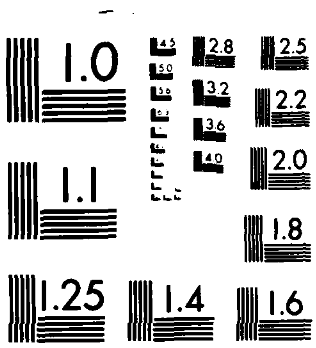
AD-A127 159 ADAPTIVE DETECTION THRESHOLD OPTIMIZATION FOR MULTI-TARGET TRACKING IN CLUTTER(U) BOLT BERANEK AND NEWMAN INC CAMBRIDGE MA S GELFAND ET AL. 31 JAN 83 1/1

UNCLASSIFIED BBN-5249 N00014-80-C-0270 F/G 12/1 NL

END
DATE
FILMED
DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

Bolt Beranek and Newman Inc.

(12)

bbn

Report No. 5249

Adaptive Detection Threshold Optimization For Multi-Target Tracking In Clutter

Saul Gelfand, Thomas E. Fortmann, and Yaakov Bar-Shalom

31 January 1983

Contract N00014-80-C-0270
Task NR 274-310

Prepared for:
Naval Analysis Program
Office of Naval Research

Report No. 5249

ADAPTIVE DETECTION THRESHOLD OPTIMIZATION FOR MULTI-TARGET TRACKING IN CLUTTER

Saul Gelfand, Thomas E. Fortmann, and Yaakov Bar-Shalom

31 January 1983

Prepared by:

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

Prepared for:

Naval Analysis Program
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Attention: Code 411SP

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: BBN Report No. 5249	2. GOVT ACCESSION NO.: AD-A127 154	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): ADAPTIVE DETECTION THRESHOLD OPTIMIZATION FOR MULTI-TARGET TRACKING IN CLUTTER		5. TYPE OF REPORT & PERIOD COVERED: Progress Report February 1982 - January 1983
		6. PERFORMING ORG. REPORT NUMBER: BBN Report No. 5249
7. AUTHOR(s): Saul Gelfand, Thomas E. Fortmann, and Yaakov Bar-Shalom		8. CONTRACT OR GRANT NUMBER(s): N00014-80-C-0270
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Bolt Beranek and Newman Inc. 10 Moulton Street Cambridge, MA 02238		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 65152N NR 274-331
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Analysis Program (Code 411SP) Office of Naval Research Arlington, VA 22217		12. REPORT DATE: 31 January 1983
		13. NUMBER OF PAGES: 56
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report): UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: N.A.

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

detection thresholds, multi-target tracking, data association, data correlation, surveillance, clutter, ROC curves, tracking performance, signal processing, optimization

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

The problem of selecting signal processing parameters, particularly detection thresholds, so as to optimize downstream tracking performance is examined further. Numerical simulations are used to establish the validity of certain approximations made previously. These simulations suggest that steady-state analysis is inadequate, and an adaptive threshold optimization scheme is proposed as an alternative. Finally, the original derivation of the Probabilistic Data Association Filter (PDAF), upon which the present work is founded, is augmented to account for finite gate size.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE OF CONTENTS

	Page
ABSTRACT	1
1. INTRODUCTION	2
2. VALIDATION OF APPROXIMATIONS	8
2.1 Results	10
2.2 Deterministic Gains	20
2.3 Results	22
2.4 Steady State	22
2.5 Conclusions	23
3. DETECTION THRESHOLD OPTIMIZATION	33
3.1 Prior Threshold Optimization	34
3.2 Posterior Threshold Optimization	37
3.3 Simulation Notes and Results	41
3.4 Conclusions	42
4. GATE SIZE OPTIMIZATION AND REVISED PDAF DERIVATION	43

4.1 Gate Size in PDAF Derivation 43
4.2 Approximate Computation of $P^g_{k|k-1}$ 50
4.3 Conclusions 53

ABSTRACT

The problem of selecting signal processing parameters, particularly detection thresholds, so as to optimize downstream tracking performance is examined further. Numerical simulations are used to establish the validity of certain approximations made previously. These simulations suggest that steady-state analysis is inadequate, and an adaptive threshold optimization scheme is proposed as an alternative. Finally, the original derivation of the Probabilistic Data Association Filter (PDAF), upon which the present work is founded, is augmented to account for finite gate size.

1. INTRODUCTION

A critical but well-understood issue in tracking problems involves the <u>inaccuracy</u> of the measurement data, which is typically modeled as additive random noise with known mean and covariance [1-5]. In many tracking problems, particularly those arising in surveillance, there is an equally critical but less-understood <u>uncertainty</u> in the <u>origin</u> of the received data, which may (or may not) include measurements from the target(s) of interest, interfering targets, or random clutter (false alarms). This leads to the problem of data association or data correlation, which has been attacked on a number of fronts [6-14] and surveyed in [15-17]. In this situation, tracking performance depends not only upon the noise covariances, but upon the amount of uncertainty in measurement origin. In some of the approaches cited above [6-10], this dependence is explicit and is characterized in terms of the detection probability P_D and false alarm probability P_F.

As shown in Figure 1, measurement data are typically provided to the tracks by some sort of signal processing and detection algorithm, where the probabilities of detection and false alarm are controlled by the selection of a detection threshold. In a previous stage of this project [20], we established a quantitative relationship between this threshold and the state error covariance matrix in the tracker downstream. More specifically, it was shown that if the tracking is done with an extended Kalman-Bucy filter modified to use probabilistic data association (PDA) [6-8,15], then its conditional covariance update equation (stochastic Riccati equation)[1]

$$\underline{P}_{k|k} = \underline{P}_{k|k-1} - (1-\beta_0)\underline{W}_k S_k \underline{W}_k' + \underline{W}_k\left(\sum_{j=1}^{m} \beta_j \underline{\nu}_j \underline{\nu}_j' - \underline{\nu}\underline{\nu}'\right)\underline{W}_k' \tag{1}$$

can be approximated by the deterministic equation

---

[1] The notation is all defined in [20]

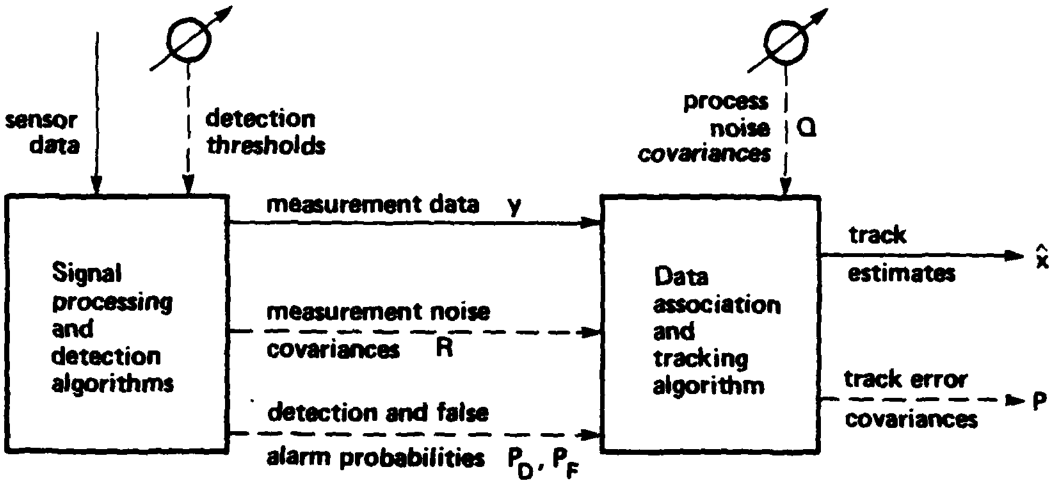


Figure 1. Tracking system block diagram

$$\underline{P}_{k|k} = \underline{P}_{k|k-1} - q_2(\underline{S}_k;P_D,P_F)\underline{W}_k\underline{S}_k\underline{W}_k' \tag{2}$$

where the scalar quantity q_2, which lies between 0 and 1, depends upon $\underline{S}_k$ (via the gate volume $V = c_M g^M |\underline{S}|^{1/2}$), P_D, and P_F. Although $q_2(\underline{S}_k;P_D,P_F)$ is defined by an infinite sum involving nested integrals, it has been evaluated using Monte Carlo integration and reduced to a table look-up procedure that facilitates numerical evaluation of (2).

In [20], (2) was iterated to a steady-state value $\bar{\underline{P}}(P_D,P_F)$ for a particular tracking example, and the steady-state RMS position error $\bar{e}(P_D,P_F)$ was displayed on a contour plot with coordinate axes P_D and P_F, called a <u>tracker operating characteristic</u> (TOC). A typical set of TOC contours is shown in Figure 2, and a corresponding set of <u>receiver operating characteristic</u> (ROC) curves appears in Figure 3.

Each ROC curve represents a locus of possible operating points for the detector/receiver that is providing measurements to the tracker: a particular setting of the <u>detection threshold</u> selects a point along the curve and determines the values of P_D and P_F that will affect the tracking performance via (2). Thus, if the appropriate ROC curve is superimposed on the TOC contours [see the dashed line in Figure 2], the dependence of tracking performance on detection threshold may be determined graphically. The point marked ■, for example, represents an approximately "optimal" choice.

The purpose of this paper is threefold. First, we validate the procedures described in [20] with some more extensive numerical simulations. Second, we propose some adaptive detection threshold optimization schemes. Third, we revise the PDAF derivation in [6] to account accurately for finite gate size.

In Section 2, we validate the procedures described in [20] by comparing the deterministic approximation to the stochastic Riccati equation, the PDAF-

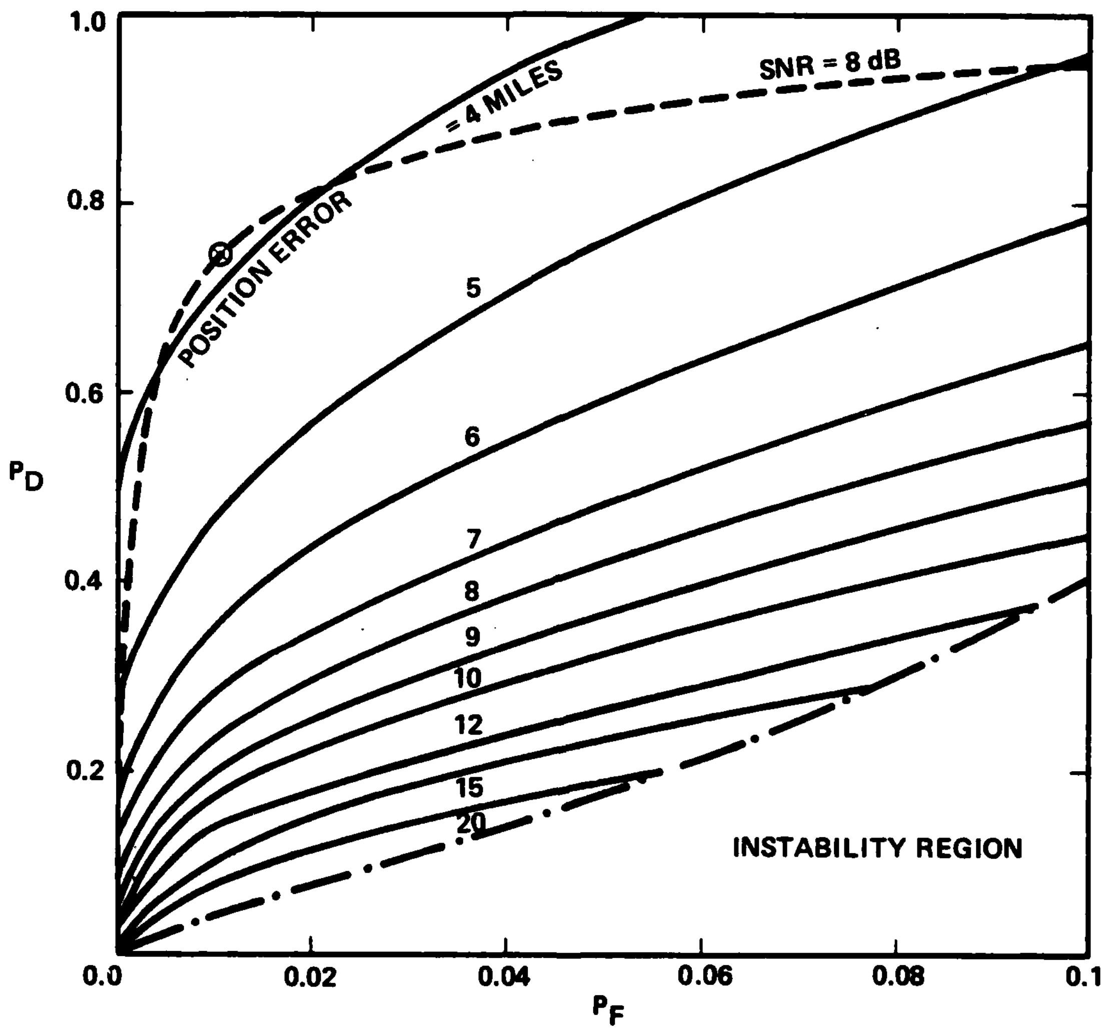


Figure 2. Tracker operating characteristic (TOC) contours

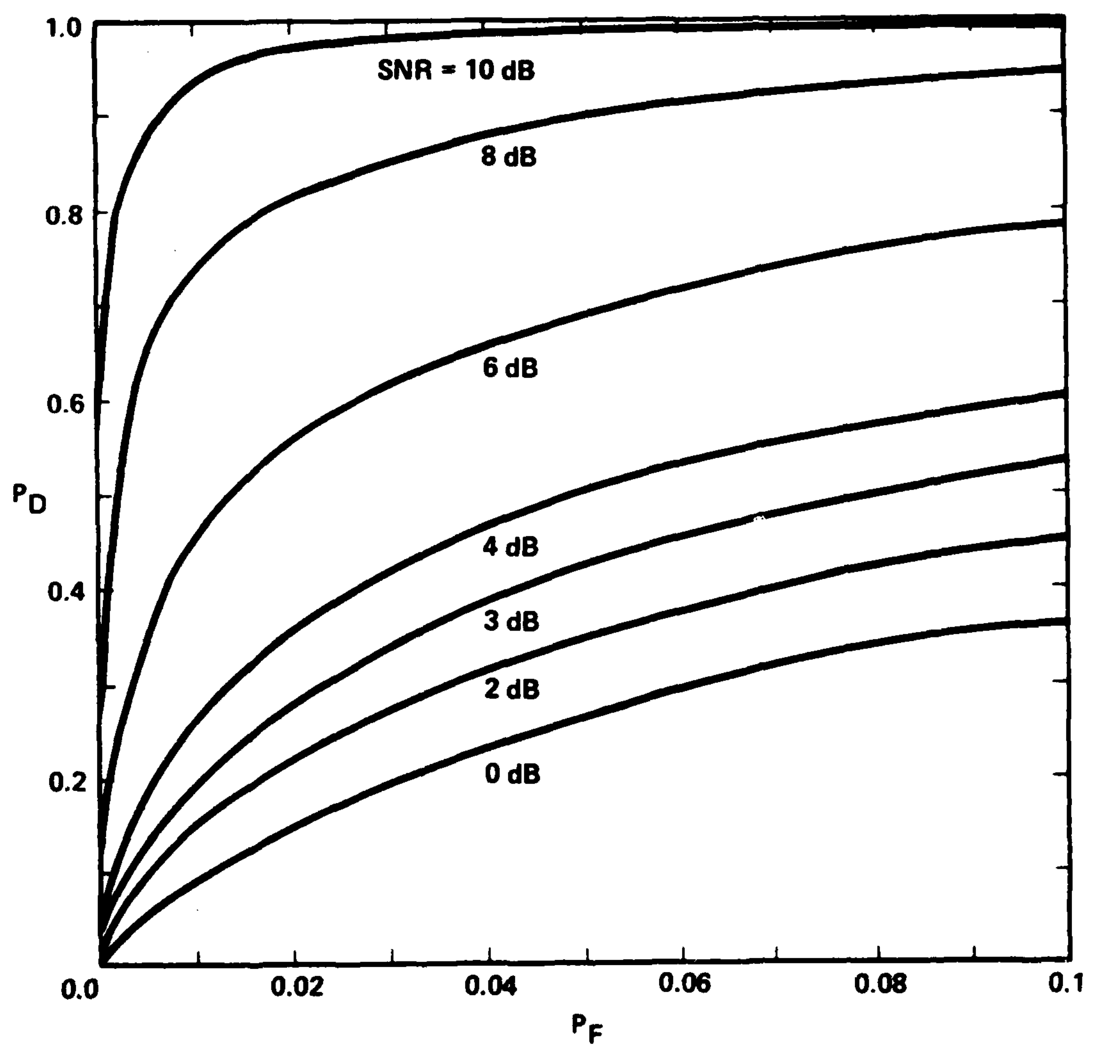


Figure 3. Receiver operating characteristic (ROC) curves

calculated error covariance from the stochastic Riccati equation, and the true PDAF error covariance. Simulation results show that all quantities are comparable for a reasonable range of P_D,P_F. Outside this range, PDAF performance degrades substantially and the computed quantities are inconsistent, especially when velocity errors are considered. Furthermore, a filter with gains based on the deterministic approximation is comparable to the PDAF within a reasonable range of P_D,P_F. Finally, we give several reasons why steady-state values of the deterministic approximation may not be useful in detection threshold optimization.

In Section 3, we propose some adaptive threshold optimization schemes which do not require iteration of the deterministic approximation to convergence. In particular, we give prior and posterior algorithms which minimize the mean square estimation error over detection thresholds which depend on observations up to the previous and current iteration, respectively.

In Section 4, we revise the PDAF derivation from [15] to account for finite gate size. It turns out that the conditional mean update is the same as the PDAF, but the conditional covariance update is increased by an additional term. The derivations in this section are important for theoretical completeness and may be of practical significance under certain operating conditions.

2. VALIDATION OF APPROXIMATIONS

There are several reasons for obtaining a measure of PDAF performance as a function of parameters (e.g., P_D, P_F) for joint detection/tracking problems, including:

1. efficient allocation of communication and computational resources
2. optimization of parameters for improved performance

A computationally efficient method of obtaining such a measure has been proposed in [20] and involves iterating a deterministic approximation to the stochastic Riccati equation. We shall now proceed to validate this procedure.

Using the notation of [20], we let $\hat{x}_{k|k}$ be the output of the PDAF, $P_{k|k}$ the PDAF-calculated conditional error covariance, $P^t_{k|k}$ the true PDAF conditional error covariance, and $P^d_{k|k}$ the output of the deterministic approximation to the stochastic Riccati equation, i.e.,

$$\left.\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + W_k \nu \\ \hat{x}_{k|k-1} &= F\hat{x}_{k-1|k-1} \; ; \qquad \hat{x}_{0|0} \triangleq E\{x_0\} = 0 \end{aligned}\right\} \tag{3}$$

$$\left.\begin{aligned} P_{k|k} &= P_{k|k-1} - (1-\beta_0) W_k S_k W_k' + W_k \Big(\sum_{j=1}^{m} \beta_j \nu_j \nu_j' - \nu\nu' \Big) W_k' \\ S_k &\triangleq H P_{k|k-1} H' + R \; ; \qquad W_k \triangleq P_{k|k-1} S_k^{-1} \\ P_{k|k-1} &= F P_{k-1|k-1} F' + GQG' \; ; \qquad P_{0|0} \triangleq E\{x_0 x_0'\} = P_0 \end{aligned}\right\} \tag{4}$$

$$\left.\begin{aligned} P^t_{k|k} &\triangleq E\{\tilde{x}_{k|k}\tilde{x}'_{k|k}|Y^k\} \\ P^t_{k|k-1} &\triangleq E\{\tilde{x}_{k|k-1}\tilde{x}'_{k|k-1}|Y^{k-1}\} \end{aligned}\right\} \quad (5)$$

$$\left.\begin{aligned} P^d_{k|k} &= P^d_{k|k-1} - q_2(S^d_k; P_D, P_F) W^d_k S^d_k W^{d'}_k \\ S^d_k &\triangleq H P^d_{k|k-1} H' + R \; ; \qquad W^d_k \triangleq P^d_{k|k-1}(S^d_k)^{-1} \\ P^d_{k|k-1} &= F P^d_{k-1|k-1} F' + GQG' \; ; \qquad P^d_{0|0} = P_0 \end{aligned}\right\} \quad (6)$$

Consider the standard PDAF assumption:

$$p(x_k|Y^{k-1}) \sim N(\hat{x}_{k|k-1}, P_{k|k-1}) \quad (7)$$

where N denotes a normal (Gaussian) density. If this assumption is satisfied, then

$$\hat{x}_{k|k} = E\{x_k|Y^k\} \qquad \text{and} \qquad P_{k|k} = P^t_{k|k} \quad (8)$$

To validate the deterministic approximation (6) to the stochastic Riccati equation (3), we compare it with both the true PDAF error covariance and the PDAF-calculated error covariance. Since the standard PDAF assumption is generally not satisfied, the PDAF-calculated error covariance will generally not be equal to the true PDAF error covariance. In the above notation, we will compare $P^d_{k|k}$ with both $E\{P^t_{k|k}\}$ and $E\{P_{k|k}\}$.

2.1 Results

All results are for the numerical problem specified in [20]. Monte Carlo simulations were carried out to compute $\underline{P}^{d}_{k|k}$, $E\{\underline{P}^{t}_{k|k}\}$, and $E\{\underline{P}_{k|k}\}$ for the values of P_D, P_F shown in Figure 4. Sample means were computed from ten (independent) trials using random variates generated by IMSL routines. The values $M\cdot\underline{P}^{d}_{k|k}$, $M\cdot E\{\underline{P}^{t}_{k|k}\}$, and $M\cdot E\{\underline{P}_{k|k}\}$ were computed for

$$M = \text{RMS error} = \sqrt{\text{trace}(\underline{P})}$$

$$M = \text{RMS position error} = \sqrt{P_{11}+P_{22}}$$

$$M = \text{error volume} \propto \sqrt{\text{determinant}(\underline{P})}$$

$$M = \text{position error volume} \propto \sqrt{\text{determinant}\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}}$$

In Figures 5-12, we show plots of RMS error and position error vs. time. From these and the rest of the data we observed the following behavior:

1. For $P_D \geq .4$ and $P_F \leq .06$, $M\cdot\underline{P}^{d}_{k|k} \simeq M\cdot E\{\underline{P}_{k|k}\} \simeq M\cdot E\{\underline{P}^{t}_{k|k}\}$, and typically $|M\cdot E\{\underline{P}_{k|k}\}-M\cdot\underline{P}^{d}_{k|k}| < |M\cdot E\{\underline{P}^{t}_{k|k}\}-M\cdot\underline{P}^{d}_{k|k}|$.

2. For $(P_D,P_F) = (.3,.07)$ and those metrics M which reflect velocity errors (the first and third), PDAF performance degrades substantially, and $M\cdot\underline{P}^{d}_{k|k}$, $M\cdot E\{\underline{P}^{t}_{k|k}\}$, and $M\cdot E\{\underline{P}_{k|k}\}$ are inconsistent.

Because the deterministic approximation to the stochastic Riccati equation is typically comparable to the PDAF-calculated error covariance, a question naturally arises as to the performance of a PDAF with gains based on $P^d_{k|k-1}$. We next explore this issue.

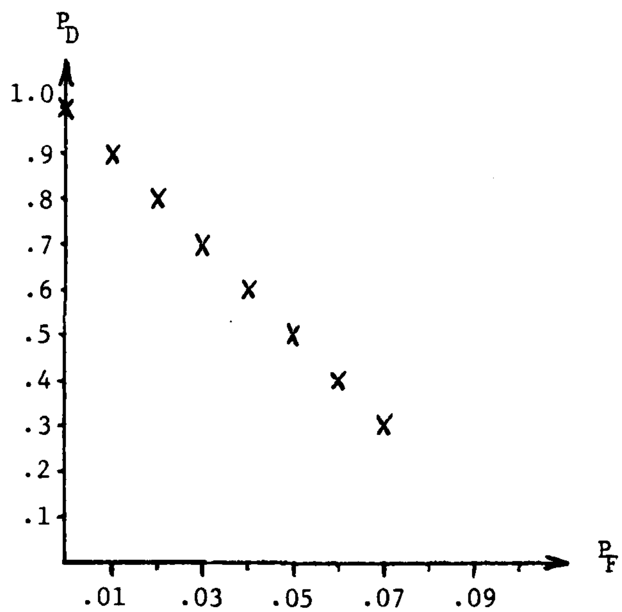


Figure 4. Values of (P_D, P_F) used in simulations

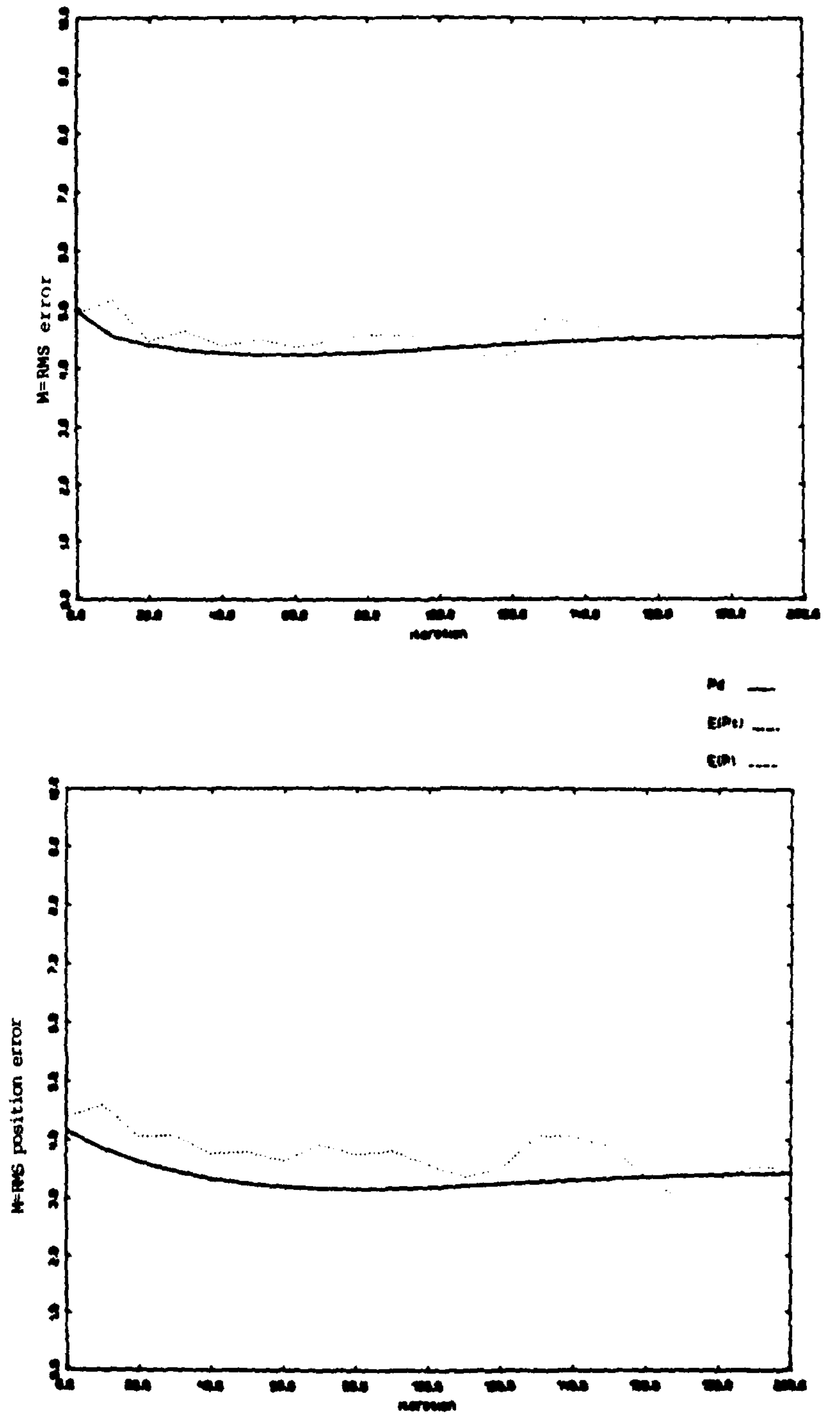


Figure 5. $M \cdot P^{d}_{k|k}$, $M \cdot E\{P^{t}_{k|k}\}$, and $M \cdot E\{P_{k|k}\}$ for $(P_D, P_F) = (1, 0)$

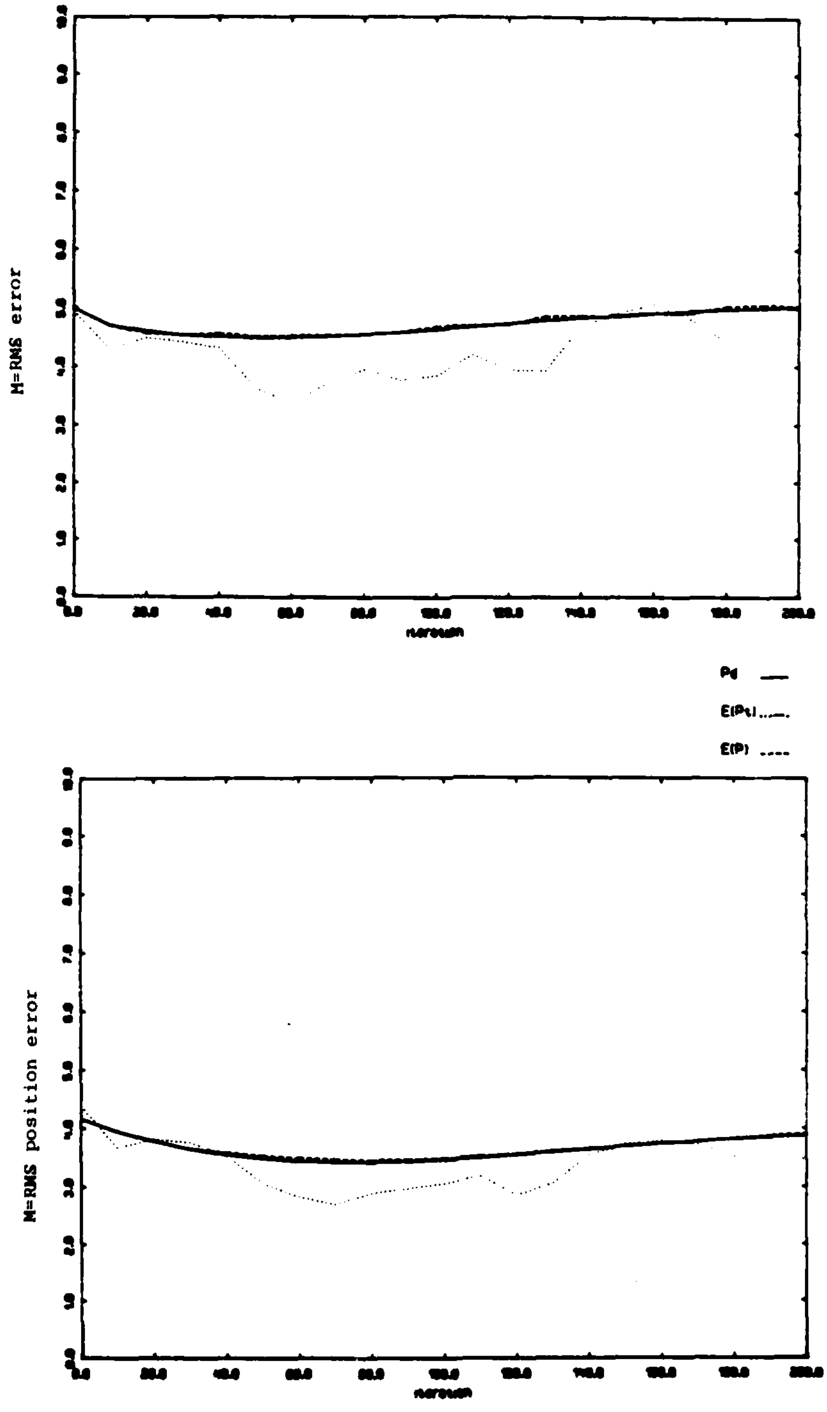


Figure 6. $M\cdot P^{d}_{k|k}$, $M\cdot E\{P^{t}_{k|k}\}$, and $M\cdot E\{P_{k|k}\}$ for $(P_D, P_F) = (.9, .01)$

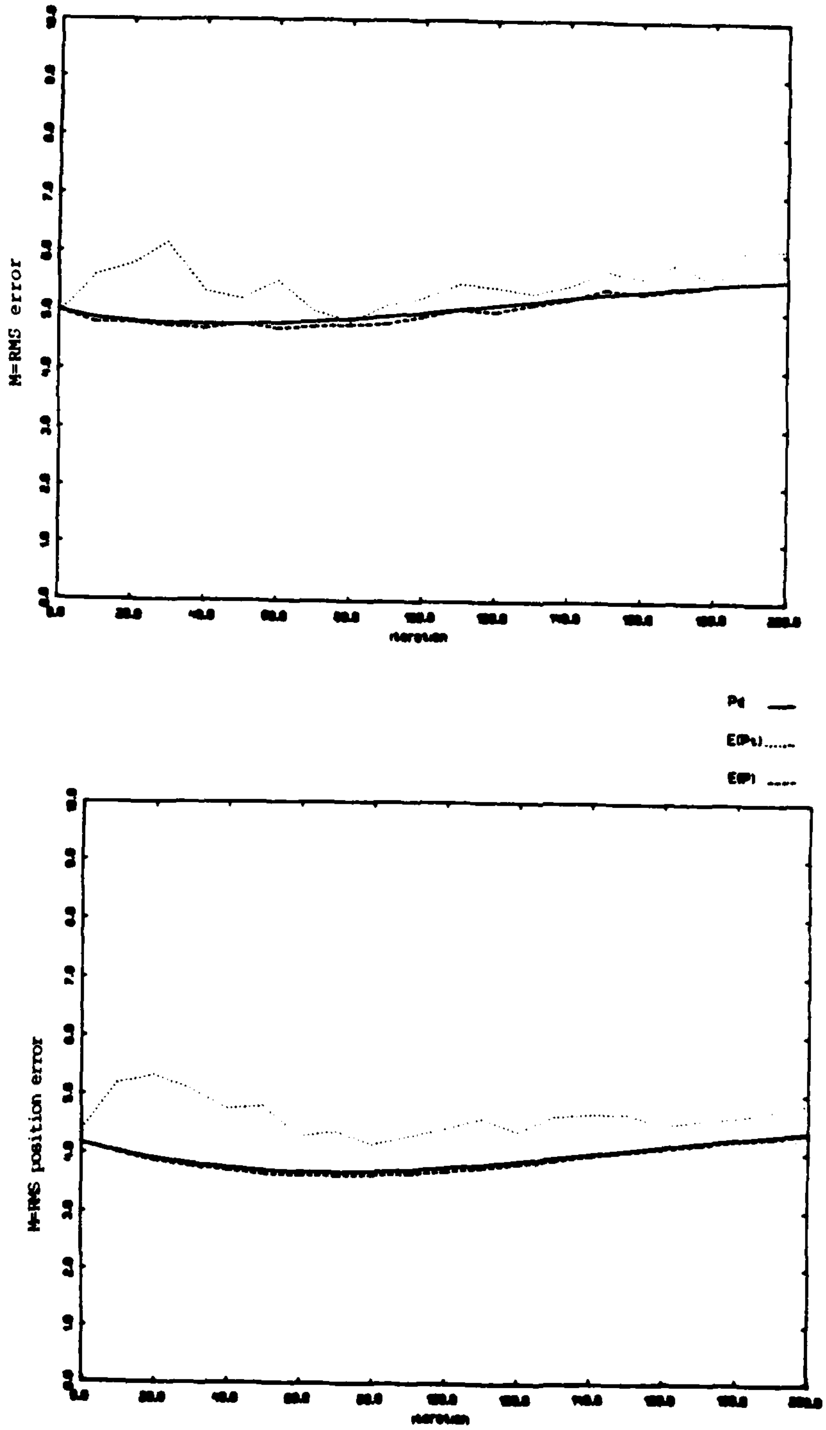


Figure 7. $M \cdot P^d_{k|k}$, $M \cdot E\{P^t_{k|k}\}$, and $M \cdot E\{P_{k|k}\}$ for $(P_D, P_F) = (.8, .02)$

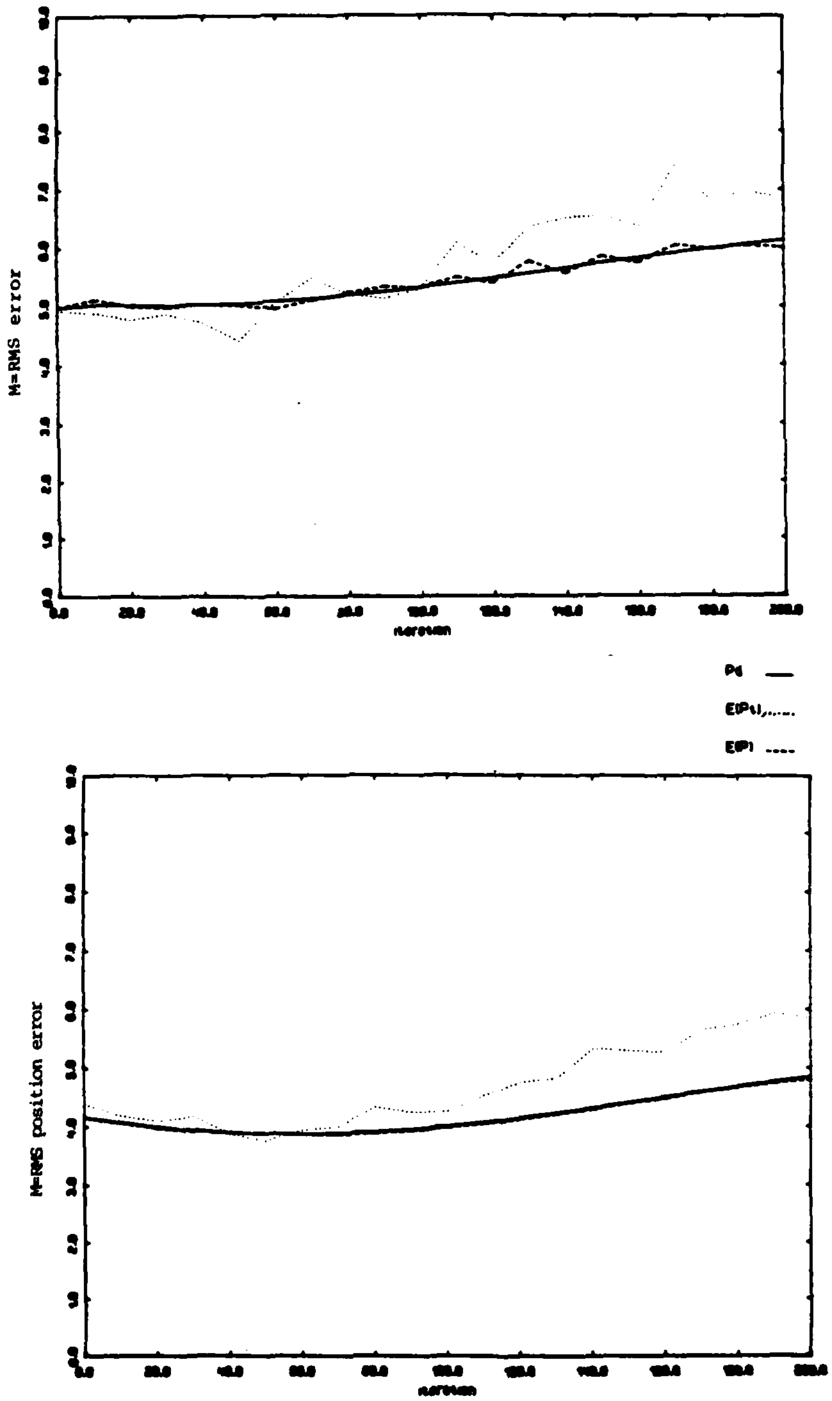


Figure 8. $M \cdot P^{d}_{k|k}$, $M \cdot E\{P^{t}_{k|k}\}$, and $M \cdot E\{P_{k|k}\}$ for $(P_D, P_F) = (.7,.03)$

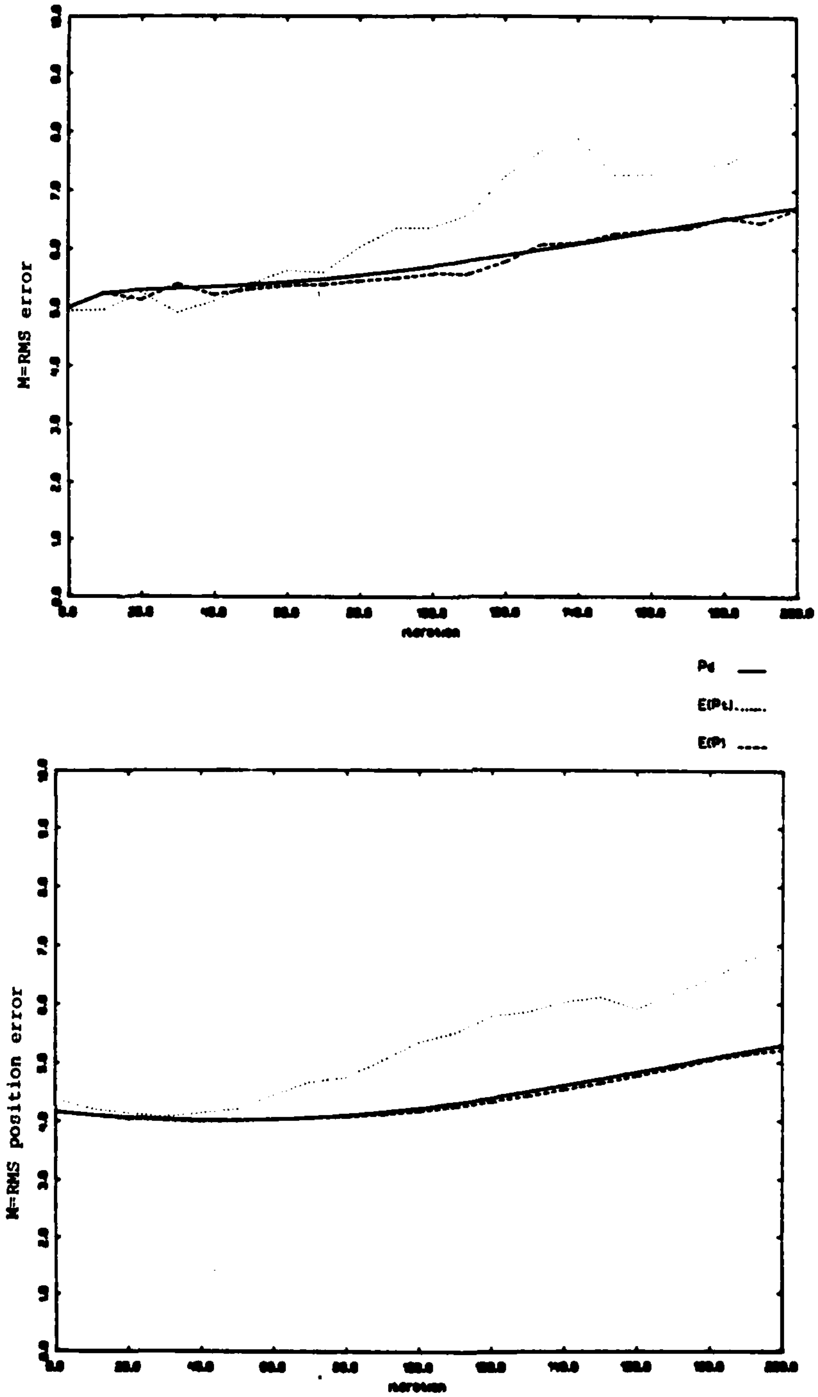


Figure 9. $M \cdot P^{d}_{k|k}$, $M \cdot E\{P^{t}_{k|k}\}$, and $M \cdot E\{P_{k|k}\}$ for (P_D, P_F) = (.6,.04)

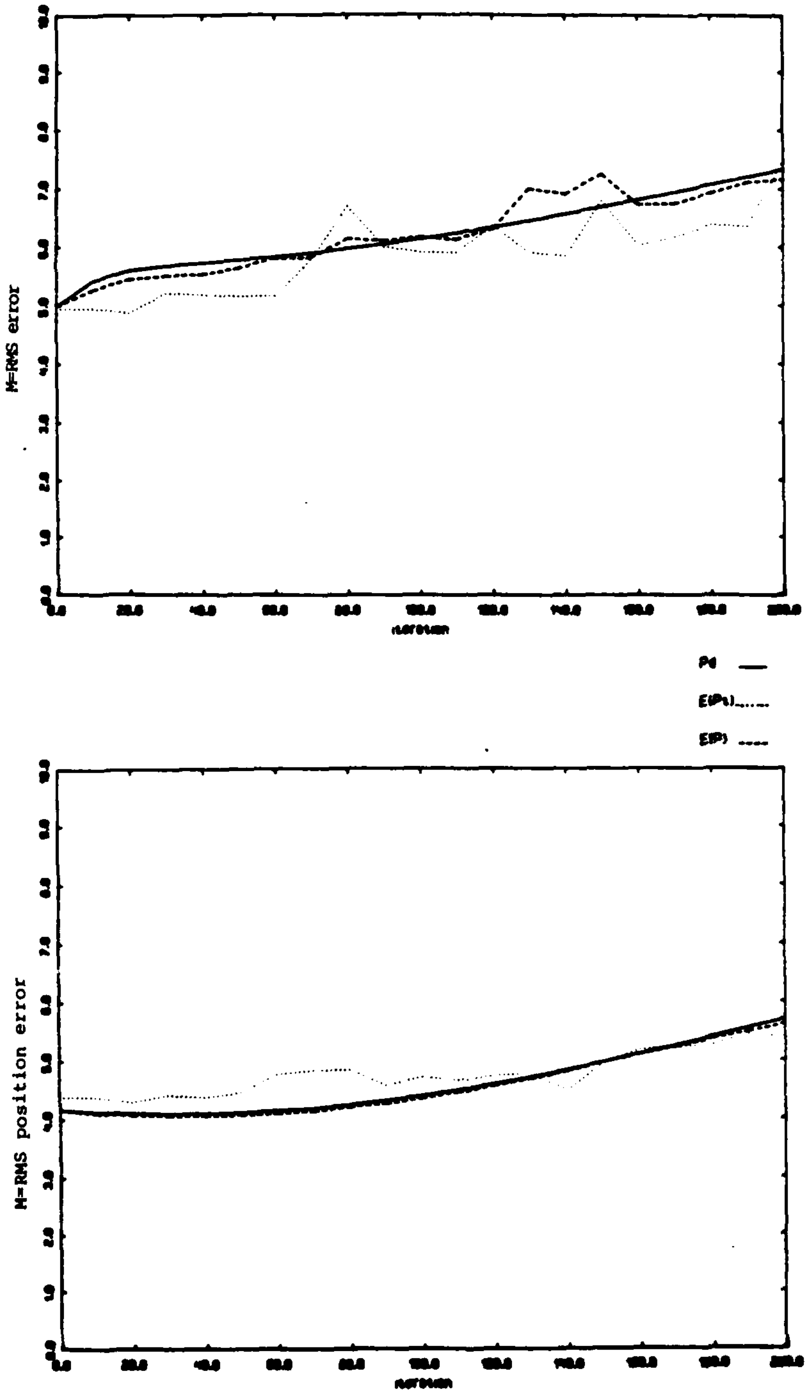


Figure 10. $M \cdot \underline{P}^{d}_{k|k}$, $M \cdot E\{\underline{P}^{t}_{k|k}\}$, and $M \cdot E\{\underline{P}_{k|k}\}$ for $(P_D, P_F) = (.5, .05)$

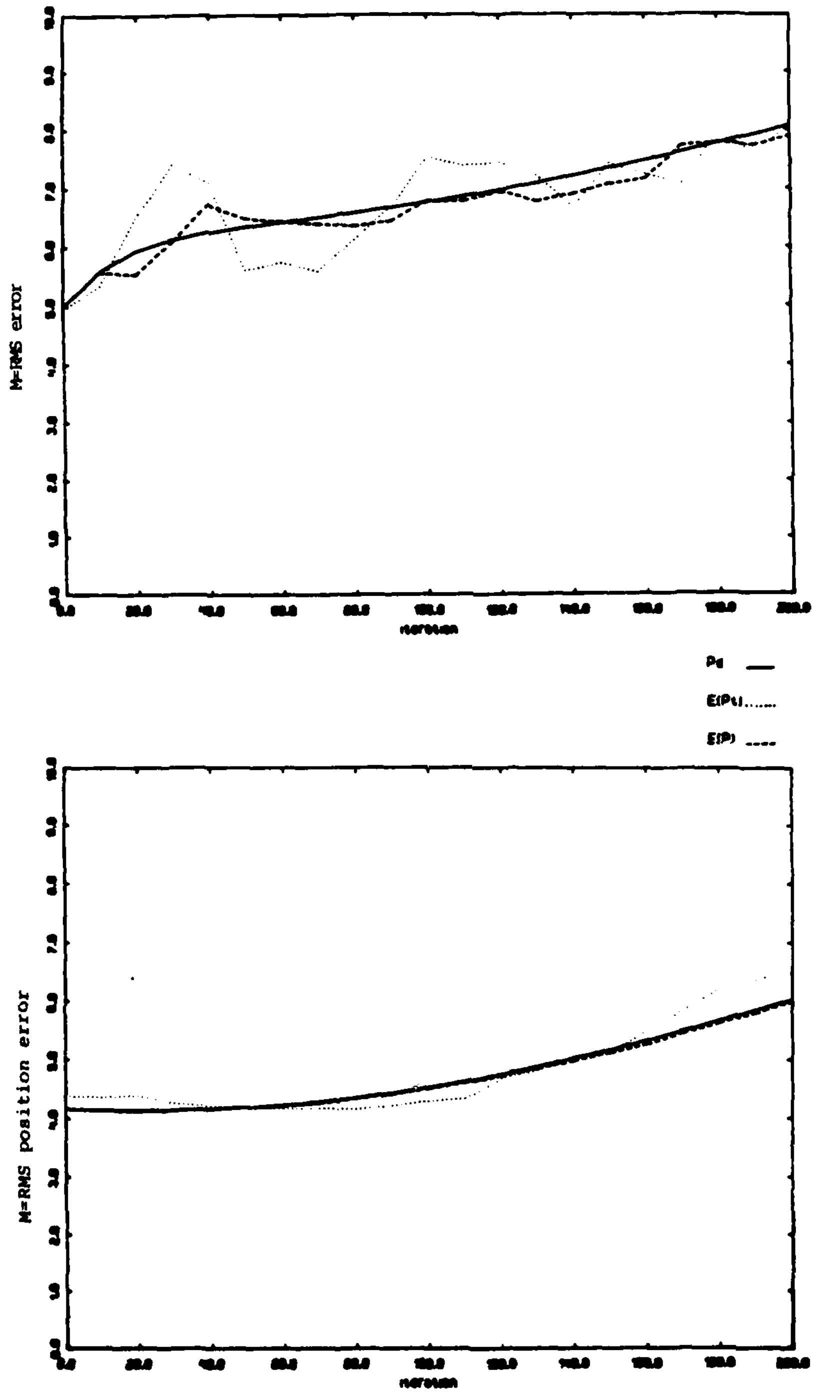


Figure 11. $M \cdot \underline{P}^{d}_{k|k}$, $M \cdot E\{\underline{P}^{t}_{k|k}\}$, and $M \cdot E\{\underline{P}_{k|k}\}$ for $(P_D, P_F) = (.4, .06)$

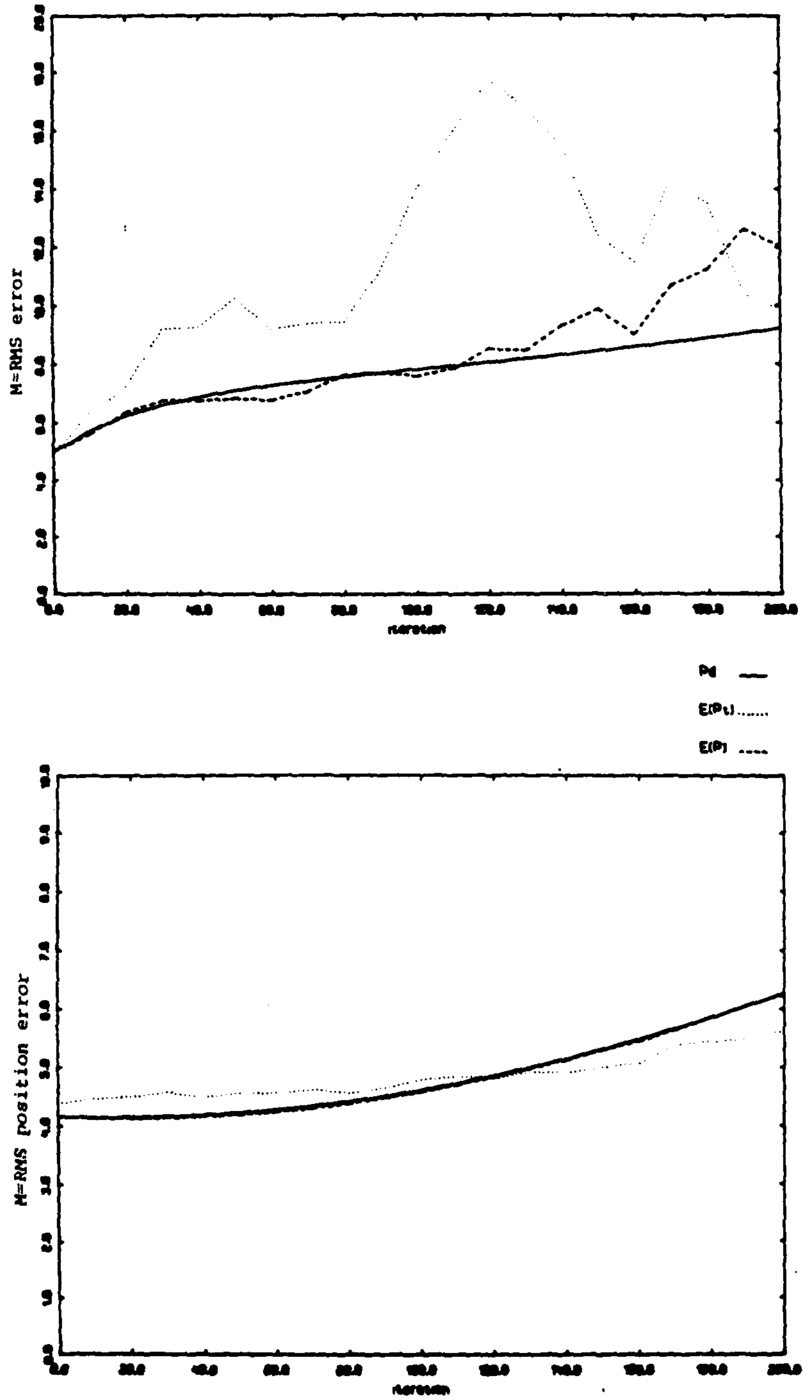


Figure 12. $M \cdot P^{d}_{k|k}$, $M \cdot E\{P^{t}_{k|k}\}$, and $M \cdot E\{P_{k|k}\}$ for $(P_D, P_F) = (.3, .07)$

2.2 Deterministic Gains

Before we present simulation results for the PDAF with gains based on $P^d_{k|k-1}$, we show that such a filter is optimal (in a Bayesian sense) under an assumption stronger than the standard PDAF assumption. Let $\hat{x}^d_{k|k}$ be the output of the PDAF with gains based on $P^d_{k|k-1}$, and $P^{dt}_{k|k}$ the true error covariance of this filter, i.e.,

$$\begin{aligned} \hat{x}^d_{k|k} &= \hat{x}^d_{k|k-1} + W^d_k \nu^d \\ \hat{x}^d_{k|k-1} &= F\hat{x}^d_{k-1|k-1} ; \qquad \hat{x}^d_{0|0} = E\{x_0\} = 0 \end{aligned} \tag{9}$$

$$\begin{aligned} P^{dt}_{k|k} &\triangleq E\{(x_k-\hat{x}^d_{k|k})(x_k-\hat{x}^d_{k|k})'\} \\ P^{dt}_{k|k-1} &\triangleq E\{(x_k-\hat{x}^d_{k|k-1})(x_k-\hat{x}^d_{k|k-1})'\} \end{aligned} \tag{10}$$

Consider the following assumption:

x_j and Y^{j-1} are jointly Gaussian for $j=1,\ldots,k$.

If this assumption is satisfied, we will show that

$$p(x_k \mid Y^{k-1}) \sim N(\hat{x}^d_{k|k-1}, P^d_{k|k-1}) \tag{11}$$

and consequently

$$\hat{x}^d_{k|k} = E\{x_k \mid Y^k\} \quad \text{and} \quad P^d_{k|k} = P^{dt}_{k|k} \tag{12}$$

Since x_j and Y^{j-1} jointly Gaussian for j=1,...,k implies that $p(x_j|Y^{j-1})$ is Gaussian for j=1,...,k, and consequently

$$p(x_j \mid Y^{j-1}) \sim N\{\hat{x}_{j|j-1}, P_{j|j-1}\}, \; j=1,\ldots,k, \tag{13}$$

it is sufficient to show that

$$\hat{x}^d_{k|k-1} = \hat{x}_{k|k-1} \quad \text{and} \quad P^d_{k|k-1} = P_{k|k-1} \tag{14}$$

We proceed by induction. First note that $P^d_{1|\emptyset} = P_{1|\emptyset}$. Next, for any j=1,...,k-1 assume that $P^d_{j|j-1} = P_{j|j-1}$. Thus

$$\begin{aligned} P^d_{j|j} &= P^d_{j|j-1} - q_2(S^d_j; P_D, P_F) W^d_j S^d_j W^{d'}_j \\ &= P_{j|j-1} - q_2(S_j; P_D, P_F) W_j S_j W'_j = E\{P_{j|j} \mid Y^{j-1}\} = E\{P_{j|j}\} \end{aligned} \tag{15}$$

and so

$$P^d_{j+1|j} = FP^d_{j|j}F' + GQG' = FE\{P_{j|j}\}F' + GQG' = E\{P_{j+1|j}\} = P_{j+1|j} \tag{16}$$

where the last equality follows from x_{j+1}, Y^j jointly Gaussian and a standard Kalman filtering argument. Thus, by induction, we have $P^d_{j|j-1} = P_{j|j-1}$ for j=1,...,k, and consequently $\hat{x}^d_{k|k-1} = \hat{x}_{k|k-1}$ and $P^d_{k|k-1} = P_{k|k-1}$ as required.

2.3 Results

In Figures 13-20, we show plots of $M \cdot P^{dt}_{k|k}$ and $M \cdot E\{P^{t}_{k|k}\}$ vs. time for M = RMS error and position error. From these and the rest of the data, we observed the following behavior:

1. For $P_D \geq .5$ and $P_F \leq .05$, $M \cdot P^{dt}_{k|k} \simeq ME\{P^{t}_{k|k}\}$.

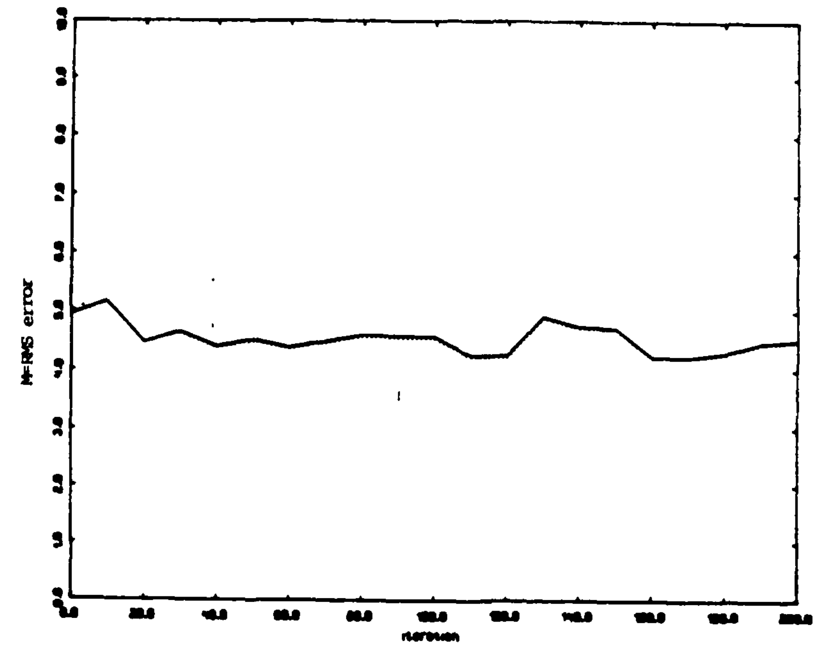
2. For $(P_D,P_F) = (.4,.06)$ and those metrics M which reflect velocity errors (the first and third), the performance of the PDAF with gains based on $P^{d}_{k|k-1}$ degrades substantially, but normal PDAF performance does not; for $(P_D,P_F) = (.3,.07)$ and the same metrics, the performance of both filters degrades substantially.

2.4 Steady State

In the ROC-TOC approach developed in [20], the deterministic approximation to the stochastic Riccati equation is iterated to convergence. We denote this steady-state value as $\bar{P}(P_D,P_F)$. In previous simulations [20], we have shown that $\bar{P}$ exists except for a region in the P_D,P_F plane where the deterministic approximation to the stochastic Riccati equation is unstable and $P^{d}_{k|k}$ diverges. However, even when $\bar{P}$ exists, convergence can be slow. For example, when $(P_D,P_F) = (.6,.04)$, approximately 500 iterations are needed for convergence (see Figure 21). Also, both existence and convergence rate are numerically sensitive to the initial covariance $P^{d}_{0|0}$. For these reasons, it is questionable whether $\bar{P}$ should be used to optimize such parameters as P_D,P_F, and alternative adaptive approaches are developed in Section 3.

2.5 Conclusions

To validate the deterministic approximation to the stochastic Riccati equation, we compared the deterministic approximation with both the true PDAF error covariance and the PDAF-calculated error covariance. Simulation results show that all quantities are comparable for a reasonable range of P_D, P_F. Outside this range, PDAF performance degrades significantly and the computed quantities are inconsistent, especially when velocity errors are considered. Since the deterministic approximation is typically comparable to the PDAF-calculated error covariance, the question arises as to the performance of a PDAF with gains based on the deterministic approximation. We first showed that such a filter is optimal (in a Bayesian sense) under an assumption stronger than that under which the PDAF is derived. To evaluate the filter, we compared the true error covariance of the PDAF with gains based on the deterministic approximation with the true PDAF error covariance. Simulation results show that these quantities are comparable for a reasonable range of P_D, P_F. Finally, we considered steady-state issues. In view of the slow convergence of $P^d_{k|k}$, it appears that steady-state values of the deterministic approximation should not be used to optimize such parameters as P_D, P_F.



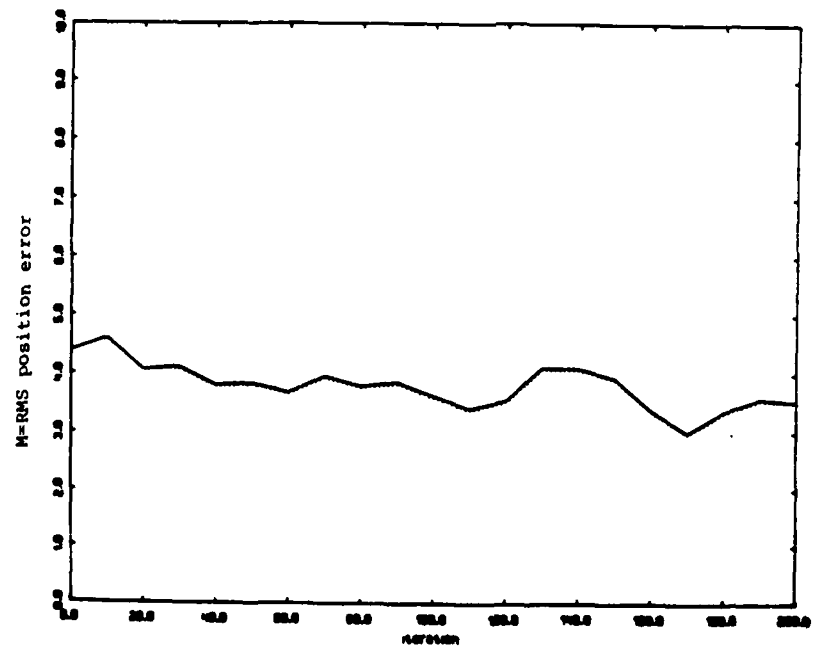


Figure 13. $M \cdot E^{dt}_{k|k}$ and $M \cdot E\{E^{t}_{k|k}\}$ for $(P_D, P_F) = (1,0)$

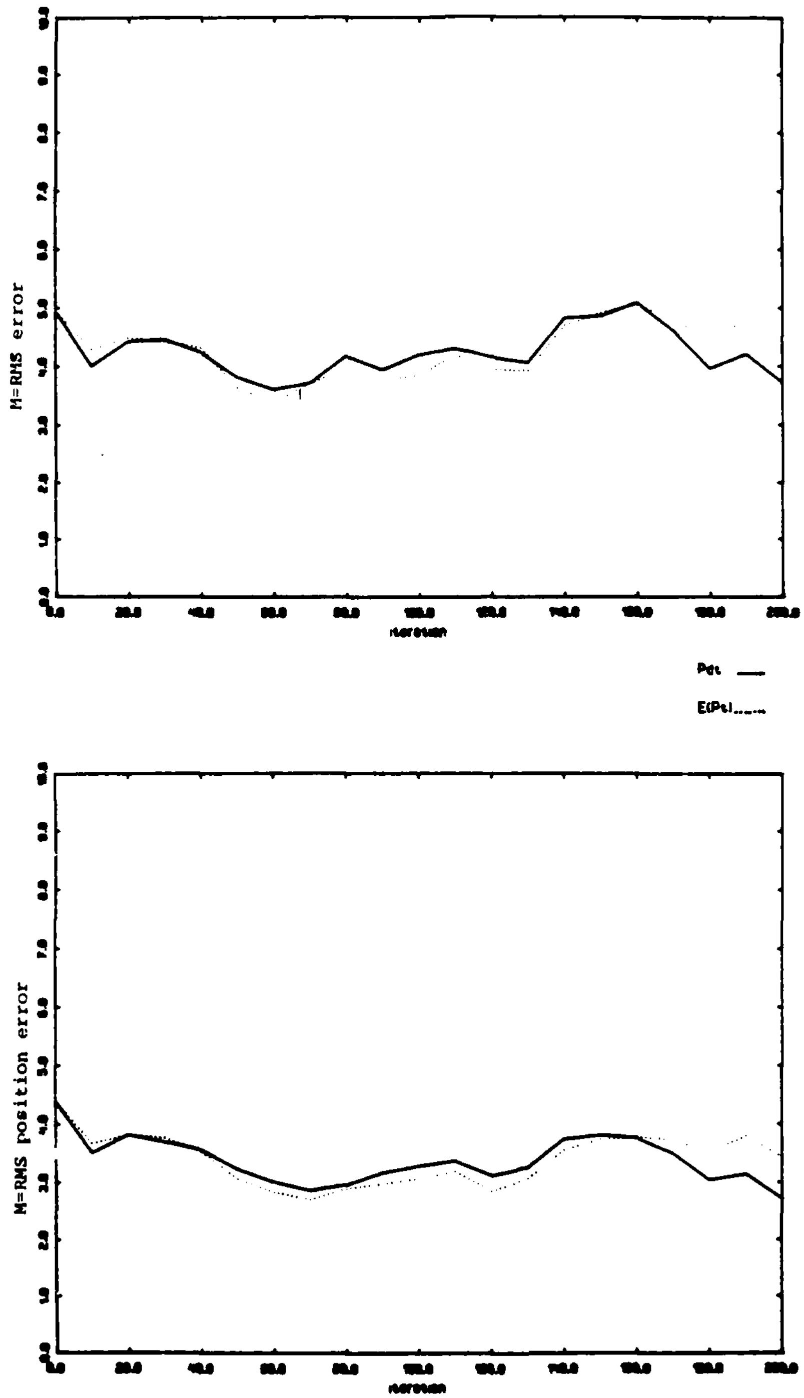


Figure 14. $M \cdot P^{dt}_{k|k}$ and $M \cdot E\{P^{t}_{k|k}\}$ for $(P_D, P_F) = (.9, .01)$

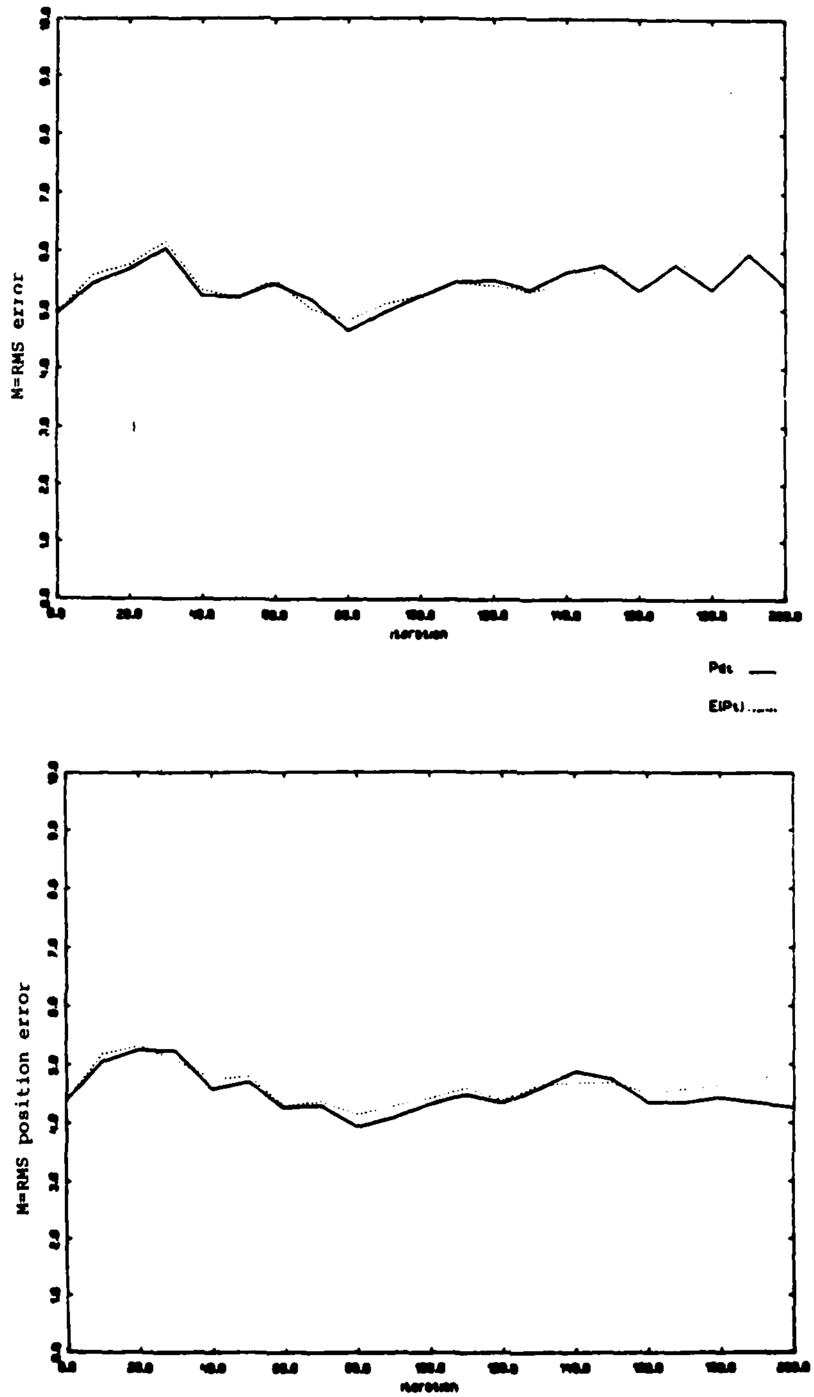


Figure 15. $M \cdot E^{dt}_{k|k}$ and $M \cdot E\{E^{t}_{k|k}\}$ for (P_D, P_F) = (.8,.02)

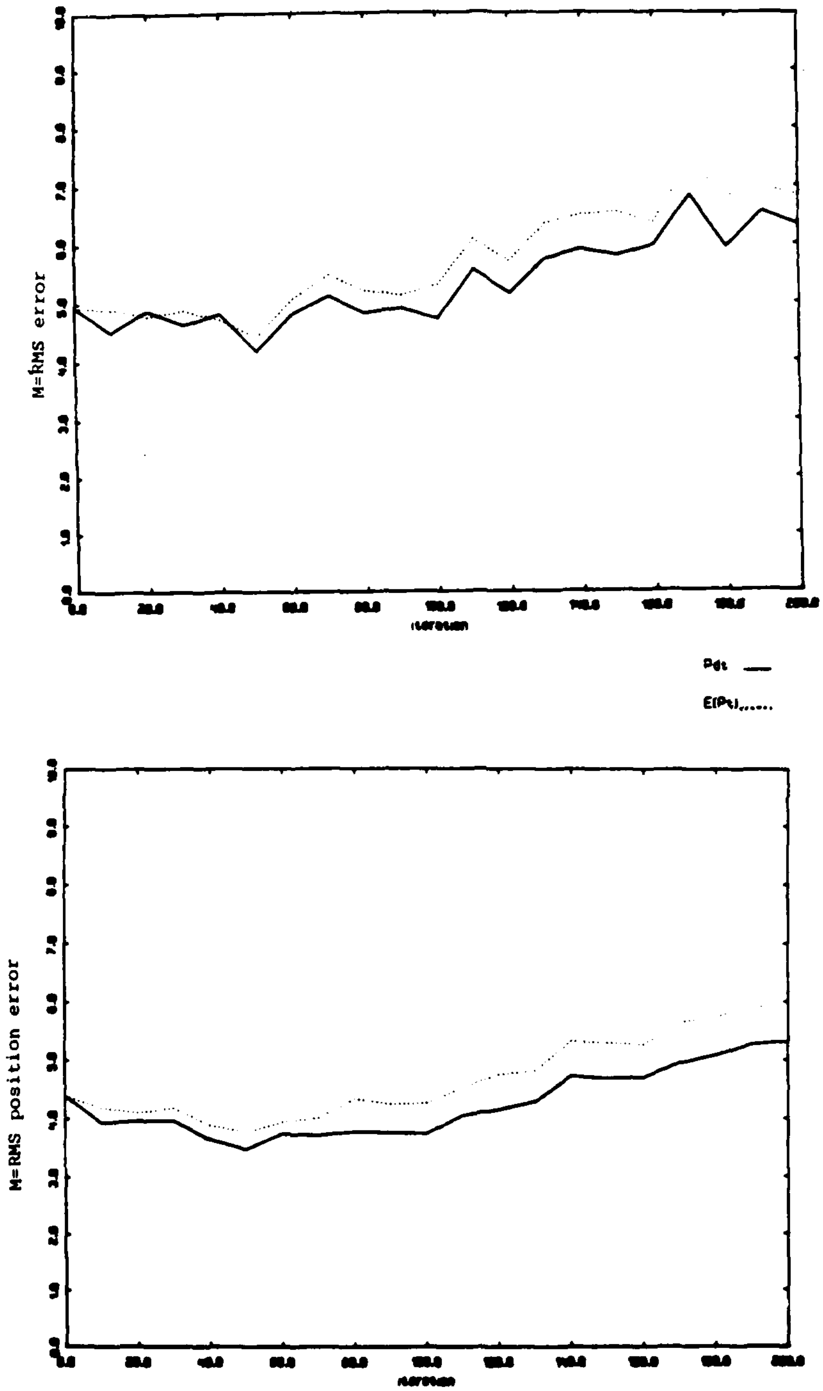


Figure 16. $M \cdot P^{dt}_{k|k}$ and $M \cdot E\{P^{t}_{k|k}\}$ for $(P_D, P_F) = (.7, .03)$

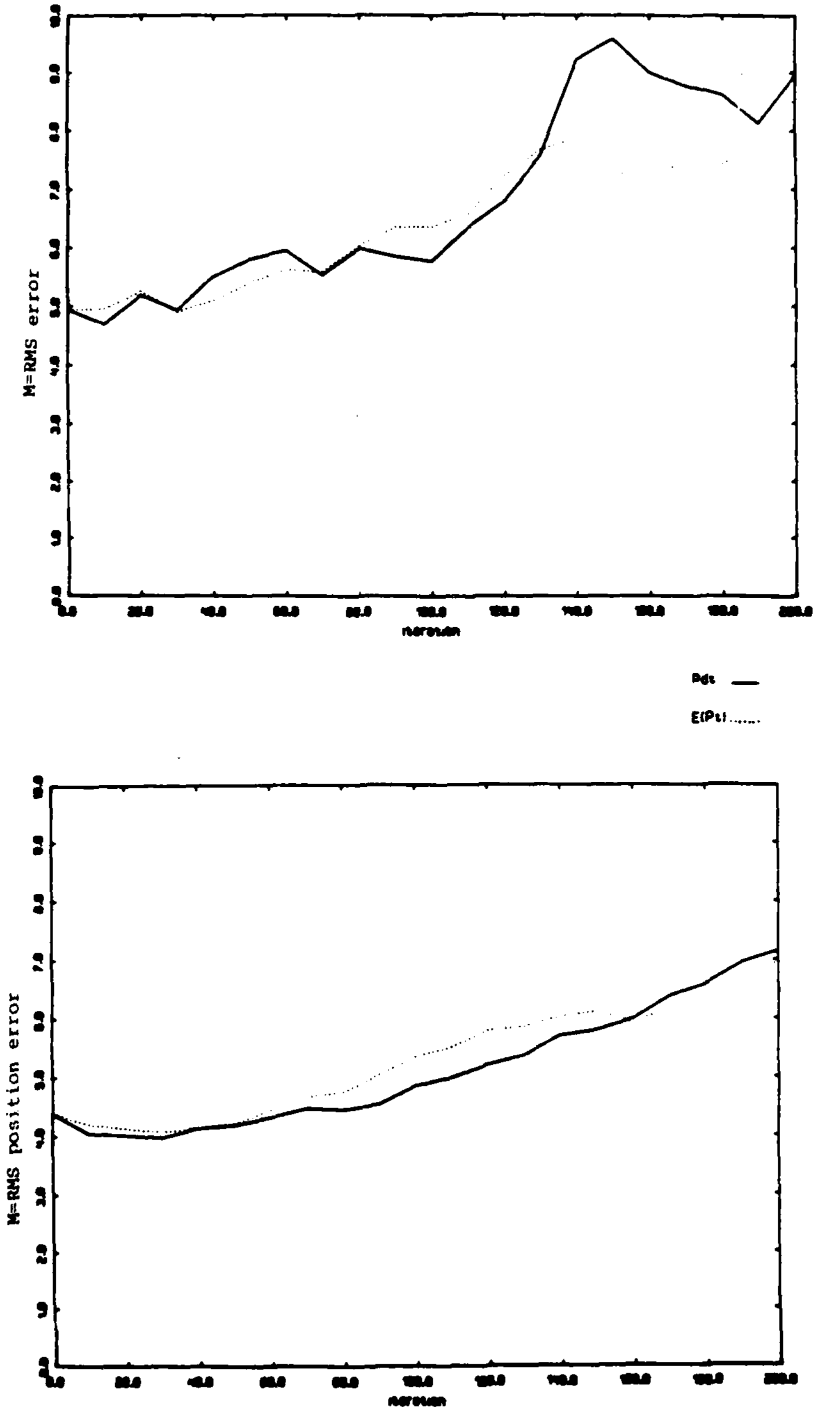


Figure 17. $M \cdot P^{dt}_{k|k}$ and $M \cdot E\{P^{t}_{k|k}\}$ for $(P_D, P_F) = (.6, .04)$

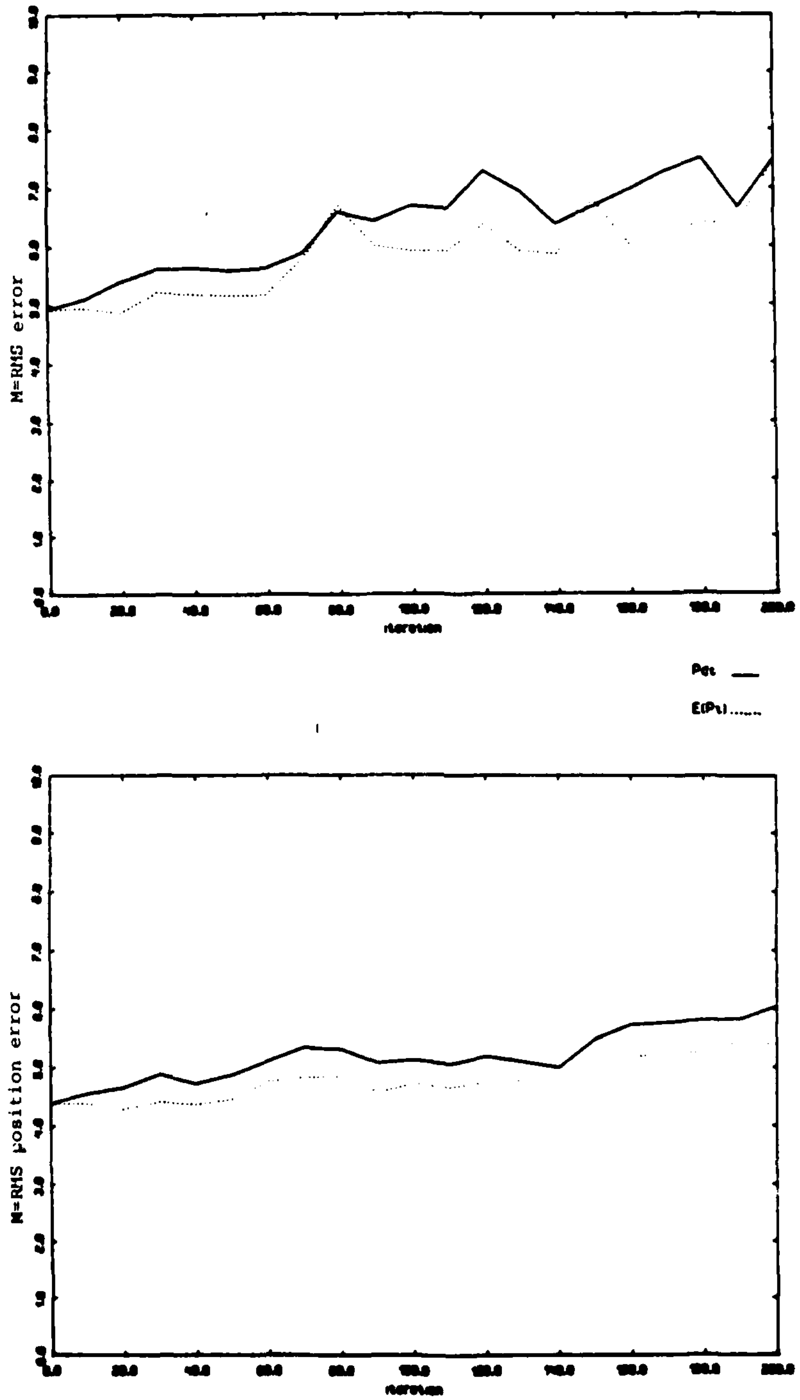


Figure 18. $M \cdot P^{dt}_{k|k}$ and $M \cdot E\{P^{t}_{k|k}\}$ for $(P_D, P_F) = (.5, .05)$

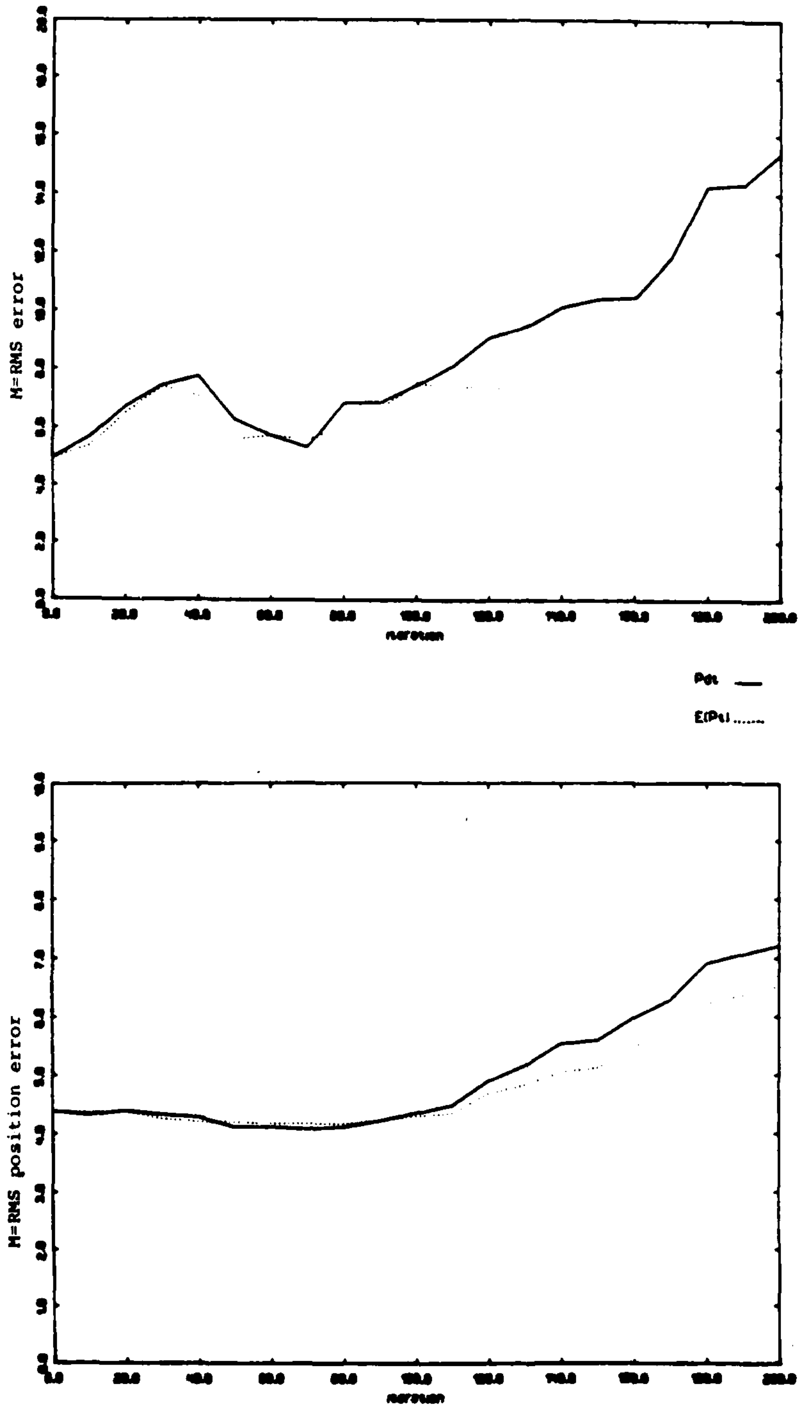


Figure 19. $M \cdot E^{dt}_{k|k}$ and $M \cdot E\{E^{t}_{k|k}\}$ for $(P_D, P_F) = (.4, .06)$

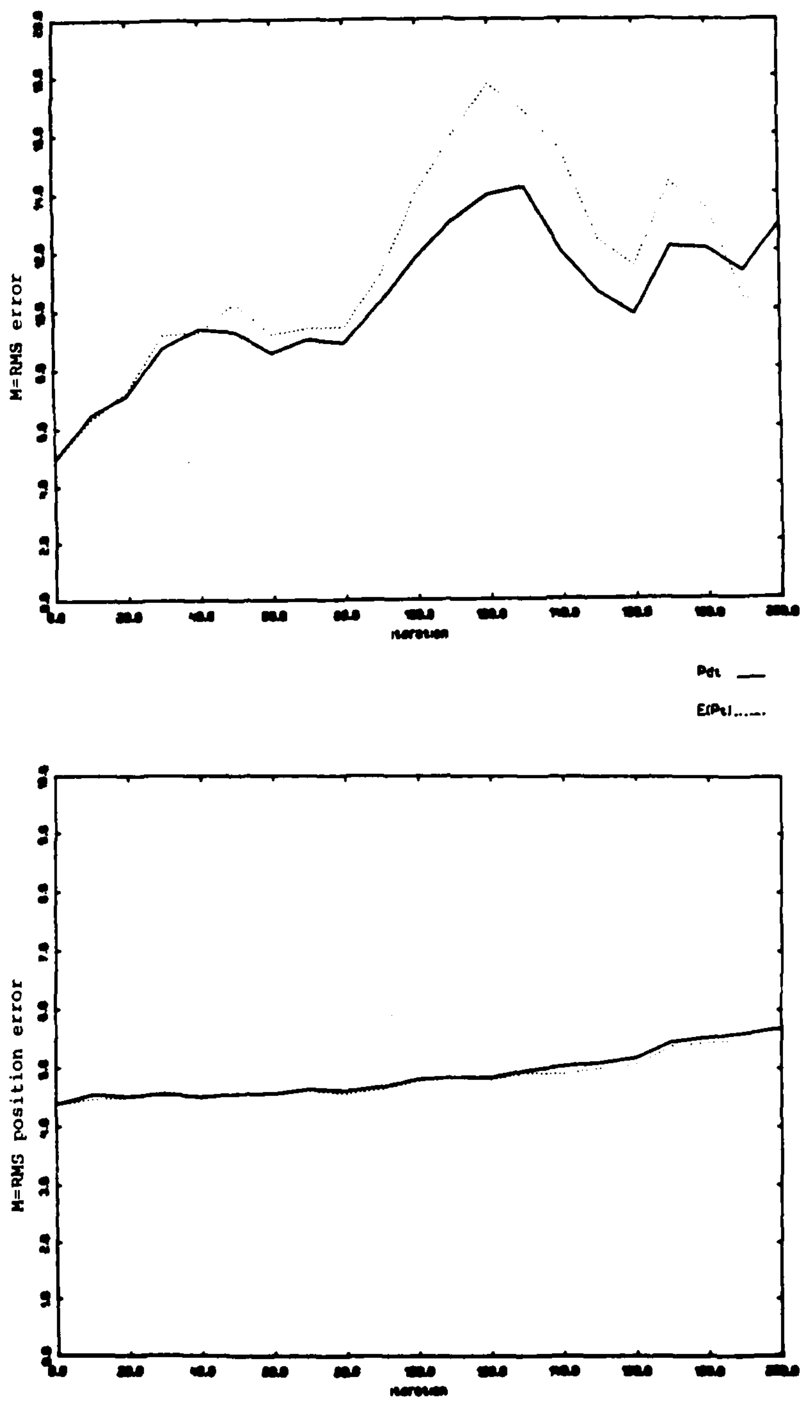


Figure 20. $M \cdot P^{dt}_{k|k}$ and $M \cdot E\{P^{t}_{k|k}\}$ for $(P_D, P_F) = (.3, .07)$

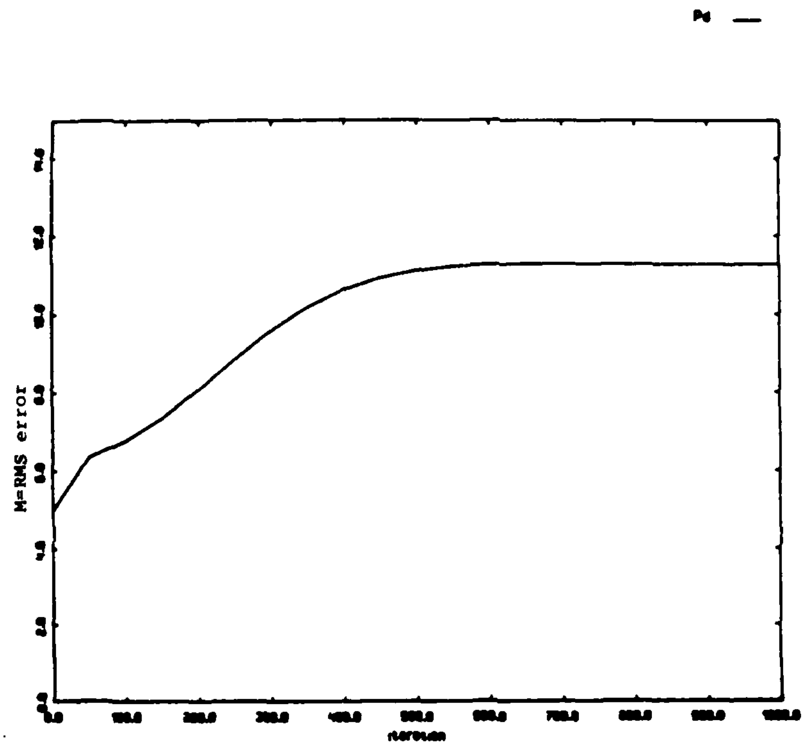


Figure 21. Convergence of $M \cdot P^{d}_{k|k}$

3. DETECTION THRESHOLD OPTIMIZATION

In the ROC-TOC approach developed in [20], the deterministic approximation to the stochastic Riccati equation is iterated to convergence. From Section 2, we denote this steady-state value as $\bar{P}(P_D,P_F)$. Then some metric on $\bar{P}$ is optimized over the values of P_D,P_F which satisfy a ROC constraint. There are two significant problems with this procedure (see Section 2):

1. $\bar{P}(P_D,P_F)$ does not exist for certain values of P_D,P_F
2. Even when $\bar{P}(P_D,P_F)$ exists, convergence can be slow

Thus we are led to examine adaptive approaches to the problem of detection threshold optimization. The adaptive approaches we shall consider are all time-dependent, but vary in data dependence.

Let $\hat{x}_{k|k-1}$ be an estimate of x_k given Y^{k-1}, and $P_{k|k-1}$ the corresponding conditional error covariance. We make the standard PDAF assumption:

$$p(x_k \mid Y^{k-1}) \sim N(\hat{x}_{k|k-1}, P_{k|k-1}) \tag{17}$$

where N denotes a normal (Gaussian) density. Thus, if

$$\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + W_k \nu \\ P_{k|k} &= P_{k|k-1} - (1-\beta_0) W_k S_k W_k' + W_k \Big(\sum_{j=1}^{m} \beta_j \nu_j \nu_j' - \nu \nu' \Big) W_k' \end{aligned} \tag{18}$$

then

$$\hat{x}_{k|k} = E\{x_k \mid Y^k\} \quad \text{and} \quad P_{k|k} = E\{\tilde{x}_{k|k}\tilde{x}'_{k|k} \mid Y^k\} \tag{19}$$

Note that $P_{k|k}$ depends on P_D, P_F and $P_{k|k-1}$. Since $P_{k|k}$ is a function of Y^k, P_D, P_F can (in general) be taken as functions of Y^j for $j \leq k$ but not for $j > k$. In the sequel, we give optimal schemes for choosing P_D, P_F as functions of Y^{k-1} and Y^k, which we refer to as prior and posterior threshold optimization, respectively.

In the following derivations, we ignore certain technical questions and assume that all extrema exist. Let $\underset{\sim}{Y}^j$ be the space of realizations of Y^j, $j=1,2,\ldots$, $\underset{\sim}{R}$ the set of (P_D, P_F) which satisfy the ROC constraint, and $(\underset{\sim}{P}_D, \underset{\sim}{P}_F)$ a mapping from $\underset{\sim}{Y}^j$ into $\underset{\sim}{R}$.

3.1 Prior Threshold Optimization

We want to solve the following functional minimization problem:

(P1) minimize $E\{|x_k - \hat{x}_{k|k}|^2\}$

over $(\underset{\sim}{P}_D, \underset{\sim}{P}_F)$: $\underset{\sim}{Y}^{k-1} \longrightarrow \underset{\sim}{R}$

First, we have

$$E\{|x_k - \hat{x}_{k|k}|^2\} = E\{\text{tr } E\{P_{k|k} | Y^{k-1}\}\} \tag{20}$$

Since $E\{P_{k|k}|Y^{k-1}\} \geq 0 \;\forall\; Y^{k-1}$, $\text{tr } E\{P_{k|k}|Y^{k-1}\} \geq 0 \;\forall\; Y^{k-1}$, and consequently (P1) reduces to the pointwise minimization:

(P2) for any given Y^{k-1}

minimize tr $E\{\underset{\sim}{P}_{k|k}|Y^{k-1}\}$

over $(P_D, P_F) \in \underset{\sim}{R}$

Let (P_D^*, P_F^*) be the optimal value of (P_D, P_F) at Y^{k-1} from (P2). Then

$$(\underset{\sim}{P}_D^*, \underset{\sim}{P}_F^*)(Y^{k-1}) = (P_D^*, P_F^*) \tag{21}$$

is optimal for (P1). Now

$$\text{tr } E\{\underset{\sim}{P}_{k|k}|Y^{k-1}\} \simeq \text{tr } \underset{\sim}{P}_{k|k-1} - q_2(\underset{\sim}{S}_k; P_D, P_F)\,\text{tr}(\underset{\sim}{W}_k \underset{\sim}{S}_k \underset{\sim}{W}_k') \tag{22}$$

Since $\underset{\sim}{W}_k \underset{\sim}{S}_k \underset{\sim}{W}_k' \geq \emptyset$, $\text{tr}(\underset{\sim}{W}_k \underset{\sim}{S}_k \underset{\sim}{W}_k') \geq \emptyset$, and since q_2 is the only term that depends on P_D, P_F, tr $E\{\underset{\sim}{P}_{k|k}|Y^{k-1}\}$ will be minimized when $q_2(\underset{\sim}{S}_k; P_D, P_F)$ is maximized. Thus our problem becomes

(P3) for any given Y^{k-1},

maximize $q_2(\underset{\sim}{S}_k; P_D, P_F)$

over $(P_D, P_F) \in \underset{\sim}{R}$

We now propose an algorithm for solving (P3). First, we parameterize P_D, P_F such that

$$\underset{\sim}{R} = \{(P_D(\lambda), P_F(\lambda)) : \lambda \in \Lambda\} \tag{23}$$

where Λ is a closed interval. Since $\underset{\sim}{R}$ corresponds to a ROC constraint, we can always choose λ to be P_D, P_F, or the detection threshold. Such a parameterization reduces (P3) to a line minimization:

(P4) maximize $q_2 \cdot (P_D, P_F)(\lambda)$

over $\lambda \in \Lambda$

where the composition $\cdot$ is defined by

$$q_2 \cdot (P_D, P_F)(\lambda) \triangleq q_2(\underset{\sim}{S}; P_D(\lambda), P_F(\lambda)) \tag{24}$$

We propose to use the golden-section search [21] to solve (P4). This requires that $q_2 \cdot (P_D, P_F)$ be unimodal. The following conditions are sufficient:

1. q_2 is strictly convex $\cap$
2. $P_D \leq \bar{P}_D$ and $P_F \geq \bar{P}_F \implies q_2(\underset{\sim}{S}; P_D, P_F) \leq q_2(\underset{\sim}{S}; \bar{P}_D, \bar{P}_F)$
3. $P_D(\lambda)$ is convex $\cap$ and $P_F(\lambda)$ is convex $\cup$

This follows since for every $\lambda_1, \lambda_2 \in \Lambda$, $\lambda_1 \neq \lambda_2$,

$$\begin{aligned} &\alpha\, q_2 \cdot (P_D, P_F)(\lambda_1) + (1-\alpha) q_2 \cdot (P_D, P_F)(\lambda_2) \\ &\quad < q_2[\underset{\sim}{S};\ \alpha P_D(\lambda_1) + (1-\alpha) P_D(\lambda_2),\ \alpha P_F(\lambda_1) + (1-\alpha) P_F(\lambda_2)] \\ &\quad \leq q_2[\underset{\sim}{S};\ P_D(\alpha\lambda_1 + (1-\alpha)\lambda_2),\ P_F(\alpha\lambda_1 + (1-\alpha)\lambda_2)] \\ &\quad = q_2 \cdot (P_D, P_F)(\alpha\lambda_1 + (1-\alpha)\lambda_2) \qquad \forall\, \alpha \in (0,1) \end{aligned} \tag{25}$$

and so $q_2^{\cdot}(P_D,P_F)$ is strictly convex $\cap$ and thus unimodal. If $\lambda=P_F$, then since $P_D(\lambda)$ is strictly convex $\cap$, the conditions simplify to

1. q_2 is convex $\cap$
2. For every P_F, $q_2(\underline{S};\cdot,P_F)$ is monotone strictly increasing

and similarly for $\lambda=P_D$. These conditions seem reasonable, although we do not give proofs. We can also verify that $q_2^{\cdot}(P_D,P_F)$ is unimodal numerically.

3.2 Posterior Threshold Optimization

We want to solve the following functional minimization problem:

(P5) minimize $E\{|\underline{x}_k-\hat{\underline{x}}_{k|k}|^2\}$

over $(\underline{P}_D,\underline{P}_F): \underline{Y}^k \longrightarrow \underline{R}$

Proceeding as in the prior case this problem is reduced to the pointwise minimization:[2]

(P6) for any given Y^k,

minimize tr $\underline{P}_{k|k}$

over $(P_D,P_F) \in \underline{R}$

---

[2]Note that the conditioning here is actually on a function of (P_D,P_F), i.e., $\{<(P_D,P_F),Y^k(P_D,P_F)>: (P_D,P_F) \in \underline{R}\}$, as opposed to a value, say $Y^k(P_D,P_F)$.

Let (P_D^*, P_F^*) be the optimal value of (P_D, P_F) at Y^k from (P6). Then

$$(\underline{P}_D^*, \underline{P}_F^*)(Y^k) = (P_D^*, P_F^*) \tag{26}$$

is optimal for (P5). Now

$$\mathrm{tr}\ \underline{P}_{k|k} = \mathrm{tr}\ \underline{P}_{k|k-1} - \{ \sum_{j=1}^{m} \beta_j [\mathrm{tr}(\underline{W}_k \underline{S}_k \underline{W}_k') - |\underline{W}_k \underline{\nu}_j|^2] + \sum_{i,j=1}^{m} \beta_i \beta_j [(\underline{W}_k \underline{\nu}_i)' (\underline{W}_k \underline{\nu}_j)] \} \tag{27}$$

Thus tr $\underline{P}_{k|k}$ will be minimized when the quantity in $\{\cdot\}$ is maximized. Hence our problem becomes:

(P7) for any given Y^k,

maximize $$\sum_{j=1}^{m} \beta_j [\mathrm{tr}(\underline{W}_k \underline{S}_k \underline{W}_k') - |\underline{W}_k \underline{\nu}_j|^2] + \sum_{i,j=1}^{m} \beta_i \beta_j [(\underline{W}_k \underline{\nu}_i)' (\underline{W}_k \underline{\nu}_j)]$$

over $(P_D, P_F) \in \underline{R}$

We now propose an algorithm for solving (P7). First, we parameterize P_D, P_F as in the prior approach. However, instead of convexity constraints, we require that $P_D(\lambda)$ (or equivalently, $P_F(\lambda)$) be monotone on Λ (for convenience, assume $P_D(\lambda)$ monotone increasing on $\Lambda = [\lambda_{min}, \lambda_{max}]$). Then there exist $\lambda_1, \ldots, \lambda_{m^*}$ such that

$$\lambda_{min} = \lambda_\emptyset < \lambda_1 < \cdots < \lambda_{m^*} < \lambda_{m^*+1} = \lambda_{max} \tag{28}$$

and

$$m = i \quad \text{for} \quad \lambda \in [\lambda_i, \lambda_{i+1}), \quad i = \emptyset, 1, \ldots, m^* \tag{29}$$

(see Figure 22). This divides (P7) into m*+1 local maximizations of the form

$$\text{maximize} \quad \sum_{j=1}^{m} \rho_j [\mathrm{tr}(\underline{W}_k \underline{S}_k \underline{W}_k') - |\underline{W}_k \underline{\nu}_j|^2]$$
$$+ \sum_{i,j=1}^{m} \rho_i \rho_j [(\underline{W}_k \underline{\nu}_i)'(\underline{W}_k \underline{\nu}_j)]$$

over $\lambda \in [\lambda_i, \lambda_{i+1})$

which we may write as

$$\text{(P8)} \quad \text{maximize} \left[\frac{d_1}{b(\lambda)+a} + \frac{d_2}{(b(\lambda)+a)^2} \right]$$

over $\lambda \in [\lambda_i, \lambda_{i+1})$

where

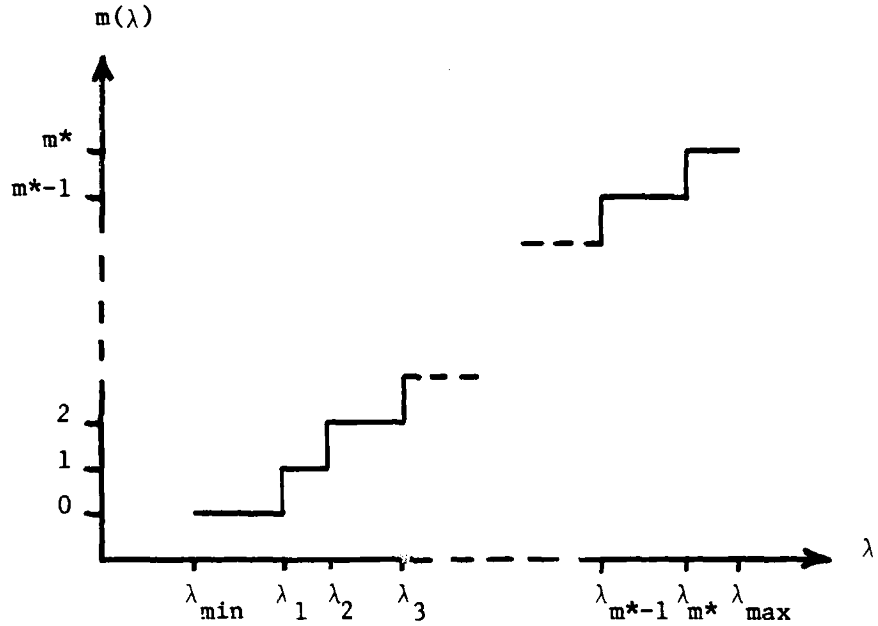


Figure 22. Dependence of m on parameter λ

$$b(\lambda) \triangleq (2\pi)^{M/2}(V_G/c_M g^M V_C)\, P_F(\lambda)\,[(1-P_D(\lambda)P_G)/P_D(\lambda)]$$

$$a \triangleq \sum_{j=1}^{m} \exp\{-\underline{\nu}_j' S_k^{-1} \underline{\nu}_j/2\}$$

$$d_1 \triangleq \sum_{j=1}^{m} \exp\{-\underline{\nu}_j' S_k^{-1} \underline{\nu}_j/2\}\,[\mathrm{tr}(\underline{W}_k S_k \underline{W}_k') - |\underline{W}_k \underline{\nu}_j|^2]$$

$$d_2 \triangleq \sum_{i,j=1}^{m} \exp\{-\underline{\nu}_i' S_k^{-1} \underline{\nu}_i/2\} \exp\{\underline{\nu}_j' S_k^{-1} \underline{\nu}_j/2\}\,[(\underline{W}_k \underline{\nu}_i)\cdot(\underline{W}_k \underline{\nu}_j)]$$

Note that if d_1 and d_2 have the same sign (say +), then (P8) simplifies to

(P8') minimize $P_F(\lambda)[(1-P_D(\lambda)P_G)/P_D(\lambda)]$

over $\lambda \in [\lambda_i, \lambda_{i+1})$

In practice, $\lambda_1, \ldots, \lambda_{m^*}$ would not be known exactly and $[\lambda_i, \lambda_{i+1})$ would be approximated from within by a closed interval. Finally, we note that this posterior procedure is very attractive when the local optimizations [(P8) or (P8')] can be done analytically.

3.3 Simulation Notes and Results

The simulation of the prior and posterior adaptive threshold optimization algorithms involve some subtleties which we shall now discuss. First note that for purposes of comparison and computational efficiency we would like to use the same Monte Carlo data generated for the fixed (P_D, P_F) simulations in Section 2. In the prior case, for every Monte Carlo trial there will in general be an optimal (P_D, P_F) which is different from the fixed (P_D, P_F).

Thus, for every Monte Carlo trial we need to evaluate a single realization of Y^k at two different values of (P_D, P_F). This can be done in two ways:

1. Resetting the seed in the generation of Y^k
2. Using an intermediate result from a single generation of Y^k (and saving computation)

Since we are ultimately concerned with Bernoulli and Poisson random variates, approach 2 is feasible (we omit the details). In the posterior case, we require (and not just for purposes of comparison and computational efficiency) a single realization of Y^k as a function of P_D, P_F, or some other λ. It turns out that approach 2 is not only feasible, but exactly specifies the values $\lambda_1, \ldots, \lambda_{m*}$ discussed above. Unfortunately, time was not available to run these simulations.

3.4 Conclusions

We have carefully posed the prior and posterior detection threshold optimization problems, and have given algorithms for their solution. The prior algorithm performs a single line search to optimize q_2, which can be evaluated by a look-up procedure. The posterior algorithm performs multiple line searches which can sometimes be done analytically. Certain subtleties in the simulation of these algorithms are pointed out, and the simulations turn out to be neither as difficult nor as expensive as originally thought. Unfortunately, time did not permit testing on either real or simulated data. It is expected that both schemes will outperform any time-independent optimization. It is also expected that posterior will do better than prior, although it may be more expensive computationally. The posterior algorithm is most attractive when the local minimizations can be done analytically.

4. GATE SIZE OPTIMIZATION AND REVISED PDAF DERIVATION

Gate size, or equivalently P_G, is a parameter in the PDAF equations. Hence, P_G could be subjected to the same analysis given P_D, P_F above. Specifically, we could obtain a measure of PDAF performance as a function of P_G as in Section 2, and then optimize over P_G using similar schemes to those in Section 3. However, we choose not to pursue these directions because:

1. PDAF performance is approximately constant for large enough gate size ($g \geq 4$ or $P_G \geq .99$)
2. PDAF performance is approximately monotonically increasing with P_G, and in any case, joint optimization over P_D, P_F and P_G is considerably more complicated than just over P_D, P_F.

For these reasons, we fix P_G in a time-independent way ($g = 4$ or $P_G \simeq 1$ for this problem) and optimize only over P_D, P_F as in Section 3. However, while considering the gate size issue, we observed that the PDAF exhibited certain counter-intuitive behavior. We next address this problem.

4.1 Gate Size in PDAF Derivation

Consider the case where $\underline{H} = \underline{I}$, $m = 0$, $P_D = 1$, $\underline{P}_{k|k-1} >> \underline{R}$ and $P_G \simeq 1$ $(g \geq 4)$. Since $m = 0$, the PDAF sets $\hat{\underline{x}}_{k|k} = \hat{\underline{x}}_{k|k-1}$ and $\underline{P}_{k|k} = \underline{P}_{k|k-1}$. But we typically have

$$16 < (\tilde{\underline{x}}_{k|k-1}+\underline{v}_k)'\underline{S}_k^{-1}(\tilde{\underline{x}}_{k|k-1}+\underline{v}_k) \simeq \tilde{\underline{x}}_{k|k}'\underline{P}_{k|k-1}^{-1}\tilde{\underline{x}}_{k|k} \tag{30}$$

which implies that $\underline{P}_{k|k} >> \underline{P}_{k|k-1}$, a contradiction.

Before examining the PDAF derivation to reconcile this contradiction, we give a simple example which clearly shows how $P_{k|k} > P_{k|k-1}$ is possible for a Gaussian problem with finite gate. Consider the problem of estimating a scalar $x \sim N(\emptyset, p)$ from an observation y given by

$$y = x+v \tag{31}$$

where $v \sim N(\emptyset, r)$ and x,v are independent. It is well known that

$$\mathrm{var}(x|y) \leq \mathrm{var}(x) \tag{32}$$

However, for $p >> r$ and $g \geq 4$, we have

$$\mathrm{var}(x|\ |y|>g\sqrt{p+r}) \simeq \mathrm{var}(x|\ |x|> g\sqrt{p}) >> \mathrm{var}(x) \tag{33}$$

The point is that conditioning on an event as opposed to a jointly Gaussian random variable can increase the conditional variance. It should be clear that this simple problem has all the relevant features of the case described above.

Turning to the PDAF derivation, it eventually became clear that gate size was not dealt with in a consistent manner or, alternatively, that certain unstated assumptions were made. We next show how to update the conditional mean and covariance when gate size is explicitly accounted for and also give assumptions for which the PDAF yields the same results. In the sequel, we use notation as developed in Section 1, except as noted.

We start by partitioning the event $\chi_\emptyset$ into the events

$\boldsymbol{\chi}_{\emptyset d}$: correct measurement is not detected

$\boldsymbol{\chi}_{\emptyset g}$: correct measurement is detected but is not in the gate

Note that

$$P\{\boldsymbol{\chi}_{\emptyset d}|Y^k\} = P\{\boldsymbol{\chi}_{\emptyset d}|\boldsymbol{\chi}_{\emptyset},Y^k\}P\{\boldsymbol{\chi}_{\emptyset}|Y^k\}$$
$$= P\{\boldsymbol{\chi}_{\emptyset d}|\boldsymbol{\chi}_{\emptyset}\}P\{\boldsymbol{\chi}_{\emptyset}|Y^k\} = \frac{1-P_D}{1-P_DP_G}P\{\boldsymbol{\chi}_{\emptyset}|Y^k\} \tag{34}$$

and similarly

$$P\{\boldsymbol{\chi}_{\emptyset g}|Y^k\} = \frac{(1-P_G)P_D}{1-P_DP_G} P\{\boldsymbol{\chi}_{\emptyset}|Y^k\} \tag{35}$$

To compute

$$\begin{aligned} \hat{x}_{k|k} &= E\{x_k|Y^k\} \\ P_{k|k} &= E\{\tilde{x}_{k|k}\tilde{x}'_{k|k}|Y^k\} \end{aligned} \tag{36}$$

we need $p(x_k|Y^k)$. But

$$p(x_k|Y^k) = P\{\boldsymbol{\chi}_{\emptyset d}|Y^k\}p(x_k|\boldsymbol{\chi}_{\emptyset d},Y^k) + P\{\boldsymbol{\chi}_{\emptyset g}|Y^k\}p(x_k|\boldsymbol{\chi}_{\emptyset g},Y^k)$$
$$+ \sum_{i=1}^{m}P\{\boldsymbol{\chi}_i|Y^k\}p(x_k|\boldsymbol{\chi}_i,Y^k) \tag{37}$$

Previous computation of $P\{\chi_i | Y^k\}$, $i=0,1,\ldots,m$, was correct, so using standard notation we have

$$p(\underline{x}_k | Y^k) = \beta_{0d} p(\underline{x}_k | \chi_{0d}, Y^k) + \beta_{0g} p(\underline{x}_k | \chi_{0g}, Y^k) + \sum_{i=1}^{m} \beta_i p(\underline{x}_k | \chi_i, Y^k) \qquad (38)$$

where

$$\beta_{0d} = \frac{1-P_D}{1-P_D P_G}\beta_0 \; ; \quad \beta_{0g} = \frac{(1-P_G)P_D}{1-P_D P_G}\beta_0 \qquad (39)$$

Also,

$$\begin{aligned} p(\underline{x}_k \mid \chi_{0d}, Y^k) &\sim N(\hat{\underline{x}}_{k|k-1}, \underline{P}_{k|k-1}) \\ p(\underline{x}_k \mid \chi_i, Y^k) &\sim N(\hat{\underline{x}}_{k|k-1} + \underline{W}_k \underline{\nu}_i, \underline{P}_{k|k-1} - \underline{W}_k \underline{S}_k \underline{W}_k'), \quad i=1,\ldots,m \end{aligned} \qquad (40)$$

Let $\underline{G}_k$ be the gate at the k-th iteration, i.e.,

$$\underline{G}_k = \{\underline{\nu} : \underline{\nu}' \underline{S}_k^{-1} \underline{\nu} \le g^2\} \qquad (41)$$

From Bayes rule, we have

$$p(\underline{x}_k \mid \chi_{\emptyset g}, Y^k) = p(\underline{x}_k \mid \chi_{\emptyset g}, Y^{k-1}, m) = p(\underline{x}_k \mid \chi_{\emptyset g}, Y^{k-1})$$

$$= p(\underline{x}_k \mid (\underline{H}\hat{\underline{x}}_{k|k-1}+\underline{v}_k) \notin \underline{G}_k, Y^{k-1})$$

$$= P\{(\underline{H}\hat{\underline{x}}_{k|k-1}+\underline{v}_k) \notin \underline{G}_k \mid \underline{x}_k, Y^{k-1}\} p(\underline{x}_k \mid Y^{k-1})/(1-P_G)$$

$$= P_{\underline{v}_k}\{(\underline{H}\hat{\underline{x}}_{k|k-1}+\underline{v}_k)'\underline{S}_k^{-1}(\underline{H}\hat{\underline{x}}_{k|k-1}+\underline{v}_k) > g^2\} p(\underline{x}_k|Y^{k-1})/(1-P_G) \quad (42)$$

where

$$p(\underline{v}_k) \sim N(\emptyset,\underline{R})$$
$$p(\underline{x}_k|Y^{k-1}) \sim N(\hat{\underline{x}}_{k|k-1}, \underline{P}_{k|k-1})$$

Note that $P_{\underline{v}_k}\{\cdot\}$ is in general very hard to evaluate (even numerically).

At this point, we have expressed $p(\underline{x}_k|Y^k)$ in terms of known quantities. We now use $p(\underline{x}_k|Y^k)$ to generate $\hat{\underline{x}}_{k|k}$ and $\underline{P}_{k|k}$. First consider $\hat{\underline{x}}_{k|k}$. We have

$$\hat{\underline{x}}_{k|k} = \int \underline{x}_k p(\underline{x}_k|Y^k)\,d\underline{x}_k$$
$$= \beta_{\emptyset d}\hat{\underline{x}}_{k|k-1} + \beta_{\emptyset g}\int \underline{x}_k p(\underline{x}_k|\chi_{\emptyset g}, Y^k)\,d\underline{x}_k$$
$$+ \sum_{i=1}^{m} \beta_i(\hat{\underline{x}}_{k|k-1}+\underline{W}_k\underline{\tilde{y}}_i) \quad (43)$$

But

$$\int \underline{x}_k p(\underline{x}_k|\chi_{\emptyset g}, Y^k)\,d\underline{x}_k = \hat{\underline{x}}_{k|k-1} \quad (44)$$

or equivalently

$$\int \tilde{x}_{k|k-1} p_{x_k}(\tilde{x}_{k|k-1}+\hat{x}_{k|k-1}|\mathbf{x}_{\emptyset g},Y^k)\,d\tilde{x}_{k|k-1} = \emptyset \tag{45}$$

This key result can be proved as follows. It is sufficient to show that for every $\tilde{x}_{k|k-1}$,

$$p_{x_k}(\tilde{x}_{k|k-1}+\hat{x}_{k|k-1}) = p_{x_k}(-\tilde{x}_{k|k-1}+\hat{x}_{k|k-1}) \tag{46}$$

and

$$P_{v_k}\{(H\tilde{x}_{k|k-1}+v_k)'S_k^{-1}(H\tilde{x}_{k|k-1}+v_k) > g^2\}$$

$$= P_{v_k}\{(-H\tilde{x}_{k|k-1}+v_k)'S_k^{-1}(-H\tilde{x}_{k|k-1}+v_k) > g^2\} \tag{47}$$

The first equation follows immediately since

$$p(x_k) \sim N(\hat{x}_{k|k-1},P_{k|k-1}) \tag{48}$$

and the second follows after using

$$p(v_k) \sim N(\emptyset,R) \tag{49}$$

(we omit the details). Thus we have

$$\hat{x}_{k|k} = \beta_{\emptyset d}\hat{x}_{k|k-1} + \beta_{\emptyset g}\hat{x}_{k|k-1} + \sum_{i=1}^{m}\beta_i(\hat{x}_{k|k-1}+W_k\nu_i)$$

$$= \hat{x}_{k|k-1} + W_k\nu \tag{50}$$

which is the <u>same</u> conditional mean update as the PDAF. Next consider $P_{k|k}$. Omitting the details, we have

$$\begin{aligned} P_{k|k} &= \int \tilde{x}_{k|k}\tilde{x}'_{k|k}p(x_k|Y^k)\,dx_k \\ &= P_{k|k-1} + \beta_{\emptyset g}(P^g_{k|k-1}-P_{k|k-1}) - (1-\beta_\emptyset)W_kS_kW'_k \\ &\quad + W_k(\sum_{j=1}^{m}\beta_j\nu_j\nu'_j-\nu\nu')W'_k \end{aligned} \tag{51}$$

where

$$P^g_{k|k-1} = \int \tilde{x}_{k|k-1}\tilde{x}'_{k|k-1}p(x_k|\chi_{\emptyset g},Y^k)\,dx_k \tag{52}$$

Comparing this recursion with the stochastic Riccati equation, we see there is an <u>additional term</u> $\beta_{\emptyset g}(P^g_{k|k-1}-P_{k|k-1})$, which increases the conditional covariance update. In view of our earlier remarks, this is not unexpected.

In general, $P^g_{k|k-1}$ cannot be expressed analytically in terms of $P_{k|k-1}$. It follows that if no approximation is made, the corresponding filter is infinite-dimensional. If we assume that

$$p(x_k | \chi_\emptyset, Y^k) = p(x_k | Y^{k-1}) \tag{53}$$

then we get the PDAF. More importantly, the PDAF is a good approximation if and only if $\beta_{\emptyset g}\mathrm{tr}(P^g_{k|k-1} - P_{k|k-1})$ is suitably small. In general, $\beta_{\emptyset g}$ decreases with increasing gate size, while $\mathrm{tr}(P^g_{k|k-1} - P_{k|k-1})$ increases with increasing gate size. Consequently, to justify using the PDAF, gate size must be chosen large enough so that $\beta_{\emptyset g}\mathrm{tr}(P^g_{k|k-1} - P_{k|k-1})$ is suitably small, and not just $1-\beta_{\emptyset g} \simeq P_G \simeq 1$. We next develop a method to compute $P^g_{k|k-1}$ approximately.

4.2 Approximate Computation of $P^g_{k|k-1}$

We start by approximately computing $P^g_{k|k-1}$ for two extreme cases. We have

$$P^g_{k|k-1} = E\{\tilde{x}_{k|k-1}\tilde{x}'_{k|k-1} \mid \chi_{\emptyset g}, Y^k\}$$

$$= E\{\tilde{x}_{k|k-1}\tilde{x}'_{k|k-1} \mid (H\tilde{x}_{k|k-1}+v_k)'S_k^{-1}(H\tilde{x}_{k|k-1}+v_k) > g^2,\ Y^{k-1}\}$$

$$\approx \begin{cases} E\{\tilde{x}_{k|k-1}\tilde{x}'_{k|k-1} \mid v_k'R^{-1}v_k > g^2, Y^{k-1}\} = P_{k|k-1}, & HP_{k|k-1}H' \ll R \\ E\{\tilde{x}_{k|k-1}\tilde{x}'_{k|k-1} \mid \tilde{x}_{k|k-1}H'(HP_{k|k-1}H')^{-1}H\tilde{x}_{k|k-1} > g^2,\ Y^{k-1}\}, & HP_{k|k-1}H' \gg R \end{cases} \tag{54}$$

In order to evaluate the $HP_{k|k-1}H' \gg R$ case, we are faced with the problem of computing the conditional covariance of a Gaussian random variable $\tilde{x}_{k|k-1}$

given that $\underline{x}_{k|k-1}$ lies outside a hyperellipsoid centered at its conditional mean $(=\underline{\emptyset})$. Furthermore, the principal axes do not in general coincide with the eigenvectors of $\underline{P}_{k|k-1}$ (see Figure 23a). We proceed in four steps:

1. Perform a nonsingular linear transformation on $\underline{x}_{k|k-1}$ which transforms $\underline{P}_{k|k-1}$ to $\underline{I}$ (see Figure 23b).
2. Replace the hyperellipsoid in the new coordinate system with a rectangular approximation (see Figure 23c).
3. Compute the conditional covariance of $\underline{x}_{k|k-1}$ given that it lies outside the rectangle.
4. Perform the inverse transformation of step 1 to get an approximation to $\underline{P}^g_{k|k-1}$.

We omit the details and give the final result:

$$\underline{P}^g_{k|k-1} \simeq \underline{Q}'\underline{\Lambda}^{1/2}\overline{\underline{Q}}\overline{\underline{P}}^g_{k|k-1}\overline{\underline{Q}}'\underline{\Lambda}^{1/2}\underline{Q} \quad \text{when} \quad \underline{H}\underline{P}_{k|k-1}\underline{H}' >> \underline{R} \qquad (55)$$

where

$\underline{Q} = [\underline{q}_1 \ldots \underline{q}_n]$ are orthonormal eigenvectors of $\underline{P}_{k|k-1}$

$\underline{\Lambda} = \text{diag}[\lambda_1 \ldots \lambda_n]$ are the corresponding eigenvalues

$\underline{Z} = \underline{\Lambda}^{1/2}\underline{Q}'\underline{H}'(\underline{H}\underline{P}_{k|k-1}\underline{H}')^{-1}\underline{H}\underline{Q}\underline{\Lambda}^{1/2}$

$\overline{\underline{Q}} = [\overline{\underline{q}}_1 \ldots \overline{\underline{q}}_n]$ are orthonormal eigenvectors of $\underline{Z}$

$\overline{\underline{\Lambda}} = \text{diag}[\overline{\lambda}_1 \ldots \overline{\lambda}_n]$ are the corresponding eigenvalues

$$\overline{\underline{P}}^g_{k|k-1} = \text{diag}\left[\frac{1-I(g^2\overline{\lambda}_1^{-1}/2,3/2)}{2\Phi(g\overline{\lambda}_1^{-1})}, \ldots, \frac{1-I(g^2\overline{\lambda}_n^{-1}/2,3/2)}{2\Phi(g\overline{\lambda}_n^{-1})}\right]$$

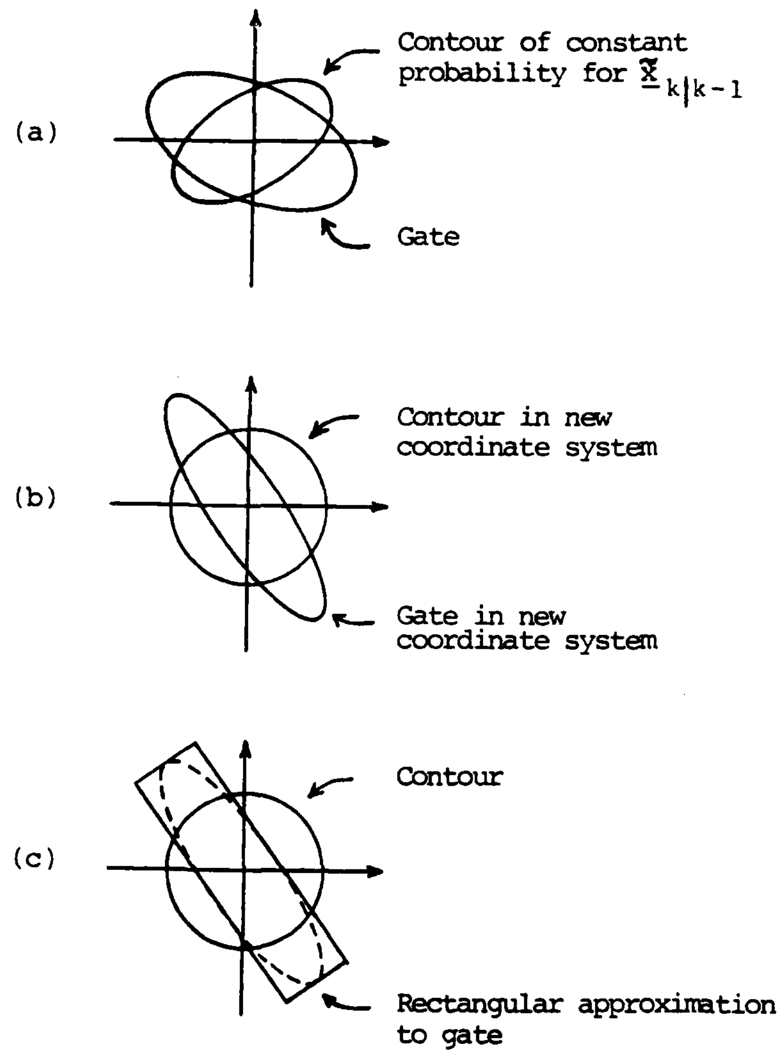


Figure 23. Approximate computation of $\underline{P}^{g}_{k|k-1}$ when $\underline{H}\underline{P}_{k|k-1}\underline{H}' >> \underline{R}$

$$I(x,p) \triangleq [1/\Gamma(p)] \int_{\emptyset}^{x} e^{-u} u^{p-1}\, du \qquad \text{(incomplete gamma function)}$$

$$\Phi(x) \triangleq [1/\sqrt{2\pi}] \int_{x}^{\infty} e^{-u^2/2}\, du \qquad \text{(error function)}$$

Since $\underline{Z}$ is in general not invertible, $\bar{\underline{Q}}$ is formed by augmenting the orthonormal eigenvectors of $\underline{Z}$ ($\underline{\Lambda}$ and $\underline{P}^{g}_{k|k-1}$ are changed appropriately). We note that IMSL routines exist for the computation of $I(x,p)$ and $\Phi(x)$.

In the general case where $\underline{HP}_{k|k-1}\underline{H}' \simeq \underline{R}$, $\underline{P}^{g}_{k|k-1}$ is approximated by interpolating between the two extreme cases in a reasonable manner.

4.3 Conclusions

We started by pointing out that gate size or, equivalently P_G, could be subjected to the same analysis given P_D, P_F in Sections 2 and 3. Specifically, we could obtain a measure of performance as a function of P_G and then optimize over P_G. For various reasons it appears uninteresting and/or unproductive to continue in this direction. However, in considering the gate size issue, it was observed that the PDAF exhibited certain counter-intuitive behavior. Upon examination of the PDAF derivation it became clear that gate size was not dealt with in a consistent manner, or alternatively, that certain unstated assumptions were made. We have shown how to update the conditional mean and covariance when gate size is explicitly accounted for and have also given assumptions under which the PDAF yields the same results. It turns out that the conditional mean update is the same as in the PDAF, but the conditional covariance update is increased by an additional term. A method is developed to approximate this additional term by considering the extreme cases where process noise >> measurement noise and conversely. The derivations in this section are important for theoretical completeness, and may also be of practical importance under certain operating conditions (e.g., if one selects a moderate gate size for computational reasons). No simulations were run for lack of time and we leave this as a future task.

REFERENCES

1. R. E. Kalman and R. S. Bucy, "New Results in Filtering and Prediction Theory," Trans. ASME: J. Basic Eng., Vol. 83, March 1961, pp. 95-108.

2. I. B. Rhodes, "A Tutorial Introduction to Estimation and Filtering," IEEE Trans. Auto. Control, Vol. AC-16, December 1971, pp. 688-706.

3. A. H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, 1970.

4. P. S. Maybeck, Stochastic Models, Estimation, and Control -- Volume 1, Academic Press, 1979.

5. B. D. O. Anderson, Optimal Filtering, Prentice-Hall, 1979.

6. Y. Bar-Shalom and E. Tse, "Tracking in a Cluttered Environment with Probabilistic Data Association," Automatica, Vol. 11, September 1975, pp. 451-460.

7. T. E. Fortmann and S. Baron, "Problems in Multi-Target Sonar Tracking," Proc. 1978 IEEE Conf. on Decision and Control, San Diego, California, January 1979.

8. T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-Target Tracking Using Joint Probabilistic Data Association," IEEE Journal of Oceanic Engineering, Vol. OE-8, July 1983, pp. ??, (also appeared in Proc. 1980 IEEE Conf. on Decision and Control, Albuquerque, New Mexico, December 1980).

9. R. Singer, R. Sea, and K. Housewright, "Derivation and Evaluation of Improved Tracking Filters for use in Dense Multitarget Environments," IEEE Trans. Info. Theory, Vol. IT-20, July 1974, pp. 423-432.

10. D. B. Reid, "An Algorithm for Tracking Multiple Targets," IEEE Trans. Auto. Control, Vol. AC-24, December 1979, pp. 843-854.

11. E. Taenzer, "Tracking Multiple Targets Simultaneously with a Phased Array Radar," IEEE Trans. Aerospace and Electronic Systems, Vol. AES-16, September 1980, pp. 604-614.

12. C. L. Morefield, "Application of 0-1 Integer Programming to Multitarget Tracking Problems," IEEE Trans. Auto. Control, Vol. AC-22, June 1977, pp. 302-312.

13. D. L. Alspach, "A Gaussian Sum Approximation to the Multitarget Identification - Tracking Problem," Automatica, Vol. 11, May 1975, pp. 285-296.

14. R. W. Sittler, "An Optimal Data Association Problem in Surveillance Theory," IEEE Trans. Mil. Electron., Vol. MIL-8, April 1964, pp. 125-139.

15. Y. Bar-Shalom, "Tracking Methods in a Multi-Target Environment," IEEE Trans. Auto. Control, Vol. AC-23, August 1978, pp. 618-626.

16. H. L. Wiener, W. W. Willman, I. R. Goodman, and J. H. Kullback, "Naval Ocean-Surveillance Correlation Handbook, 1978," NRL Report 8340, Naval Research Lab, October 1979.

17. I. R. Goodman, H. L. Wiener, and W. W. Willman, "Naval Ocean-Surveillance Correlation Handbook, 1979," NRL Report 8402, Naval Research Laboratory, September 1980.

18. T. E. Fortmann and Y. Bar-Shalom, "Modification of the Likelihood Function to Account for Probabilistic Data Association," BBN Report 3964A (revised), Bolt Beranek and Newman Inc., November 1979, Contract N00039-78-C-0296.

19. A. D. Whalen, Detection of Signals in Noise, Academic Press, 1971.

20. T. E. Fortmann, Y. Bar-Shalom, M. Scheffe, and S. Gelfand, "Detection Thresholds for Multi-Target Tracking in Clutter," Proc. 1981 IEEE Conference on Decision and Control, San Diego, California, December 1981.

21. D.P. Berteskas, "Notes on Nonlinear Programming and Discrete-Time Optimal Control," tech. report LIDS-R-919, Massachusetts Institute of Technology, July 1979.

END

DATE
FILMED

5 — 83

DTIC