AFOSR-TR-97-0376

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
12 March 97

3. REPORT TYPE AND DATES COVERED
Final Technical: 12/01/93 - 11/30/96

4. TITLE AND SUBTITLE
A Disturbance Attenuation Approach to Missile Guidance and Control

5. FUNDING NUMBERS
G - F49620-94-1-0084

6. AUTHORS
Jason L. Speyer

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Regents of the University of Califronia
Mechanical and Aerospace Engineering Department, 38-137 Engr. IV, Box 951597
University of California, Los Angeles
Los Angeles, CA 90095-1597

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
110 Duncan Avenue, Room B115
Bolling AFB, DC 20332-8080
attn: Dr. Marc Q. Jacobs

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION IS UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
Solutions to stochastic control problems currently do not produce mechanizable missile control laws. To fill the gap without making excessive structural assumptions, it is suggested that the disturbance attenuation problem be extended to nonlinear guidance and control problems. This deterministic approach, which does not include all the properties of the stochastic control solution, does have characteristics which are quite appealing. For example, certain classes of partial information disturbance attenuation problems can be solved numerically essentially because variation methods are available. Over the three year period of the grant, robust and adaptive guidance and control laws which are mechanizable with near-future computer technology are developed which can meet system objectives in the presence of large uncertainties, system structural changes, and nonlinearities. Of particular importance emerging from our focus on disturbance attenuation is a new structure for adaptive control, new detection filters for detection and identifying structural changes, and methodologies for including system nonlinearities. Finally, a new multiple-hypothesis adaptive estimator, using a single linear filter whose parameters are changing according to the on-line computation of the probability of each hypothesis conditioned on the residual history, is shown to have better or equivalent performance to the current bank of filters concept with dramatic decrease in computation and ease of implementation.

14. SUBJECT TERMS
disturbance attenuation; adaptive control; differential games; robust control; detection filters; nonlinear estimation and control

15. NUMBER OF PAGES
46

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UL	UL	UL	UNLIMITED

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)

Final Technical Report

A DISTURBANCE ATTENUATION APPROACH TO MISSILE GUIDANCE AND CONTROL

Grant No. F49620-94-1-0084

Jason L. Speyer

Mechanical and Aerospace Engineering Department
University of California, Los Angeles
Los Angeles, California 90024-1597

Contents

A Disturbance Attenuation Approach To Missile Guidance And Control

1. Objective of the Research Effort

Solutions to stochastic control problems currently do not produce mechanizable missile control laws. To fill the gap without making excessive structural assumptions, it is suggested that the disturbance attenuation problem be extended to nonlinear guidance and control problems. This deterministic approach, which does not include all the properties of the stochastic control solution, does have characteristics which are quite appealing. For example, certain classes of partial information disturbance attenuation problems can be solved numerically essentially because variation methods are available. Over the three year period of the grant, robust and adaptive guidance and control laws which are mechanizable with near-future computer technology are developed which can meet system objectives in the presence of large uncertainties, system structural changes, and nonlinearities. Of particular importance emerging from our focus on disturbance attenuation is a new structure for adaptive control, new detection filters for detection and identifying structural changes, and methodologies for including system nonlinearities. Finally, a new multiple-hypothesis adaptive estimator, using a single linear filter whose parameters are changing according to the on-line computation of the probability of each hypothesis conditioned of the residual history, is shown to have better or equivalent performance to the current bank of filters concept with dramatic decrease in computation and ease of implementation.

2. Status of the Research Effort

The general missile guidance problem can be formulated as the minimization of some statistical function of the terminal miss subject to a nonlinear stochastic system composed of missile and target dynamics, a noisy nonlinear measurement of the state space, and in-flight constraints on the system structure and aerodynamics. In the stochastic control context, the terminal miss distance is described by a probability density function. The object is to shape this density function to give the desired performance. For example, the expected value of the miss distance could be the performance index. However, it may be desirable to allow the mean miss to increase somewhat if the dispersion about the mean can be decreased. The addition of an intelligent target significantly increases the complexity of the problem. Solution to this stochastic control problem by dynamic programming is not currently feasible. Only for the simplest class of stochastic control problems are solutions available. This class includes the linear-quadratic-Gaussian (LQG) and linear-exponential-Gaussian (LEG) problems. In Section 2.1, for the discrete-time LEG problem, the solution of this problem for both the centralized and decentralized information structures are shown to reduce to that of a class of differential games equivalent to that used to solve the disturbance attenuation problem.

Since the solution to the stochastic control problem is currently not feasible, our approach to robust and adaptive guidance and control is to consider a disturbance attenuation formulation which produces a coherent framework for developing and combining guidance laws and nonlinear estimators without excessive structural assumptions. This deterministic approach to uncertainty based on differential game theory allows certain classes of partial information problems to be solved by variational methods which are characterized by ordinary differential equations rather that by the partial differential equation of dynamic programming, thereby substantially reducing the numerical complexity. As significant amount of progress has been made in understanding and exploiting the relationship between minimax and saddle point strategies of these differential games for particular classes of nonlinear dynamic systems. For example, in Section 2.2, a synthesis technique for the design of robust flight control and missile autopilots in the presence of system parameter and initial conditions uncertainties was developed via a disturbance attenuation approach. Whereas the objective of robust controller design is to make the system insensitive to the parameter uncertainty, our efforts also focused on a new class of adaptive control techniques in which the value of the uncertain parameters are to be estimated. In Section 2.3, it is shown that the structure of this new adaptive control scheme is the result of formulating a disturbance attenuation problem for a particular class of nonlinear systems whose solution is obtained without *any* approximation. A global solution is obtained and must be contrasted with much of the nonlinear $\mathcal{H}_\infty$ results which assumes that the scheme operates *locally* about some equilibrium condition. The class of nonlinearities considered is that of a linear system where the coefficient matrix of the control is assumed to be a linear function of an unknown parameter. To bring these mathematical abstractions to engineering practice, a significant effort was made to apply this new adaptive control scheme to the development of an adaptive flight control system for a high angle-of-attach aircraft such as the F-18 HARV (High Angle-of-attach Research Vehicle).

Since target motion detection is important for miss distance reduction in missile guidance, two techniques were developed and discussed in Section 2.4. The first is based on a disturbance attenuation problem in order to achieve a degree of robustness. In the limit, when the transmission bound is brought to zero, the classical detection filter is recovered. In the second approach a multiple hypothesis adaptive estimator for tracking in the presence of target motion is developed. By replacing the old concept of using a bank of filters each tuned to a particular hypothesis, a single filter is used where the parameters are changed according to the on-line computation of the probability of each hypothesis conditioned on the residual history.

In Section 2.5, robust missile guidance laws with bearing-only measurements based on the dissapative inequality are shown to satisfy a disturbance attenuation bound if certain modifications to the guidance law are made and certain conditions are satisfied. Furthermore, in a different approach, over a finite-time interval, for a class of linear time-varying dynamical systems, an approximate nonlinear optimal estimator and control scheme are derived based on an expansion technique for solving the disturbance attenuation problem. Finally, in Section 2.6 using an expansion technique, the midcourse guidance problem for tactical and ballistic missile intercepts is solved where velocity on target is maximized subject to miss distance and orientation constraints.

In the following subsections the work on this grant are briefly described. The papers [2, 6, 7, 9, 10, 13, 14, 15, 16, 17, 18] represent the work performed on this grant. Since [10] is in the early stages of the review process, and because of its importance, it is given in the Appendix.

2.1 Some Relationships Between Stochastic Control and Disturbance Attenuation

For the discrete time LEG problem, where the dynamics and measurements are linear with additive Gaussian uncertainty and the performance index is the expected value of the exponential of a quadratic function, the solution of this problem is equivalent to that of a differential game [1,2]. Note that in [2] the extension of the LEG problem to a class of decentralized information structures is presented. Similarly, for linear systems the disturbance attenuation problem reduces to a differential game [3]. The appealing aspect of disturbance attenuation problems is that they are a deterministic methodology to handle stochastic problems. Therefore, the disturbance attenuation approach is being extended to classes of nonlinear problems. It should be noted that the deterministic disturbance attenuation formulation can be obtained from a stochastic control formulation through large deviation theory [4]. Although many important properties such as dual control appear lost, the disturbance attenuation approach still allows a unified theory for the coherent development of guidance laws assuming partial information. The functions of guidance and estimation may be integrated together without gross assumptions such as the certainty equivalence principle, although aspects of this concept occur naturally.

2.2 Robust Game Theoretic Synthesis in the Presence of Uncertain Initial States

Many controller design problems are multicriteria problems, i.e., there is more than one performance criterion to be considered. For example, in the bench mark problem [5], the controller is required to satisfy certain design specifications on the settling time, peak value of the control effort, and the sensitivity of the output to the measurement noise. Despite a significant research effort into the reduction of the sensitivity of the output to the process and measurement disturbances, for example, in $\mathcal{H}_\infty$ or linear-quadratic-Gaussian (LQG) control theories, improvement of transient performance, measured as settling time and peak amplitude of output due to nonzero initial states, draws little attention in control theory for linear time-invariant (LTI) systems. The usual LTI control problem has been concerned with optimization problems with a performance criterion on the sensitivity of output to the input process and measurement disturbances, with the result that various weighting strategies have been introduced to improve transient performances of the closed-loop system. However, only experienced designers have a good feel for the relation of the weighting matrices to the transient behavior of the closed-loop system.

Strict performance requirements on the controller design often demand an exact model of the system to be controlled. However, uncertainties are usually involved in the mathematical description of physical systems. As a result, a closed-loop system in the presence of system uncertainty often has poor performance or even becomes unstable when the controllers are optimized with respect to a nominal model of the system. Hence, robustness of performance with respect to system uncertainty must also be considered in controller design. Many physical systems have real-valued structured uncertainty. For these systems, robust controller design based upon an unstructured uncertainty model, as in the $\mathcal{H}_\infty$ control, or on a complex-valued structured uncertainty model, as in μ -synthesis, may produce conservative results. These observations lead to the introduction of a constrained disturbance attenuation problem (CDAP). The idea in this problem can be roughly stated as follows: A controller is to be found that minimized the worst-cast transmission of initial state to the output out of a set of controllers that provide internal stability and achieve a prescribed level of disturbance attenuation for all real parameter uncertainties in a given set. The transmission of a nonzero initial state to the output is closely related to the transient response of the system. Hence, transient behavior of the control system as well as attenuation of disturbances are directly considered in CDAP, in contrast to the usual disturbance attenuation problem or $\mathcal{H}_\infty$ control problem in which only attenuation of disturbances is considered. A game theoretic approach is employed for the design of the optimal controller to solve the CDAP [6]. Initial state, process and measurement disturbances, and unknown system parameters are considered as the maximizing players, whereas the control is treated as the minimizing player in a minimax game problem. The game cost criterion consists of a positive term composed of a quadratic norm on the system output and negative terms composed of a quadratic norm on the disturbances. Instead of being penalized in the game cost criterion, initial state and uncertain system parameters are restricted to lie on or within the prescribed sets (here they are multidimensional ellipsoids). In contrast to the results in [3], the cost criterion is nonseparable and this formulation produces multiple worst initial states for a given linear controller. These differences make it difficult to use the simple method of [3] based upon two algebraic Riccati equations in solving the game problem. As a result, a linear fixed-order dynamic compensator based upon partial information is assumed. The controller is viewed as a function of a control parameter vector.

2.3 An Adaptive Controller Based on Disturbance Attenuation

To understand the underlying issue of generalizing the disturbance attenuation problem to nonlinear dynamical systems, we have examined the problem of linear systems with uncertain parameters in the control coefficient matrix [7]. First, the minimax problem is reduced to a one sided control problem whose numerical solution for real-time control is challenging. To simplify the numerical difficulties of finding a nonlinear feedback controller, a dynamic programming approach is used. In the dynamic programming approach, the control problem decomposes into a sum of an optimal return function representing a full information game problem and an optimal accumulation function representing an associated estimation problem. The worst case state, found by maximizing the sum with respect to the current state, is then used in the full information controller. Previous results required that the worst case state to be a singleton [8]. We show [7] that the existence of a solution to the minimax problem guarantees a unique saddle control and the validity of the dynamic programming approach, even when the worst case state is *not* a singleton. Note that interesting adaptive control problems now can be explicitly solved producing implemetable controllers, whereas solutions to their stochastic counterpart are still not available.

An alternate and more direct approach is given in [9] where the minimax problem is reduced to an equivalent full information control problem. The full information state space dynamics are composed of an estimator equation and its associated Riccati equation. The equivalence between the minimax adaptive controller and a saddle point certainty equivalence adaptive controller is shown via Hamilton-Jacobi-Bellman theory where the optimal return function is differentiable. Points of discontinuity in the partial derivatives of the optimal return function are shown to form a manifold of Darboux points, from which multiple global optimal trajectories emanate. Therefore, the Dynamic Programming approach can be extended to be valid over the entire phase space, and the uniqueness of the value of the optimal return function is guaranteed. We finally show that with additional assumptions the finite-time problem can be extended to infinite time.

To begin to understand the implementation issues of this new adaptive controller, the longitudinal mode of the F-18 was controlled. Since a global maximum of the sum of the optimal return function and the optimal accumulation function is required, only the most important parameter, the moment coefficient due to elevator deflection, is used as the adversary. The coefficient of the other control variable, thrust vectoring, is assumed known. As shown [10], for initial conditions, remarkable performance is obtained over current adaptive controllers. In regimes where the scalar parameter is quite unknown the control emphasizes thrust vectoring and reduces the elevator deflection to be almost zero. As more information is obtained the controller begins to use more elevator deflection and less thrust vectoring which is eventually faded to zero. This is to be contrasted with standard adaptive controllers in which substantial elevator deflection is used early even though the parameter is the wrong sign. Therefore, the initial response is in the wrong direction. The inherent conservative, but intelligent, performance of the new controller is associated with the two worst cast states in which usually only one is a global maximum. In the beginning the global worst case state dictates a conservative policy where the elevator deflection is made small and the thrust vectoring dominates the response. At some time, say t_c, both worst case states produce identical cost. Here, there is a switch from the conservative policy to one similar to that of the standard adaptive controller.

2.4 Detection and Identification of Structural Changes

Accurate target motion defection and identification are important in reducing terminal miss of homing missiles. In [11] a detection filter approach to detecting and identifying target acceleration was used when the relative position can be measured using angle and range measurements. In [12] a detection filter approach to fault detection is developed from an eigensystem assignment approach. The detection filter is an observer whose gains are constructed such that the presence of a fault induces in the measurement residuals an invariant (fixed) direction. Therefore, during a sudden maneuver of the target, although the detection filter's estimate of the state of the target is inaccurate, the target's acceleration direction can be determined by examining the detection filter residuals of the position measurement. These fault detectors are made robust to variations in the system parameters as well as exogenous inputs by modification of the detection filter structure or by reformulating as a disturbance attenuation problem. During this grant period an alternative approach, called the multiple-hypothesis adaptive filter, was also developed which allows a dramatic reduction in computation over that of the bank of filter concept.

Detection filters form a class of linear observers that produce residuals with known and fixed directional characteristics in response to a set of design fault directions. In practical applications, reliable fault isolation takes place only when the detection space structure is insensitive to system parameter variations. A left-eigenvector assignment approach is developed [13] that allows for the application of existing results relating supremal controllability subspaces and ill-conditioned eigenvectors. Modifications to the detection filter structure that yield improved eigenvector conditioning and sensitivity to system parameter variations are applied to an aircraft accelerometer fault detection filter.

An alternate approach to detection filter design is given in [14] where the fault detection process is approximated with a disturbance attenuation problem. The solution to this problem leads to an $\mathcal{H}_\infty$ filter which bound the transmission of all exogenous signals save the fault to be detected. In the limit, when the transmission bound is brought to zero, a complete transmission block is achieve by embedding the nuisance inputs into an unobservable, invariant subspace. As this is the same subspace structure seen in some types of detection filters, one can then make the claim that the game theoretic filter asymptotically becomes a detection filter. One can also make use of this subspace structure to reduce the order of the limiting game theoretic filter by factoring this invariant subspace. The resulting lower dimensional signal will then be sensitive only to the failure to be detected. This approach also extends detection filter design to time-varying problems.

In [15, Appendix], a new algorithm for adaptive estimation of time-varying parameters in certain classes of linear stochastic dynamic systems has been developed. The algorithm is based on an adaptive Kalman filter whose hypothesized parameters are modified at each stage by generating the probability of each hypothesis, conditioned on the residual history and a given probability of transition. It is then shown that when a particular hypothesis is true, the expected value of the corresponding posterior probability conditioned on the residual history converges to unity. Moreover, the expected value of the norm of the difference between the constructed error covariance and the true posteriori error covariance is shown to converge to a lower bound close to zero. By invoking an information function, the filter is also shown to be robust with respect to modeling errors. A few numerical simulations have been performed to evaluate this algorithm against the backdrop of the multiple model adaptive estimation scheme.

2.5 Approaches to Disturbance Attenuation for Nonlinear System

Robust nonlinear filters and guidance laws are investigated using the deterministic methodology of disturbance attenuation. In our first approach we show that using the dissapative inequality, the disturbance attenuation bound is satisfied if certain modifications are made to the guidance law and certain conditions are satisfied. In a different approach where the nonlinearities are assumed small in the dynamic system, an approximate nonlinear optimal estimator and control scheme are derived based on an expansion technique for solving the disturbance attenuation problem.

In [16], for a class of nonlinear time-varying system, a stable nonlinear game-theoretic filter (GTF), which solves the disturbance attenuation problem, is obtained with a dissipative approach. For this special class of nonlinear functions, called modifiable nonlinearities, the estimation error propagates almost linearly. A sufficient condition to render the dissipative inequality satisfied for the GTF with respect to the worst strategy for the process and sensor disturbances is derived. The filter gain is obtained from a Riccati differential equation (RDE) which results from bounding the dissipative inequality.

In bounding the dissipative inequality, the design parameters are scaling coefficients multiplying the weighting matrices in the RDE. Then, a sufficient condition for the GTF to be asymptotically stable is derived which is also shown as the sufficient condition of the GTF to be dissipative. Furthermore, the disturbance attenuation property is implied by the dissipativity of the GTF. Next, for a class of dynamical system with modifiable nonlinear measurement functions but linear dynamics, an implementable stabilizing time-varying, nonlinear game-theoretic controller (GTC) is derived by bounding the dissipative inequality of the feedback control system. This class includes the terminal phase of the missile guidance problem. The structure of this GTC is assumed from a natural modification of the corresponding linear quadratic game problem to accommodate the above mentioned modifiable nonlinearity. Sufficient conditions for both the dissipativity and the internal stability of the overall measurement feedback control system are obtained provided two coupled RDE are satisfied. These RDEs are somewhat modified from those obtained from the corresponding linear quadratic problem.

In [17], over a finite-time interval, for a class of linear time-varying dynamical systems with a small nonlinearity, an approximate nonlinear optimal estimation scheme is derived based on a deterministic game-theoretic criterion. Using the calculus of variation approach, this game-theoretic criterion is first maximized by the process disturbance and initial state vectors. The resulting optimality condition is expanded with respect to a small parameter ε and the expression of the worst case state and the Lagrange Multiplier vectors are determined. Subsequently, the approximate game-theoretic estimator is derived by minimizing each term in the series of cost criterion over the corresponding element of state estimate vector expansion. The estimator Riccati differential equations (RDE) necessary for the first and higher order correction terms are the same as in the zeroth-order case. The first-order and higher-order correction terms are computed on-line based on nonlinear functions evaluated along the minimax trajectory of the zeroth-order state estimate which has to be updated, through a backward integration, as each new measurement becomes available. The infinite-order approximate minimax estimator is shown to be a priori disturbance attenuating. The Nth-order approximate minimax estimator achieves disturbance attenuation with an incremental increase in the bound proportional to ε^{N+1}.

2.6 Near-Optimal Missile Guidance For Tactical and Theater Defense

In [18], a perturbation procedure is applied to the problem of finding an optimal control for a ballistic missile interceptor. Certain forces, such as thrust and gravity, are assumed to dominate the equations of motion. The optimal control problem is integrable if the remaining forces are neglected; the approximate effects of the neglected forces can be calculated noniteratively and added to the solution. For certain trajectories, however, the aerodynamic forces are not negligible. Including the aerodynamics directly in the dominant dynamics destroys the analytical solution upon which the procedure depends. Instead, approximations of the aerodynamic forces are included through narrow pulse functions. This technique produces a good approximation to the optimal control and is computationally more efficient than previous methods. Extensions to previous work are also made to account for the interceptor's coast phase and terminal constraints. The near-optimal guidance law is used to produce intercept trajectories against a number of target trajectories. The approximate trajectories compare well with numerically-generate optimal trajectories.

2.7 References

1. P. Whittle, *Risk-Sensitive Optimal Control*, John Wiley and Sons, New York, 1990.
2. C.-H. Fan, J.L. Speyer, and C.R. Jaensch, "Centralized and Decentralized Solution of the Linear-Exponential Gausssian Problem," *IEEE Trans. on Automatic Control*, October 1994.
3. I. Rhee and J.L. Speyer, "A Game Theoretic Approach to Approach to a Finite-time Disturbance Attenuation Problem," *IEEE Trans. on Automatic Control*, Sept. 1991.
4. M.R. Jones, J.S. Baras, and R.J. Elliot, "An Adaptive Controller Based on Disturbance Attenuation," *IEEE Trans. on Automatic Control*, Vol. 40, No. 7, July 1995.
5. I. Rhee and J.L. Speyer, "Application of a Game Theoretic Controller to a Benchmark Problem," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 15, No. 5, September-October 1992.
6. S. Hong and J.L. Speyer, "Robust Control Synthesis with Uncertain Initial Parameters", *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 4, July-August 1995.
7. D.F. Chichka and J.L. Speyer, "An Adaptive Controller Based on Disturbance Attenuation," *IEEE Trans. on Automatic Control*, Vol. 40, No. 7, July 1995.
8. T. Basar and P. Barhard, *$\mathcal{H}_\infty$-Optimal Control and Related Minimax Design Problems*, Birkhäuser, Boston, 1991.
9. J. Yoneyama, J.L. Speyer and C.H. Dillon, "Robust Adaptive Control Problem for Linear Systems with Unknown Parameters," to be published in *Automatica*.
10. C.H. Dillon and J.L. Speyer, "Disturbance Attenuation Approach to Adaptive Control: A Longitudinal Flight Control Example," Proceedings of the AIAA Guidance, navigation and Control Conference, AIAA paper No. 96-3830, July 29-31, 1993.
11. G.A. Bowman and J.L. Speyer, "Detection Filters for Missile Tracking," AIAA Guidance and Control Conference, August 1987.
12. J.E. White and J.L. Speyer, "Detection Filter Design: Spectral Theory and Algorithms," *IEEE Trans. on Automatic Control*, AC-32(7), July 1987.
13. R.K. Douglas and J.L. Speyer, "Robust Fault Detection Filter Design," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 1, January-February 1996.
14. W.H. Chung and J.L. Speyer, "A Game Theoretic Fault Detection Filter," to be published in the *IEEE Trans. on Automatic Control*.
15. D.P. Malladi and J.L. Speyer, "A New Approach to Multiple Model Adaptive Estimation," Submitted to the *IEEE Trans. on Automatic Control*.

16. J. Jang and J.L. Speyer, "A Class of Nonlinear Game-Theoretic Filters and Controllers Using a Dissipative Approach," Proceedings of the IEEE Conference on Decision and Control, December 1994.

17. J. Jang and J.L. Speyer, "Nonlinear Approximate Game-Theoretic Estimation," Proceedings of the IEEE Conference on Decision and Control, December 1996.

18. J.J. Dougherty and J.L. Speyer, "A Near-Optimal Guidance Law for Ballistic Missile Interception," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 2, March-April 1997.

3. Publications

3.1 Journal Publications

1. C.-H. Fan, J.L. Speyer and C.R. Jaensch, "Centralized and Decentralized Solution of the Linear-Exponential Gaussian Problem," *IEEE Trans. on Automatic Control,* October 1994.
2. S. Hong and J.L. Speyer, "Robust Control Synthesis with Uncertain Initial Parameters," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 4, July-August 1995.
3. D.F. Chichka and J.L. Speyer, "An Adaptive Controller Based on Disturbance Attenuation," *IEEE Trans. on Automatic Control,* Vol. 40, No. 7, July 1995.
4. J. Yoneyama, J.L. Speyer and C.H. Dillon, "Robust Adaptive Control Problem for Linear Systems and Unknown Parameters," to be published in *Automatica.*
5. R.K. Douglas and J.L. Speyer, "Robust Fault Detection Filter Design," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 1, January-February 1996.
6. W.H. Chung and J.L. Speyer, "A Game Theoretic Fault Detection Filter," to be published in the *IEEE Trans. on Automatic Control.*
7. J.J. Dougherty and J.L. Speyer, "A New-Optimal Guidance Law for Ballistic Missile Interception," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 2, March-April 1997.

3.2 Papers Presented in Conference Proceedings

1. C.H. Dillon and J.L. Speyer, "Disturbance Attenuation Approach to Adaptive Control: A Longitudinal Flight Control Example," Proceedings of the AIAA Guidance, navigation and Control Conference, AIAA paper No. 96-3830, July 29-31, 1993.
2. J. Jang and J.L. Speyer, "A Class of Nonlinear Game-Theoretic Filters and Controllers Using a Dissipative Approach," Proceedings of the IEEE Conference on Decision and Control, December 1994.
3. J. Jang and J.L. Speyer, "Nonlinear Approximate Game-Theoretic Estimation," Proceedings of the IEEE Conference on Decision and Control, December 1996.

4. Research Professional Personnel

4.1 Professional Personnel

Graduate Students:

David Chichka (Graduated)
Jinsheng Jang (Graduated)
Jun Yoneyama (Graduated)
John Dougherty (Graduated)
John Bain
Charles Dillon
Durga Malladi
Walter Chung

Post-Doctoral Researcher:

Randal Douglas

4.2 Doctoral Degree Awarded and Dissertation Title

1. David Francis Chichka, "A Disturbance Attenuation Approach to a Class of Uncertain Systems", 1994.
2. John Joseph Dougherty, "Near-Optimal Guidance Algorithms for Intercept Vehicles," 1995.
3. Jinsheng Jang, "Nonlinear Approximate Game-Theoretic Estimation and Control," 1996.
4. Jun Yoneyama, "Asymptotic Analysis of Risk-Sensitive Control for Linear Systems with Uncertain Parameters," 1996.

Readers may contact the UCLA Science and Engineering Library, Engineering and Math Sciences Collection for information on obtaining the theses listed above or may refer to the papers in Section 3 of this report.

Engineering and Math Sciences Collection (SEL/EMS)

8251 Boelter Hall
(310) 825-4951
http://www.library.ucla.edu/libraries/sel

Appendix

"A New Approach to Multiple Model Adaptive Estimation"
Submitted to the *IEEE Transactions on Automatic Control*

A New Approach to Multiple Model Adaptive Estimation

Durga P. Malladi * Jason L. Speyer †

Department of Mechanical and Aerospace Engineering
University of California, Los Angeles, CA 90095
email : durga@seas.ucla.edu, speyer@seas.ucla.edu
phone : (310)206-0898, (310)206-4451
fax : (310)206-2302

November 5, 1996

Abstract

A new algorithm for adaptive estimation of time-varying parameters in certain classes of linear stochastic dynamic systems has been developed. The algorithm is based on an adaptive Kalman filter (AKF) whose hypothesized parameters are modified at each stage by generating the probability of each hypothesis, conditioned on the residual history and a given probability of transition. It is then shown that when a particular hypothesis is true, the expected value of the corresponding posterior probability conditioned on the residual history coverges to unity. Moreover, the expected value of the norm of the difference between the constructed error covariance and the true posteriori error covariance is shown to converge to a lower bound close to zero. By invoking an information function, the filter is also shown to be robust with respect to modeling errors. A few numerical simulations have been performed to evaluate this algorithm against the backdrop of the multiple model adaptive estimation (MMAE) scheme.

1 Introduction

We consider a class of adaptive estimation problems wherein the unknown system model may correspond to one of the specified models and the model uncertainty is summarized as a time-varying parametric uncertainty. In particular, we concern ourselves with estimation in linear stochastic systems with time-varying parameters. Earlier attempts to tackle this problem resulted in the development of the Multiple Model Adaptive Estimation (MMAE) algorithm, first proposed by [1] and later generalized by [2] to form the framework of *partitioned algorithms* . This algorithm addresses the most basic adaptive estimation problem, viz, estimation in a linear stochastic system with time-invariant parametric uncertainty. It is a joint estimation and system identification algorithm consisting of a bank of Kalman filters, each "matched" to one hypothesis and an identification subsystem, which maybe construed as the dynamics of

*Graduate Student Researcher, University of California, Los Angeles
†Professor, School of Engineering and Applied Science, University of California, Los Angeles

a sub-optimal multiple hypothesis Wald's Sequential Probability Ratio Test (WSPRT). We denote the underlying dynamics of the WSPRT by F^w_{ki} , which is defined as the posterior probability of hypothesis $\mathcal{H}_i$ conditioned on the residual history upto t_k. The use of F^w_{ki} is motivated by the implicit assumption that we are dealing with a time-invariant parametric uncertainty. However, as stated in [3], there is no rigorous proof that the posterior probability associated with the true model will converge to unity. Moreover, apart from being computationally intensive, this algorithm suffers from ***beta dominance*** [4], which arises out of incorrect system modeling and leads to irregular residuals.

Recently, there have been efforts to improve the performance of the MMAE algorithm [5]. Recall that the recursive relation for the generation of F^w_{ki} does not allow for transitions from one hypothesis to another : It can be shown that if the conditional probability of a particular hypothesis becomes unity/zero, it stays at unity/zero irrespective of what the correct hypothesis is. To avoid this, the recursive relation was modified by upper and lower bounding the conditional probabilities of all hypotheses. Secondly, in an affort to remove beta dominance, the conditional density functions were altered by removing the covariance term from the denominator. The probabilities still sum to one, though the "density" functions are no longer scaled. However, there appears to be no rigorous theoretical justification for both these procedures.

We develop a new algorithm based on a *single* adaptive Kalman filter wherein the time-varying parameters are updated by feeding back the posterior probability of each hypothesis conditioned on the residual process. It is then shown that the expected value of the true posterior probability coverges to unity and, under certain assumptions, the expected value of the norm of the difference between the constructed error covariance and the true posteriori error covariance converges to a lower bound. It is also shown that in the presence of modeling errors, the filter converges to the hypothesis which maximizes a certain information function. We also make a comment about the extension of the MMAE algorithm by using the dynamics F^s_{ki} of a multiple hypothesis Shiryayev sequential probability ratio test (MHSSPRT), which explicitly allows for transitions to occur. This paper is orgainzed as follows. In section 2, we form the framework of the time-varying parameter estimation problem. In section 3, we highlight the salient features of the MMAE scheme. In section 4, we develop the AKF algorithm and in section 5, we derive the properties of this scheme. In section 6, we compare the two algorithms in various numerical simulations. Finally, in section 7, we conclude by summarizing the AKF algorithm and its theoretical properties.

2 Problem Statement

Consider a linear time-varying stochastic system :

$$x_{k+1} = A_k \cdot x_k + b_k + w_k \tag{1}$$

$$y_k = C_k \cdot x_k + d_k + v_k \tag{2}$$

wherein $x_k \in \Re^n$ is the state, $y_k \in \Re^s$ is the measurement, $b_k \in \Re^n$ and $d_k \in Re^s$ are bias vectors, while A_k and C_k are given matrices of appropriate dimensions. Under each hypothesis $\mathcal{H}_i$, the process noise $\{w_k\}$ and measurement noise $\{v_k\}$ sequences are white, with the following statistics :

$$v_k \sim \mathcal{N}(0, V_i) \qquad A_k = A_{ki} \qquad b_k = b_i \tag{3}$$

$$w_k \sim \mathcal{N}(0, W_i) \qquad C_k = C_{ki} \qquad d_k = d_i \tag{4}$$

Note that instead of parameterizing the noise statistics and other matrices, we have hypothesized them. Clearly they are equivalent.

Now, as a particular application, we can reduce the problem of detection and isolation of the occurence of a change in a correlated measurement sequence, by assuming an ARMA model for the measurement process. Assuming the AR-coefficients to be time-varying, we can formulate a state-space equivalent of the ARMA process as :

$$x_{k+1} = A_k \cdot x_k + b_k + w_k \tag{5}$$

$$y_k = C_k \cdot x_k + d_k + v_k \tag{6}$$

wherein $y_k \in \Re^s$ is the measurement, $C_k = [\, y_{k-1} | \dots | y_{k-n} \,]$ is the measurement matrix, $x_k \in \Re^n$ are the AR-coefficients, A_k is a given matrix and b_k and d_k are appropriate bias vectors. Again, from (3)-(4), the process and sensor noise sequences are white with different statistics under different hypotheses.

In any case, the problem maybe stated as follows : Identify the current system model in minimum time, by detecting and isolating a change in the measurement process. As stated earlier, all existing algorithms have an identification subsystem imbedded in them : this subsystem maybe construed as the recursive relation for F^w_{ki} , which assumes that *no change* occurs in the measurement process when the test is in progress. However, in our AKF algorithm, we explicitly model the probability of a transition from one hypothesis to another, thereby allowing for time-varying hypotheses and using the recursive relation for F^s_{ki} [6]. We also extend the MMAE algorithm to time-varying hypotheses by using this F^s_{ki} instead of the bounded F^w_{ki}. Finally, we develop sufficient conditions for the convergence of this adaptive filter structure.

3 MMAE Algorithm

The Multiple Model Adaptive Estimation algorithm and its variations are widely used for parameter estimation in linear stochastic systems [3][4]. Let there be $L + 1$ linear, discrete-time stochastic dynamic systems, each generating measurements corrupted by white noise : therefore, the available measurement sequence maybe assumed to correspond to one of the m different hypotheses. The sensor and process noise statistics vary with each hypothesis. One can then construct a bank of $L+1$ discrete-time Kalman filters, each "matched" to one hypothesis, generating a white residual process, *provided* the corresponding hypothesis is the true one. The residual process becomes the input to the recursive relation for F^w_{ki}, which generates the posterior probability of each hypothesis, conditioned on the measurement sequence. This leads to a neat parallel structure, shown in figure 1.

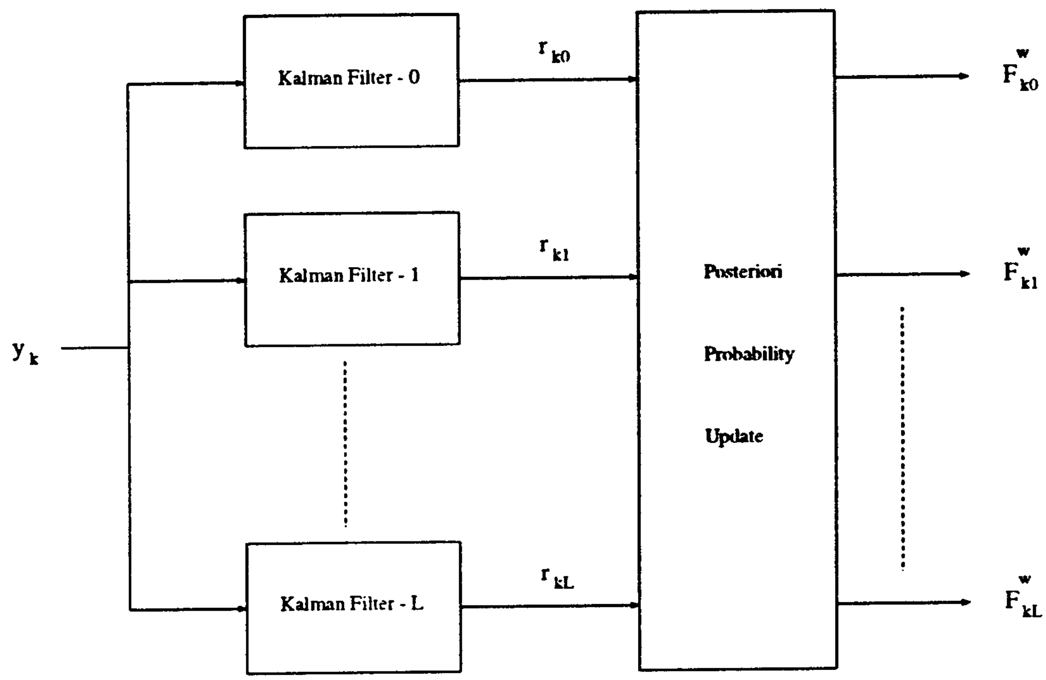

Figure 1: Multiple Model Adaptive Estimation - Lainiotis Filters

The update equations for a generic Kalman Filter for hypothesis $\mathcal{H}_i$ become :

$$S_{ki} = C_{ki} \cdot M_{ki} \cdot {C_{ki}}^T + V_i \qquad (7)$$

$$K_{ki} = M_{ki} \cdot {C_{ki}}^T \cdot {S_{ki}}^{-1} \qquad (8)$$

$$\hat{x}_{ki} = \bar{x}_{ki} + K_{ki} \cdot [\, y_k - C_{ki} \cdot \bar{x}_{ki} - d_{ki} \,] \qquad (9)$$

$$P_{ki} = [\, I - K_{ki} \cdot C_{ki} \,] \cdot M_{ki} \qquad (10)$$

$$r_{ki} = y_k - C_{ki} \cdot \bar{x}_{ki} - d_{ki} \qquad (11)$$

$$\mathcal{R}_{ki} \triangleq [\, r_{1i} \ldots r_{ki} \,] \qquad (12)$$

$$\mathcal{R}_k \triangleq [\, \mathcal{R}_{k1} \ldots \mathcal{R}_{kM} \,] \qquad (13)$$

wherein M_{ki} is the apriori error covariance matrix, P_{ki} is the posteriori covariance matrix, $\bar{x}_{ki}$ is the apriori state estimate and $\hat{x}_{ki}$ is the posteriori state estimate at time t_k. The propagation equations become :

$$\bar{x}_{k+1,i} = A_{ki} \cdot \hat{x}_{ki} + b_{ki} \tag{14}$$

$$M_{k+1,i} = A_{ki} \cdot P_{ki} \cdot A_{ki}{}^T + W_i \tag{15}$$

In the cited literature and in the figure shown, F^w_{ki} is generated. However, allowing transitions from one hypothesis to another, we generate F^s_{ki} :

$$F^s_{kj} \triangleq P\,(\,\mathcal{H}_j\,/\,\mathcal{R}_k\,)$$

The overall posteriori state estimate and error covariance become :

$$x^*_k = \sum_j x_{kj} \cdot F^s_{kj} \tag{16}$$

$$P^*_k = \sum_j \{\, P_{kj} + (\,x^*_k - x_{kj}\,) \cdot (\,x^*_k - x_{kj}\,)^T \,\} \cdot F^s_{kj} \tag{17}$$

We note the following :

- The noise characteristics of each filter are time-invariant.

- As the number of hypotheses increases, the algorithm becomes computationally intensive, as one needs to compute all the time-varying filter gains. To alleviate this problem, sometimes the steady-state gains of each Kalman filter are used, instead of the time-varying gains [3]. But this can lead to convergence to the wrong hypothesis.

- There is no rigorous proof that in the posterior probability associated with the correct hypothesis will converge to unity.

- The recursive relation for F^s_{ki} or F^w_{ki} assumes that the residual sequence is conditionally independent, but when $\mathcal{H}_i$ is true, $\mathcal{R}_{kj}$ is *not* conditionally independent $\forall\ j \neq i$. Hence the generated F^w_{ki} or F^s_{ki} is always wrong no matter what the correct hypothesis is.

- Under certain circumstances [3], the algorithm leads to the convergence to the wrong hypothesis. This phenomenon has been termed as *beta dominance* in [4].

3.1 Beta Dominance

Let $\mathcal{H}_i$ be true : Then, one would expect the residual process r_{ki} to be small while the residuals of the "mismatched" Kalman filters to be large. If for some unknown reason (choice of wrong noise statistics, for example) this doesn't happen, it can be shown that the posterior probability of $\mathcal{H}_i$ might actually decrease, depending upon $S_{kj}\ \forall\ j$. Refer appendix A for the proof.

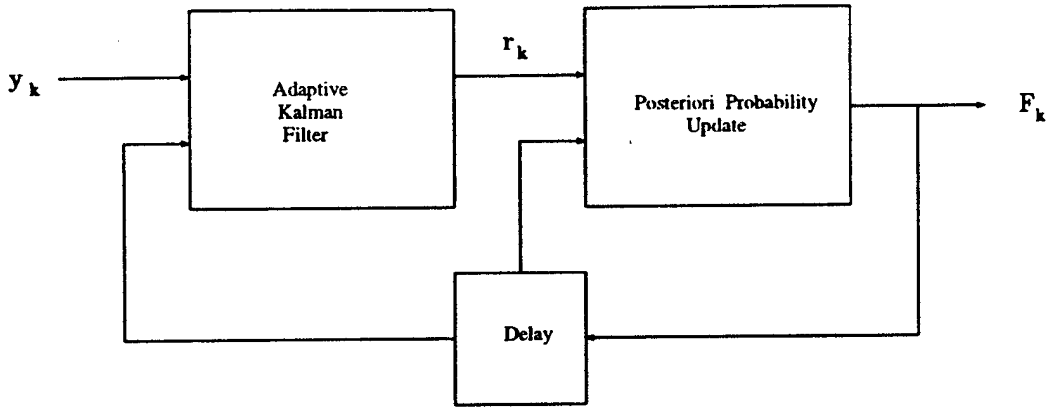

Figure 2: Adaptive Kalman Filter Algorithm

4 Adaptive Kalman Filter Algorithm

We formulate an algorithm based on a structure which uses a single adaptive Kalman filter in conjunction with the recursive relation for F^s_{ki}. Consider figure 2 : An *approximate* posterior probability F_{ki} of each hypothesis conditioned on the residual history is generated and fed back to the filter. All the bias vectors and system matrices, including the process and sensor noise statistics, are updated in the following way :

$$v_k \sim \mathcal{N}(0, V_k) \qquad w_k \sim \mathcal{N}(0, W_k) \qquad \mathcal{R}_k \triangleq [\, r_0\, r_1 \ldots r_k \,] \tag{18}$$

$$F_{kj} \triangleq \text{approximate } F^s_{kj} \qquad F_k \triangleq [\, F_{k0}\;\; F_{k1} \ldots F_{kL} \,]^T \tag{19}$$

$$A_k = \sum_j F_{kj} \cdot A_j \qquad b_k = \sum_j F_{kj} \cdot b_j \qquad W_k = \sum_j F_{kj} \cdot W_j \tag{20}$$

$$C_k = \sum_j F_{k-1,j} \cdot C_j \qquad d_k = \sum_j F_{k-1,j} \cdot d_j \qquad V_k = \sum_j F_{k-1,j} \cdot V_j \tag{21}$$

We derive sufficient conditions for the convergence of F_{kj} to F^s_{kj} in the next section. As mentioned earlier. the structure of the MMAE algorithm never permits the exact calculation of F^s_{ki} or F^w_{ki}. Note that A_k , b_k and W_k are updated using F_{kj}, as it is already available. That is not the case for C_k , d_k and V_k. However, this does not produce any difference in the theoretical results presented later on in Section 5. The filter equations remain the same except that we remove the subscript i from equations (7)-(15). Therefore :

$$S_k = C_k \cdot M_k \cdot C_k^{\,T} + V_k \tag{22}$$

$$K_k = M_k \cdot C_k^{\,T} \cdot S_k^{\,-1} \tag{23}$$

$$\hat{x}_k = \bar{x}_k + K_k \cdot [\, y_k - C_k \cdot \bar{x}_k - d_k \,] \tag{24}$$

$$P_k = [\, I - K_k \cdot C_k \,] \cdot M_k \tag{25}$$

$$r_k = y_k - C_k \cdot \bar{x}_k - d_k \tag{26}$$

$$\bar{x}_{k+1} = A_k \cdot \hat{x}_k + b_k \tag{27}$$

$$M_{k+1} = A_k \cdot P_k \cdot A_k^T + W_k \tag{28}$$

The true error covariance, $\tilde{M}_k$, of this sub-optimal state estimate is *not* computed in this algorithm as it requires the knowledge of the correct hypothesis : instead, as shown later on, it is approximated by M_k. Of course, if $\mathcal{H}_i$ is true, we can compute $\tilde{M}_k$ in the following way :

$$\begin{array}{lll} \hat{e}_k \triangleq x_k - \hat{x}_k & \bar{e}_k \triangleq x_k - \bar{x}_k & \\ m_k \triangleq E\{\, x_k \,/\, \mathcal{R}_k\} & \hat{m}_k^e \triangleq E\{\, \hat{e}_k \,/\, \mathcal{R}_k\} & \bar{m}_k^e \triangleq E\{\, \bar{e}_k \,/\, \mathcal{R}_k\} \\ X_k \triangleq E\{\, x_k \cdot x_k^T \,/\, \mathcal{R}_k\} & \hat{E}_k \triangleq E\{\, x_k \cdot \hat{e}_k^T \,/\, \mathcal{R}_k\} & \bar{E}_k \triangleq E\{\, x_k \cdot \bar{e}_k^T \,/\, \mathcal{R}_k\} \\ \bar{P}_k \triangleq E\{\, \hat{e}_k \cdot \hat{e}_k^T \,/\, \mathcal{R}_k\} & \bar{E}_k \triangleq E\{\, x_k \cdot \bar{e}_k^T \,/\, \mathcal{R}_k\} & \end{array} \tag{29}$$

Clearly :

$$\hat{e}_k = (I - K_k \cdot C_k) \cdot \bar{e}_k - K_k \cdot (C_{ki} - C_k) \cdot x_k - K_k \cdot (d_{ki} - d_k) - K_k \cdot v_k \tag{30}$$

$$\bar{e}_{k+1} = A_k \cdot \hat{e}_k + (A_{ki} - A_k) \cdot x_k + (b_{ki} - b_k) + w_k \tag{31}$$

The difference equations become :

$$\begin{aligned} \tilde{P}_k = {} & (I - K_k \cdot C_k) \cdot \tilde{M}_k \cdot (I - K_k \cdot C_k)^T + K_k \cdot (C_{ki} - C_k) \cdot X_k \cdot (C_{ki} - C_k)^T \cdot K_k^T \\ & + K_k \cdot (d_{ki} - d_k) \cdot (d_{ki} - d_k)^T \cdot K_k^T + K_k \cdot V_i \cdot K_k^T \\ & - 2 \cdot (I - K_k \cdot C_k) \cdot \bar{E}_k^T \cdot (C_{ki} - C_k)^T \cdot K_k^T - 2 \cdot (I - K_k \cdot C_k) \cdot \bar{m}_k^e \cdot (d_{ki} - d_k)^T \cdot K_k^T \\ & + 2 \cdot K_k \cdot (C_{ki} - C_k) \cdot m_k \cdot (d_{ki} - d_k)^T \cdot K_k^T \end{aligned} \tag{32}$$

$$\begin{aligned} \tilde{M}_{k+1} = {} & A_k \cdot \tilde{P}_k \cdot A_k^T + (A_{ki} - A_k) \cdot X_k \cdot (A_{ki} - A_k)^T + (b_{ki} - b_k) \cdot (b_{ki} - b_k)^T + W_i \\ & + 2 \cdot A_k \cdot \hat{E}_k^T \cdot (A_{ki} - A_k)^T + 2 \cdot A_k \cdot \hat{m}_k^e \cdot (b_{ki} - b_k)^T \\ & + 2 \cdot (A_{ki} - A_k) \cdot m_k \cdot (b_{ki} - b_k)^T \end{aligned} \tag{33}$$

$$X_{k+1} = A_{ki} \cdot X_k \cdot A_{ki}^T + b_{ki} \cdot b_{ki}^T + W_i + 2 \cdot A_{ki} \cdot m_k \cdot b_{ki}^T \tag{34}$$

$$\hat{E}_k = \bar{E}_k \cdot (I - K_k \cdot C_k)^T - X_k \cdot (C_{ki} - C_k)^T \cdot K_k^T - m_k \cdot K_k \cdot (d_{ki} - d_k) \tag{35}$$

$$\begin{aligned} \bar{E}_{k+1} = {} & A_{ki} \cdot \hat{E}_k \cdot A_k^T + A_{ki} \cdot X_k \cdot (A_{ki} - A_k)^T + A_k \cdot \hat{m}_k^e \cdot b_{ki}^T \\ & + A_{ki} \cdot m_k \cdot (b_{ki} - b_k)^T + (A_{ki} - A_k) \cdot m_k \cdot b_{ki}^T + b_{ki} \cdot (b_{ki} - b_k)^T + W_i \end{aligned} \tag{36}$$

$$m_{k+1} = A_{ki} \cdot m_k + b_{ki} \tag{37}$$

$$\hat{m}_k^e = (I - K_k \cdot C_k) \cdot \bar{m}_k^e - K_k \cdot (C_{ki} - C_k) \cdot m_k - K_k \cdot (d_{ki} - d_k) \tag{38}$$

$$\bar{m}_{k+1}^e = A_k \cdot \hat{m}_k^e + (A_{ki} - A_k) \cdot m_k + (b_{ki} - b_k) \tag{39}$$

The initial conditions become :

$$X_0 = \tilde{P}_0 + \hat{x}_0 \cdot \hat{x}_0^T \qquad \hat{E}_0 = \tilde{P}_0$$
$$m_0 = \hat{x}_0 \qquad \hat{m}_0^e = 0 \qquad ♣$$

These equations are computationally intensive, but can be computed off-line to assess the performance of the AKF algorithm for each specific application. In certain cases, the off-line computation becomes necessary to analyze the steady state behavior of the AKF algorithm, in particular, its convergence to the correct hypothesis. Given $\mathcal{H}_i$, the true distribution of r_k is :

$$\tilde{f}_{ki}(r_k) \triangleq f_k(r_k / \mathcal{H}_i, \mathcal{R}_{k-1}) \tag{40}$$
$$E[r_k / \mathcal{H}_i, \mathcal{R}_{k-1}] = (C_{ki} - C_k) \cdot \bar{x}_k + (d_{ki} - d_k) \triangleq \tilde{b}_{ki} \tag{41}$$
$$E[r_k \cdot r_k^T / \mathcal{H}_i, \mathcal{R}_{k-1}] = C_{ki} \cdot \tilde{M}_k \cdot C_{ki}{}^T + V_i \triangleq \tilde{S}_{ki} \tag{42}$$

Therefore :

$$(r_k / \mathcal{H}_i, \mathcal{R}_{k-1}) \sim \mathcal{N}(\tilde{b}_{ki}, \tilde{S}_{ki}) \tag{43}$$

However, since the correct hypothesis is unknown, the AKF algorithm is *designed* to approximate the true error covariance $\tilde{M}_k$ with the error "covariance" M_k, computed on-line from (25) and (28). In section 5, this assumption is justified by showing that under certain conditions, $E\{M_k / \mathcal{H}_i\} \rightarrow M_{ki}$, wherein M_{ki} is the apriori error covariance corresponding to the i^{th} filter in the MMAE algorithm. It is also shown that $E\{M_k / \mathcal{H}_i\} \rightarrow M_k^*$, wherein M_k^* is the exact apriori error covariance derived from (17) and $E\{M_k / \mathcal{H}_i\} \rightarrow \tilde{M}_k$. At each t_k, we *assume* :

$$E[r_k / \mathcal{H}_i, \mathcal{R}_{k-1}] = 0$$
$$E[r_k \cdot r_k^T / \mathcal{H}_i, \mathcal{R}_{k-1}] = C_{ki} \cdot M_k \cdot C_{ki}{}^T + V_i \triangleq \Lambda_{ki} \tag{44}$$

Under the gaussian assumption, we explicitly *construct* the density function $f_{ki}(\cdot)$ as :

$$f_{ki}(r_k) \triangleq \text{Approx. } \tilde{f}_{ki}(r_k)$$
$$(r_k / \mathcal{H}_i, \mathcal{R}_{k-1}) \sim \mathcal{N}(0, \Lambda_{ki}) \tag{45}$$

We make a crucial observation that $\{r_k\}$ is no longer an independent residual process. Now, F_{ki}^s requires the knowledge of the density functions $\tilde{f}_{ki}(\cdot)$: since we approximate these functions with $f_{ki}(\cdot)$, F_{ki}^s has been approximated by F_{ki}.

Note that, by removing the parallel structure of the MMAE approach while retaining the time-varying filter gain, the AKF algorithm is computationally less intensive, especially if the number of hypotheses is large. Further, following the argument used in appendix A, it can be shown that *beta dominance* cannot exist in this structure as there is only one residual process here. In the next section, we address issues concerning the convergence and robustness of this algorithm.

5 Performance of AKF Algorithm

We now consider the convergence properties of this algorithm. We propose to prove the following :

- If hypothesis $\mathcal{H}_i$ is true. then the generated posterior probability of $\mathcal{H}_j$ (called F_{kj}) decreases monotonically $\forall\ j \neq i$.
- If hypothesis $\mathcal{H}_i$ is true, then the posterior error covariance P_k of the AKF algorithm converges to P_{ki} of the "correct" filter from the MMAE algorithm.
- If hypothesis $\mathcal{H}_i$ is true, but is not included in the set of probable hypotheses $\Theta \stackrel{\triangle}{=} \{\ \mathcal{H}_j\ \}$, then, the generated posterior probability converges to that hypothesis $\mathcal{H}_m \in \Theta$ which maximizes a particular information function.

5.1 Underlying Assumptions of MHSSPRT

We briefly rederive the MHSSPRT [6] by defining the following notation :

$$\begin{aligned}
\pi_i &= P(\ \mathcal{H}_i\ ) \\
\bar{p}_i &= \text{Apriori probability of change from } \mathcal{H}_0 \text{ to } \mathcal{H}_i \text{ from } t_k \text{ to } t_{k+1},\ \forall\, k \\
f_{ki}(\cdot) &= \text{Approximate probability density function of } r_k \text{ conditioned on } \mathcal{H}_i \text{ and } \mathcal{R}_{k-1} \\
f_{k0}(\cdot) &= \text{Approximate probability density function of } r_k \text{ conditioned on } \mathcal{H}_0 \text{ and } \mathcal{R}_{k-1} \\
L+1 &= \text{Number of hypotheses} \\
\theta_i &= \text{Time of occurence of } \mathcal{H}_i
\end{aligned}$$

At Stage t_1 :

$$P(\theta_i \leq t_1/r_1) = \frac{P(r_1/\theta_i \leq t_1)\cdot P(\theta_i \leq t_1)}{P(r_1)} \tag{46}$$

$$P(r_1) = \sum_{i=1}^{L} P(r_1/\theta_i \leq t_1)\cdot P(\theta_i \leq t_1) + P(r_1/\theta_i > t_1)\cdot P(\theta_i > t_1)$$

$$P(\theta_i \leq t_1) = P(\theta_i \leq t_0) + P(\theta_i = t_1/\theta_i > t_0) \tag{47}$$

$$= \pi_i + \bar{p}_i \cdot (1-\pi_i) \tag{48}$$

$$P(r_1/\theta_i \leq t_1) = f_{1i}(r_1)\cdot dr_1 \tag{49}$$

Strictly, (49) denotes the probability that the measurement lies between r_1 and r_1+dr_1 given the occurrence of $\mathcal{H}_i$ at or before t_1.

$$\sum_{i=1}^{L} P(\theta_i > t_1) = 1 - \sum_{i=1}^{L} P(\theta_i \leq t_1) \tag{50}$$

From (46), we get :

$$F_{1,i} = \frac{[\pi_i + \tilde{p}_i \cdot (1-\pi_i)] \cdot f_{1i}(r_1)}{\sum_{i=1}^{L}[\pi_i + \tilde{p}_i \cdot (1-\pi_i)] \cdot f_{1i}(r_1) + [1 - \sum_{i=1}^{L} \pi_i + \tilde{p}_i \cdot (1-\pi_i)] \cdot f_{10}(r_1)} \tag{51}$$

At Stage 2 :

$$\begin{aligned} P(\theta_i \leq t_2/\mathcal{R}_2) &= \frac{P(\mathcal{R}_2/\theta_i \leq t_2) \cdot P(\theta_i \leq t_2)}{P(\mathcal{R}_2)} && (52)\\ P(r_2/\theta_i \leq t_2\ ,\ r_1) &= f_{2i}(r_2) \cdot dr_2 && (53)\\ P(r_1/\theta \leq t_2) &= \frac{P(\theta_i \leq t_2/r_1) \cdot P(r_1)}{P(\theta_i \leq t_2)} && (54)\\ P(\mathcal{R}_2) &= P(r_2/r_1) \cdot P(r_1) && (55) \end{aligned}$$

Since from (45), we know the density function $f_k\ (\,r_k\ /\ \mathcal{H}_i\ ,\ \mathcal{R}_{k-1}\,)\ \forall\ k$, from (52) :

$$P(\theta_i \leq t_2/\mathcal{R}_2) = \frac{P(r_2/\theta_i \leq t_2\ ,\ r_1) \cdot P(r_1/\theta_i \leq t_2) \cdot P(\theta_i \leq t_2)}{P(\mathcal{R}_2)} \tag{56}$$

Now, from (53),(54),(55), we have :

$$\begin{aligned} P(\theta_i \leq t_2/\mathcal{R}_2) &= \frac{f_{2i}(r_2) \cdot P(\theta_i \leq t_2/r_1) \cdot dr_2}{P(r_2/r_1)} && (57)\\ P(\theta_i \leq t_2/r_1) &= P(\theta_i \leq t_1/r_1) + P(\theta_i = t_2/\theta_i > t_0, r_1) \\ &= F_{1,i} + \tilde{p}_i \cdot (1 - F_{1,i}) && (58)\\ P(r_2/r_1) &= \sum_{i=1}^{L} P(r_2\ /\ \theta_i \leq t_2\ ,\ r_1) \cdot P(\theta_i \leq t_2/r_1) + P(r_2/\theta_i\ ,\ r_1 > t_2) \cdot P(\theta_i > t_2/r_1) \\ &= \sum_{i=1}^{L}[F_{1,i} + \tilde{p}_i \cdot (1 - F_{1,i})] \cdot f_{2i}(r_2) \cdot dr_2 \\ &\quad +[1 - \sum_{i=1}^{L} F_{1,i} + \tilde{p}_i \cdot (1 - F_{1,i})] \cdot f_{20}(r_2) \cdot dr_2 && (59) \end{aligned}$$

Clearly, by induction, we can now write the recursive relation for $F_{k+1,i}$ in terms of $F_{k,i}$:

$$\begin{aligned} F_{k+1,i} &= \frac{[F_{k,i} + \tilde{p}_i \cdot (1 - F_{k,i})] \cdot f_{k+1,i}(r_{k+1})}{\sum_{i=1}^{L}[F_{k,i} + \tilde{p}_i \cdot (1 - F_{k,i})] \cdot f_{k+1,i}(r_{k+1}) + [1 - \sum_{i=1}^{L} F_{k,i} + \tilde{p}_i \cdot (1 - F_{k,i})] \cdot f_{k+1,0}(r_{k+1})} \\ F_{0,i} &= \pi_i && (60) \end{aligned}$$

Nowhere have we made any assumptions about the independence of the residual process. From (45), we explicitly construct $f_{ki}(\cdot)\ \forall\ i$ at each t_k. This approximates F_{ki}^s by F_{ki} as mentioned in the earlier section. In the next section, we derive sufficient conditions for the convergence of F_{ki} and the associated error covariance.

5.2 Convergence of the Posterior Probability

We seek to prove that when $\mathcal{H}_i$ is true, the posterior probabilities of all hypotheses $\mathcal{H}_j \ \forall \ j \neq i$ decrease. We define the following :

$$\mathcal{F} \triangleq \{ f (r / \mathcal{H}) \ : \ \mathcal{H} \in \Theta \} \tag{61}$$

$$\mathcal{J}_{ji} (k) \triangleq E [\ln\{ f_{kj} (r_k) \} / \mathcal{H}_i , \mathcal{R}_{k-1}] \tag{62}$$

$$\rho_{jmi} \triangleq \max_k E [[\frac{f_{kj} (r_k)}{f_{km} (r_k)}]^t / \mathcal{H}_i , \mathcal{R}_{k-1}] \qquad \text{for some } t \in (0,1)$$

$$\triangleq \max_k \rho_{jmi} (\mathcal{R}_{k-1}) \tag{63}$$

Assumption 1 *The family of density functions $\mathcal{F}$ is identifiable, i.e :*

$$f (r / \mathcal{H}_i) = f (r / \mathcal{H}_j) \quad \text{iff } \theta_i = \theta_j \quad \forall \ r \tag{64}$$

This assumption is invoked to prove Assertion 1.

Assertion 1 *In view of assumption 1, beta dominance cannot exist in the AKF algorithm.*

Proof : Refer Appendix B1.

In an effort to illustrate the classes of problems for which $\mathcal{J}_{ji}$ is an information function, we consider a time-varying ARMA process of order n, wherein the measurement noise is different for each hypothesis. Therefore under each hypothesis $\mathcal{H}_i$, the process noise $\{w_k\}$ and measurement noise $\{v_k\}$ sequences are white, with the following statistics :

$$\begin{aligned} v_k &\sim \mathcal{N} (0 , V_i) & A_k &= A_k & b_k &= b_k \\ w_k &\sim \mathcal{N} (0 , W_k) & C_k &= C_k & d_k &= d_k \end{aligned} \tag{65}$$

wherein $C_k = [y_{k-1} | \ldots | y_{k-n}]$ is the measurement matrix. We now prove the following lemma.

Lemma 1 *Let $\mathcal{H}_i$ be true. Then, for the ARMA process shown in (65) :*

- *If $V_i > V_k$ and $\tilde{M}_k \geq M_k$, then $\tilde{M}_{k+1} > M_{k+1} \quad \forall \ k$.*
- *If $V_i < V_k$ and $\tilde{M}_k \leq M_k$, then $\tilde{M}_{k+1} < M_{k+1} \quad \forall \ k$.*

Proof : From (32), (33), (25) and (28)

$$\begin{aligned} \tilde{P}_k &= (I - K_k \cdot C_k) \cdot \tilde{M}_k \cdot (I - K_k \cdot C_k)^T + K_k \cdot V_i \cdot K_k^T \\ P_k &= (I - K_k \cdot C_k) \cdot M_k \cdot (I - K_k \cdot C_k)^T + K_k \cdot V_k \cdot K_k^T \\ \tilde{M}_{k+1} &= A_k \cdot \tilde{P}_k \cdot A_k^T + W_k \\ M_{k+1} &= A_k \cdot P_k \cdot A_k^T + W_k \end{aligned}$$

Therefore :

$$\tilde{P}_k - P_k = (I - K_k \cdot C_k) \cdot (\tilde{M}_k - M_k) \cdot (I - K_k \cdot C_k)^T + K_k \cdot (V_i - V_k) \cdot K_k^T$$
$$\tilde{M}_{k+1} - M_{k+1} = A_k \cdot (\tilde{P}_k - P_k) \cdot A_k$$

Clearly :

$$V_i > V_k \, , \, \tilde{M}_k \geq M_k \quad \Rightarrow \quad \tilde{M}_{k+1} > M_{k+1} \qquad \forall \, k$$
$$V_i < V_k \, , \, \tilde{M}_k \leq M_k \quad \Rightarrow \quad \tilde{M}_{k+1} < M_{k+1} \qquad \forall \, k \qquad \clubsuit$$

As a consequence, from (22),(42) and (44) :

$$V_i > V_j \quad \Rightarrow \quad \Lambda_{kj} < \Lambda_{ki} < \tilde{S}_{ki} \qquad \forall \, k$$
$$V_i < V_j \quad \Rightarrow \quad \Lambda_{kj} > \Lambda_{ki} > \tilde{S}_{ki} \qquad \forall \, k \tag{66}$$

We could consider another class of problems similar to (65) wherein the hypotheses differ only in the process noise statistics, i.e , under $\mathcal{H}_i$:

$$w_k \sim \mathcal{N}(0, W_i) \tag{67}$$

We now prove the follwing lemma.

Lemma 2 *Let $\mathcal{H}_i$ be true. Then, for the process shown in (67) :*

- *If $W_i > W_k$ and $\tilde{M}_k \geq M_k$, then $\tilde{M}_{k+1} > M_{k+1} \quad \forall \, k$.*
- *If $W_i < W_k$ and $\tilde{M}_k \leq M_k$, then $\tilde{M}_{k+1} < M_{k+1} \quad \forall \, k$.*

Proof : The proof is very similar to the one of Lemma 1. From (32), (33), (25) and (28)

$$\tilde{P}_k = (I - K_k \cdot C_k) \cdot \tilde{M}_k \cdot (I - K_k \cdot C_k)^T + K_k \cdot V_k \cdot K_k^T$$
$$P_k = (I - K_k \cdot C_k) \cdot M_k \cdot (I - K_k \cdot C_k)^T + K_k \cdot V_k \cdot K_k^T$$
$$\tilde{M}_{k+1} = A_k \cdot \tilde{P}_k \cdot A_k^T + W_i$$
$$M_{k+1} = A_k \cdot P_k \cdot A_k^T + W_k$$

Therefore :

$$\tilde{P}_k - P_k = (I - K_k \cdot C_k) \cdot (\tilde{M}_k - M_k) \cdot (I - K_k \cdot C_k)^T$$
$$\tilde{M}_{k+1} - M_{k+1} = A_k \cdot (\tilde{P}_k - P_k) \cdot A_k + (W_i - W_k)$$

Clearly :

$$W_i > W_k\,,\ \tilde{M}_k \geq M_k \quad \Rightarrow \quad \tilde{M}_{k+1} > M_{k+1} \qquad \forall\, k$$

$$W_i < W_k\,,\ \tilde{M}_k \leq M_k \quad \Rightarrow \quad \tilde{M}_{k+1} < M_{k+1} \qquad \forall\, k \qquad \clubsuit$$

So, from (22),(42) and (44) :

$$W_i > W_j \quad \Rightarrow \quad \Lambda_{kj} < \Lambda_{ki} < \tilde{S}_{ki} \qquad \forall\, k$$

$$W_i < W_j \quad \Rightarrow \quad \Lambda_{kj} > \Lambda_{ki} > \tilde{S}_{ki} \qquad \forall\, k$$

These results are used in proving the following assertion.

Assertion 2 *If the state estimate bias is sufficiently small and*

$$\Lambda_{kj} < \Lambda_{ki} < \tilde{S}_{ki} \qquad or \qquad \Lambda_{kj} > \Lambda_{ki} > \tilde{S}_{ki} \qquad \forall\, k$$

then $\mathcal{J}_{ji}$ is an information function. Therefore, when $\mathcal{H}_i$ is true :

$$\mathcal{J}_{ii}\,(\,k\,) > \mathcal{J}_{ji}\,(\,k\,) \qquad \forall\ j \neq i\ ,\ \forall\ k \tag{68}$$

Proof : Refer Appendix B2.

Remarks Assertion 2 assumes that convergence to the wrong hypothesis has occurred and proves that under certain conditions, the filter cannot remain in the wrong hypothesis. The conditions spelt out are sufficient but not necessary. Moreover for the processes shown in (65) and (67), the state estimate is unbiased and from Lemmas 1 and 2, Assertion 2 is always valid.

Lemma 3 *Let $\mathcal{H}_i$ be true and $\mathcal{H}_m \in \Theta$ be such that :*

$$\mathcal{J}_{mi}\,(\,k\,) \triangleq \max_j\ \mathcal{J}_{ji}\,(\,k\,) \qquad \forall\ \mathcal{H}_j \in \Theta \qquad \forall\ k \tag{69}$$

$$Then \qquad \rho_{jmi} < 1 \qquad \forall\ j \neq m \tag{70}$$

Proof : Refer Appendix B3.

For the classes of problems that we consider, in view of Assertion 2 and Lemma 3 :

$$\text{If } \mathcal{H}_i \in \Theta \quad \Rightarrow \quad \mathcal{H}_i \equiv \mathcal{H}_m$$

$$\Rightarrow \quad \rho_{jii} < 1 \tag{71}$$

From now on, we only consider the classes of problems for which Assertion 2 is valid. We now proceed to prove the following theorem.

Theorem 1 *Let the family $\mathcal{F}$ be identifiable and $\mathcal{H}_i \in \Theta$ be true. Then :*

$$E\,[\,F_{k,j}\,/\,\mathcal{H}_i\,] \;<\; E\,[\,F_{k-1,j}\,/\,\mathcal{H}_i\,] \qquad \forall\; j \neq i \tag{72}$$

Proof : Let hypothesis $\mathcal{H}_i \in \Theta$ be true. Since $F_{kj} \leq 1$, we have :

$$F_{kj} \leq F_{kj}^t \qquad \text{for some } t \in (0,1) \tag{73}$$

This is a crucial observation and is invoked to use $\rho_{jii} < 1$ from (71). All the terms in the denominator of the recursive relation (60) are postive : therefore we get

$$F_{kj} \;\leq\; \frac{\{\,F_{k-1,j} + \tilde{p}_j \cdot (\,1 - F_{k-1,j}\,)\,\} \cdot f_{kj}\,(r_k)}{\{\,F_{k-1,i} + \tilde{p}_i \cdot (\,1 - F_{k-1,i}\,)\,\} \cdot f_{ki}\,(r_k)} \tag{74}$$

$$\phi_{ji} \;\triangleq\; \frac{\{\,F_{k-1,j} + \tilde{p}_j \cdot (\,1 - F_{k-1,j}\,)\,\}}{\{\,F_{k-1,i} + \tilde{p}_i \cdot (\,1 - F_{k-1,i}\,)\,\}} \tag{75}$$

We note that the density functions $f_{kj}(\cdot)$ are the approximate conditional density functions as defined in (45). Assume that $\tilde{p}_j = \tilde{p} \;\; \forall\, j$. Let

$$F_{k-1,j} \;\geq\; F_{k-1,i} \qquad \forall\; j \neq i\,,\, \forall\; k \,\leq\, N$$

Then :

$$\begin{aligned}
\phi_{ji} &= \frac{\tilde{p} + (\,1-\tilde{p}\,) \cdot F_{k-1,j}}{\tilde{p} + (\,1-\tilde{p}\,) \cdot F_{k-1,j}} \\
&= \frac{F_{k-1,j}}{F_{k-1,i}} \cdot \frac{(\,1-\tilde{p}\,) + \frac{\tilde{p}}{F_{k-1,j}}}{(\,1-\tilde{p}\,) + \frac{\tilde{p}}{F_{k-1,i}}} \\
&\leq \frac{F_{k-1,j}}{F_{k-1,i}}
\end{aligned} \tag{76}$$

From (73),(75).(76) :

$$\begin{aligned}
F_{kj} &\leq [\,\frac{F_{k-1,j}}{F_{k-1,i}}\,]^t \cdot [\,\frac{f_{kj}\,(r_k)}{f_{ki}\,(r_k)}]^t \\
&\leq [\,\frac{F_{k-2,j}}{F_{k-2,i}}\,]^t \cdot [\,\frac{f_{k-1,j}\,(r_{k-1})}{f_{k-1,i}\,(r_{k-1})}]^t \cdot [\,\frac{f_{kj}\,(r_k)}{f_{ki}\,(r_k)}]^t \\
&\leq [\,\frac{\pi_j}{\pi_i}\,]^t \cdot [\,\frac{f_{1j}\,(r_1)}{f_{1i}\,(r_1)}]^t \ldots [\,\frac{f_{kj}\,(r_k)}{f_{ki}\,(r_k)}]^t
\end{aligned} \tag{77}$$

We note that $f_{ki}\,(\,r_k\,)$ are the constructed density functions $f_k\,(\,r_k\,/\,\mathcal{H}_i\,,\,\mathcal{R}_{k-1}\,)$ in (42). Taking the expectation conditioned on $\mathcal{H}_i$:

$$\begin{aligned}
E\,[\,F_{kj}\,/\,\mathcal{H}_i\,] &\leq [\,\frac{\pi_j}{\pi_i}\,]^t \cdot E\,\{[\,\frac{f_{1j}\,(r_1)}{f_{1i}\,(r_1)}]^t \ldots [\,\frac{f_{kj}\,(r_k)}{f_{ki}\,(r_k)}]^t \,/\, \mathcal{H}_i\,\} \\
&= [\,\frac{\pi_j}{\pi_i}\,]^t \cdot \int [\,\frac{f_{1j}\,(r_1)}{f_{1i}\,(r_1)}]^t \ldots [\,\frac{f_{kj}\,(r_k)}{f_{ki}\,(r_k)}]^t \cdot \tilde{f}_{ki}\,(\,\mathcal{R}_k\,/\,\mathcal{H}_i\,) \cdot d\mathcal{R}_k \\
&\triangleq [\,\frac{\pi_j}{\pi_i}\,]^t \cdot I_k
\end{aligned} \tag{78}$$

wherein I_k is the integral. Now :

$$\begin{aligned}
\tilde{f}_{ki}(\mathcal{R}_k / \mathcal{H}_i) &= \tilde{f}_{ki}(r_k / \mathcal{R}_{k-1}, \mathcal{H}_i) \cdot \tilde{f}_{k-1,i}(r_{k-1} / \mathcal{R}_{k-2}, \mathcal{H}_i) \ldots \tilde{f}_{1i}(r_1 / \mathcal{H}_i) \\
I_k &= \int [\frac{f_{1j}(r_1)}{f_{1i}(r_1)}]^t \ldots [\frac{f_{k-1,j}(r_{k-1})}{f_{k-1,i}(r_{k-1})}]^t \cdot \tilde{f}_{k-1,i}(\mathcal{R}_{k-1} / \mathcal{H}_i) \cdot d\mathcal{R}_{k-1} \\
&\cdot \int [\frac{f_{kj}(r_k)}{f_{ki}(r_k)}]^t \cdot \tilde{f}_{ki}(r_k / \mathcal{R}_{k-1}, \mathcal{H}_i) \cdot dr_k \\
&= \int [\frac{f_{1j}(r_1)}{f_{1i}(r_1)}]^t \ldots [\frac{f_{k-1,j}(r_{k-1})}{f_{k-1,i}(r_{k-1})}]^t \cdot \tilde{f}_{k-1,i}(\mathcal{R}_{k-1} / \mathcal{H}_i) \cdot d\mathcal{R}_{k-1} \cdot \rho_{jii}(\mathcal{R}_{k-1}) \\
&\leq I_{k-1} \cdot \rho_{jii}
\end{aligned}$$

Therefore, from (78) :

$$E[F_{k,j} / \mathcal{H}_i] \leq [\frac{\pi_j}{\pi_i}]^t \cdot \rho_{jii}^k \tag{79}$$

From Assertion 2 and Lemma 3, $\rho_{jii} < 1$. Hence, the posterior probability rapidly decreases as $k \to \infty$ and the assumption of $F_{k-1,j} \geq F_{k-1,i} \ \forall \ j \neq i, \forall \ k$ is no longer valid. Let $F_{k-1,j} \leq F_{k-1,i} \ \forall \ j \neq i, \forall \ k \geq N$. Then :

$$\begin{aligned}
\phi_{ji} &= \frac{F_{k-1,j}}{F_{k-1,i}} \cdot \frac{[1 - \tilde{p} \cdot (1/F_{k-1,j} - 1)}{1 - \tilde{p} \cdot (1/F_{k-1,j} - 1)]} \\
&\leq \frac{F_{k-1,j}}{F_{k-1,i}} \cdot [1 - \tilde{p} \cdot (1/F_{k-1,j} - 1)] \\
&\leq \frac{F_{k-1,j}}{F_{k-1,i}} + \tilde{p} \cdot \frac{(1 - F_{k-1,j})}{F_{k-1,i}} \\
&\leq \frac{F_{k-1,j}}{F_{k-1,i}} + \tilde{p}
\end{aligned} \tag{80}$$

Following the argument used above, we get :

$$\begin{aligned}
F_{kj}^t &\leq [\frac{F_{k-1,j}}{F_{k-1,i}} + \tilde{p}]^t \cdot [\frac{f_{kj}(r_k)}{f_{ki}(r_k)}]^t \\
&\leq [[\frac{F_{k-1,j}}{F_{k-1,i}}]^t + \tilde{p}^t] \cdot [\frac{f_{kj}(r_k)}{f_{ki}(r_k)}]^t \\
[\frac{F_{k-1,j}}{F_{k-1,i}}]^t &\leq [[\frac{F_{k-2,j}}{F_{k-2,i}}]^t + \tilde{p}^t] \cdot [\frac{f_{k-1,j}(r_{k-1})}{f_{k-1,i}(r_{k-1})}]^t
\end{aligned}$$

Therefore :

$$\begin{aligned}
\Rightarrow F_{kj}^t &\leq [\frac{F_{k-2,j}}{F_{k-2,i}}]^t \cdot [\frac{f_{kj}(r_k) \cdot f_{k-1,j}(r_{k-1})}{f_{ki}(r_k) \cdot f_{k-1,i}(r_{k-1})}]^t + \tilde{p}^t \cdot [\frac{f_{kj}(r_k) \cdot f_{k-1,j}(r_{k-1})}{f_{ki}(r_k) \cdot f_{k-1,i}(r_{k-1})}]^t \\
&+ \tilde{p}^t \cdot [\frac{f_{kj}(r_k)}{f_{ki}(r_k)}]^t
\end{aligned} \tag{81}$$

Again, taking an expectation conditioned on $\mathcal{H}_i$ and from (63) :

$$E[F_{kj} / \mathcal{H}_i] \leq E\{[\frac{F_{Nj}}{F_{Ni}}]^t / \mathcal{H}_i\} \cdot \rho_{jii}^{k-N} + \tilde{p}^t \cdot [\rho_{jii} + \rho_{jii}^2 + \ldots + \rho_{jii}^{k-N}]$$

$$\triangleq \epsilon_k \tag{82}$$

$$\epsilon_{k+1} - \epsilon_k = \rho_{jii}^{k-N} \cdot E\,\{[\frac{F_{Nj}}{F_{Ni}}]^t \,/\, \mathcal{H}_i\,\} \cdot (\,\rho_{jii} - 1\,) \,+\, \rho_{jii}^{k+1-N} \cdot \tilde{p}^t \tag{83}$$

$$< 0 \qquad \forall\ k \geq N_k$$

Hence :

$$\epsilon < \ldots < \epsilon_{k+1} < \epsilon_k < \epsilon_{k-1} < \ldots \qquad \forall\ k \geq N_k \tag{84}$$

Clearly, for the WSPRT, $\tilde{p} = 0$. and hence, the lower bound $\epsilon = 0$. ♣

Corollary 1 *Let the family $\mathcal{F}$ be identifiable, $\mathcal{H}_i \notin \Theta$ be true. Then :*

$$E\,[\,F_{k,j} \,/\, \mathcal{H}_i\,] \;<\; E\,[\,F_{k-1,j} \,/\, \mathcal{H}_i\,] \qquad \forall\ j \neq m \tag{85}$$

wherein $\mathcal{H}_m$ maximizes the information function defined in (69).

Proof : The proof remains essentially the same as at appropriate places we replace the subscript i by m, except that now $\rho_{jmi} < 1$. The corollary is relevant to the robustness issues associated with any mulitple model adaptive estimation scheme.

5.3 Convergence of the Posterior Error Covariance

Let $\mathcal{H}_i$ be true. We now compare the posteriori error covariance matrices of the adaptive filter and the "true" filter matched to $\mathcal{H}_i$ in the MMAE algorithm. From (7)-(15) and (22)-(28) :

$$P_{ki} = [\,I - K_{ki} \cdot C_{ki}\,] \cdot M_{ki} \tag{86}$$

$$P_k = [\,I - K_k \cdot C_k\,] \cdot M_k \tag{87}$$

$$M_k - M_{ki} = A_{k-1} \cdot P_{k-1} \cdot A_{k-1}^T \,-\, A_{k-1,i} \cdot P_{k-1,i} \cdot A_{k-1,i}^T \,+\, [\,W_{k-1} - W_i\,] \tag{88}$$

Therefore :

$$\begin{aligned} P_k - P_{ki} &= [\,I - K_k \cdot C_k\,] \cdot [\,M_k - M_{ki}\,] \cdot [\,I - K_{ki} \cdot C_{ki}\,]^T \\ &\quad + [\,I - K_k \cdot C_k\,] \cdot M_k \cdot C_{ki}^T \cdot K_{ki}^T \,-\, K_k \cdot C_k \cdot M_{ki} \cdot [\,I - K_{ki} \cdot C_{ki}\,]^T \\ &= [\,I - K_k \cdot C_k\,] \cdot [\,M_k - M_{ki}\,] \cdot [\,I - K_{ki} \cdot C_{ki}\,]^T \,+\, G_{ki} \end{aligned} \tag{89}$$

Now from the definitions of K_k and K_{ki} in (8),(23) :

$$\begin{aligned} G_{ki} &= K_{ki} \cdot C_{ki} \cdot M_k \,-\, K_k \cdot C_k \cdot M_k \cdot C_{ki}^T \cdot K_{ki}^T \,-\, K_k \cdot C_k \cdot M_{ki} \\ &\quad + K_{ki} \cdot C_{ki} \cdot M_{ki} \cdot C_k^T \cdot K_k^T \end{aligned}$$

$$
\begin{aligned}
&= M_{ki} \cdot C_{ki}^T \cdot S_{ki}^{-1} \cdot C_{ki} \cdot M_k \;-\; M_k \cdot C_k^T \cdot S_k^{-1} \cdot C_k \cdot M_k \cdot C_{ki}^T \cdot S_{ki}^{-1} \cdot C_{ki} \cdot M_{ki} \\
&\quad - M_k \cdot C_k^T \cdot S_k^{-1} \cdot C_k \cdot M_{ki} \;+\; M_{ki} \cdot C_{ki}^T \cdot S_{ki}^{-1} \cdot C_{ki} \cdot M_{ki} \cdot C_k^T \cdot S_k^{-1} \cdot C_k \cdot M_k \\
&= M_{ki} \cdot [\, C_{ki}^T \cdot S_{ki}^{-1} \cdot C_{ki} \;-\; C_k \cdot S_k^{-1} \cdot C_k \;-\; C_{ki}^T \cdot S_{ki}^{-1} \cdot C_{ki} \cdot (\, M_k - M_{ki} \,) \cdot C_k^T \cdot S_k^{-1} \cdot C_k \,] \cdot M_k
\end{aligned}
$$

Now :

$$
C_{ki}^T \cdot S_{ki}^{-1} \cdot C_{ki} \;-\; C_k \cdot S_k^{-1} \cdot C_k \;=\; C_{ki}^T \cdot (\, S_{ki}^{-1} - S_k^{-1} \,) \cdot C_k \;-\; (\, C_k^T \cdot S_k^{-1} + C_{ki}^T \cdot S_{ki}^{-1} \,) \cdot [\, C_k - C_{ki} \,]
$$

Therefore :

$$
\begin{aligned}
G_{ki} &= M_{ki} \cdot C_{ki}^T \cdot [\, S_{ki}^{-1} - S_k^{-1} - S_{ki}^{-1} \cdot C_{ki} \cdot (\, M_k - M_{ki} \,) \cdot C_k^T \cdot S_k^{-1} \,] \cdot C_k \cdot M_k \\
&\quad - M_{ki} \cdot (\, C_k^T \cdot S_k^{-1} + C_{ki}^T \cdot S_{ki}^{-1} \,) \cdot [\, C_k - C_{ki} \,] \cdot M_k \\
&= M_{ki} \cdot C_{ki}^T \cdot S_{ki}^{-1} \cdot [\, S_k - S_{ki} - C_{ki} \cdot (\, M_k - M_{ki} \,) \cdot C_k^T \,] \cdot S_k^{-1} \cdot C_k \cdot M_k \\
&\quad - M_{ki} \cdot (\, C_k^T \cdot S_k^{-1} + C_{ki}^T \cdot S_{ki}^{-1} \,) \cdot [\, C_k - C_{ki} \,] \cdot M_k \\
&= K_{ki} \cdot [\, C_k \cdot M_k \cdot C_k^T + V_k - C_{ki} \cdot M_{ki} \cdot C_{ki}^T - V_i - C_{ki} \cdot M_k \cdot C_k^T + C_{ki} \cdot M_{ki} \cdot C_k^T \,] \cdot K_k \\
&\quad - M_{ki} \cdot (\, C_k^T \cdot S_k^{-1} + C_{ki}^T \cdot S_{ki}^{-1} \,) \cdot [\, C_k - C_{ki} \,] \cdot M_k \\
&= K_{ki} \cdot [\, V_k - V_i \,] \cdot K_k \;+\; K_{ki} \cdot [\, C_k - C_{ki} \,] \cdot (\, M_k \cdot C_k^T + C_{ki} \cdot M_{ki} \,) \cdot K_k^T \\
&\quad - M_{ki} \cdot (\, C_k^T \cdot S_k^{-1} + C_{ki}^T \cdot S_{ki}^{-1} \,) \cdot [\, C_k - C_{ki} \,] \cdot M_k \qquad (90)
\end{aligned}
$$

Simlarly from (88) :

$$
\begin{aligned}
M_k - M_{ki} &= A_{k-1} \cdot [\, P_{k-1} - P_{k-1,i} \,] \cdot A_{k-1,i}^T \;+\; (\, A_{k-1,i} \cdot P_{k-1,i} + A_{k-1} \cdot P_{k-1} \,) \cdot [\, A_{k-1} - A_{k-1,i} \,]^T \\
&\quad + [\, W_{k-1} - W_i \,]
\end{aligned}
$$

From (88)-(90) :

$$
\begin{aligned}
P_k - P_{ki} &= [\, I - K_k \cdot C_k \,] \cdot A_{k-1} \cdot [\, P_{k-1} - P_{k-1,i} \,] \cdot A_{k-1i}^T \cdot [\, I - K_{ki} \cdot C_{ki} \,]^T \\
&\quad + [\, I - K_k \cdot C_k \,] \cdot (\, A_{k-1,i} \cdot P_{k-1,i} + A_{k-1} \cdot P_{k-1} \,) \cdot [\, A_{k-1} - A_{k-1,i} \,]^T \cdot [\, I - K_{ki} \cdot C_{ki} \,]^T \\
&\quad + [\, I - K_k \cdot C_k \,] \cdot [\, W_{k-1} - W_i \,] \cdot [\, I - K_{ki} \cdot C_{ki} \,]^T \;+\; K_{ki} \cdot [\, V_k - V_i \,] \cdot K_k \\
&\quad + K_{ki} \cdot [\, C_k - C_{ki} \,] \cdot (\, M_k \cdot C_k^T + C_{ki} \cdot M_{ki} \,) \cdot K_k^T \\
&\quad - M_{ki} \cdot (\, C_k^T \cdot S_k^{-1} + C_{ki}^T \cdot S_{ki}^{-1} \,) \cdot [\, C_k - C_{ki} \,] \cdot M_k \qquad (91)
\end{aligned}
$$

The filter equations are :

$$
\begin{aligned}
\hat{x}_{ki} &= [\, I - K_{ki} \cdot C_{ki} \,] \cdot A_{k-1,i} \cdot \hat{x}_{k-1,i} \;+\; K_{ki} \cdot (\, y_k - C_{ki} \cdot b_{k-1,i} - d_{ki} \,) \\
\hat{x}_k &= [\, I - K_k \cdot C_k \,] \cdot A_{k-1} \cdot \hat{x}_{k-1} \;+\; K_k \cdot (\, y_k - C_k \cdot b_{k-1} - d_k \,)
\end{aligned}
$$

Denote the state transition matrices as :

$$\Phi_i\,(k,k-1) \;\triangleq\; [\,I - K_{ki}\cdot C_{ki}\,]\cdot A_{k-1,i}$$

$$\Phi\,(k,k-1) \;\triangleq\; [\,I - K_k\cdot C_k\,]\cdot A_{k-1}$$

$$\begin{aligned}\Psi_{ki} \;\triangleq\; & [\,I - K_k\cdot C_k\,]\cdot(\,A_{k-1,i}\cdot P_{k-1,i} + A_{k-1}\cdot P_{k-1}\,)\cdot[\,A_{k-1} - A_{k-1,i}\,]^T\cdot[\,I - K_{ki}\cdot C_{ki}\,]^T \\ & + [\,I - K_k\cdot C_k\,]\cdot[\,W_{k-1} - W_i\,]\cdot[\,I - K_{ki}\cdot C_{ki}\,]^T \;+\; K_{ki}\cdot[\,V_k - V_i\,]\cdot K_k \\ & + K_{ki}\cdot[\,C_k - C_{ki}\,]\cdot(\,M_k\cdot C_k^T + C_{ki}\cdot M_{ki}\,)\cdot K_k^T \\ & - M_{ki}\cdot(\,C_k^T\cdot S_k^{-1} + C_{ki}^T\cdot S_{ki}^{-1}\,)\cdot[\,C_k - C_{ki}\,]\cdot M_k \end{aligned} \tag{92}$$

Therefore :

$$\begin{aligned}\delta P_{ki} \;&\triangleq\; P_k - P_{ki} \\ &= \Phi\,(k,k-1)\cdot\delta P_{k-1,i}\cdot\Phi_i^T\,(k,k-1) \;+\; \Psi_{ki}\end{aligned} \tag{93}$$

We now prove the following theorem :

Theorem 2 *If the system in (1)-(4) is uniformly completely controllable and uniformly completely observable, and if* $\{\Psi_{ki}\}$ *is uniformly bounded and decreasing, then :*

$$E\,\{\|\delta P_{ki}\|\,/\,\mathcal{H}_i\} \;\le\; \mathcal{L}_{ki}$$

wherein :

$$\mathcal{L}_i < \ldots < \mathcal{L}_{k+1,i} < \mathcal{L}_{ki} < \mathcal{L}_{k-1,i} < \ldots \qquad \forall\; k \;\ge\; N_k \tag{94}$$

Proof : From (93)

$$\begin{aligned}\delta P_{ki} \;&=\; \Phi\,(k,0)\cdot\delta P_{0i}\cdot\Phi_i^T\,(k,0) \;+\; \sum_{l=1}^{k}\Phi\,(k,l)\cdot\Psi_{li}\cdot\Phi_i^T\,(k,l) \\ \Rightarrow\; E\,\{\delta P_{ki}\,/\,\mathcal{H}_i\} \;&=\; \Phi\,(k,0)\cdot\delta P_{0i}\cdot\Phi_i^T\,(k,0) \;+\; \sum_{l=1}^{k}\Phi\,(k,l)\cdot E\,\{\Psi_{li}\,/\,\mathcal{H}_i\}\cdot\Phi_i^T\,(k,l)\end{aligned} \tag{95}$$

Since the system is uniformly completely controllable and observable :

$$\begin{aligned}\|\Phi\,(k,l)\| \;&\le\; C_1\cdot e^{-c_2\cdot(k-l)} \\ \|\Phi_i\,(k,l)\| \;&\le\; C_3\cdot e^{-c_4\cdot(k-l)} \qquad \forall\; C_1\,,\,C_2\,,\,c_3\,,\,c_4 \;>\; 0\end{aligned} \tag{96}$$

Further :

$$\begin{aligned}\|W_{k-1} - W_i\| \;&=\; \|\sum_{j=0}^{L-1} F_{k-1,j}\cdot W_j \;-\; W_i\| \\ &\le\; \|(\,1 - F_{k-1,i}\,)\cdot W_a \;-\; (\,1 - F_{k-1,i}\,)\cdot W_i\| \qquad \text{wherein } \mathrm{W_a} \;=\; \max_j \mathrm{W_j} \\ &\le\; F_{k-1,a}\cdot\|W_a - W_i\|\end{aligned} \tag{97}$$

Similarly :

$$
\begin{aligned}
\|A_{k-1} - A_i\| &\leq F_{k-1,b} \cdot \|A_b - A_i\| \qquad \text{wherein } \mathrm{A_b} = \max_{\mathrm{j}} \mathrm{A_j} \\
\|V_k - V_i\| &\leq F_{k-1,c} \cdot \|V_c - V_i\| \qquad \text{wherein } \mathrm{V_c} = \max_{\mathrm{j}} \mathrm{V_j} \\
\|C_k - C_i\| &\leq F_{k-1,d} \cdot \|C_d - C_i\| \qquad \text{wherein } \mathrm{C_d} = \max_{\mathrm{j}} \mathrm{C_j}
\end{aligned}
\tag{98}
$$

From (92),(97),(98) :

$$
E\,\{\|\Psi_{ki}\| \,/\, \mathcal{H}_i\} \;\leq\; \|\Psi\| \cdot \|\epsilon_{ki}\| \tag{99}
$$

wherein from Theorem 1 and (84), $\{\epsilon_{ki}\}$ is monotonically decreasing $\forall\ k \geq N_k$ and bounded from below, and Ψ is some matrix defined from (92). Clearly, from (95)-(99) :

$$
\begin{aligned}
E\,\{\|\delta P_{ki}\| \,/\, \mathcal{H}_i\} &\leq \|C_1\| \cdot \|\delta P_{0i}\| \cdot \|C_3\| \cdot e^{-(c_2+c_4)\cdot k} + \|C_1\| \cdot \|\Psi\| \cdot [\sum_{l=1}^{k} e^{-(c_2+c_4)\cdot(k-l)} \cdot \|\epsilon_{li}\|\,] \cdot \|C_3\| \\
&\triangleq \mathcal{L}_{ki} \\
\mathcal{L}_{k+1,i} - \mathcal{L}_{ki} &= \|C_1\| \cdot \|\delta P_{0i}\| \cdot \|C_3\| \cdot e^{-(c_2+c_4)\cdot k} \cdot [\, e^{-(c_1+c_2)} - 1\,] \\
&\quad + \|C_1\| \cdot \|\Psi\| \cdot \|\epsilon_{1i}\| \cdot \|C_3\| \cdot e^{-(c_2+c_4)\cdot k} \\
&\quad + \|C_1\| \cdot \|\Psi\| \cdot [\sum_{l=1}^{k} e^{-(c_2+c_4)\cdot(k-l)} \cdot \{\|\epsilon_{l+1,i}\| - \|\epsilon_{li}\|\}\,] \cdot \|C_3\| \\
\Rightarrow\ \mathcal{L}_{k+1,i} - \mathcal{L}_{ki} &< 0 \qquad \forall\ k \geq N_k
\end{aligned}
$$

Hence :

$$
\mathcal{L}_i < \ldots < \mathcal{L}_{k+1,i} < \mathcal{L}_{ki} < \mathcal{L}_{k-1,i} < \ldots \qquad \forall\ k \geq N_k \qquad \clubsuit
$$

Remarks : We note that the lower bound is governed by the factor $\tilde{p}$ of the MHSSPRT as it controls the lower bound of the sequence $\{\epsilon_{ki}\}$, as shown in Theorem 1. This theorem, based upon our adaptive filter structure, shows that the apriori "covariance" assumed for the conditional density function of the residual $f_{ki}\,(\,r_k\,)$ approaches the true apriori error covariance, i.e, $E\,\{\,M_k \,/\, \mathcal{H}_i\,\} \rightarrow M_{ki}$. As a special case, we consider the WSPRT, wherein $\tilde{p} = 0$. From Theorem 1, it can be seen :

$$
\begin{aligned}
\epsilon_k &\sim \epsilon \cdot \rho^k \qquad \text{wherein } \rho < 1 \\
\Rightarrow\ E\,\{\|\delta P_{ki}\| \,/\, \mathcal{H}_i\,\} &\leq \|C_1\| \cdot \|\delta P_{0i}\| \cdot \|C_3\| \cdot e^{-(c_2+c_4)\cdot k} \\
&\quad + \|C_1\| \cdot \|\Psi\| \cdot [\frac{\rho^{k+1} \cdot e^{(c_2+c_4)\cdot(k+1)} - e^{(c_2+c_4)\cdot k}}{\rho \cdot e^{c_2+c_4} - 1}] \cdot \|C_3\| \\
\Rightarrow\ E\,\{\|\delta P_{ki}\| \,/\, \mathcal{H}_i\,\} &= 0 \quad \text{as } \mathrm{k} \rightarrow \infty
\end{aligned}
\tag{100}
$$

So, for the WSPRT, $\mathcal{L}_i = 0$. The adaptive filter converges *exactly* to the "true" filter of the MMAE scheme.

The assumptions in the AKF algorithm may be justified in yet another way, by looking at the exact expressions for the overall state estimate and posteriori error covariance, as developed in the MMAE algorithm. Recall from (16) and (17) :

$$
\begin{aligned}
x_k^* &= \sum_j x_{kj} \cdot F_{kj} \\
P_k^* &= \sum_j \{ P_{kj} + (x_k^* - x_{kj}) \cdot (x_k^* - x_{kj})^T \} \cdot F_{kj}
\end{aligned}
$$

Let :

$$
\delta P_k^* \triangleq P_k^* - P_k \tag{101}
$$

We now show that the expected value of δP_k^* conditioned on any hypothesis decreases as $k \to \infty$.

Theorem 3 *If the system in (1)-(4) is uniformly completely controllable and uniformly completely observable, then :*

$$
E \{ \| \delta P_k^* \| \, / \, \mathcal{H}_i \} \leq \mathcal{L}_{ki}^* \qquad \forall \, \mathcal{H}_i
$$

wherein :

$$
\mathcal{L}_i^* < \ldots < \mathcal{L}_{k+1,i}^* < \mathcal{L}_{ki}^* < \mathcal{L}_{k-1,i}^* < \ldots \qquad \forall \, k \geq N_k
$$

Proof : From the above equations

$$
\begin{aligned}
\delta P_k^* &= \sum_j \{ \{ P_{kj} - P_k \} + (x_k^* - x_{kj}) \cdot (x_k^* - x_{kj})^T \} \cdot F_{kj} \\
&= \sum_j \{ \delta P_{kj} + (x_k^* - x_{kj}) \cdot (x_k^* - x_{kj})^T \} \cdot F_{kj} \\
&= \sum_{j \neq i} \{ \delta P_{kj} + (x_k^* - x_{kj}) \cdot (x_k^* - x_{kj})^T \} \cdot F_{kj} \\
&\quad + \{ \delta P_{ki} + (x_k^* - x_{ki}) \cdot (x_k^* - x_{ki})^T \} \cdot F_{ki}
\end{aligned} \tag{102}
$$

Taking the norm and using an analysis similar to (98) :

$$
\begin{aligned}
\delta P_k^* &\leq \sum_{j \neq i} \{ \| \delta P_{kj} \| + \| (x_k^* - x_{kj}) \cdot (x_k^* - x_{kj})^T \| \} \cdot F_{kj} \\
&\quad + \{ \| \delta P_{ki} \| + \| (x_k^* - x_{ki}) \cdot (x_k^* - x_{ki})^T \| \} \cdot F_{ki} \\
\| \delta P_{kj} \| &= \| P_{kj} - P_k \| \\
&= \| P_{kj} - P_{ki} + P_{ki} - P_k \| \\
&\leq \| \delta P_{kji} \| + \| \delta P_{ki} \| \leq \| \delta P_{ji} \| + \| \delta P_{ki} \| \qquad j \neq i
\end{aligned} \tag{103}
$$

$$
\begin{aligned}
\|\delta X_{kj}\| &\triangleq \|(x_k^* - x_{kj}) \cdot (x_k^* - x_{kj})^T\| \\
&\leq (1 - F_{kj})^2 \cdot \|x_{km} - x_{kj}\|^2 \qquad\qquad m \neq j \\
&\leq (1 - F_{kj})^2 \cdot \|\delta X_j\| \qquad\qquad (104)
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \qquad \delta P_k^* &\leq \sum_{j \neq i} \{ \|\delta P_{ki}\| + \|\delta P_{ji}\| + (1 - F_{kj})^2 \cdot \|\delta X_j\| \} \cdot F_{kj} \\
&\quad + \{ \|\delta P_{ki}\| + (1 - F_{ki})^2 \cdot \|\delta X_i\| \} \cdot F_{ki} \qquad\qquad (105)
\end{aligned}
$$

From Theorem 1 and Theorem 2, $\forall\ k \geq N_k$:

$$
\begin{aligned}
E\{F_{k+1,j} / \mathcal{H}_i\} &\leq E\{F_{kj} / \mathcal{H}_i\} \\
E\{\|\delta P_{k+1,i}\| / \mathcal{H}_i\} &\leq E\{\|\delta P_{ki}\| / \mathcal{H}_i\} \\
E\{F_{k+1,i} / \mathcal{H}_i\} &\geq E\{F_{ki} / \mathcal{H}_i\} \qquad\qquad (106)
\end{aligned}
$$

Therefore :

$$
\begin{aligned}
E\{\delta P_k^* / \mathcal{H}_i\} &\leq E\{ [\sum_{j \neq i} (\|\delta P_{ki}\| + \|\delta P_{ji}\| + (1 - F_{kj})^2 \cdot \|\delta X_j\|) \cdot F_{kj}] / \mathcal{H}_i \} \\
&\quad + E\{ [(\|\delta P_{ki}\| + (1 - F_{ki})^2 \cdot \|\delta X_i\|) \cdot F_{ki}] / \mathcal{H}_i \} \\
&= E\{ [\sum_{j \neq i} (\|\delta P_{ji}\| + (1 - F_{kj})^2 \cdot \|\delta X_j\|) \cdot F_{kj}] / \mathcal{H}_i \} \\
&\quad + E\{ [(1 - F_{ki})^2 \cdot \|\delta X_i\| \cdot F_{ki}] / \mathcal{H}_i \} + E\{ \|\delta P_{ki}\| / \mathcal{H}_i \} \\
&\triangleq \mathcal{L}_{ki}^*
\end{aligned}
$$

Now :

$$
\begin{aligned}
\mathcal{L}_{k+1,i}^* - \mathcal{L}_{ki}^* &= \sum_{j \neq i} \|\delta P_{ji}\| \cdot E\{ (F_{k+1,j} - F_{kj}) / \mathcal{H}_i \} + E\{ (\|\delta P_{k+1,i}\| - \|\delta P_{ki}\|) / \mathcal{H}_i \} \\
&\quad + \sum_j \|\delta X_j\| \cdot E\{ ((1 - F_{k+1,j})^2 \cdot F_{k+1,j}) - ((1 - F_{kj})^2 \cdot F_{kj}) / \mathcal{H}_i \} \qquad\qquad (107)
\end{aligned}
$$

Consider the function :

$$
\begin{aligned}
E\{ ((1 - F_{kj})^2 \cdot F_{kj}) / \mathcal{H}_i \} &= E\{ F_{kj} / \mathcal{H}_i \} + E\{ F_{kj}^3 / \mathcal{H}_i \} - 2 \cdot E\{ F_{kj}^2 / \mathcal{H}_i \} \\
&\leq 2 \cdot E\{ F_{kj} / \mathcal{H}_i \} - 2 \cdot E\{ F_{kj}^2 / \mathcal{H}_i \} \\
\text{Now} \quad E\{ F_{kj}^2 / \mathcal{H}_i \} &\geq (E\{ F_{kj} / \mathcal{H}_i \})^2 \\
\Rightarrow \quad E\{ ((1 - F_{kj})^2 \cdot F_{kj}) / \mathcal{H}_i \} &\leq 2 \cdot [E\{ F_{kj} / \mathcal{H}_i \} - (E\{ F_{kj} / \mathcal{H}_i \})^2] \\
&\triangleq \varphi(E\{ F_{kj} / \mathcal{H}_i \})
\end{aligned}
$$

Since $\varphi(\cdot)$ is an increasing function in the interval $[0, 1/2)$ and a decreasing function in the interval $(1/2, 1]$, from (106) and (107), we have :

$$\begin{aligned} E\{((1-F_{k+1,j})^2 \cdot F_{k+1,j}) / \mathcal{H}_i\} &\leq E\{((1-F_{kj})^2 \cdot F_{kj}) / \mathcal{H}_i\} \qquad \forall\ j \neq i \\ E\{((1-F_{k+1,i})^2 \cdot F_{k+1,i}) / \mathcal{H}_i\} &\leq E\{((1-F_{ki})^2 \cdot F_{ki}) / \mathcal{H}_i\} \\ \Rightarrow \quad \mathcal{L}^*_{k+1,i} - \mathcal{L}^*_{ki} &\leq 0 \qquad k \geq N_k \end{aligned} \tag{108}$$

Hence :

$$\mathcal{L}^*_i < \ldots < \mathcal{L}^*_{k+1,i} < \mathcal{L}^*_{ki} < \mathcal{L}^*_{k-1,i} < \ldots \qquad \forall\ k \geq N_k \qquad \clubsuit$$

Clearly the posteriori "covariance" of the AKF algorithm approaches the *exact* posteriori error covariance as computed in the MMAE algorithm.

Finally, we analyze the difference between the assumed error "covariance" P_k and the exact error covariance $\tilde{P}_k$. Let $\mathcal{H}_i$ be true. From (25),(32) and (33) :

$$\delta\tilde{P}_{ki} \triangleq P_k - \tilde{P}_k \tag{109}$$

$$\begin{aligned} &= (I - K_k \cdot C_k) \cdot A_k \cdot (P_{k-1} - \tilde{P}_{k-1}) \cdot A_k^T \cdot (I - K_k \cdot C_k)^T + \tilde{\Psi}_k \\ &= \Phi(k, k-1) \cdot \delta\tilde{P}_{k-1} \cdot \Phi(k, k-1)^T + \tilde{\Psi}_k \end{aligned} \tag{110}$$

wherein :

$$\begin{aligned} \tilde{\Psi}_k = \; & K_k \cdot \{ (V_k - V_i) - (C_k - C_{ki}) \cdot X_k \cdot (C_k - C_{ki})^T + (d_k - d_{ki}) \cdot (d_k - d_{ki})^T \\ & + 2 \cdot (C_k - C_{ki}) \cdot m_k \cdot (d_{ki} - d_k)^T \} \cdot K_k^T \\ & - (I - K_k \cdot C_k) \cdot \{ 2 \cdot \bar{E}_k^T \cdot (C_k - C_{ki})^T - 2 \cdot \bar{m}_k^e \cdot (d_k - d_{ki})^T \} \cdot K_k^T \\ & + (I - K_k \cdot C_k) \cdot \{ (W_{k-1} - W_i) - (A_{k-1} - A_{k-1,i}) \cdot X_{k-1} \cdot (A_{k-1} - A_{k-1,i})^T \\ & - (b_{k-1,i} - b_{k-1}) \cdot (b_{k-1,i} - b_{k-1})^T + 2 \cdot A_{k-1} \cdot \hat{E}_{k-1}^T \cdot (A_{k-1} - A_{k-1,i})^T \\ & + 2 \cdot A_{k-1} \cdot \hat{m}_{k-1}^e \cdot (b_{k-1} - b_{k-1,i})^T \\ & - 2 \cdot (A_{k-1} - A_{k-1,i}) \cdot m_{k-1} \cdot (b_{k-1} - b_{k-1,i})^T \} \cdot (I - K_k \cdot C_k)^T \end{aligned}$$

We now prove that the expected value of $\delta\tilde{P}_{ki}$ conditioned on $\mathcal{H}_i$ decreases as $k \to \infty$.

Theorem 4 *If the system in (1)-(4) is uniformly completely controllable and uniformly completely observable, then :*

$$E\{\|\delta\tilde{P}_{ki}\| / \mathcal{H}_i\} \leq \tilde{\mathcal{L}}_{ki} \qquad \forall\ \mathcal{H}_i$$

wherein :

$$\tilde{\mathcal{L}}_i < \ldots < \tilde{\mathcal{L}}_{k+1,i} < \tilde{\mathcal{L}}_{ki} < \tilde{\mathcal{L}}_{k-1,i} < \ldots \qquad \forall\ k \geq N_k$$

Proof : From (110)

$$\begin{aligned} \delta\tilde{P}_{ki} &= \Phi(k,0)\cdot\delta\tilde{P}_{0i}\cdot\Phi^T(k,0) + \sum_{l=1}^{k}\Phi(k,l)\cdot\tilde{\Psi}_{li}\cdot\Phi^T(k,l) \\ \Rightarrow\quad E\{\delta\tilde{P}_{ki}\,/\,\mathcal{H}_i\} &= \Phi(k,0)\cdot\delta\tilde{P}_{0i}\cdot\Phi^T(k,0) + \sum_{l=1}^{k}\Phi(k,l)\cdot E\{\tilde{\Psi}_{li}\,/\,\mathcal{H}_i\}\cdot\Phi^T(k,l) \end{aligned}$$

The rest of the proof follows that of Theorem 2. ♣

This concludes our analysis to justify the structure of the AKF algorithm. Under $\mathcal{H}_i$, we derived sufficient conditions for the convergence of F_{ki} to F^s_{ki} and P_k to $\tilde{P}_k$, P_{ki} , and P^*_k . In the next section, we test the AKF algorithm in a few numerical simulations

6 Simulations

6.1 Example 1

Consider a scalar dynamic system :

$$\begin{aligned} x_{k+1} &= A_k\cdot x_k + b_k + w_k \\ y_k &= C_k\cdot x_k + d_k + v_k \end{aligned}$$

wherein under each hypothesis :

$$\begin{aligned} \mathcal{H}_0 \;:\; & A_k = -0.5 \quad b_k = 0.00 \quad C_k = 1.00 \quad d_k = 0.00 \\ & v_k \sim \mathcal{N}(0\,,\,1.0) \quad w_k \sim \mathcal{N}(0\,,\,0.001) \\ \mathcal{H}_1 \;:\; & A_k = -0.6 \quad b_k = 0.25 \quad C_k = 1.25 \quad d_k = 0.25 \\ & v_k \sim \mathcal{N}(0\,,\,2.0) \quad w_k \sim \mathcal{N}(0\,,\,0.001) \\ \mathcal{H}_2 \;:\; & A_k = -0.7 \quad b_k = 0.50 \quad C_k = 1.50 \quad d_k = 0.50 \\ & v_k \sim \mathcal{N}(0\,,\,3.0) \quad w_k \sim \mathcal{N}(0\,,\,0.001) \end{aligned}$$

We compared the Adaptive Kalman Filter to the MMAE algorithm. In the MMAE approach, F^u_{ki} was replaced by F^s_{ki} to allow for transitions from one hypothesis to another. Of course, from our earlier discussion, it is clear that the recursive relation is not strictly F^s_{ki} but an approximation to it. In order to design the AKF algorithm, it is essential to consider scenarios when a particular hypothesis is true and the filter is "matched" to the wrong hypothesis. An off-line computation of the true residual error covariance was conducted for all scenarios. It is seen from figure 3 that when $\mathcal{H}_i$ is true and the filter is matched to $\mathcal{H}_j$, either $\Lambda_{kj} < \Lambda_{ki} < \tilde{S}_{ki}$ or $\Lambda_{kj} > \Lambda_{ki} > \tilde{S}_{ki}$. Moreover the matrix B_{ki} in the exponential term

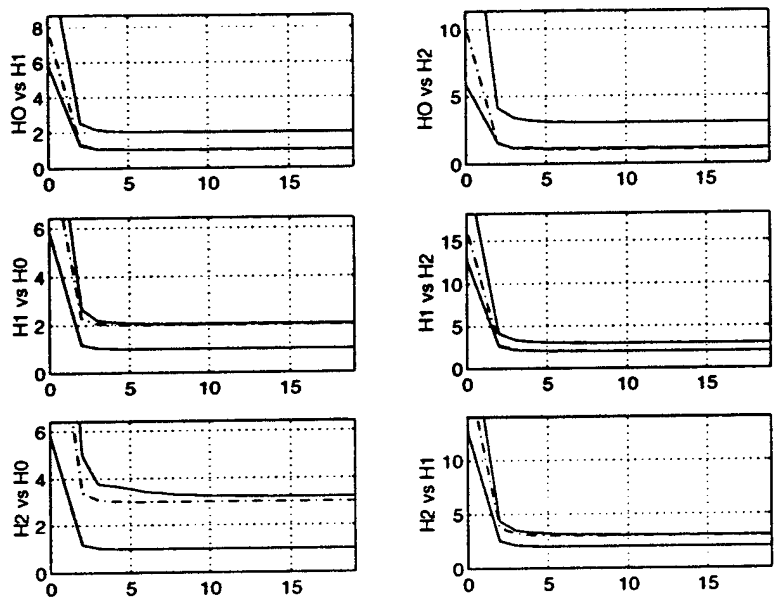

Figure 3: Off-line computation of Λ_{kj} , Λ_{ki} and $\tilde{S}_{ki}$: $\mathcal{H}_i$ vs $\mathcal{H}_j$ denotes $\mathcal{H}_i$ is true while the filter is matched to $\mathcal{H}_j$: Λ_{ki} is shown by the dotted line

is always positive definite and so, from Appendix B2, the system satisfies Assertion 2. This would imply that the filter cannot remain matched to $\mathcal{H}_j$.

We now test the AKF algorithm. At t = 40 sec, the hypothesis was changed from $\mathcal{H}_0$ to $\mathcal{H}_1$. The posterior probabilities of the three hypotheses are shown in figure 4. The bold line denotes the AKF approach while the dotted line denotes the MMAE approach. However, the computational time taken by the MMAE approach is much larger than the AKF approach. These plots have been averaged over ten different realizations.

Figure 5 shows the normed differences between the posteriori error covariance matrix of the AKF and each of the Lainiotis filters. For $t \leq 40$ seconds, $\mathcal{H}_0$ is true : As proved in Theorem 2, $E\ \{\ \|\delta P_{k0}\| / \mathcal{H}_0\ \} \rightarrow 0$ while $E\ \{\ \|\delta P_{k1}\|\ /\ \mathcal{H}_0\ \}$ and $E\ \{\ \|\delta P_{k2}\|\ /\ \mathcal{H}_0\ \}$ are high. For $t > 40$ seconds, $\mathcal{H}_1$ is true : Therefore $E\ \{\ \|\delta P_{k1}\|\ /\ \mathcal{H}_1\ \} \rightarrow 0$ while $E\ \{\ \|\delta P_{k0}\|\ /\ \mathcal{H}_1\ \}$ and $E\ \{\ \|\delta P_{k2}\|\ /\ \mathcal{H}_1\ \}$ are high.

6.2 Example 2

Consider another dynamic system wherein under each hypothesis :

$$\mathcal{H}_0 \ : \ A_k = 0.5 \qquad b_k = 0.00 \qquad C_k = 1.00 \qquad d_k = 0.00$$

$$v_k \sim \mathcal{N}(0, 1.0) \qquad w_k \sim \mathcal{N}(0, 0.001)$$

$$\mathcal{H}_1 \ : \ A_k = 0.6 \qquad b_k = 0.25 \qquad C_k = 1.25 \qquad d_k = 0.25$$

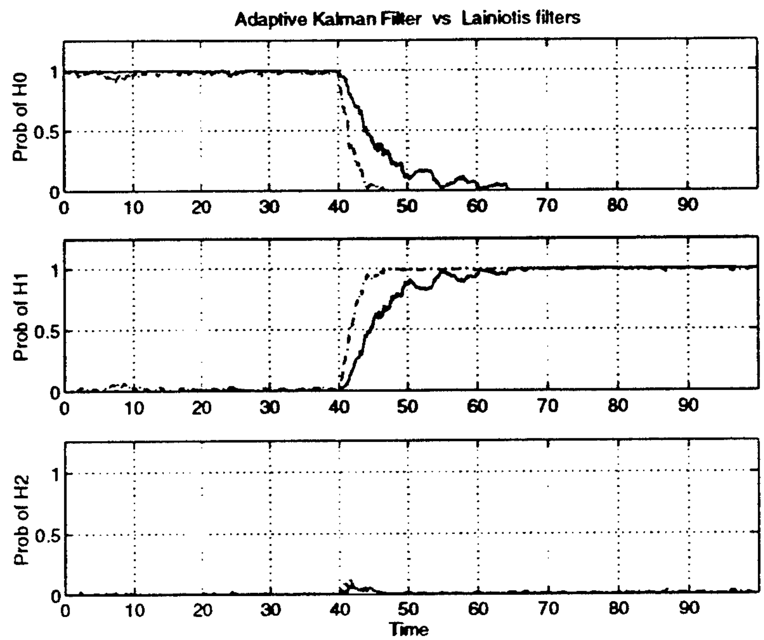

Figure 4: Adaptive Kalman Filter Performance - Change from $\mathcal{H}_0$ to $\mathcal{H}_1$

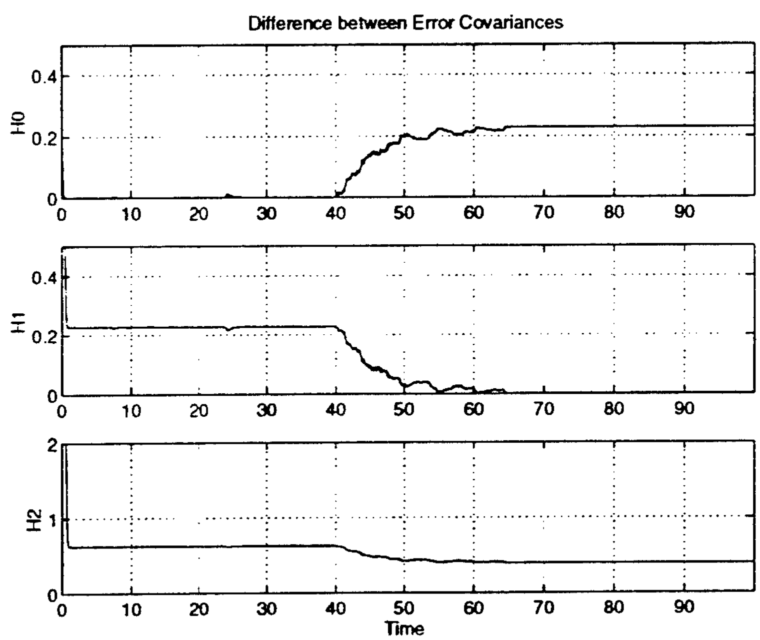

Figure 5: $E\{\ \|\delta P_{ki}\|\ \}$ vs t_k

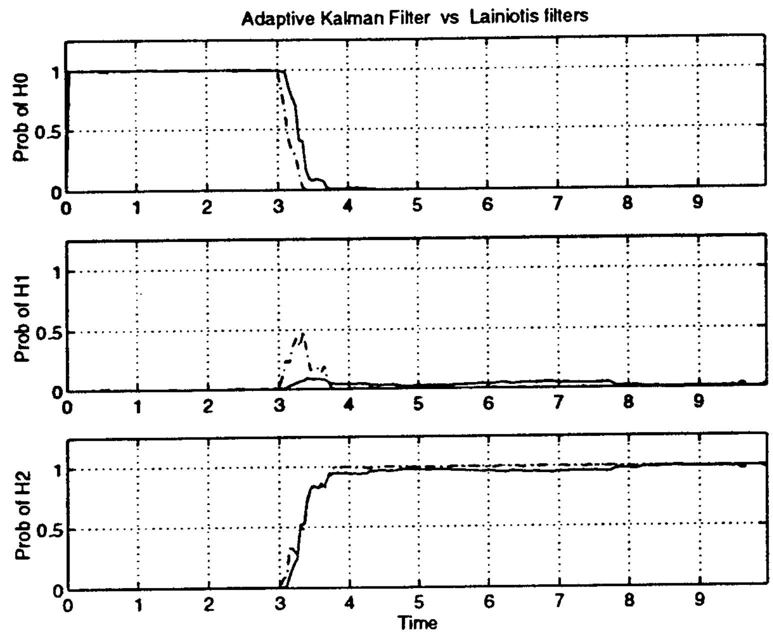

Figure 6: Adaptive Kalman Filter Performance - Change from $\mathcal{H}_0$ to $\mathcal{H}_2$

$$v_k \sim \mathcal{N}(0, 1.5) \qquad w_k \sim \mathcal{N}(0, 0.001)$$

$$\mathcal{H}_2 \; : \; A_k = 0.7 \qquad b_k = 0.50 \qquad C_k = 1.50 \qquad d_k = 0.50$$

$$v_k \sim \mathcal{N}(0, 2.0) \qquad w_k \sim \mathcal{N}(0, 0.001)$$

Again, we compared the Adaptive Kalman Filter to the MMAE algorithm. At $t = 3$ sec, the hypothesis was changed from $\mathcal{H}_0$ to $\mathcal{H}_2$. The posterior probabilities of the three hypotheses are shown in figure 6. The plots have been averaged over ten different realizations.

Figure 7 shows the normed differences between the posteriori error covariance matrix of the AKF and each of the Lainiotis filters.

6.3 Example 3

Consider 3 hypotheses wherein :

$$\mathcal{H}_0 \; : \; v_k \sim \mathcal{N}(0, 1.0)$$

$$\mathcal{H}_1 \; : \; v_k \sim \mathcal{N}(0, 1.5)$$

$$\mathcal{H}_2 \; : \; v_k \sim \mathcal{N}(0, 2.0)$$

A fourth order ARMA measurement process was simulated thus :

$$y_k = 0.1 \cdot [\, y_{k-1} - y_{k-2} + y_{k-3} - y_{k-4} \,] + v_k$$

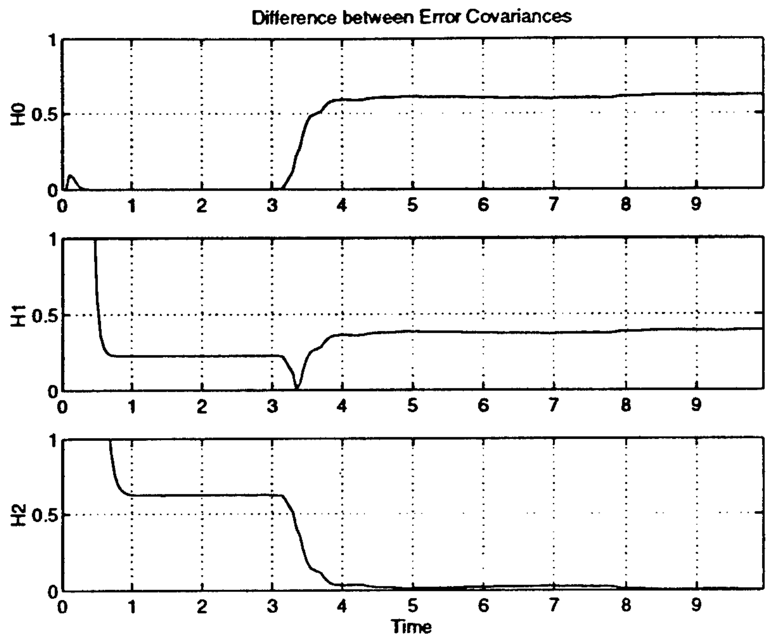

Figure 7: $E\ \{\ \|\delta P_{ki}\|\ \}$ vs t_k

Since the order of the ARMA process is typically unknown, a fifth order ARMA model was assumed for the measurement process. The assumed system model is the same as (1)-(4) and $\forall\ \mathcal{H}_j$:

$$A = I \qquad W_j = 0.001 * I \qquad d = 0$$

$$C_k = [\, y_{k-1} | \cdots | y_{k-5} \,] \qquad b = [0\ 0\ 0\ 0\ 0]^T$$

Recall from Lemma 1 and Appendix B2 that for ARMA processes, Assertion 2 is always valid thereby obviating any off-line computation. At $t = 40$ sec, the hypothesis was changed from $\mathcal{H}_0$ to $\mathcal{H}_1$. The posterior probabilities of the three hypotheses are shown in figure 8. The plots have been averaged over ten different realizations.

Figure 9 shows the normed differences between the posteriori error covariance matrix of the AKF and each of the Lainiotis filters.

6.4 Example 4

For the same system, the hypothesis was changed from $\mathcal{H}_0$ to $\mathcal{H}_2$ at $t = 40$ seconds. The posterior probabilities of the three hypotheses are shown in figure 10. Again the plots have been averaged over ten different realizations.

Figure 11 shows the normed differences between the posteriori error covariance matrix of the AKF and each of the Lainiotis filters.

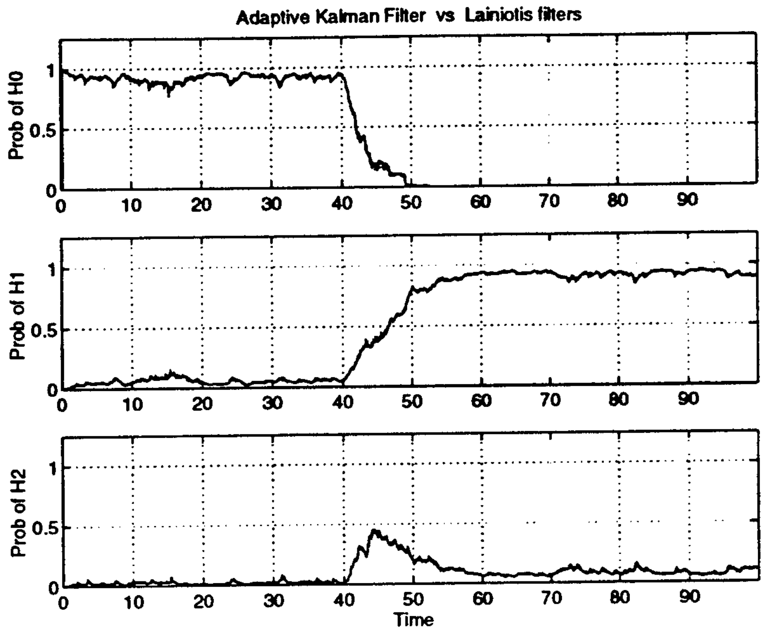

Figure 8: Adaptive Kalman Filter Performance - Change from $\mathcal{H}_0$ to $\mathcal{H}_1$

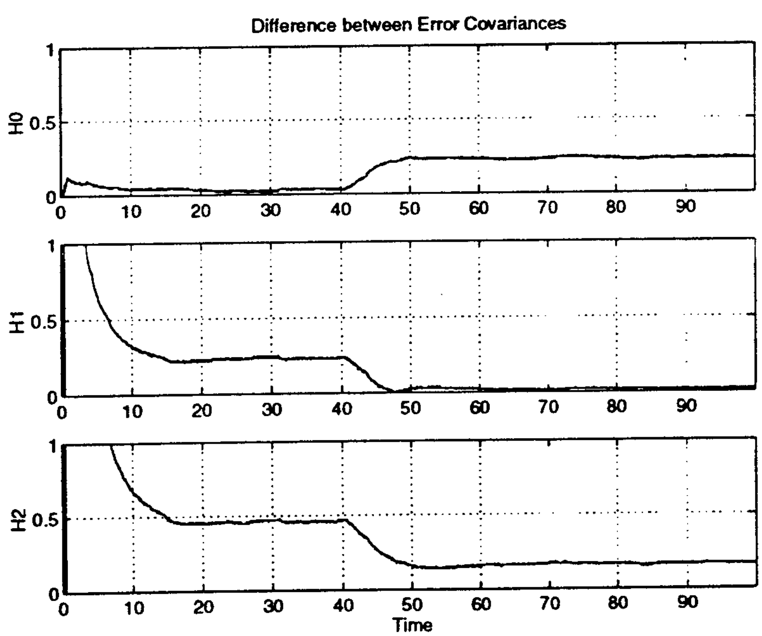

Figure 9: $E\{\ \|\delta P_{ki}\|\ \}$ vs t_k

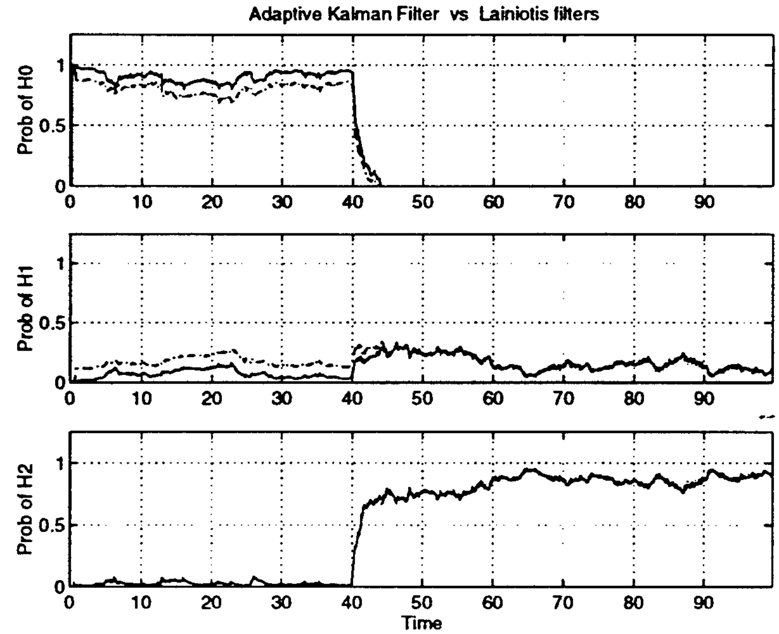

Figure 10: Adaptive Kalman Filter Performance - Change from $\mathcal{H}_0$ to $\mathcal{H}_2$

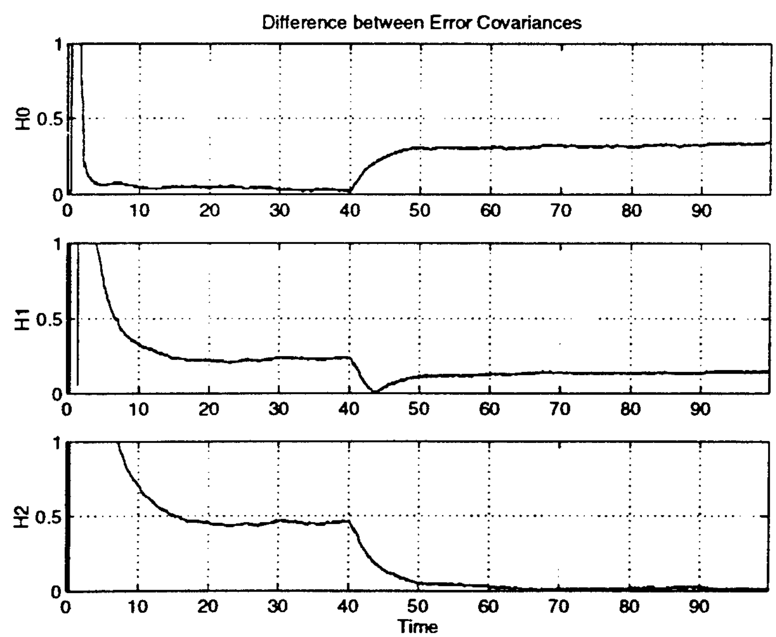

Figure 11: $E\ \{\ \|\delta P_{ki}\|\ \}$ vs t_k

7 Conclusions

An AKF algorithm and sufficient conditions for its convergence have been developed for adaptive estimation in linear time-varying stochastic dynamic systems. In the simulated examples, it is seen to perform on par with the modified MMAE algorithm, while significantly reducing the computational intensity. It has also been shown that for a class of problems, the expected value of the true posterior probability conditioned on the residual history coverges to unity. In its most general form, an off-line computation is necessary to investigate the convergence of the true posterior probability. Under assumptions of uniform complete controllability and observability, the expected value of the norm of the difference between the constructed error covariance and the true posteriori error covariance converges to a lower bound. This lower bound is determined by the apriori probability of change from one hypothesis to another in the MHSSPRT. In the presence of modeling errors, the AKF algorithm has been shown to coverge to the hypothesis which maximizes a particular information function, while the MMAE algorithm might show beta dominance.

8 Acknowledgements

This work was sponsored by NASA-Ames Research Center Cooperative Agreement No. NCC 2-374, supplement 19, Air Force Office of Scientific Research Grant F49620-94-1-0084, and California Department of Transportation Agreement No. 65H 978, MOU 126.

Appendix A

From (60), for the MMAE scheme :

$$\begin{aligned}
\phi_{ki} &\triangleq F_{k,i} + \tilde{p}_i \cdot (1 - F_{k,i}) \\
f_{k+1,i}(r_{k+1,i}) &= \frac{1}{\{2 \cdot \pi\}^{s/2} \cdot \|S_{k+1,i}\|^{1/2}} \cdot exp\left\{-1/2 \cdot r_{k+1,i} \cdot S^{-1}_{k+1,i} \cdot r_{k+1,i}\right\} \\
&\triangleq \beta_{k+1,i} \cdot \alpha_{k+1,j} \\
F_{k+1,i} &= \frac{\phi_{ki} \cdot f_{k+1,i}(r_{k+1,i})}{\sum_{j=0}^{m} \phi_{kj} \cdot f_{k+1,j}(r_{k+1,j})} \\
&= \frac{\phi_{ki} \cdot \beta_{k+1,i} \cdot \alpha_{k+1,i}}{\sum_{j=0}^{m} \phi_{kj} \cdot \beta_{k+1,j} \cdot \alpha_{k+1,j}} \\
F_{k+1,i} - F_{ki} &= \frac{\phi_{ki} \cdot (1 - F_{ki})\beta_{k+1,i} \cdot \alpha_{k+1,i} - \sum_{j \neq i} \phi_{kj} \cdot \beta_{k+1,j} \cdot \alpha_{k+1,j} \cdot F_{ki}}{\sum_{j=0}^{m} \phi_{kj} \cdot \beta_{k+1,j} \cdot \alpha_{k+1,j}}
\end{aligned}$$

If $\mathcal{H}_i$ is true, then we would expect :

$$\alpha_{k+1,j} \approx 0 \quad \forall\ j \neq i$$

$$\Rightarrow\ F_{k+1,i} - F_{ki} \geq 0$$

However, if for some unknown reason, $\alpha_{kj} \approx \alpha \quad \forall\, j$ for a prolonged sequence of measurements :

$$F_{k+1,i} - F_{ki} = \frac{(1-\tilde{p}) \cdot \sum_{j\neq i} F_{kj} \cdot \{\beta_{k+1,i} - \beta_{k+1,j}\} \cdot F_{ki} + \tilde{p} \cdot \sum_{j\neq i} \{\beta_{k+1,i} - \beta_{k+1,j}\} \cdot F_{ki}}{\sum_{j=0}^{m} \phi_{kj} \cdot \beta_{k+1,j}}$$

If $\beta_{ki} < \beta_{kj} \quad \forall j \neq i$, then the posterior probability corresponding to the dominant β increases irrespective of the true hypothesis.

Appendix B1

For the AKF algorithm, there is only *one* residual process. Hence, if for some reason, $\alpha_{kj} \approx \alpha \quad \forall\, j$ for a prolonged sequence of measurements :

$$\Rightarrow \; r_{k+1}^{T} \cdot [\, S_{k+1,i} - S_{k+1,j} \,] \cdot r_{k+1} = 0$$

$$\Rightarrow \; S_{k+1,i} \equiv S_{k+1,j} \qquad \forall\, k$$

This violates the identifiability assumption of the family $\mathcal{F}$. Moreover, since now $\beta_{ki} \equiv \beta_{kj} \;\; \forall\, j$, one β cannot dominate over the other.

Appendix B2

From (62) :

$$\mathcal{J}_{ji}(k) \triangleq E[\, \ln\{ f_{kj}(r_k) \} \,/\, \mathcal{H}_i \,,\, \mathcal{R}_{k-1} \,]$$

Let $\mathcal{H}_i$ be true. Then :

$$\mathcal{J}_{ji}(k) - \mathcal{J}_{ii}(k) = \int \ln\{\frac{f_{kj}(r_k)}{f_{ki}(r_k)}\} \cdot \tilde{f}_{ki}(r_k)\; dr_k$$

$$\text{Since } \ln x \leq x - 1$$

$$\begin{aligned}
\mathcal{J}_{ji}(k) - \mathcal{J}_{ii}(k) &\leq \int \{\frac{f_{kj}(r_k)}{f_{ki}(r_k)} - 1\} \cdot \tilde{f}_{ki}(r_k)\; dr_k \\
&\leq \{\int \{\frac{f_{kj}(r_k)}{f_{ki}(r_k)}\} \cdot \tilde{f}_{ki}(r_k)\; dr_k\} - 1 \qquad \forall\, j \neq i \\
&\triangleq I_k - 1
\end{aligned}$$

From (41), (42) and (44) :

$$I_k = \frac{\|\Lambda_{ki}\|^{1/2} \cdot \|\Lambda_k\|^{1/2}}{\|\Lambda_{kj}\|^{1/2} \cdot \|\tilde{S}_k\|^{1/2}} \cdot exp\,\{ \frac{-1}{2} \cdot a_{ki} \}$$

$$\begin{aligned}
\text{wherein} \quad a_{ki} &\triangleq \tilde{b}_{ki}^{T} \cdot [\, \tilde{S}_{ki} + (\Lambda_{kj}^{-1} - \Lambda_{ki}^{-1})^{-1} \,]^{-1} \cdot \tilde{b}_{ki} \\
&= \tilde{b}_{ki}^{T} \cdot B_{ki} \cdot \tilde{b}_{ki} \\
\Lambda_k^{-1} &\triangleq \tilde{S}_k^{-1} + \Lambda_{kj}^{-1} - \Lambda_{ki}^{-1} \\
&> 0 \;\; \text{for the integral to exist}
\end{aligned}$$

Now :

$$\text{Either} \quad \Lambda_{kj} < \Lambda_{ki} < \tilde{S}_{ki} \quad \text{or} \quad \Lambda_{kj} > \Lambda_{ki} > \tilde{S}_{ki} \quad \forall k$$

Clearly, the integral always exists and $\Lambda_k > 0$. Since the bias terms are small, we neglect the exponential term a_{ki} : however, we do note that if

$$\begin{aligned} \Lambda_{kj} < \Lambda_{ki} &\Rightarrow B_{ki} > 0 \\ &\Rightarrow a_{ki} > 0 \quad \forall\ b_{ki} \\ &\Rightarrow exp\{ \frac{-1}{2} \cdot a_{ki} \} < 1 \end{aligned}$$

Hence in certain cases, the bias terms need not be small. Anyway, removing the exponential term from I_k, we can show that :

$$\begin{aligned} I_k &= \| \Lambda_{ki}^{-1} \cdot \Lambda_{kj} + \Lambda_{ki}^{-1} \cdot \tilde{S}_k - \Lambda_{ki}^{-1} \cdot \Lambda_{kj} \cdot \Lambda_{ki}^{-1} \cdot \tilde{S}_k \|^{-1/2} \\ &\leq 1 \end{aligned}$$

with the equality sign if and only if $\Lambda_{ki} = \Lambda_{kj}$ or $\Lambda_{ki} = \tilde{S}_k \quad \forall k$. The former situation violates the identifiability assumption while the latter assumes that $M_k = \tilde{M}_k \ \forall\ k$, in which case the algorithm has already converged. Therefore :

$$\mathcal{J}_{ji}(k) - \mathcal{J}_{ii}(k) \leq 0$$

Now, the equality sign in holds good if and only if $f_{ki}(\cdot) = f_{kj}(\cdot)$ almost everywhere. Since we assumed the family $\mathcal{F}$ to be identifiable, the $\mathcal{J}_{ji}(k)$ is strictly less than $\mathcal{J}_{ii}(k) \ \forall\ j \neq i$ and $\forall\ k$.

Appendix B3

The proof follows the analysis in [7]. Let $\mathcal{H}_i \notin \Theta$ be true. We first prove that whenever $\mathcal{J}_{ji}(k) < \mathcal{J}_{mi}(k) \ \forall\ k$:

$$E\,[\,\{\frac{f_{kj}(r_k)}{f_{km}(r_k)}\}^t \,/\, \mathcal{H}_i,\ \mathcal{R}_{k-1}\,] \triangleq \rho_{jmi}(\mathcal{R}_{k-1}) < 1 \qquad \text{for some } t \in (0,1) \qquad \forall\ k$$

By definition :

$$\begin{aligned} \mathcal{J}_{ji}(k) - \mathcal{J}_{mi}(k) &= E\,[\,\ln\{\frac{f_{kj}(r_k)}{f_{km}(r_k)}\} \,/\, \mathcal{H}_i,\ \mathcal{R}_{k-1}\,] \\ &= E\,[\,\lim_{t \to 0} (\{\frac{f_{kj}(r_k)}{f_{km}(r_k)}\}^t - 1) \cdot t^{-1} \,/\, \mathcal{H}_i,\ \mathcal{R}_{k-1}\,] \end{aligned}$$

Using the Lebesgue dominated convergence theorem, the limit and expectation may be interchanged. Therefore, for any $\delta \in (0,1)$, there exists a $t \in (0,1)$ such that :

$$
\begin{aligned}
\lim_{t \to 0} t^{-1} \cdot \left(E \left[\left\{ \frac{f_{kj}(r_k)}{f_{km}(r_k)} \right\}^t / \mathcal{H}_i, \mathcal{R}_{k-1} \right] - 1 \right) &\leq \left[\mathcal{J}_{ji}(k) - \mathcal{J}_{mi}(k) \right] \cdot (1 - \delta) \\
E \left[\left\{ \frac{f_{kj}(r_k)}{f_{km}(r_k)} \right\}^t / \mathcal{H}_i, \mathcal{R}_{k-1} \right] &\leq 1 + t \cdot (1 - \delta) \cdot \left[\mathcal{J}_{ji}(k) - \mathcal{J}_{mi}(k) \right] \\
&< 1
\end{aligned}
$$

The same analysis can be carried out $\forall\ k$. In other words, for *any* realization of $\mathcal{R}_k$:

$$
\begin{aligned}
\rho_{jmi}(\mathcal{R}_{k-1}) &< 1 \qquad \forall\ k \\
\Rightarrow \rho_{jmi} &\stackrel{\triangle}{=} \max_k \rho_{jmi}(\mathcal{R}_{k-1}) \\
&< 1
\end{aligned}
$$

References

[1] D.T.Magill. Optimal adaptive estimation of sampled processes. *IEEE Transactions on Automatic Control*, 10:434–439, 1965.

[2] Demetrios G. Lainiotis. Partitioning : A unified framework for adaptive systems : Estimation. *Proceedings of the IEEE*, 64:1126–1143, 1976.

[3] Michael Athans et al. The stochastic control of the f-8c aircraft using a multiple model adaptive control method : Equilibrium flight. *IEEE Transactions on Automatic Control*, 5:768–780, 1977.

[4] Timothy E. Menke and Peter S. Maybeck. Sensor/actuator failure detection in the vista f-16 by multiple model adaptive estimation. *IEEE Transactions on Aerospace and Electronic Systems*, 31:1218–1228, 1995.

[5] Peter S. Maybeck and Peter D. Hanlon. Performance enhancement of a multiple model adaptive estimator. *IEEE Transactions on Aerospace and Electronic Systems*, 31:1240–1253, 1995.

[6] Durga P. Malladi and Jason L. Speyer. A generalized shiryayev sequential probability ratio test for change detection and isolation. *sent to the IEEE Transactions on Automatic Control*, 1996.

[7] Louis A. Liporace. Variance of bayes estimates. *IEEE Transactions on Information Theory*, 17:665–669, 1971.