1} \tag{C-13}$$

$$\underline{P}_{i-1}^{-1}(i-N+1) = \underline{P}_{i}^{-1}(i-N+1) + \underline{\Phi}^{-T}\underline{H}^{T}\underline{R}^{-1}\underline{H}\underline{\Phi}^{-1}$$
$$- \underline{\Phi}^{-T}\left[\underline{M}^{-1} + \underline{H}^{T}\underline{R}^{-1}\underline{H}\right]\underline{\Phi}^{-1}\underline{G}\left[\underline{Q}^{-1} + \underline{G}^{T}\underline{\Phi}^{-T}\underline{P}_{i-1}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1}$$
$$\cdot\ \underline{G}^{T}\underline{\Phi}^{-T}\left[\underline{M}^{-1} + \underline{H}^{T}\underline{R}^{-1}\underline{H}\right]\underline{\Phi}^{-1}$$
$$+ \underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\left[\underline{Q}^{-1} + \underline{G}^{T}\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1} \underline{G}^{T}\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1} \tag{C-14}$$

The negative term in (C-14) is now expanded into four individual terms by separating the two multiplications by $(\underline{M}^{-1} + \underline{H}^T\underline{R}^{-1}\underline{H})$ into their component parts. One such term duplicates the last term in (C-14) except that the large inverse within it has $\underline{M}^{-1}(i\text{-}N)$ replaced by $\underline{P}_{i-1}^{-1}(i\text{-}N)$; this inverse will be expanded by a form of the matrix inversion lemma to allow a direct cancellation of the last term in (C-14). The particular form of the lemma is developed as

$$\underline{P} = (\underline{M}^{-1} + \underline{H}^T\underline{R}^{-1}\underline{H})^{-1} = \underline{M} - \underline{M}\underline{H}^T(\underline{H}\underline{M}\underline{H}^T + \underline{R})^{-1}\,\underline{H}\,\underline{M}$$

$$= \underline{M} - \underline{K}\underline{H}\underline{M}$$

$$= \underline{M} - \underline{P}\underline{H}^T\underline{R}^{-1}\underline{H}\underline{M} \qquad \text{(C-15)}$$

The term to be transformed is

$$\left[\underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{P}_{i-1}^{-1}\,\underline{\Phi}^{-1}\underline{G}\right]^{-1} =$$

$$= \left[\underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G} + \underline{G}^T\underline{\Phi}^{-T}\underline{H}^T\underline{R}^{-1}\underline{H}\underline{\Phi}^{-1}\underline{G}\right]^{-1} \qquad \text{(C-16)}$$

To use (C-15), make the identification

$$\text{"}\underline{M}^{-1}\text{"} = \underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}$$

$$\text{"}\underline{H}^T\text{"} = \underline{G}^T\underline{\Phi}^{-T}\underline{H}$$

$$\text{"}\underline{R}^{-1}\text{"} = \underline{R}^{-1}$$

and apply the lemma to express (C-16) as:

$$\left[\underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{P}_{i-1}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1} = \left[\underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1}$$

$$- \left[\underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{P}_{i-1}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1}\underline{G}^T\underline{\Phi}^{-T}\underline{H}^T\underline{R}^{-1}\underline{H}\underline{\Phi}^{-1}\underline{G} \cdot$$

$$\cdot \left[\underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1} \qquad \text{(C-17)}$$

Equation (C-17) is put into the expanded form of (C-14), and terms rearranged, to obtain

$$\underline{P}_i^{-1}(i-N+1) = \underline{P}_{i-1}^{-1}(i-N+1) - \underline{\Phi}^{-T}\underline{H}^T\underline{R}^{-1}\underline{H}\underline{\Phi}^{-1} + \underline{C}_i(i-N+1) \qquad \text{(C-18)}$$

$$\underline{C}_i(i-N+1) = \underline{\Phi}^{-T}\underline{H}^T\underline{R}^{-1}\underline{H}\underline{\Phi}^{-1}\underline{G}\left[\ \right]^{-1}\underline{G}^T\underline{\Phi}^{-T}\underline{P}_{i-1}^{-1}\underline{\Phi}^{-1}$$

$$+ \underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\left[\ \right]^{-1}\underline{G}^T\underline{\Phi}^{-T}\underline{H}^T\underline{R}^{-1}\underline{H}\underline{\Phi}^{-1}\left[\underline{I} - \underline{\chi}\underline{G}^T\right]^T \qquad \text{(C-19)}$$

where $[\ ]^{-1}$ represents $\left[\underline{Q}^{-1} + \underline{G}^T\underline{\Phi}^{-T}\underline{P}_{i-1}^{-1}\underline{\Phi}^{-1}\underline{G}\right]^{-1}$, and $\underline{\chi}$ is the matrix defined by equation (2.1.62), using $\underline{M}^{-1}$ in (2.1.61):

$$\underline{\chi} = \underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G}\left[\underline{G}^T\underline{\Phi}^{-T}\underline{M}^{-1}\underline{\Phi}^{-1}\underline{G} + \underline{Q}^{-1}\right]^{-1} \qquad \text{(C-20)}$$

From equation (C-19), $\underline{C}_i(i-N+1)$ will be positive semidefinite as long as the positive semidefiniteness of a matrix is not destroyed by multiplying it by $\lfloor \underline{I} - \underline{\chi}(i-N)\underline{G}^T(i-N)\rfloor^T$. This will be demonstrated in a more general context after an intermediate development.

Now equations (C-13) and (C-14) are propagated one step to obtain $\underline{P}_i^{-1}(i-N+2)$ and $\underline{P}_{i-1}^{-1}(i-N+2)$. After considerable algebraic manipulation, the result can be expressed as:

$$\underline{P}_i^{-1}(i-N+2) = \underline{P}_{i-1}^{-1}(i-N+2) - \underline{\Phi}^{-2T}\underline{H}^T\underline{R}^{-1}\underline{H}\underline{\Phi}^{-2} + \underline{C}_i(i-N+2) \tag{C-21}$$

where $\underline{\Phi}^{-2T}$ denotes $\underline{\Phi}^T(i-N,i-N+2)$. All negative terms appearing in $\underline{C}_i(i-N+2)$ can be combined with similar terms to form the groupings $[\underline{I} - \underline{\chi}(i-N)\underline{G}^T(i-N)]\underline{Y}$, $[\underline{I} - \underline{\chi}(i-N+1)\underline{G}^T(i-N+1)]\underline{Y}$, or transposes of these, where $\underline{Y}$ is some positive semidefinite matrix. Therefore, the positive semidefiniteness of $\underline{C}_i(i-N+2)$ depends upon $[\underline{I} - \underline{\chi}(j)\underline{G}^T(j)]$ being positive semidefinite. However, this is just the term used to propagate $\underline{P}^{-1}(j)$ to $\underline{M}^{-1}(j+1)$:

$$\underline{M}^{-1}(j+1) = \underline{\Phi}^T(j,j+1)\underline{P}^{-1}(j)\underline{\Phi}(j,j+1)[\underline{I} - \underline{\chi}(j)\underline{G}^T(j)]^T$$

$$= [\underline{I} - \underline{\chi}(j)\underline{G}^T(j)]\ \underline{\Phi}^T(j,j+1)\underline{P}^{-1}(j)\underline{\Phi}(j,j+1) \tag{C-22}$$

From section 2.1 it is known that $\underline{M}^{-1}(j+1)$ equals $[\underline{\Phi}^T(j,j+1)\underline{P}^{-1}(j)\underline{\Phi}(j,j+1)]$ if $\underline{Q} = \underline{0}$. Now, adding a finite amount of dynamic driving noise in a given sample period can neither increase $\underline{M}^{-1}(j+1)$ nor decrease it to zero. Therefore,

$$\underline{0} < [\underline{I} - \underline{\chi}(j)\underline{G}^T(j)] \leq \underline{I} \qquad \text{for } \underline{Q}(j) < \infty \tag{C-23}$$

and so $\underline{C}_i(i-N+2)$ is positive semidefinite. By continuing this development for N steps, the conclusion can be reached that, since (C-23) is true for all j, that there is a positive semidefinite $\underline{C}(i) = \underline{C}_i(i)$ such that:

$$\tilde{\underline{P}}^{-1}(i) = \tilde{\underline{M}}^{-1}(i) + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i) + \underline{C}(i)$$

$$- \underline{\Phi}^T(i-N,i)\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i) \tag{C-24}$$

thereby proving the result sought.

Since $\underline{\tilde{P}}^*(i) \geq \underline{\tilde{P}}(i)$, one can conclude that

$$\det\left[\underline{\tilde{P}}^*(i)\right] \geq \det\left[\underline{\tilde{P}}(i)\right] \qquad \text{(C-25)}$$

$$\text{tr}\left[\underline{\tilde{P}}^*(i)\right] \geq \text{tr}\left[\underline{\tilde{P}}(i)\right] \qquad \text{(C-26)}$$

$$\tilde{P}^*_{kk}(i) \geq \tilde{P}_{kk}(i) \qquad \text{(C-27)}$$

Equation (C-27) states that every diagonal term of $\underline{\tilde{P}}^*(i)$ is greater than or equal to the corresponding diagonal term of $\underline{\tilde{P}}(i)$. In general, however, no such statement can be made about the off-diagonal terms.

It would be desirable to derive an approximate update for the state in a form similar to that of equation (2.1.92), which pertains to the case of no dynamic noise in the system. The most apparent approximation, in the same vein as that described for the covariance, would be to replace $\underline{\mathcal{Y}}^{-1}(i,i-N+1)$ in equation (2.1.92) by $\underline{\tilde{P}}(i)$. This yields an expression for the term to be subtracted due to $\underline{z}(i-N)$ as

$$\underline{\tilde{P}}(i)\underline{\Phi}^T(i-N,i)\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\left[\underline{z}(i-N) - \underline{z}^*(i-N)\right]$$

where $\underline{z}^*(i-N)$ is the term subtracted from $\underline{z}(i-N)$ in the brackets of (2.1.92). Attempts at validating this approximation theoretically, as by comparison with the term due to $\underline{z}(i-N)$ in equation (4.2.24), have proven to be inconclusive. Even if adequate, the approximate form requires considerable computation, and the benefits of this procedure do not outweigh the disadvantages enough to make further evolution of the idea fruitful. If it is desired to use a fixed-length state estimator on-line, the first of the two on-line conceptualizations of section 5.2 could be employed. This implementation makes a parameter estimate only every N sample periods, and uses the standard

recursive filter propagation equations from the "initial conditions" given by equations (3.4.40) to (3.4.43), thereby circumventing the problem of evaluating the appropriate terms to be subtracted.

Appendix D: Recursive State Estimation Equations for $\ell n\ f(\underline{Z}|\underline{x},\underline{a})$ or $\ell n\ f(\underline{Z}_N|\underline{x},\underline{a})$ with $\underline{Q} = \underline{0}$

This appendix will develop the approximate state recursions for a likelihood function of $\ell n\ f(\underline{Z}_N(i)|\underline{x}(i),\underline{a})$. An analogous, but less complex, development readily follows for the case of $\ell n\ f(\underline{Z}(i)|\underline{x}(i),\underline{a})$, so it will not be presented.

From section 3.2.3, the state estimate is

$$\underline{x}^*(i)\Big|_{\underline{a}\to\underline{a}^*(i)} = \underline{\mathcal{J}}^{-1}(i,i-N+1;\underline{a}^*(i)) \sum_{j=i-N+1}^{i} \underline{\Phi}^T(j,i;\underline{a}^*(i))\underline{H}^T(j)\underline{R}^{-1}(j) \cdot$$

$$\cdot\left\{\underline{z}(j) + \underline{H}(j) \sum_{k=j+1}^{i} \underline{\Phi}(j,k;\underline{a}^*(i))\underline{B}(k-1;\underline{a}^*(i))\underline{u}(k-1)\right\} \tag{D-1}$$

where

$$\underline{\mathcal{J}}(i,i-N+1;\underline{a}^*(i)) = \sum_{j=i-N+1}^{i} \underline{\Phi}^T(j,i;\underline{a}^*(i))\underline{H}^T(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i;\underline{a}^*(i)) \tag{D-2}$$

To implement this as a full-scale recursion, one could solve for $\underline{x}^*(i;\underline{a}^*(i-1))$ in an N-step propagation when going from time (i-1) to time i, simultaneously evaluating the terms needed for the parameter likelihood equations. A scoring iteration can then be performed upon these equations to update the parameter estimate. However, a number of "local iterations" of this entire process may be required before the estimates converge onto the true maximum likelihood values $\underline{a}^*(i)$ and $\underline{x}^*(i;\underline{a}^*(i))$. Clearly, this is not acceptable for on-line procedures. Even a single N-step propagation every sample period, to define an approximate $\underline{x}^*(i;\underline{a}^*(i-1))$ would require too much computation time. What would be desirable would be a single-step propagation to approximate $\underline{x}^*(i;\underline{a}^*(i-1))$ from the values $\underline{x}^*(i-1;\underline{a}^*(i-2))$ and

$\underline{a}^*(i-1)$ available from the previous sample time. Since the parameters are essentially constant over N sample intervals, this method should provide adequate estimation for on-line use.

First, investigate the information matrix update: it is desired to obtain $\underline{\mathcal{J}}(i,i-N+1;\underline{a}^*(i-1))$ approximately from a one-step propagation of $\underline{\mathcal{J}}(i-1,i-N;\underline{a}^*(i-2))$, where

$$\underline{\mathcal{J}}(i,i-N+1;\underline{a}^*(i-1)) = \sum_{j=i-N+1}^{i} \underline{\Phi}^T(j,i;\underline{a}^*(i-1))\underline{H}^T(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i;\underline{a}^*(i-1)) \tag{D-3}$$

$$\underline{\mathcal{J}}(i-1,i-N;\underline{a}^*(i-2)) = \sum_{j=i-N}^{i-1} \underline{\Phi}^T(j,i-1;\underline{a}^*(i-2))\underline{H}^T(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i-1;\underline{a}^*(i-2)) \tag{D-4}$$

Equation (D-3) can be written as

$$\begin{aligned}\underline{\mathcal{J}}(i,i-N+1;\underline{a}^*(i-1)) &= \underline{\Phi}^T(i-1,i;\underline{a}^*(i-1)) \cdot \\ &\quad \cdot\left[\sum_{j=i-N+1}^{i-1} \underline{\Phi}^T(j,i-1;\underline{a}^*(i-1))\underline{H}^T(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i-1;\underline{a}^*(i-1))\right]\cdot \\ &\quad \cdot \underline{\Phi}(i-1,i;\underline{a}^*(i-1)) \\ &\quad + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)\end{aligned}$$

$$\begin{aligned}&\cong \underline{\Phi}^T(i-1,i;\underline{a}^*(i-1))\left[\sum_{j=i-N}^{i-1} \underline{\Phi}^T(j,i-1;\underline{a}^*(i-2)\underline{H}^T(j)\underline{R}^{-1}(j)\underline{H}(j)\underline{\Phi}(j,i-1;\underline{a}^*(i-2))\right. \\ &\quad \left.-\underline{\Phi}^T(i-N,i-1;\underline{a}^*(i-1))\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i-1;\underline{a}^*(i-1))\right]\cdot \\ &\quad \cdot \underline{\Phi}(i-1,i;\underline{a}^*(i-1)) \\ &\quad + \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)\end{aligned}$$

$$= \underline{\Phi}^T(i-1,i;\underline{a}^*(i-1))\underline{\mathcal{Q}}(i-1,i-N;\underline{a}^*(i-2))\underline{\Phi}(i-1,i;\underline{a}^*(i-1))$$
$$+ \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)$$
$$- \underline{\Phi}^T(i-N,i;\underline{a}^*(i-1))\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i;\underline{a}^*(i-1)) \quad (D-5)$$

The approximation used to obtain this equation is, explicitly, that

$$\underline{\Phi}(j,i-1;\underline{a}^*(i-1)) \cong \underline{\Phi}(j,i-1;\underline{a}^*(i-2)) \quad (D-6)$$

for j=(i-N), ... (i-1). It would be exact if $\underline{\mathcal{Q}}(i-1,i-N;\underline{a}^*(i-2))$ were instead $\underline{\mathcal{Q}}(i-1,i-N;\underline{a}^*(i-1))$, but the approximation is necessary to obtain a usable one-step recursive relation.

Now consider the state estimates

$$\underline{x}^*(i;\underline{a}^*(i-1)) = \underline{\mathcal{Q}}^{-1}(i,i-N+1;\underline{a}^*(i-1))\sum_{j=i-N+1}^{i}\underline{\Phi}^T(j,i;\underline{a}^*(i-1))\underline{H}^T(j)\underline{R}^{-1}(j)\cdot$$
$$\cdot\left\{\underline{z}(j)+\underline{H}(j)\sum_{k=j+1}^{i}\underline{\Phi}(j,k;\underline{a}^*(i-1))\underline{B}(k-1;\underline{a}^*(i-1))\underline{u}(k-1)\right\} \quad (D-7)$$

$$\underline{x}^*(i-1;\underline{a}^*(i-2)) = \underline{\mathcal{Q}}^{-1}(i-1,i-N;\underline{a}^*(i-2))\sum_{j=i-N}^{i-1}\underline{\Phi}^T(j,i-1;\underline{a}^*(i-2))\underline{H}^T(j)\underline{R}^{-1}(j)\cdot$$
$$\cdot\left\{\underline{z}(j)+\underline{H}(j)\sum_{k=j+1}^{i-1}\underline{\Phi}(j,k;\underline{a}^*(i-2))\underline{B}(k-1;\underline{a}^*(i-2))\underline{u}(k-1)\right\} \quad (D-8)$$

Equation (D-7) can be written in terms of (D-8) as:

$$\underline{x}^*(i;\underline{a}^*(i-1)) = \underline{\mathcal{Q}}^{-1}(i,i-N+1;\underline{a}^*(i-1))\underline{\Phi}^T(i-1,i;\underline{a}^*(i-1))\cdot$$

$$\cdot\left[\sum_{j=i-N+1}^{i-1}\underline{\Phi}^T(j,i-1;\underline{a}^*(i-1))\underline{H}^T(j)\underline{R}^{-1}(j)\cdot\right.$$

$$\cdot\left\{\underline{z}(j)+\underline{H}(j)\sum_{k=j+1}^{i-1}\underline{\Phi}(j,k;\underline{a}^*(i-1))\underline{B}(k-1;\underline{a}^*(i-1))\underline{u}(k-1)\right.$$

$$\left.\left.+\ \underline{H}(j)\underline{\Phi}(j,i;\underline{a}^*(i-1))\underline{B}(i-1;\underline{a}^*(i-1))\underline{u}(i-1)\right\}\right]$$

$$+\ \underline{\mathcal{Q}}^{-1}(i,i-N+1;\underline{a}^*(i-1))\underline{H}^T(i)\underline{R}^{-1}(i)\underline{z}(i)$$

$$=\underline{\mathcal{Q}}^{-1}(i,i-N+1;\underline{a}^*(i-1))\underline{\Phi}^T(i-1,i;\underline{a}^*(i-1))\left[\underline{\mathcal{Q}}(i-1,i-N;\underline{a}^*(i-1))\underline{x}^*(i-1;\underline{a}^*(i-1))\right.$$

$$-\ \underline{\Phi}^T(i-N,i-1;\underline{a}^*(i-1))\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{z}(i-N)$$

$$-\ \underline{\Phi}^T(i-N,i-1;\underline{a}^*(i-1))\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\cdot$$

$$\left.\cdot\sum_{k=i-N+1}^{i-1}\underline{\Phi}(i-N,k;\underline{a}^*(i-1))\underline{B}(k-1;\underline{a}^*(i-1))\underline{u}(k-1)\right]$$

$$+\ \underline{\mathcal{Q}}^{-1}(i,i-N+1;\underline{a}^*(i-1))\left[\underline{\mathcal{Q}}(i,i-N+1;\underline{a}^*(i-1))-\underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)\right]\cdot$$

$$\cdot\ \underline{B}(i-1;\underline{a}^*(i-1))\underline{u}(i-1)$$

$$+\ \underline{\mathcal{Q}}^{-1}(i,i-N+1;\underline{a}^*(i-1))\underline{H}^T(i)\underline{R}^{-1}(i)\underline{z}(i)$$

(D-9)

But, it is known that, from an exact development similar to that leading to (D-5),

$$\underline{\mathcal{Q}}(i-1,i-N;\underline{a}^*(i-1)) = \underline{\Phi}^T(i,i-1;\underline{a}^*(i-1))\Big[\underline{\mathcal{Q}}(i,i-N+1;\underline{a}^*(i-1))$$

$$- \underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)$$

$$+ \underline{\Phi}^T(i-N,i;\underline{a}^*(i-1))\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i;\underline{a}^*(i-1))\Big]\cdot$$

$$\cdot \underline{\Phi}(i,i-1;\underline{a}^*(i-1)) \qquad (D-10)$$

Substitute this into equation (D-9) to obtain, with all matrices understood to be evaluated with $\underline{a}^*(i-1)$

$$\underline{x}^*(i;\underline{a}^*(i-1)) = \Big[\underline{\mathcal{Q}}^{-1}(i,i-N+1)\underline{\Phi}^T(i-1,i)\underline{\Phi}^T(i,i-1)\underline{\mathcal{Q}}(i,i-N+1)\Big]\cdot$$

$$\cdot \underline{\Phi}(i,i-1)\underline{x}^*(i-1;\underline{a}^*(i-1))$$

$$- \underline{\mathcal{Q}}^{-1}(i,i-N+1)\Big[\underline{\Phi}^T(i-1,i)\underline{\Phi}^T(i,i-1)\Big]\cdot$$

$$\cdot\Big\{\underline{H}^T(i)\underline{R}^{-1}(i)\underline{H}(i)\underline{\Phi}(i,i-1)\underline{x}^*(i-1;\underline{a}^*(i-1)) + \underline{\Phi}(i-N,1)\cdot$$

$$\cdot\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\underline{H}(i-N)\underline{\Phi}(i-N,i)\underline{\Phi}(i,i-1)\underline{x}^*(i-1;\underline{a}^*(i-1))\Big\}$$

$$- \underline{\mathcal{Q}}^{-1}(i,i-N+1)\underline{\Phi}^T(i-1,i)\underline{\Phi}^T(i-N,i-1)\cdot$$

$$\cdot\underline{H}^T(i-N)\underline{R}^{-1}(i-N)\Big\{\underline{z}(i-N)+\underline{H}(i-N)\sum_{k=i-N+1}^{i-1}\underline{\Phi}(i-N,k)\underline{B}(k-1)\underline{u}(k-1)\Big\}$$

$$+ \underline{B}(i-1)\underline{u}(i-1)$$

$$- \underline{\mathcal{Q}}^{-1}(i,i-N+1)\underline{H}^T(i)\underline{R}^{-1}(i)\Big\{\underline{H}(i)\underline{B}(i-1)\underline{u}(i-1)-\underline{z}(i)\Big\} \qquad (D-11)$$

where the two terms in brackets are equal to $\underline{I}$. This equation can be rearranged to form an exact equation for $\underline{x}^*(i;\underline{a}^*(i-1))$ in terms of $\underline{x}^*(i-1;\underline{a}^*(i-1))$. To obtain a useful recursion,

make the same approximation as before, namely equation (D-6), and also

$$\underline{B}(j;\underline{a}^*(i-1)) \cong \underline{B}(j;\underline{a}^*(i-2)) \qquad \text{(D-12)}$$

for $j=(i-N), \ldots (i-1)$, so that $\underline{x}^*(i-1;\underline{a}^*(i-2))$ can be substituted for $\underline{x}^*(i-1;\underline{a}^*(i-1))$ to obtain the approximate recursion

$$\begin{aligned}\underline{x}^*(i;\underline{a}^*(i-1)) \cong\ & \underline{\Phi}(i,i-1;\underline{a}^*(i-1))\underline{x}^*(i-1;\underline{a}^*(i-2)) \\ & + \underline{B}(i-1;\underline{a}^*(i-1))\underline{u}(i-1) \\ & + \underline{\mathcal{Q}}^{-1}(i,i-N+1;\underline{a}^*(i-1))\Big[\underline{H}^T(i)\underline{R}^{-1}(i)[\underline{z}(i) - \underline{z}^*(i)] \\ & - \underline{\Phi}^T(i-N,i;\underline{a}^*(i-1))\underline{H}^T(i-N)\underline{R}^{-1}(i-N)[\underline{z}(i-N) - \underline{z}^*(i-N)]\Big]\end{aligned} \qquad \text{(D-13)}$$

where

$$\begin{aligned}\underline{z}^*(i) =\ & \underline{H}(i)\Big[\underline{\Phi}(i,i-1;\underline{a}^*(i-1))\underline{x}^*(i-1;\underline{a}^*(i-2)) \\ & + \underline{B}(i-1;\underline{a}^*(i-1))\underline{u}(i-1)\Big]\end{aligned}$$

$$\begin{aligned}\underline{z}^*(i-N) =\ & \underline{H}(i-N)\Big[\underline{\Phi}(i-N,i-1;\underline{a}^*(i-1))\underline{x}^*(i-1;\underline{a}^*(i-2)) \\ & - \sum_{k=i-N+1}^{i-1} \underline{\Phi}(i-N,k;\underline{a}^*(i-1))\underline{B}(k-1;\underline{a}^*(i-1))\underline{u}(k-1)\Big]\end{aligned}$$

and where the information matrix satisfies the recursion given by (D-5). The computational advantage of these equations is that they update the state estimate immediately after a measurement, without requiring a new parameter estimate to do so. Subsequently, a new estimate of $\underline{a}$ can be calculated in between measurements to use at the next sample

time. Furthermore, the terms $\underline{z}^*(i)$ and $\underline{z}^*(i-N)$, interpretted as the expected values of the measurements at times i and (i-N) respectively, given the value $\underline{x}^*(i-1;\underline{a}^*(i-2))$, can be computed between sample times (i-1) and i, and be ready for use at the time the measurement $\underline{z}(i)$ is to be incorporated into the estimate. Finally, the summation in the $\underline{z}^*(i-N)$ can be replaced with a recursion, as given by equation (2.1.94).

A higher order recursive approximation can be generated by using $\underline{x}^*(i-1;\underline{a}^*(i-1))$ to generate $\underline{x}^*(i;\underline{a}^*(i-1))$ and $\underline{a}^*(i)$, then using $\underline{a}^*(i)$ to re-evaluate $\underline{x}^*(i)$ as $\underline{x}^*(i;\underline{a}^*(i))$, and then continuing in a similar manner. This provides a form of local iteration, but requires more computation, and practical experience with this technique often reveals that the benefits gained do not warrant the additional computer load.

Appendix E: Evaluation of $\underline{J}[\,i\,,\underline{a}_t\,]$ for the Likelihood Function $\ell n\, f(\underline{Z}(i)|\underline{a})$

In sections 3.4 and 4.1, there was need to derive an expression for $\underline{J}[\,i\,,\underline{a}_t\,]$ for a likelihood function of $\ell n\, f(\underline{Z}(i)|\underline{a})$. This is a p-by-p matrix, whose k-ℓ-th entry is

$$J_{k\ell}[\,i\,,\underline{a}_t\,] = E\left\{s_k[\underline{Z}(i),\underline{a}_t]\; s_\ell[\underline{Z}(i),\underline{a}_t]\,\middle|\,\underline{a}_t\right\} \qquad \text{(E-1)}$$

where the ℓ-th component of the score is defined as

$$s_\ell[\underline{Z}(i),\underline{a}_t] = \sum_{j=1}^{i}\left[\frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell}\,\underline{H}^T(j)\underline{A}^{-1}(j)\underline{r}(j) - \tfrac{1}{2}\,\mathrm{tr}\left\{\left[\underline{A}^{-1}(j)-\underline{A}^{-1}(j)\underline{r}(j)\underline{r}^T(j)\underline{A}^{-1}(j)\right]\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}\right] \qquad \text{(E-2)}$$

It will be necessary in this derivation to evaluate the form $E\{[\underline{\beta}^T\underline{C}\,\underline{\beta}][\underline{\beta}^T\underline{C}'\underline{\beta}]\}$ for $\underline{\beta}$ equal to a zero mean Gaussian vector with covariance $\underline{B}$, and for $\underline{C}$ and $\underline{C}'$ arbitrary weighting matrices. To simplify the evaluation, let $\underline{\alpha}$ be a vector, each of whose components are independent Gaussian vectors with mean zero and variance one. In other words, $\underline{\alpha}$ is a zero mean Gaussian vector with covariance $\underline{I}$, defined in terms of $\underline{\beta}$ by the transformation

$$\underline{\beta} = \sqrt{\underline{B}}\;\underline{\alpha} \qquad \text{(E-3)}$$

This transformation is valid, since, if $\underline{\beta}$ is zero mean and Gaussian, so is $\underline{\alpha}$, and

$$E\{\underline{\beta}\,\underline{\beta}^T\} = \sqrt{\underline{B}}\;E\{\underline{\alpha}\,\underline{\alpha}^T\}\;\sqrt{\underline{B}}^T = \sqrt{\underline{B}}\sqrt{\underline{B}}^T = \underline{B} \qquad \text{(E-4)}$$

as desired. Using this transformation, one can write:

$$[\underline{\beta}^T \underline{C} \underline{\beta}]\ [\underline{\beta}^T \underline{C}' \underline{\beta}] = [\underline{\alpha}^T (\sqrt{\underline{B}}^T \underline{C} \sqrt{\underline{B}})\, \underline{\alpha}]\ [\underline{\alpha}^T (\sqrt{\underline{B}}^T \underline{C}' \sqrt{\underline{B}})\, \underline{\alpha}]$$

$$= [\underline{\alpha}^T \underline{A}\, \underline{\alpha}]\ [\underline{\alpha}^T \underline{A}' \underline{\alpha}] \qquad \text{(E-5)}$$

Thus the term to be evaluated is

$$E\left\{ [\underline{\alpha}^T \underline{A}\, \underline{\alpha}]\ [\underline{\alpha}^T \underline{A}' \underline{\alpha}] \right\} = \sum_i \sum_j \sum_k \sum_\ell E\left\{ \lambda_i A_{ij} \alpha_j \alpha_k A'_{k\ell} \alpha_\ell \right\} \qquad \text{(E-6)}$$

Because the first and third central moments of a Gaussian variable are zero, the term within the expectation on the right-hand side of equation (E-6) can only be nonzero when the indices are pairwise equal, or all equal. Now, coordinates can be rotated to make $\underline{A}$ or $\underline{A}'$ diagonal, without affecting the independence of the components of $\underline{\alpha}$ since its covariance is $\underline{I}$, to obtain

$$E\left\{ [\underline{\alpha}^T \underline{A}\, \underline{\alpha}]\ [\underline{\alpha}^T \underline{A}' \underline{\alpha}] \right\} = E\left\{ \sum_i \sum_j \lambda_i \lambda_j'\, \alpha_i^2 \alpha_j^2 \right\}$$

$$= \sum_i \sum_j \lambda_i \lambda_j'\, E\left\{ \alpha_i^2 \alpha_j^2 \right\} \qquad \text{(E-7)}$$

where λ_i and λ_j' are, respectively, eigenvalues of $\underline{A}$ and $\underline{A}'$. Now if i is not equal to j, the expectation in (E-7) for α_i and α_j independent is $E\{\alpha_i^2\}\ E\{\alpha_j^2\} = \sigma^2 \sigma^2 = 1$, whereas for i equal to j, $E\{\alpha_i^4\} = 3\sigma^2 = 3$. Thus, (E-7) becomes

$$E\left\{ [\underline{\alpha}^T \underline{A}\, \underline{\alpha}]\ [\underline{\alpha}^T \underline{A}' \underline{\alpha}] \right\} = \sum_i \sum_j (\lambda_i \lambda_j' + 2\, \lambda_i \lambda_i' \delta_{ij})$$

$$= \sum_i \lambda_i \sum_j \lambda_j' + 2 \sum_i \sum_j \lambda_i \lambda_i' \delta_{ij}$$

$$= (\text{tr}\ \underline{A})\ (\text{tr}\ \underline{A}') + 2\ \text{tr}\ \underline{A}\, \underline{A}' \qquad \text{(E-8)}$$

since $\text{tr}\ \underline{A} = \sum_i \lambda_i$, $\text{tr}\ \underline{A}' = \sum_i \lambda_i'$, and $\text{tr}\ \underline{A}\, \underline{A}' = \sum_i \lambda_i \lambda_i'$. The terms in (E-8) can be evaluated by means of the definition of $\underline{A}$ and $\underline{A}'$ given in equation (E-5):

$$\text{tr } \underline{A} = \text{tr } [\sqrt{\underline{B}}^T \underline{C} \sqrt{\underline{B}}] = \text{tr } [\underline{C} \sqrt{\underline{B}} \sqrt{\underline{B}}^T]$$
$$= \text{tr } \underline{C}\,\underline{B} \qquad \text{(E-9)}$$

and similarly for tr $\underline{A}'$, and

$$\text{tr } \underline{A}\,\underline{A}' = \text{tr } [\sqrt{\underline{B}}^T \underline{C} \sqrt{\underline{B}}\ \sqrt{\underline{B}}^T \underline{C}' \sqrt{\underline{B}}]$$
$$= \text{tr } \underline{C}\,\underline{B}\,\underline{C}'\,\underline{B} \qquad \text{(E-10)}$$

It is thereby concluded that, for $\underline{\beta}$ a zero mean Gaussian vector with covariance $\underline{B}$, that the following relation holds for arbitrary $\underline{C}$ and $\underline{C}'$:

$$E\left\{ [\underline{\beta}^T \underline{C}\ \underline{\beta}] [\underline{\beta}^T \underline{C}' \underline{\beta}] \right\} = (\text{tr } \underline{C}\,\underline{B})\ (\text{tr } \underline{C}'\,\underline{B}) + 2 \text{ tr } \underline{C}\,\underline{B}\,\underline{C}'\,\underline{B} \qquad \text{(E-11)}$$

This is the relation needed in the following derivation.

Section 4.1 showed that, provided the parameter vector assumes its true value, $\underline{a}_t$, the measurement residuals are normally distributed random variables and independent from one sample time to the next:

$$E\left\{ \underline{r}(j) \middle| \underline{a}_t \right\} = E\left\{ [\underline{z}(j) - \underline{H}(j)\underline{\bar{x}}(j)] \middle| \underline{a}_t \right\} = \underline{0} \qquad \text{(E-12)}$$

$$E\left\{ \underline{r}(j)\underline{r}^T(k) \middle| \underline{a}_t \right\} = \underline{A}(j)\,\delta_{jk} \qquad \text{(E-13)}$$

where $\underline{A}(j)$ signifies the term $[\underline{H}(j)\underline{M}(j;\underline{a}_t)\underline{H}^T(j) + \underline{R}(j)]$. This fact and equation (E-12) provide the means of evaluating the conditional information matrix.

For simplicity, the term $E\left\{ s^1{}_k[\underline{Z}(j), \underline{a}_t]\, s^1{}_\ell [\underline{Z}(j), \underline{a}_t] \middle| \underline{a}_t \right\}$ will be evaluated. By use of equation (E-13), it can readily be shown that:

$$E\left\{ s_k[\underline{Z}(i),\underline{a}_t]\ s_\ell\ [\underline{Z}(i),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} =$$

$$= \sum_{j=1}^{i} E\left\{ s^1{}_k\,[\underline{Z}(j),\underline{a}_t]\ s^1{}_\ell\ [\underline{Z}(j),\underline{a}_t] \,\middle|\, \underline{a}_t \right\} \qquad \text{(E-14)}$$

so that the general result is in fact achieved. Thus, the term to be evaluated is $E\left\{[\phi_k + \psi_k]\ [\phi_\ell + \psi_\ell] \,\middle|\, \underline{a}_t\right\}$, where

$$\phi_\ell = \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell}\, \underline{H}^T(j)\underline{A}^{-1}(j)\underline{r}(j) \qquad \text{(E-15)}$$

$$\psi_\ell = -\tfrac{1}{2}\ \mathrm{tr}\left\{\left[(\underline{A}^{-1}(j) - \underline{A}^{-1}(j)\underline{r}(j)\underline{r}^T(j)\underline{A}^{-1}(j)\right]\ \frac{\partial \underline{A}(j)}{\partial a_\ell}\right\} \qquad \text{(E-16)}$$

By straightforward algebra and using the fact that tr $(\underline{XY})$ = tr $(\underline{XY}^T)$ = tr$(\underline{YX})$ and that $\underline{x}^T\underline{A}\underline{y}$ = tr$(\underline{A}\underline{y}\underline{x}^T)$, it is readily shown that

$$E\left\{ \phi_k\, \phi_\ell \,\middle|\, \underline{a}_t \right\} = \mathrm{tr}\left[\underline{A}^{-1}(j)\underline{H}(j) E\left\{ \frac{\partial \underline{\bar{x}}(j)}{\partial a_k}\ \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell} \,\middle|\, \underline{a}_t \right\} \underline{H}^T(j) \right. \qquad \text{(E-17)}$$

Similarly, with the additional knowledge that the first and third order moments of $\underline{r}(j)$ are zero, the cross terms become

$$E\left\{ \phi_k\, \psi_\ell \,\middle|\, \underline{a}_t \right\} = E\left\{ \psi_k\, \phi_\ell \,\middle|\, \underline{a}_t \right\} = 0 \qquad \text{(E-18)}$$

Finally, $E\left\{ \psi_k\, \psi_\ell \,\middle|\, \underline{a}_t \right\}$ becomes

$$E\left\{ \psi_k\, \psi_\ell \,\middle|\, \underline{a}_t \right\} = (\tfrac{1}{4}\text{-}\tfrac{1}{4}\text{-}\tfrac{1}{4})\ \mathrm{tr}\left\{ \underline{A}^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_k} \right\} \mathrm{tr}\left\{ A^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_\ell} \right\}$$

$$+ \tfrac{1}{4}\, E\left\{ \underline{r}^T(j)\underline{A}^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_k}\ \underline{A}^{-1}(j)\underline{r}(j) \cdot \right.$$

$$\left. \cdot\ \underline{r}^T(j)\underline{A}^{-1}(j)\ \frac{\partial \underline{A}(j)}{\partial a_\ell}\ \underline{A}^{-1}(j)\underline{r}(j) \,\middle|\, \underline{a}_t \right\} \qquad \text{(E-19)}$$

Now equation (E-11) can be utilized by identifying $\underline{\beta} = \underline{r}$, $\underline{C} = \underline{A}^{-1}\, \partial\underline{A}/\partial a_k\, \underline{A}^{-1}$, $\underline{C}' = \underline{A}^{-1}\, \partial\underline{A}/\partial a_\ell\, \underline{A}^{-1}$, $E\{\underline{\beta}\,\underline{\beta}^T|\underline{a}_t\} = \underline{A}$, to yield

$$E\left\{\psi_k \psi_\ell \,\middle|\, \underline{a}_t\right\} = -\tfrac{1}{4}\ \mathrm{tr}\left\{\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_k}\right\}\ \mathrm{tr}\left\{\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}$$

$$+\tfrac{1}{2}\ \mathrm{tr}\left\{\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_k}\,\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}$$

$$+\tfrac{1}{4}\ \mathrm{tr}\left\{\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_k}\right\}\ \mathrm{tr}\left\{\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\}$$

$$=\tfrac{1}{2}\ \mathrm{tr}\left\{\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_k}\,\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_\ell}\right\} \qquad \text{(E-20)}$$

Combining equations (E-17) to (E-20) results in

$$E\left\{s^1{}_k\,[\underline{Z}(j),\underline{a}_t]\ s^1{}_\ell\,[\underline{Z}(j),\underline{a}_t]\,\middle|\,\underline{a}_t\right\} =$$

$$=\tfrac{1}{2}\ \mathrm{tr}\left[\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_k}\,\underline{A}^{-1}(j)\,\frac{\partial \underline{A}(j)}{\partial a_\ell}\right.$$

$$\left.+\,2\,\underline{A}^{-1}(j)\underline{H}(j)E\left\{\frac{\partial \underline{\bar{x}}(j)}{\partial a_k}\,\frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell}\,\middle|\,\underline{a}_t\right\}\underline{H}^T(j)\right] \qquad \text{(E-21)}$$

Appendix F: Computational Requirements

In Chapter V, a hypothetical problem was posed, and the number of multiplications, additions, subtractions, and inversions ascertained for various estimator implementations. This appendix develops these evaluations in terms of the dimensions of the vector quantities involved. These dimensions are:

n = dimension of state $\underline{x}$
r = dimension of control $\underline{u}$
s = dimension of dynamic noise $\underline{w}$
m = dimension of measurement $\underline{z}$
p = dimension of parameter vector $\underline{a}$

It will be convenient to separate p into $p = p_A + p_B$, where p_B is the number of parameters whose effects are confined to $\underline{B}$, whereas p_A corresponds to the parameters that affect both $\underline{\Phi}$ and $\underline{B}$ in general.

For these evaluations, symmetry is not exploited fully: only $\underline{J}$ is calculated in lower triangular form. Other symmetric matrices are evaluated fully, and only when a matrix and its transpose are both to be evaluated does transpostion replace the reevaluation of a term. Terms that recur in a number of equations are assumed to be stored rather than regenerated.

The required computations will be divided according to the categories of section 3.4: propagation between measurement times of the state, the score, and the conditional information matrix; measurement update of $\hat{\underline{x}}$, $\underline{s}$, and $\underline{J}$; $\underline{s}$ and $\underline{J}$ terms to be added at the end of the N-step recursion; formation of $\underline{s}$, $\underline{J}$, and $\underline{a}^*$. To make the listings concise, M, A, S, and I will denote number of multiplications, additions, subtractions, and inversions respectively.

The following evaluations are for the on-line conceptualization that processes a parameter estimate every

sample period. Alterations for the other forms will be discussed subsequently.

(1) Time Propagation of State

M: $2n^3 + n^2\ (m+s+1) + n\ (m^2+m+r+s) + m^2$

A: $2n^3 + n^2\ (m+s) + n\ (m^2+r+s^2-s-2) - m$

S: m

I: $1\ (m \times m)$

(2) Time Propagation of Score

M: $m^2 + p_A \left[6n^3+n^2\ (m+2) + n\ (m^2+m+r+1) + (m^2+1)\right]$
$+ p_B \left[n^2 + n\ (m+r+1)\right]$

A: $p_A \left[6n^3 + n^2\ (m-4) + n\ (m^2+r) - (m+2)\right]$
$+ p_B \left[n^2 + n\ (m+r) - (m+1)\right]$

S: $m^2 + p_A$

I: 0

(3) Time Propagation of Conditional Information Matrix

M: $2n^3 + 3n^2\ r + n r^2 + p_A \left[4n^3 + 7n^2 r + n r^2 + m\right]$
$+ p_B \left[2n^3 + 5n^2 r + n r^2\right]$
$+ \tfrac{1}{2}\ p_A\ (p_A+1) \left[6n^3 + n^2\ (m+5r) + n\ (m^2+r^2) + (m^3+2m^2+1)\right]$
$+ \tfrac{1}{2}\ p_B\ (p_B+1) \left[2n^3 + n^2\ (m+3r) + n\ (m^2+r^2) + 2m^2\right]$
$+ \tfrac{1}{2}\ p_A p_B \left[8n^3 + n^2\ (2m+8r) + n\ (2m^2+2r^2) + (m^3+4m^2+1)\right]$

A: $2n^3 + n^2\,(3r-1) + n\,(r^2-2r) + p_A\left[4n^3 + n^2\,(7r-3) + n\,(r^2-4r) + (m^3-m^2)\right]$
$+ p_B\left[2n^3 + n^2\,(5r-2) + n\,(r^2-3r)\right]$
$+ \tfrac{1}{2}\,p_A\,(p_A+1)\left[6n^3 + n^2\,(m+5r-1) + n\,(m^2-m+r^2-3r) + (m^3-m)\right]$
$+ \tfrac{1}{2}\,p_B\,(p_B+1)\left[2n^3 + n^2\,(m+3r-1) + n\,(m^2-m+r^2-2r) + (m^2-m)\right]$
$+ \tfrac{1}{2}\,p_Ap_B\left[8n^3 + n^2\,(2m+8r-2) + n\,(2m^2-2m+2r^2-5r) + (m^3+m^2-2m)\right]$

S: 0

I: 0

(4) Measurement Update of State

M: $2n^3 + 4n^2m + n\,(m^2+m)$

A: $2n^3 + n^2\,(3m-3) + n\,(m^2-m+1) - m$

S: n^2

I: 0

(5) Measurement Update of Score

M: $p_A\left[2n^3 + n^2\,(m+1) + nm\right] + p_B\left[n^2\right]$

A: $p_A\left[2n^3 + n^2\,(m-1) + n - m\right] + p_B\left[n^2 - n\right]$

S: 0

I: 0

(6) Measurement Update of Conditional Information Matrix

M: $n^2m + nm^2 + p_A\left[n^3 + n^2m\right] + p_B\left[n^2\right]$
$+ \tfrac{1}{2}\,p_A\,(p_A+1)\left[2n^3 + 5n^2m + nm^2\right] + \tfrac{1}{2}\,p_B\,(p_B+1)\left[2n^3\right]$
$+ \tfrac{1}{2}\,p_Ap_B\left[4n^3 + 5n^2m + nm^2\right]$

A: $n^2m + n\ (m^2-m) + p_A\left[n^3 + n^2\ (m-1)\right] + p_B\left[n^3 - n^2\right]$
$+ \frac{1}{2}\ p_A\ (p_A+1)\left[2n^3 + n^2\ (5m-2) + n\ (m^2-5m)\right]$
$+ \frac{1}{2}\ p_B\ (p_B+1)\left[2n^3 - 2n^2\right]$
$+ \frac{1}{2}\ p_Ap_B\left[4n^3 + n^2\ (5m-4) + n\ (m^2-5m)\right]$

S: 0

I: 0

(7) Score Contribution at End of Interval

M: $p_A\left[n^3 + 1\right]$

A: $p_A\left[n^3 - n^2 + n - 1\right]$

S: 0

I: 1 (n x n)

(8) Conditional Information Matrix Contribution at End of Interval

M: $p_A\left[n^3\right] + \frac{1}{2}\ p_A\ (p_A+1)\left[2n^3 + 2\right] + \frac{1}{2}\ p_B\ (p_B+1)\left[n^3 + 2\right]$
$+ p_Ap_B\left[n^3 + 2\right]$

A: $p_A\left[n^3 - n^2\right] + \frac{1}{2}\ p_A\ (p_A+1)\left[2n^3 - 2n^2 + 2n\right]$
$+ \frac{1}{2}\ p_B\ (p_B+1)\left[n^3 - n^2 + n\right] + p_Ap_B\left[n^3 - n^2 + n\right]$

S: 0

I: 0

(9) Formation of Score and Conditional Information Matrix

M: 0

A: $p + \frac{1}{2}\ p\ (p+1)$

S: $p + \frac{1}{2}\ p\ (p+1)$

I: 0

(10) Formation of New Parameter Estimate

M: p^2

A: p^2

S: 0

I: 1 (p x p)

For the on-line estimator that estimates the parameter less frequently, the number of subtractions in (9) becomes zero, and (7) through (10) are processed only at parameter estimation times. The requirements for the full-scale estimator are obtained by multiplying all values in (1) through (6) by N, and number (9) is multiplied by N, or (N-1) if (7) and (8) are not included in the calculations.

If a precomputed control history is used, the following additions are made to number (3):

M: $n^2 + 2nr + p_A \left[n^2 (m+3) + 2nr \right] + p_B \left[n^2 (m+2) + 2nr \right]$

A: $n^2 + n (r-1) + p_A \left[2n^2 + n (r-1) \right] + p_B \left[n^2 + n (r-1) \right]$

S: n^2

I: 0

On the other hand, if feedback control is used, nr multiplications and r (n-1) additions are added to number (1), and $\left[n^2 r (p+1) + nr^2 \right]$ multiplications added to number (3).

If the conditional information matrix is evaluated by means of the approximation given by equations (5.1.2) and (5.1.3), then number (3) is replaced by:

M: $p_A \left[m^3 \right] + \frac{1}{2} p_A (p_A+1) \left[m^3 + m^2 + m + 1 \right] + \frac{1}{2} p_B (p_B+1) \left[m^2 + m \right] + p_A p_B \left[m^2 + m \right]$

$$A: \quad p_A\left[m^3-m^2\right] + \tfrac{1}{2}\, p_A\,(p_A+1)\left[m^3\right] + \tfrac{1}{2}\, p_B\,(p_B+1)\left[m^2-1\right] + p_A p_B\left[m^2-1\right]$$

$$S: \quad 0$$

$$I: \quad 0$$

Furthermore, number (6) is removed entirely and number (8) is replaced by

$$M: \quad p_A n^3 + \tfrac{1}{2}\, p_A\,(p_A+1)\left[n^3+n^2+n+1\right] + \tfrac{1}{2}\, p_B\,(p_B+1)\left[n^2+n\right] + p_A p_B\left[n^2+n\right]$$

$$A: \quad p_A\left[n^3-n^2\right] + \tfrac{1}{2}\, p_A\,(p_A+1)\, n^3 + \tfrac{1}{2}\, p_B\,(p_B+1)\left[n^2-1\right] + p_A p_B\left[n^2-1\right]$$

$$S: \quad 0$$

$$I: \quad 0$$

When only weighted least squares terms are included, the following number of computations are subtracted from the totals cited previously. From number (2) are subtracted

$$M: \quad m^2 + p_A\left[6n^3 + n^2 m + nm^2 + (m^2+1)\right]$$

$$A: \quad p_A\left[6n^3 + n^2\,(m-6) + nm^2 + (m^2-1)\right]$$

$$S: \quad m^2 + p_A$$

Subtracted from (5) are

$$M: \quad p_A\left[2n^3 + n^2 m + nm\right]$$

$$A: \quad p_A\left[2n^3 + n^2\,(m-2) + n - m\right]$$

For the $\underline{J}$ matrix, the approximation of equation (5.1.2) is used, but the following number of computations would be subtracted from the modified number (3):

M: $\frac{1}{2}\ p_A\ (p_A+1)\ \left[3m^3 + 1\right]$

A: $\frac{1}{2}\ p_A\ (p_A+1)\ \left[3m^3 - 3m^2 + m\right]$

Finally, (7) and (8) are removed entirely.

If precomputations are employed, as suggested in section 5.4, then the following subtractions are made. From (1):

M: $2n^3 + n^2\ (m+s) + n\ (m^2+s^2)$

A: $2n^3 + n^2\ (m+s-2) + n\ (m^2-m+s^2-s)$

I: 1 (m x m)

From (2):

M: $p_A \left[6n^3 + n^2m + nm^2\right]$

A: $p_A \left[6n^3 + n^2\ (m-6) + n\ (m^2-m) - m^2\right]$

From (4):

M: $2n^3 + 4n^2m + nm^2$

A: $2n^3 + n^2\ (3m-3) + n\ (m^2-2m)$

S: n^2

From (5):

M: $p_A \left[2n^3 + n^2m\right]$

A: $p_A \left[2n^3 + n^2\ (m-2) - nm\right]$

Furthermore, numbers (3), (6), (7), and (8) would be removed; the number of additions and subtractions in (9) would both become p; and the inversion in number (10) would be removed. If precomputations and weighted least squares were both utilized, the results would be the same as above, except

that the results for (2) and (5) would be as described previously for the weighted least squares case.

When canonical state space is used, the same procedure is used, but further reductions result from the known positions of zeroes and ones (the ones reduce the number of multiplications only) in the system matrices, as discussed in section 5.5.

Appendix G: An Approximate Closed Form Estimator of Uncertain Entries in $\underline{\Phi}$ or $\underline{B}$

The technique proposed in section 5.3.1 yielded a parameter estimate by solving the equations

$$\sum_{j=i-N+1}^{i} \frac{\partial \underline{\bar{x}}^T(j)}{\partial a_\ell} \underline{H}^T(j)\underline{A}^{-1}(j) \left[\underline{z}(j)-\underline{H}(j)\underline{\bar{x}}(j)\right] = 0 \quad \text{(G-1)}$$

for ℓ equal to 1, 2, ... p. Since no general closed form solution for $\underline{a}$ exists for these equations, an iterative solution was required. This section will show that additional approximations can be introduced to generate a closed form solution for uncertain parameters that are explicit entries in the $\underline{\Phi}(i,i-1)$ or $\underline{B}(i-1)$ matrices.

Assume for the moment that one component of $\underline{\Phi}$, $\Phi_{k\ell}$, is the uncertain parameter to be estimated. Equation (G-1) could then be written as

$$\Phi^*_{k\ell}(i) = \arg \min_{\Phi_{k\ell}} \left\{ \sum_{j=i-N+1}^{i} \underline{r}^T(j)\underline{A}^{-1}(j)\underline{r}(j) \right\} \quad \text{(G-2)}$$

where $\underline{A}^{-1}(j)$ is not treated as a function of $\Phi_{k\ell}$. If $\underline{K}(j)$ were not treated as a function of $\Phi_{k\ell}$, then this would be equivalent to minimizing the N-step sum of

$$\left[\underline{K}(j)\underline{r}(j)\right]^T \left[\underline{K}(j)\underline{A}(j)\underline{K}^T(j)\right]^{-1} \left[\underline{K}(j)\underline{r}(j)\right] =$$

$$= \left[\underline{\hat{x}}(j)-\underline{\bar{x}}(j)\right]^T \left[\underline{K}(j)\underline{A}(j)\underline{K}^T(j)\right]^{-1} \left[\underline{\hat{x}}(j)-\underline{\bar{x}}(j)\right] \quad \text{(G-3)}$$

if the required inverse existed. Unfortunately, $\underline{K}(j)\underline{A}(j)\underline{K}^T(j)$ is an n-by-n matrix with rank of m (at most), so that the inverse does not generally exist since the measurements are often of lower dimension than the state. In propagating from

time (i-1) to time i, $\Phi_{k\ell}(i,i-1)$ directly affects only the k-th component of the state vector; so, to first order, the norm minimization of the vector $[\underline{\hat{x}}(j)-\underline{\bar{x}}(j)]$ can be replaced with the minimization of the weighted square of scalars $[\hat{x}_k(j)-\bar{x}_k(j)]$. Thus, an approximation to equation (G-2) would be

$$\Phi^*_{k\ell}(i) = \arg\min_{\Phi_{k\ell}} \left\{ \sum_{j=i-N+1}^{i} W_k(j)\left[\hat{x}_k(j)-\bar{x}_k(j)\right]^2 \right\} \qquad \text{(G-4)}$$

where $W_k(j)$ is an appropriate scalar weighting factor, such as

$$W_k(j) = \frac{1}{\underline{q}_k^T \underline{K}(j)\underline{A}(j)\underline{K}^T(j)\underline{q}_k} \qquad \text{(G-5)}$$

The minimization of the difference between state estimates before and after a measurement can be motivated as follows. Assume the system is at time instant i and that an optimal estimate (prediction) of the parameters $\underline{a}$ established at the previous sample time has been used to propagate the state estimate to the current time. Before the measurement $\underline{z}(i)$ is incorporated, the best state estimate would be

$$\underline{\bar{x}}(i) = \underline{\Phi}^*(i,i-1)\underline{\hat{x}}(i-1) + \underline{B}^*(i-1)\underline{u}(i-1) \qquad \text{(G-6)}$$

Now $\underline{z}(i)$ is used to obtain $\underline{\hat{x}}(i)$ from $\underline{\bar{x}}(i)$. Once these calculations are made, the difference between $\underline{\bar{x}}(i)$ and $\underline{\hat{x}}(i)$ can be regarded as an indication of the rate of divergence of the state $\underline{x}$ and the estimate of it due to faulty knowledge of the parameters in $\underline{\Phi}$ and $\underline{B}$. (This difference is also caused by the error in the state estimate $\underline{\hat{x}}(i-1)$, the driving noise $\underline{w}(i-1)$, and the measurement noise $\underline{v}(i)$.) One means of defining an optimal $\underline{\Phi}^*(i+1,i)$ and $\underline{B}^*(i)$ would be to choose the values $\underline{\tilde{\Phi}}$ and $\underline{\tilde{B}}$ which, had they been used in the

previous propagations, would have resulted in the values of $\hat{\underline{x}}(j)$ being closest to $\tilde{\underline{\Phi}}\ \hat{\underline{x}}(j-1) + \tilde{\underline{B}}\ \underline{u}(j-1)$ for $j = (i-N+1), \ldots i$. Basic to this technique is the concept of the "divergence" of $[\hat{\underline{x}}(j)-\bar{\underline{x}}(j)]$ over each step providing information about incorrectly estimated parameters. This is substantiated by the fact that the driving and measurement noises are independent from sample to sample, while the parameters are consistent within the N-step interval.

A full-scale estimate would require an N-step propagation of $\bar{\underline{x}}(j)$ and $\hat{\underline{x}}(j)$ for each step. Similar to the on-line implementations of section 5.2 in which the residuals were calculated only once and stored for later use, approximate the full-scale method by considering $\hat{\underline{x}}(j)$ invariant once calculated. Thus, the parameter estimator takes the N most recent values of $\hat{\underline{x}}(j)$, which represent a combination of information from the dynamic model and the real system, and finds a new dynamic model which produces state estimates $\tilde{\underline{x}}(j)$ that most nearly duplicate these $\hat{\underline{x}}(j)$ values.

Assuming that $\Phi_{k\ell}$ is the uncertain parameter, equation (5.3.37) would be solved by satisfying

$$\sum_{j=i-N+1}^{i} W_k(j)\ [\hat{x}_k(j) - \tilde{x}_k(j)]\ \frac{\partial \tilde{x}_k(j)}{\partial \Phi_{k\ell}} = 0 \qquad \text{(G-7)}$$

where $\tilde{x}_k(j)$ is the k-th component of $\tilde{\underline{x}}(j)$,

$$\begin{aligned}\tilde{x}_k(j) &= \underline{q}_k^T\ [\tilde{\underline{\Phi}}\ \hat{\underline{x}}(j-1) + \underline{B}(j-1)\underline{u}(j-1)] \\ &= \Phi_{k1}\hat{x}_1(j-1) + \ldots + \tilde{\Phi}_{k\ell}\hat{x}_\ell(j-1) + \ldots + \Phi_{kn}\hat{x}_n(j-1) \\ &\quad + B_{k1}u_1(j-1) + \ldots + B_{kr}u_r(j-1) \qquad \text{(G-8)}\end{aligned}$$

and thus, $\partial\tilde{x}_k(j)/\partial\Phi_{k\ell}$ is simply $\hat{x}_\ell(j-1)$. Substituting these expressions into equation (G-7) and rearranging yields the estimate $\tilde{\Phi}_{k\ell}$ at time i as:

$$\tilde{\Phi}_{k\ell}(i) = \frac{1}{\sum_{j=i-N+1}^{i} W_k(j)\hat{x}_\ell^{\,2}(j-1)} \left[\sum_{j=i-N+1}^{i} W_k(j)\hat{x}_\ell(j-1) \cdot \right.$$

$$\cdot \Big(\hat{x}_k(j) - \Phi_{k1}\hat{x}_1(j-1) - \ldots - \Phi_{k\,(\ell-1)}x_{\ell-1}(j-1)$$

$$- \Phi_{k\,(\ell+1)}\hat{x}_{\ell+1}(j-1) - \ldots - \Phi_{kn}\hat{x}_n(j-1) - B_{k1}u_1(j-1) - \ldots$$

$$\left. - B_{kr}u_r(j-1) \Big) \right] \qquad \text{(G-9)}$$

If the uncertain parameter were instead an entry in $\underline{B}$, $B_{k\ell}$, a similar development would generate the estimate

$$\tilde{B}_{k\ell}(i) = \frac{1}{\sum_{j=i-N+1}^{i} W_k(j)u_\ell^{\,2}(j-1)} \left[\sum_{j=i-N+1}^{i} W_k(j)u_\ell(j-1) \cdot \right.$$

$$\cdot \Big(\hat{x}_k(j) - \Phi_{k1}\hat{x}_1(j-1) - \ldots - \Phi_{kn}\hat{x}_n(j-1) - B_{k1}u_1(j-1) - \ldots$$

$$- B_{k(\ell-1)}u_{\ell-1}(j-1) - B_{k\,(\ell+1)}u_{\ell+1}(j-1) - \ldots$$

$$\left. - B_{kr}u_r(j-1) \Big) \right] \qquad \text{(G-10)}$$

These equations are valid when there is only one uncertain parameter in the k-th row of $\underline{\Phi}$ and $\underline{B}$; if there is more than one uncertain parameter in their k-th rows, then Cramer's Rule can be used to solve for their values simultaneously. However, many problems will require only one or two key parameters to be estimated, so the simpler forms of (G-9) and (G-10) would often be used. Numerical inaccuracies may be a problem if all $x_\ell(j-1)$ or $u_\ell(j-1)$ values over an N-step interval are very small, but the same would be true of other estimation techniques as well.

For this formulation, the storage requirements would be satisfied by providing two running sum registers, one for each summation term in equations (G-9) and (G-10), if the

parameter estimate were made every N sample periods. If the estimate is to be made more frequently, then the individual terms in the summations must be stored separately (or as N'-step sums, if the parameter is estimated every N' samples) as well. Thus, the memory requirements are no greater than those of the implementations of section 5.2.

Equations (G-9) and (G-10) can also be expressed recursively. Two successive estimates using (G-9), assuming a parameter estimate is made every sample period, would be

$$\tilde{\Phi}_{k\ell}(i-1) = \frac{1}{f_{k\ell}(i-1)} \, g_{k\ell}(i-1) \qquad \text{(G-11a)}$$

$$\tilde{\Phi}_{k\ell}(i) = \frac{1}{f_{k\ell}(i)} \, g_{k\ell}(i) \qquad \text{(G-11b)}$$

where

$$f_{k\ell}(i) = f_{k\ell}(i-1) + W_k(i)\hat{x}_\ell^{\,2}(i-1) - W_k(i-N)\hat{x}_\ell^{\,2}(i-N-1) \qquad \text{(G-12)}$$

$$g_{k\ell}(i) = g_{k\ell}(i-1) + W_k(i)d_{k\ell}(i) - W_k(i-N)d_{k\ell}(i-N) \qquad \text{(G-13)}$$

$$d_{k\ell}(j) = \Big[\hat{x}_k(j) - \Phi_{k1}\hat{x}_1(j-1) - \ldots - \Phi_{k(\ell-1)}\hat{x}_{\ell-1}(j-1) - \Phi_{k(\ell+1)}\hat{x}_{\ell+1}(j-1) - \ldots - \Phi_{kn}\hat{x}_n(j-1) - B_{k1}u_1(j-1) - \ldots - B_{kr}u_r(j-1)\Big]\,\hat{x}_\ell(j-1) \qquad \text{(G-14)}$$

Putting equations (G-11a) and (G-11b) into equation (G-13) yields the desired relation:

$$\tilde{\Phi}_{k\ell}(i) = \frac{f_{k\ell}(i-1)}{f_{k\ell}(i)} \tilde{\Phi}_{k\ell}(i-1)$$

$$+ \frac{1}{f_{k\ell}(i)} \left[W_k(i) d_{k\ell}(i) - W_k(i-N) d_{k\ell}(i-N) \right] \quad \text{(G-15)}$$

The startup procedure, up to time N, would entail

$$f_{k\ell}(i) = f_{k\ell}(i-1) + W_k(i) \hat{x}_\ell^2(i-1) \quad \text{(G-16)}$$

$$\tilde{\Phi}_{k\ell}(i) = \frac{f_{k\ell}(i-1)}{f_{k\ell}(i)} \tilde{\Phi}_{k\ell}(i-1) + \frac{1}{f_{k\ell}(i)} W_k(i) d_{k\ell}(i) \quad \text{(G-17)}$$

from the initial conditions of $f_{k\ell}(0) = 0$, $\tilde{\Phi}_{k\ell}(0)$ = best apriori estimate.

Similarly, the recursive form for the estimate of the k-ℓ-th entry of $\underline{B}$ can be written as

$$f'_{k\ell}(i) = f'_{k\ell}(i-1) + W_k(i) u_\ell^2(i-1) - W_k(i-N) u_\ell^2(i-N-1) \quad \text{(G-18)}$$

$$\tilde{B}_{k\ell}(i) = \frac{f'_{k\ell}(i-1)}{f'_{k\ell}(i)} \tilde{B}_{k\ell}(i-1)$$

$$+ \frac{1}{f'_{k\ell}(i)} \left[W_k(i) d'_{k\ell}(i) - W_k(i-N) d'_{k\ell}(i-N) \right] \quad \text{(G-19)}$$

where the variable $d'_{k\ell}(j)$ is defined as:

$$d'_{k\ell}(j) = \Big[\hat{x}_k(j) - \Phi_{k1}\hat{x}_1(j-1) - \ldots - \Phi_{kn}\hat{x}_n(j-1)$$

$$- B_{k1}u_1(j-1) - \ldots - B_{k(\ell-1)}u_{\ell-1}(j-1)$$

$$- B_{k(\ell+1)}u_{\ell+1}(j-1) - \ldots - B_{kr}u_r(j-1)\Big]\, u_\ell(j-1) \qquad \text{(G-20)}$$

and the same form of startup procedure as (G-16) and (G-17) would be used up to time N.

In the case of continuous dynamics, as described by equation (3.4), the first order effect (small measurement sample period) of a change in $F_{k\ell}(t)$ would be a change in $\Phi_{k\ell}(i+1,i)$ since $\underline{\Phi}(i+1,i) = \underline{I} + \underline{F}(t_i)\ T$ to first order in T, the sample period; a similar conclusion is true for $\underline{B}$ since $\underline{B}(i) = \underline{B}(t_i)\ T$ to first order in T. Thus, if the uncertain parameter were actually $F_{k\ell}(t)$, the above technique could be used to estimate $\Phi_{k\ell}(i+1,i)$ in expectation that the k-th component of $\underline{\hat{x}}$ and $\underline{\tilde{x}}$ will offer the most sensitivity to the parameter value. Once $\tilde{\Phi}_{k\ell}(i)$ is calculated, the other entries in $\underline{\Phi}$ that are affected by $F_{k\ell}(t)$ can be evaluated approximately as precomputed functions of $\tilde{\Phi}_{k\ell}(i)$.

This technique was applied to the example of section 6.1, with the weighting factor $W_k(j)$ chosen to be unity. The resulting parameter estimate was:

$$\tilde{\Phi}_{22}(i) = \frac{1}{\sum_{j=i-N+1}^{i} \hat{x}_2{}^2(j-1)} \left[\sum_{j=i-N+1}^{i} \hat{x}_2(j-1)\,[\hat{x}_2(j) + .8\,\hat{x}_1(j-1)]\right] \qquad \text{(G-21)}$$

However, initial application of this parameter estimator was unsatisfactory, mainly due to the uncertainty in the estimate $\hat{x}_2$. The magnitude of the error in $\hat{x}_2$ is large relative to that of $\hat{x}_1$, because w is a strong noise source that directly affects $x_2(i)$, while the measurement is directly of $x_1(i)$,

which is simply $x_2(i-1)$. (The continuous analog would be the estimation of position and velocity from position measurements when a strong, uncertain acceleration is driving the system.) Consequently, a significantly more accurate estimate of $x_2(i)$ is provided by $\hat{x}_1(i+1)$, and (G-21) can be written as

$$\tilde{\Phi}_{22}(i) = \frac{1}{\sum_{j=i-N+1}^{i} \hat{x}_1^2(j)} \sum_{j=i-N+1}^{i} \hat{x}_1(j)\,[\hat{x}_1(j+1) + .8\,\hat{x}_1(j-1)] \tag{G-22}$$

Equation (G-22) cannot be evaluated until time (i+1), but since the parameters are assumed to be slowly varying, the approximation that $\underline{\tilde{\Phi}}_{22}(i) \cong \underline{\tilde{\Phi}}_{22}(i-1)$ can be used to write

$$\tilde{\Phi}_{22}(i) = \frac{1}{\sum_{j=i-N+1}^{i} \hat{x}_1^2(j-1)} \sum_{j=i-N+1}^{i} \hat{x}_1(j-1)\,[\hat{x}_1(j) + .8\,\hat{x}_1(j-2)] \tag{G-23}$$

The startup procedure would entail replacing the lower limits on the summations by one, and the first parameter estimate can be made at time $i = n + 1 = 3$.

Figure (G-1) presents a representative parameter estimate error trajectory for an estimator that utilizes a data interval length of N = 5 plus averaging over 30 steps. Comparing these results to those depicted in figure 6.1.16 for the on-line iterative estimator reveals a similar level of performance. However, it should be noted that this is a more restrictive technique, designed to estimate only independent entries of $\underline{\Phi}$ or $\underline{B}$. Moreover, the transformation to an expression in terms of $\hat{x}_1$ alone (necessary in this case for adequate performance) was conveniently accomplished because the system was modelled in phase variables; when the parameter estimate cannot be readily expressed in terms of the state variable whose coordinate direction in state space defines the minimum axis of the error ellipsoid, the performance will deteriorate.

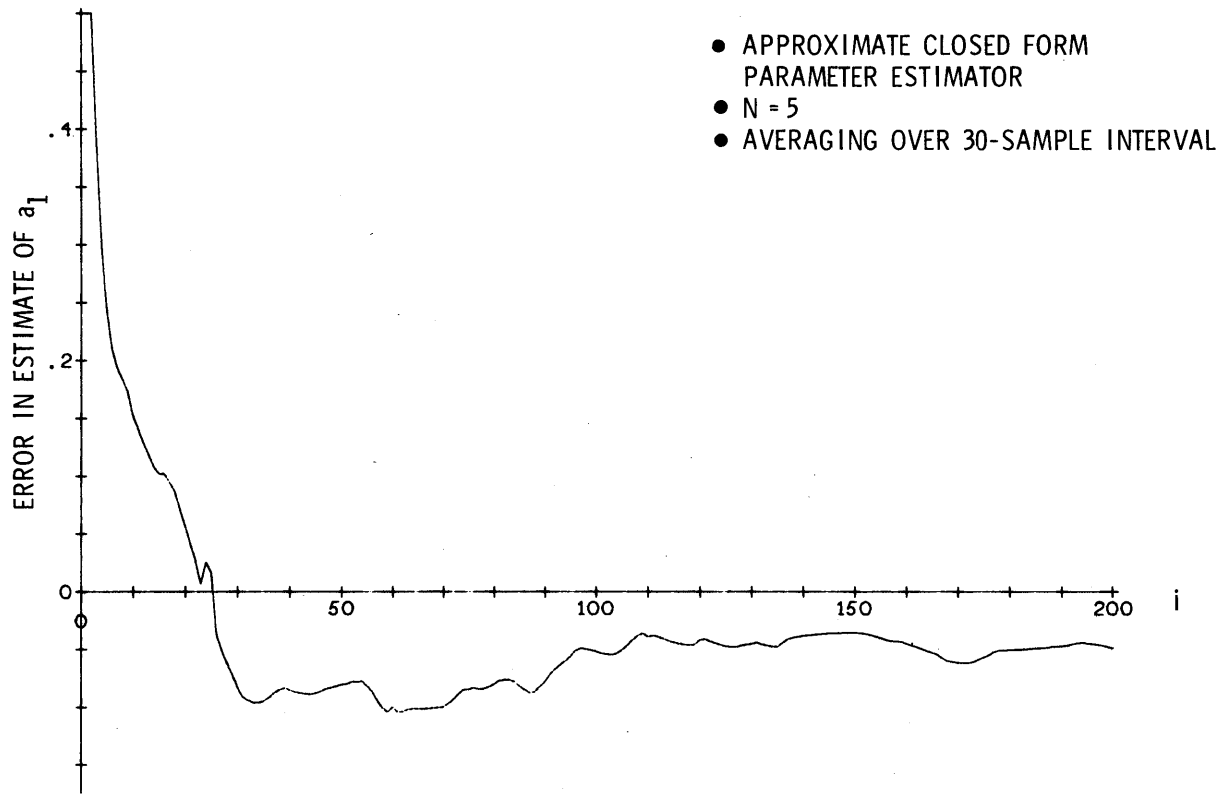


Figure G-1 Parameter Estimate Error

REFERENCES

Abramson, P. D. Jr. (1968), "Simultaneous Estimation of the State and Noise Statistics in Linear Dynamical Systems," Sc.D. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Mass., May.

Albert, A. E. and Gardner, L. A. Jr. (1967), Stochastic Approximation and Nonlinear Regression, Research Monograph No. 42, The M.I.T. Press, Cambridge, Mass.

Alonso, R. L. and Randa, G. C. (1967), "Flight - Computer Hardware Trends," Astronautics and Aeronautics, April.

Apollo Guidance Program Symbolic Listing - Information for Block 2, NAS 9-4810 (1967), working paper prepared for the Mission Planning and Analysis Division of NASA Manned Spacecraft Center, Houston, Texas, January.

Aoki, M. (1967), Optimization of Stochastic Systems - Topics in Discrete-Time Systems, Academic Press, New York, New York.

Battin, R. H. (1964), Astronautical Guidance, McGraw-Hill, New York, New York, pp. 338-339.

Bellantoni, J. F. and Dodge, K. W. (1967), "A Square Root Formulation of the Kalman - Schmidt Filter," AIAA Journal, Vol. 5, No. 7, pp. 1309-1314, July.

Brockett, R. W. (1969), Finite Dimensional Linear Systems, Electronic Systems Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., June.

Brunk, H. D. (1965), An Introduction to Mathematical Statistics, Blaisdell Publishing Co., New York, New York.

Cox, H. (1964), "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems," IEEE Transactions on Automatic Control, Vol. AC-19, No. 1, pp. 5-12, January.

Cramér, H. (1946), Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey.

Detchmendy, D. M. and Sridhar, R. (1966), "Sequential Estimation of States and Parameters in Noisy Nonlinear Dynamical Systems," Journal of Basic Engineering (Transactions of the ASME), Series D, Vol. 88, No. 2, pp. 362-368, June.

Deutsch, R. (1968), Estimation Theory, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Dyer, P. and McReynolds, S. (1969), "The Extension of Square-Root Filtering to Include Process Noise," Journal of Optimization Theory and Applications, Vol. 3, No. 6, pp. 444-459, June.

Fagin, S. L. (1964), "Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analysis of Optimal Systems," 1964 IEEE International Convention Record, New York, New York, pp. 216-240, March.

Farison, J. B. (1964), "Identification and Control of Random-Parameter Discrete Systems," Stanford Electronics Laboratories Technical Report No. 6302-4, Stanford, Calif., January.

Farison, J. B., Graham, R. E. and Shelton, R. C. (1967), "Identification and Control of Linear Discrete Systems," IEEE Transactions on Automatic Control, Vol. AC-12, No. 4, pp. 438-442, August.

Fourth Quarterly Progress Report for Contract NAS 9-10386 (1971), The Charles Stark Draper Laboratory, Cambridge, Mass., April.

Fraser, D. C. (1967), "A New Technique for the Optimal Smoothing of Data," Sc.D. Thesis at Massachusetts Institute of Technology, Instrumentation Laboratory Report T-474, Cambridge, Mass., January.

Galiana, F. D. (1969), "A Review of Basic Principles and of Available Techniques in System Identification," Power Systems Engineering Group Report 20, Massachusetts Institute of Technology, Cambridge, Mass., November.

Gantmacher, F. R. (1959), The Theory of Matrices (Vol. 1), Chelsea Publishing Co., New York, New York.

Geise, C. and McGhee, R. B. (1965), "Estimation of Nonlinear System States and Parameters by Regression Methods," Preprints of the 1965 Joint Automatic Control Conference, Rensselaer Polytechnic Institute, Troy, New York, pp. 46-55, June.

Ho, Y. (1962), "On the Stochastic Approximation Method and Optimal Filtering Theory," Journal of Mathematical Analysis and Applications, Vol. 6, pp. 152-154.

Ho, Y. (1962'), "The Method of Least Squares and Optimal Filtering Theory," The Rand Corporation, Memorandum RM-3329-PR, Santa Monica, Calif., October.

Ho, Y. and Lee, R. C. K. (1965), "Identification of Linear Dynamic Systems," Information and Control, Vol. 8, No. 1, pp. 93-110, February.

Ho, Y. and Whalen, B. H. (1963), "An Approach to the Identification and Control of Linear Dynamic Systems with Unknown Parameters," IEEE Transactions on Automatic Control, Vol. AC-8, No. 3, pp. 255-256, July.

Householder, A. S. (1953), Principles of Numerical Analysis, McGraw-Hill, New York, New York.

Isaacson, E. and Keller, H. B. (1966), Analysis of Numerical Methods, John Wiley & Sons, Inc., New York, New York.

Jazwinski, A. H. (1968), "Limited Memory Optimal Filtering," IEEE Transactions on Automatic Control, Vol. AC-13, No. 5, pp. 558-563, October.

Jazwinski, A. H. (1970), Stochastic Processes and Filtering Theory, Academic Press, New York, New York.

Kale, B. K. (1961), "On the Solution of the Likelihood Equation by Iteration Processes," Biometrika, Vol. 48, pp. 452-456.

Kalman, R. E. (1958), "Design of a Self-Optimizing Control System," Transactions of the ASME, Vol. 80, No. 2, pp. 468-478, February.

Kaminski, P. G., Bryson, A. E. Jr. and Schmidt, S. F. (1971), "Discrete Square Root Filtering: A Survey of Current Techniques," submitted for publication in IEEE Transactions on Automatic Control.

Kantorovich, L. V. and Akilov, G. P. (1964), Functional Analysis in Normed Spaces, Pergamon Press, Inc., New York, New York.

Kashyap, R. L. (1970), "Maximum Likelihood Identification of Stochastic Linear Systems," IEEE Transactions on Automatic Control, Vol. AC-15, No. 1, pp. 25-34, February.

Kerr, R. B. (1965), "Statistical, Structural, and Regression Models in Linear Process Identification," Preprints of the 1965 Joint Automatic Control Conference, Rensselaer Polytechnic Institute, Troy, New York, pp. 458-466, June.

Kumar, K. S. P. and Sridhar, R. (1964), "On the Identification of Control Systems by the Quasi-Linearization Method," IEEE Transactions on Automatic Control, Vol. AC-9, No. 2, pp. 151-154, April.

Lee, R. C. K. (1964), Optimal Estimation, Identification and Control, Research Monograph No. 28, The M.I.T. Press, Cambridge, Mass.

Levin, M. J. (1964), "Estimation of a System Pulse Transfer Function in the Presence of Noise," IEEE Transactions on Automatic Control, Vol. AC-9, No. 3, pp. 229-235, July.

Luenberger, D. G. (1966), "Observers for Multivariable Systems," IEEE Transactions on Automatic Control, Vol. AC-11, No. 2, pp. 190-197, April.

Magill, D. T. (1965), "Optimal Adaptive Estimation of Sampled Stochastic Processes," IEEE Transactions on Automatic Control, Vol. AC-10, No. 4, pp. 434-439, October.

Meditch, J. S. (1969), Stochastic Optimal Linear Estimation and Control, McGraw-Hill, New York, New York.

Mowery, V. O. (1965), "Least Squares Recursive Differential-Correction Estimation in Nonlinear Problems," IEEE Transactions on Automatic Control, Vol. AC-10, No. 4, pp. 399-407, October.

Neal, S. R. (1967), "Linear Estimation in the Presence of Error in Assumed Plant Dynamics," IEEE Transactions on Automatic Control, Vol. AC-12, No. 5, pp. 592-594, October.

Papoulis, A. (1965), Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, New York,

Pearson, J. D. (1967), "Deterministic Aspects of the Identification of Linear Discrete Dynamic Systems," IEEE Transactions on Automatic Control, Vol. AC-12, No. 6, pp. 764-766, December.

Potter, J. E. (1964), "W Matrix Augmentation," MIT Instrumentation Laboratory Memo SGA 5-64, Cambridge, Mass., January.

Rao, C. R. (1968), Linear Statistical Inference and its Applications, John Wiley & Sons, Inc., New York, New York.

Rauch, H. E., Tung, F. and Striebel, C. T. (1965), "Maximum Likelihood Estimates of Linear Dynamic Systems," AIAA Journal, Vol. 3, No. 8, pp. 1445-1450, August.

Rogers, A. E. and Steiglitz, K. (1967), "Maximum Likelihood Estimation of Rational Transfer Function Parameters," IEEE Transactions on Automatic Control, Vol. AC-12, No. 5, pp. 594-597, October.

Sage, A. P. (1968), Optimum Systems Control, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Sage, A. P. and Ellis, T. W. (1966), "Sequential Suboptimal Adaptive Control of Nonlinear Systems," Proceedings of the National Electronics Conference, Vol. 22, pp. 692-697, Chicago, Illinois, October.

Sage, A. P. and Husa, G. W. (1969), "Adaptive Filtering with Unknown Prior Statistics," Preprints of the 1969 Joint Automatic Control Conference of the American Automatic Control Council, pp. 760-769, August.

Sage, A. P. and Masters, G. W. (1966), "On-Line Estimation of States and Parameters for Discrete Nonlinear Dynamic Systems," Proceedings of the National Electronics Conference, Vol. 22, pp. 677-682, Chicago, Illinois, October.

Sakrison, D. J. (1967), "The Use of Stochastic Approximation to Solve the System Identification Problem," IEEE Transactions on Automatic Control, Vol. AC-12, No. 5, pp. 563-567, October.

Schmidt, S. F. (1970), "Computational Techniques in Kalman Filtering," Theory and Applications of Kalman Filtering, AGARDograph 139, Ch. 3, London, England, February.

Schulz, E. R. (1968), "Estimation of Pulse Transfer Function Parameters by Quasilinearization," IEEE Transactions on Automatic Control, Vol. AC-13, No. 4, pp. 424-426, August.

Schweppe, F. C. (1964), "Algorithms for Estimating a Re-Entry Body's Position, Velocity and Ballistic Coefficient in Real Time or for Post Flight Analysis," Lincoln Laboratory Grou Report 1964-64, Lexington, Mass., December.

Schweppe, F. C. (1965), "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Transactions on Information Theory, Vol. IT-11, No. 1, pp. 61-70, January.

Schweppe, F. C. (1970), Course Notes for "Uncertain Dynamic Systems," printed at Massachusetts Institute of Technology, Cambridge, Mass.

Shellenbarger, J. C. (1966), "Estimation of Covariance Parameters for an Adaptive Kalman Filter," Proceedings of the National Electronics Conference, Vol. 22, pp. 298-302, Chicago, Illinois, October.

Sims, F. L. and Lainiotis, D. G. (1969), "Recursive Algorithm for the Calculation of the Adaptive Kalman Filter Weighting Coefficients," IEEE Transactions on Automatic Control, Vol. AC-14, No. 2, pp. 215-217, April.

Speyer, J. L. and Desai, M. N. (1970), "An Algorithm for Propagating the Square Root of the Covariance Matrix," MIT Instrumentation Laboratory ATC Memo No. 1-70, Cambridge, Mass., June.

Steiglitz, K. and McBride, L. E. (1965), "A Technique for the Identification of Linear Systems," IEEE Transactions on Automatic Control, Vol. AC-10, No. 4, pp. 461-464, October.

Widnall, W. S. (1968), Applications of Optimal Control Theory to Computer Controller Design, Research Monograph No. 48, The M.I.T. Press, Cambridge, Mass.

Wilks, S. S. (1963), Mathematical Statistics, John Wiley & Sons, Inc., New York, New York.

BIOGRAPHY

Peter S. Maybeck was born February 9, 1947 in Queens, New York. While attending Baldwin High School in Baldwin, New York, he won the Alicia Patterson Scholarship, an annual award presented to the outstanding senior on Long Island. He entered M.I.T. and received the Bachelor of Science Degree in June of 1968. During this time, he became a member of Sigma Chi Fraternity and Sigma Gamma Tau honorary, and also served as President of Tau Beta Pi. Industrial experience has included summer employment with Kaman Avidyne and two summers at the M.I.T. Instrumentation Laboratory, now the Charles Stark Draper Laboratory. Partially on the basis of papers which won the James Means Memorial Prize and the Luis de Florez Prize, he was permitted by the Department of Aeronautics and Astronautics to enter directly into the doctoral program. His graduate education has been supported financially by the Fannie and John Hertz Foundation through its national fellowship program.

In June of 1968, he and Miss Beverly J. Bedsaul were married, and they and their daughter, Kristen, currently reside in Cambridge, Massachusetts. After receiving his degree, Mr. Maybeck will serve as an engineer in the Flight Control Division of the Air Force Flight Dynamics Laboratory.