

REPRESENTATION OF COVARIANCE MATRICES IN TRACK FUSION
PROBLEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MELİH GÜNAY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

NOVEMBER 2007

Approval of the thesis:

**REPRESENTATION OF COVARIANCE MATRICES IN TRACK FUSION
PROBLEMS**

submitted by **MEL H GÜNAY** in partial fulfillment of the requirements for the degree of
**Master of Science in Electrical and Electronics Engineering Department, Middle East
Technical University** by,

Prof. Dr. Canan Özgen

Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Smet Erkmen

Head of Department, **Electrical and Electronics Engineering**

Prof. Dr. Mübeccel Demirekler

Supervisor, **Electrical and Electronics Engineering Dept., METU**

Examining Committee Members:

Prof. Dr. Kemal Leblebicioğlu

Electrical and Electronics Engineering Dept., METU

Prof. Dr. Mübeccel Demirekler

Electrical and Electronics Engineering Dept., METU

Assoc. Prof. Dr. Tolga Çiloğlu

Electrical and Electronics Engineering Dept., METU

Assist. Prof. Dr. Afar Saranlı

Electrical and Electronics Engineering Dept., METU

Dr. Hüseyin Yavuz

ASELSAN Inc.

Date:

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Melih Günay

Signature :

ABSTRACT

REPRESENTATION OF COVARIANCE MATRICES IN TRACK FUSION PROBLEMS

Günay, Melih

M. S., Department of Electrical and Electronics Engineering

Supervisor: Prof. Dr. Mübeccel Demirekler

November 2007, 61 pages

Covariance Matrix in target tracking algorithms has a critical role at multi-sensor track fusion systems. This matrix reveals the uncertainty of state estimates that are obtained from different sensors. So, many subproblems of track fusion usually utilize this matrix to get more accurate results. That is why this matrix should be interchanged between the nodes of the multi-sensor tracking system. This thesis mainly deals with analysis of approximations of the covariance matrix that can best represent this matrix in order to effectively transmit this matrix to the demanding site. Kullback-Leibler (KL) Distance is exploited to derive some of the representations for Gaussian case. Also comparison of these representations is another objective of this work and this is based on the fusion performance of the representations and the performance is measured for a system of a 2-radar track fusion system.

Keywords: Covariance Matrix, Kullback-Leibler Distance, Track Fusion, K-means Vector Quantization Method, Extended Kalman Filtering

ÖZ

İZ BİRLEŞTİRME PROBLEMLERİNDE DEĞİŞİNTİ MATRİSİ GÖSTERİMİ

Günay, Melih

Yüksek Lisans, Elektrik Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Mübeccel Demirekler

Kasım 2007, 61 sayfa

Hedef takip algoritmalarından elde edilen değişinti matrisi iz birleştirme sistemlerinde önemli bir yere sahiptir. Bu matris değişik radarlardan elde edilen durum tahmin vektörlerindeki belirsizliği göstermektedir. Bu nedenle, iz birleştirme sistemleri alt problemleri daha kesin sonuçlar elde etmek için genellikle bu matrisi kullanırlar. Bundan dolayı değişinti matrisinin çoklu-sensör hedef takip sistemlerinde sistem birimleri arasında değişilmesi gerekmektedir. Bu tez değişinti matrisinin ihtiyaç duyulan tarafa gönderilebilmesi için yerini iyi bir şekilde tutabilen çeşitli gösterim şekillerinin analizini kapsamaktadır. Bu amaçla, Gaussian durum için, bazı gösterim şekillerinin elde edilebilmesi amacıyla Kullback-Leibler uzaklığından faydalanılmıştır. Bunların yanında bu gösterimlerin karşılaştırılabilmesi bu tezin diğer amaçlarından bir tanesidir. Karşılaştırma gösterimlerin iz birleştirme başarımlarına bakılarak yapılmıştır. Başarımın ölçülebilmesi için de 2 radarlı iz birleştirme simülasyon ortamı kullanılmıştır.

Anahtar Kelimeler: Değişinti Matrisi, Kullback-Leibler Uzaklığı, İz Birleştirme, K-ortalı Vektör Nicemleme Metodu, Genişletilmiş Kalman Süzgeci

To my beautiful daisy whom my big love belongs to...

ACKNOWLEDGMENTS

It is a pleasure for me to express my sincere gratitude to my thesis supervisor Prof. Dr. Mübeccel Demirekler for her belief, patience, encouragement and guidance throughout the study. More importantly than this thesis work, her life perspective has always enlightened my life and provided me confidence for solving real life problems.

I would also like to thank my company ASELSAN Inc. for supporting this thesis work and encouraging us to make scientific studies for the problems that we face off during the work. I also appreciate my company for trusting us which is so important for the success both in life and work.

I would also like to give my special thanks to my family whose patient love enabled me to complete this work.

PREFACE

This thesis is the result of the author's 3-year M.S. study under the supervision of Prof. Dr. Mübeccel Demirekler. This problem has arisen from the Sensor Fusion project work that has been studied for 3 years. In the thesis, we have tried to find a solution to the problem of representation of covariance matrix for limited bandwidth case to increase the performance of the sensor fusion system that we work on.

TABLE OF CONTENTS

PLAGIARISM	iii
ABSTRACT	iv
ÖZ	v
DEDICATION	vi
ACKNOWLEDGMENTS	vi
PREFACE	viii
TABLE OF CONTENTS	ix
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER	
1 INTRODUCTION	1
1.1 Covariance Matrix	3
1.1.1 Properties of Covariance Matrix	4
1.1.2 Role of Covariance Matrix in Track Fusion Systems	5
2 KULLBACK-LEIBLER DISTANCE	7
2.1 Definition	7
2.2 Kullback-Leibler Distance Between Two Gaussian Dis- tributions	8
3 REPRESENTATIONS AND THEIR ANALYSIS	12
3.1 Some Approximations of Covariance Matrices Utilizing Kullback-Leibler Distance	13
3.1.1 Alpha Representation (AR)	14

3.1.2	Diagonal Matrix Representation (DMR)	15
3.1.3	Eigenvalue - Eigenvector Representation (EER)	16
3.2	Other Approximations of Covariance Matrices	23
3.2.1	Whole Matrix Representation (WMR)	23
3.2.2	Block Diagonal Matrix Representation (BDMR)	24
4	EXPERIMENTS AND ANALYSIS	25
4.1	Simulation Model	25
4.1.1	Target Motion and Measurement Models . . .	27
4.1.2	Radar Tracker	29
4.1.3	Track-to-Track Fusion	31
4.2	Performance Evaluation	32
4.3	Representations of Covariances Specific to Target Track- ing and Their Quantization	36
4.3.1	Representation Methods	36
4.3.2	KL K-Means Clustering Algorithm	41
4.3.3	Results	46
5	SUMMARY, CONCLUSION AND FUTURE WORK	57
	REFERENCES	59

LIST OF TABLES

4.1	Performance of Representation Matrices without Quantization.	46
-----	--	----

LIST OF FIGURES

3.1	Kalman filter output vector	13
3.2	Schematic representation of the angle concept	21
4.1	Simulation System Model	26
4.2	Flowchart of EKF Algorithm (One Cycle),[1]	30
4.3	Sensor-to-Sensor track Fusion Scheme	31
4.4	MSE calculation process for the true covariance matrices . . .	34
4.5	MSE calculation process for the represented covariance matrices	35
4.6	AR approximation process	37
4.7	Histogram of eigenvalue average and uniform quantization lev- els for 5 bits.	38
4.8	DMR approximation process	39
4.9	BDMR approximation process	40
4.10	Generated Tracks for Training and Cross-Sections of the Cor- responding Covariances	42
4.11	Cross-Sections of initial selected covariances for WMR method	44
4.12	Cross-Sections of final codebook covariances of WMR method for 4 bits.	45

4.13	Performances of Representation Matrices	47
4.14	AR method examples	49
4.15	DMR method examples	50
4.16	BDMWMR2 method examples	51
4.17	BDMWMR1 method examples	52
4.18	BDMEER1 method examples	53
4.19	BDMEER2 method examples	54
4.20	WMR method examples	55

CHAPTER 1

INTRODUCTION

Lemma 1.1 *We have to learn to live with uncertainty.*

Theorem 1.1 *We can model uncertainty.*

Corollary 1 *We can make a living out of uncertainty.*

written by Y. Bar-Shalom and X.R. Li in their well-known book on tracking and fusion systems [1]. They pointed out the importance of the uncertainty and knowledge of it by these statements. Because of this importance, several algorithms exploiting the uncertainty have been developed to obtain more accurate results.

The relation between the uncertainty and the estimation theory have been realized by the mathematical variable 'covariance' that is defined as the deviation of the random vector around its mean. For instance, tracking systems that are based on Kalman filtering generate a covariance matrix for each state estimate. This matrix reveals the uncertainty in the estimation which gives an idea about how well the corresponding estimate is. This measure is especially functional for multisensor fusion systems who have several sensors employing Kalman-based filter trackers. Measurement decorrelation, data association and track fusion are some of the subproblems of fusion systems and solution of these problems utilizes the covariance information. Some more detailed information will be discussed later within this chapter.

In view of these matters, since the knowledge of covariance information at a fusion center is crucial for better performance, Bar-Shalom and Li's idea

above can be extended to the following corollary.

Corollary 2 *Covariance (Uncertainty) should go with the corresponding estimated state.*

In other words, covariance matrix might be as important as an estimated state and it should not be underestimated and should be shared in between the nodes of the fusion system. So, the covariance matrix transfer problem between the system nodes has become a basis for this thesis work.

In the literature, there are few studies on covariance matrix approximation and quantization techniques which is a prerequisite for sharing the covariance. [2] and [3] discussed a vector quantization scheme for the elements of the covariance matrix. They also deal with the quantization of the state estimate together with the covariance. [4] proposed an algorithm not for the covariance matrix itself but for the eigenvectors of it. In this study, codebooks for eigenvector matrices are generated and the best matrix is chosen within this book. Although the method proposed in this study is not directly related with the covariance matrix quantization, its eigenvector quantization method might be exploited in covariance quantization.

Bandwidth limitation is the main constraint for covariance matrix representation. In real life, bandwidth for covariance matrices is always much more less than the bandwidth for some other information such as the state estimate. That is why the quantization architecture for the covariance matrix should be constructed in an intelligent manner. So, this work is an analysis of efficient representation of the covariance matrices belonging to Gaussian density functions. Our objective is finding out some useful representations of the covariance matrix and to obtain explicitly the degradation corresponding to it. Degradations are measured by using the Kullback-Leibler (KL) distance or relative entropy. More clearly, this thesis basically deals with the analysis and comparison of different representations for a given covariance matrix of a Gaussian distributed random vector. The KL distance is used for getting the distance measure of how far away the representation matrix is from the actual covariance. So, optimal representation matrices can be found in terms of some

parameters of the covariance. Another objective of this thesis is improving the performance of the track fusion system by using good approximations of the covariance matrix at various stages of the fusion systems.

In this thesis, an analysis on the approximations of the covariances matrices of Gaussian density distributions has been made. To measure the performances of these approximations, a track fusion system simulator has been generated. We assumed a scenario that two radars are connected to a fusion center and sending their own track information to the center. Track information is composed of the estimated state and the approximation of the corresponding covariance matrix. Then, track fusion results of the actual and approximated covariance matrices are compared. We have generated randomly 1000 track samples with frame length of 50 for each radar and taken one sample covariance matrix for each track. Using these matrices and the state estimates, average errors of the fused tracks are calculated and compared for each approximation method.

The outline of this thesis study is as follows: Chapter 1 will go on giving some information about the properties of covariance matrices and roles of the covariance matrix at fusion systems. Chapter 2 presents information about the KL distance and reveals the equations deriving the KL distance between two Gaussian densities. Chapter 3 offers some representation methods and analyzes them according to the KL distance criteria. Simulation environment and experiments with the results are given in Chapter 4. Finally, summary and conclusion of this study and future work plan are told at Chapter 5.

To begin with, before attempting to represent a given covariance matrix, it is better to summarize its basic properties since the equations will heavily use these properties.

1.1 Covariance Matrix

This chapter discusses covariance matrix, its definition and properties. Importance of the covariance matrix for engineers is also emphasized by giving its

role at the sensor fusion systems.

Since covariance represents the variation of the random vector around the mean, it is clear to see that it is a measure of how well the mean is obtained. For instance, a considerable work on covariance matrices have been done for multisensor-multitarget tracking and sensor fusion systems. These systems require the covariance since the covariance represents the uncertainty in the state estimation and this uncertainty is used in various algorithms [1], [5].

1.1.1 Properties of Covariance Matrix

As mentioned before, the covariance represents the variation of the random vector around its mean. So, the mathematical definition of the covariance is given by the equation (1.1).

$$P = E\{(x - \hat{x})(x - \hat{x})^T\} \quad (1.1)$$

where x_k is the random vector and \hat{x}_k is its mean. Considering the mathematical definition, one can easily derive the following features of covariance:

1. symmetricity
2. positive semi-definiteness
3. all eigenvalues of a covariance matrix are real and non-negative
4. all eigenvectors are orthonormal to each other
5. a covariance matrix may be decomposed as $P = VDV^T$ where column vectors of V are eigenvectors such that $e_i \perp e_j$ for $i \neq j$ and D is a diagonal matrix containing eigenvalues of P as its diagonal elements.

In the introduction of this chapter, it is mentioned that some application areas highly use the covariance information for increasing their system performance. A bit more discussion on these applications is made in the following section:

1.1.2 Role of Covariance Matrix in Track Fusion Systems

Since multisensor-multitarget Tracking systems or sensor fusion systems deeply work on the knowledge of the covariance matrices, we will mention about the usage of the covariance for these systems. There are three basic reasons to know the covariance of an estimated state for effective target tracking and decision making: measurement noise decorrelation, data association and track fusion.

Before passing to the next section, it is also important to note that that in the sensor fusion systems, the covariance information is usually generated by the filters that utilizes Kalman filtering.

1.1.2.1 Measurement Noise Decorrelation

Trackers using Kalman filters work on the idea of independent and Gaussian measurement noise assumption, i.e., they assume that filter-input measurement and process noise sequences are uncorrelated, [6]. On the other hand, filter output data has no more independent noise but it has a dependent noise sequence. This filter output data can not be used as input to a Kalman-based fusion system because of the violation of the independence assumption. This type of usage might result in unacceptable errors. To avoid this danger, some methods are generated to convert the output state into a state that has uncorrelated measurement noise sequence. In the literature, there are some methods proposed to decorrelate the noise in tracker output data such as 'information decorrelation method' [7],[8] and 'equivalent measurement approach' [5],[9],[10]. These methods rely on the **covariance information or usage of the covariance matrices** to decorrelate the noise inside the data. For instance, at fusion systems, to compensate errors in the state estimations caused by the time delays within the system, one of these methods should be exploited to get the target position more accurately.

1.1.2.2 Data Association

Data association or track association is one of the challenging problems at multisensor-multitarget fusion systems. Since association algorithms of these systems give the decision of whether information from different sensors belongs to the same target or not, one may guess the importance of inclusion of the covariance in the association methods. For instance, one of the commonest method for the distance calculation in the association algorithms is calculating the statistical distance, i.e., Mahalanobis distance [11]. This distance equation includes the state estimations with their **corresponding covariances** and produces a distance measure between the estimates considering the uncertainties.

1.1.2.3 Track Fusion

Similar to the data association mechanism, track fusion algorithms heavily use the covariance matrix information to obtain better target positions. For example, 'weighted covariance fusion (WCF)' [1], [12], exploits the **covariance matrix** and the state estimates to obtain the fused state of the target.

CHAPTER 2

KULLBACK-LEIBLER DISTANCE

In this chapter, Kullback-Leibler (KL) measure - basic measure tool for comparison of two densities- will be defined. We will derive KL measure between two gaussian densities for future purposes of our work.

2.1 Definition

In probability theory and information theory, the Kullback-Leibler distance is defined as a distance measure for evaluating the distance between a probability distribution ' p ' and another probability distribution ' q ' [13]. While ' p ' represents data, observations, or a precise calculated probability distribution, the distribution ' q ' typically represents a theory, a model, a description or an approximation of ' p '. The KL measure or informally distance is used for various applications at engineering and mathematics disciplines.

For probability distributions ' p ' and ' q ' of a discrete variable, the KL distance between ' q ' and ' p ' is defined to be as in (2.1).

$$KL(p, q) = \sum_i p(i) \ln \frac{p(i)}{q(i)} \quad (2.1)$$

For distributions p and q of a continuous random variable, the summations turn out to be integrals, so that the distance becomes:

$$KL(p, q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx \quad (2.2)$$

where $p(x)$ and $q(x)$ denote the densities. KL distance is not a formal distance definition since it is not symmetric with respect to its argument, i.e., $KL(p, q) \neq KL(q, p)$. However, it usually represents the difference between a pdf and its approximation.

For more information on the Kullback Leibler measure, reader is referred to [13] and [14], and for some applications, to [15], [16] and [17].

2.2 Kullback-Leibler Distance Between Two Gaussian Distributions

In this section, we will calculate the Kullback Leibler distance (KLD) between two Gaussian densities $p(x)$ and $q(x)$. Therefore, we will assume that

$p(x) = N(x; \hat{x}_p, P)$ and $q(x) = N(x; \hat{x}_q, Q)$ where $N(x; a, A)$ denotes the Gaussian density with mean a and covariance A . So that, probability density functions of p and q can be written as:

$$p(x) = \frac{1}{\sqrt{(2\pi)^n |P|}} \exp\left(-\frac{1}{2}(x - \hat{x}^p)^T P^{-1}(x - \hat{x}^p)\right) \quad (2.3)$$

$$q(x_k) = \frac{1}{\sqrt{(2\pi)^n |Q|}} \exp\left(-\frac{1}{2}(x - \hat{x}^q)^T Q^{-1}(x - \hat{x}^q)\right) \quad (2.4)$$

where n is the dimension of the random vector x .

Computation of this KL distance is given below:

1. Write distance formula:

$$KL(p, q) = \int p(x) \ln \frac{p(x)}{q(x)} dx \quad (2.5)$$

2. Evaluate the ratio $\frac{p(x)}{q(x)}$:

$$\frac{p}{q} = \sqrt{\frac{|Q|}{|P|}} \exp\left(-\frac{1}{2}(x - \hat{x}^p)^T P^{-1}(x - \hat{x}^p)\right) \exp\left(\frac{1}{2}(x - \hat{x}^q)^T Q^{-1}(x - \hat{x}^q)\right) \quad (2.6)$$

3. Insert the ratio equation (2.6) into the KLD equation (2.5):

$$\begin{aligned}
KL(p, q) &= \int \frac{1}{\sqrt{(2\pi)^n |P|}} \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) \cdot \\
&\quad \left(\ln \sqrt{\frac{|Q|}{|P|}} - \frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) + \right. \\
&\quad \left. \frac{1}{2} (x - \hat{x}^q)^T Q^{-1} (x - \hat{x}^q) \right) dx
\end{aligned} \tag{2.7}$$

Then, equation (2.7) takes the following form:

$$\begin{aligned}
KL(p, q) &= \ln \sqrt{\frac{|Q|}{|P|}} - \frac{1}{2} \int \frac{(x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p)}{\sqrt{(2\pi)^n |P|}} \cdot \\
&\quad \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx + \\
&\quad \frac{1}{2} \int \frac{(x - \hat{x}^q)^T Q^{-1} (x - \hat{x}^q)}{\sqrt{(2\pi)^n |P|}} \cdot \\
&\quad \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx
\end{aligned} \tag{2.8}$$

4. Since trace of a scalar is equal to the scalar itself, equation (2.8) turns out to be the equation (2.9):

$$\begin{aligned}
KL(p, q) &= \ln \sqrt{\frac{|Q|}{|P|}} - \frac{1}{2} \int \frac{\text{tr}\{(x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p)\}}{\sqrt{(2\pi)^n |P|}} \cdot \\
&\quad \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx + \\
&\quad \frac{1}{2} \int \frac{\text{tr}\{(x - \hat{x}^q)^T Q^{-1} (x - \hat{x}^q)\}}{\sqrt{(2\pi)^n |P|}} \cdot \\
&\quad \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx
\end{aligned} \tag{2.9}$$

5. Because of the fact that $\text{tr}(AB) = \text{tr}(BA)$, we find the distance as in (2.10).

$$\begin{aligned}
KL(p, q) &= \ln \sqrt{\frac{|Q|}{|P|}} - \frac{1}{2} \int \frac{\text{tr}\{P^{-1} (x - \hat{x}^p) (x - \hat{x}^p)^T\}}{\sqrt{(2\pi)^n |P|}} \cdot \\
&\quad \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx + \\
&\quad \frac{1}{2} \int \frac{\text{tr}\{Q^{-1} (x - \hat{x}^q) (x - \hat{x}^q)^T\}}{\sqrt{(2\pi)^n |P|}} \cdot \\
&\quad \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx
\end{aligned} \tag{2.10}$$

6. Trace is a linear function. So we can write:

$$\begin{aligned}
KL(p, q) = & \ln \sqrt{\frac{|Q|}{|P|}} - \frac{1}{2} \text{tr} \left\{ P^{-1} \int \frac{(x - \hat{x}^p)(x - \hat{x}^p)^T}{\sqrt{(2\pi)^n |P|}} \right. \\
& \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx \Big\} + \\
& \frac{1}{2} \text{tr} \left\{ Q^{-1} \int \frac{(x - \hat{x}^q)(x - \hat{x}^q)^T}{\sqrt{(2\pi)^n |P|}} \right. \\
& \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx \Big\} \quad (2.11)
\end{aligned}$$

7. We know that the covariance matrix of a random vector can be obtained as in (2.12).

$$P = \int \frac{(x - \hat{x}^p)(x - \hat{x}^p)^T}{\sqrt{(2\pi)^n |P|}} \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx \quad (2.12)$$

This term is also seen in equation (2.11). As a result, (2.11) can be written as:

$$\begin{aligned}
KL(p, q) = & \frac{1}{2} \ln \frac{|Q|}{|P|} - \frac{1}{2} \underbrace{\text{tr} \{ P^{-1} P \}}_n + \\
& \frac{1}{2} \text{tr} \left\{ Q^{-1} \int \frac{(x - \hat{x}^q)(x - \hat{x}^q)^T}{\sqrt{(2\pi)^n |P|}} \right. \\
& \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx \Big\} \quad (2.13)
\end{aligned}$$

8. Focusing on the last element of the summation in (2.13), let

$$A = \int \frac{(x - \hat{x}^q)(x - \hat{x}^q)^T}{\sqrt{(2\pi)^n |P|}} \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx \quad (2.14)$$

Then, let us write A more explicitly:

$$\begin{aligned}
A = & \int (xx^T - x(\hat{x}^q)^T - \hat{x}^q x^T + \hat{x}^q (\hat{x}^q)^T) \cdot \\
& \frac{1}{\sqrt{(2\pi)^n |P|}} \exp \left(-\frac{1}{2} (x - \hat{x}^p)^T P^{-1} (x - \hat{x}^p) \right) dx \quad (2.15)
\end{aligned}$$

$$A = P + \hat{x}^p(\hat{x}^p)^T - \hat{x}^p(\hat{x}^q)^T - \hat{x}^q(\hat{x}^p)^T + \hat{x}^q(\hat{x}^q)^T \quad (2.16)$$

Simplified A is found and given as in (2.17)

$$A = P + (\hat{x}^p - \hat{x}^q)(\hat{x}^p - \hat{x}^q)^T \quad (2.17)$$

9. Finally, inserting A given as in (2.17) to the equation (2.13), we find

simplified form of KL distance between the two normal densities

$N(x; \hat{x}^p, P)$ **and** $N(x; \hat{x}^q, Q)$ as the following:

$$KL(p, q) = \frac{1}{2} \left(\ln \frac{|Q|}{|P|} - n + tr \left\{ Q^{-1} P \right\} + tr \left\{ Q^{-1} (\hat{x}^p - \hat{x}^q)(\hat{x}^p - \hat{x}^q)^T \right\} \right) \quad (2.18)$$

It is clear to see that when $p(x) = q(x) = N(x; \hat{x}, R)$, $KL(p, q) = 0$

This distance equation will be useful in quantization of covariance matrix since it is simple and gives an idea about **”how far away”** a Gaussian density is from the other one. Especially for equal-mean Gaussian distributions, we have a very simple distance equation. The KL distance will be a basis for generating new approximation methods in our study.

CHAPTER 3

REPRESENTATIONS AND THEIR ANALYSIS

As previously mentioned, there are few studies on covariance matrix approximation and quantization in the literature. [2] and [3] proposed a vector quantization scheme for the elements of the covariance matrix with the quantization of the state estimate. [4] proposed an algorithm not for the covariance matrix itself but for the eigenvector matrix of it. As far as we know there isn't any study on various representations -which are introduced in this thesis- of the covariance and their comparison. So, this thesis offers a new approach for the representation of the covariance matrix as well as their comparison.

In this chapter, some representations for the covariance matrix are proposed. While deriving them, we utilize the KL measure for method generation and for distance calculation between the covariance matrices. We use the KL distance as our cost function and try to find the most appropriate matrices that minimize this cost and these matrices minimizing the distance will be assumed to best represent the actual covariance matrix. We assume that we have two pdf's, one is the actual gaussian pdf, the other is the approximate version of the actual. We work on Gaussian densities because our objective is to represent the output of a Kalman tracker which is assumed to be a Gaussian random vector whose mean is taken as the estimated state (Figure (3.1)).

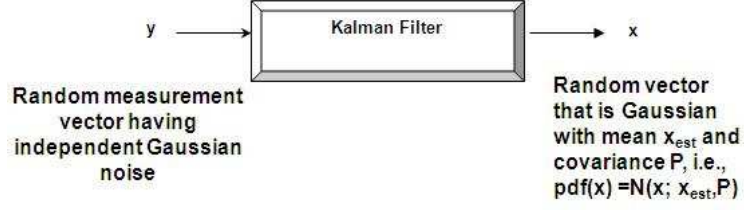


Figure 3.1: Kalman filter output vector

Since our aim is to analyze the effect of the approximations of the covariance matrix on performance of a track fusion system, we assume that mean of two densities are identical. Considering Kullback-Leibler distance, since $\hat{x}^p = \hat{x}^q$, the distance equation given by (2.18) is reduced to the following form:

$$KL(p, q) = \frac{1}{2} \left(\ln \frac{|Q|}{|P|} + \text{tr} \{ Q^{-1} P \} - n \right) \quad (3.1)$$

As a result, we have the distance equation as a function of only P and Q , i.e., $KL(p, q) = KL(P, Q)$.

Now, we will offer some representation matrices and we will make Kullback-Leibler analysis for some of them.

3.1 Some Approximations of Covariance Matrices Utilizing Kullback-Leibler Distance

In this part, representation approaches for a given $n \times n$ covariance matrix are introduced and some derivations are made regarding the optimality of them using the Kullback-Leibler distance. So, we first start with the simplest approximation and approximate a given covariance matrix by αI , called α representation.

3.1.1 Alpha Representation (AR)

Assume that α is a real positive number and $Q = \alpha I$. Our aim is to minimize the Kullback- Leibler distance in terms of α so that the best α for a given P can be found. Considering the fact that determinant of a matrix equals to the multiplication of eigenvalues, the objective function that must be minimized can be written as:

$$KL(P, Q) = \frac{1}{2} \left(\ln \frac{\alpha^n}{\prod_{k=1}^n \lambda_k^p} + \text{tr} \{ Q^{-1} P \} - n \right) \quad (3.2)$$

where λ_i^p is the i^{th} eigenvalue of P . Then,

$$KL(P, Q) = \frac{1}{2} \left(\ln \frac{\alpha^n}{\lambda_1^p \dots \lambda_n^p} + \frac{1}{\alpha} \text{tr} \{ P \} - n \right) \quad (3.3)$$

The solution of this problem is given in the next lemma.

Lemma 3.1 *The minimum value of the objective function given in (3.3) above is $KL^*(P, Q) = \frac{1}{2} \ln \frac{\lambda_{avg}^p}{\lambda_1^p \dots \lambda_n^p}$ where $\alpha^* = \lambda_{avg}^p = \frac{1}{n} \sum_{k=1}^n \lambda_k^p$*

Proof To find the optimum solution, the derivative of $KL(P, Q)$ with respect to α is equated to zero. Rewriting KL distance by considering $\text{tr} \{ P \} = \sum_{k=1}^n \lambda_k^p$:

$$KL(P, Q) = \frac{1}{2} \left(n \ln(\alpha) - \ln \left(\prod_{k=1}^n \lambda_k^p \right) + \frac{1}{\alpha} \sum_{k=1}^n \lambda_k^p \right) \quad (3.4)$$

Then taking the derivative:

$$\frac{\partial KL(P, Q)}{\partial \alpha} = \frac{1}{2} \left(\frac{n}{\alpha} - \frac{1}{\alpha^2} \sum_{k=1}^n \lambda_k^p \right) \quad (3.5)$$

$$\frac{\partial KL(P, Q)}{\partial \alpha} = 0 \implies \alpha^* = \frac{1}{n} \sum_{k=1}^n \lambda_k^p \quad (3.6)$$

Inserting α^* into the distance equation (3.3),

$$KL(\alpha^*) = \frac{1}{2} \ln \frac{\lambda_{avg}^p}{\lambda_1^p \dots \lambda_n^p} \quad (3.7)$$

It is easy to show that this solution is the global minimum of the objective function since the derivative of the function is negative for $0 < \alpha < \alpha^*$ and positive for $\alpha > \alpha^*$.

This completes the proof. It is interesting to note that optimal α is the average of eigenvalues of the covariance matrix P .

Another approach may be just representing the covariance with only a diagonal matrix. Following section exploits this method:

3.1.2 Diagonal Matrix Representation (DMR)

At this section we assume that $Q = \text{diag}(d_1, \dots, d_n)$. For this approximation of the Q matrix, we can write the objective function as:

$$KL(P, Q) = \frac{1}{2} \left(\ln \frac{d_1 \dots d_n}{\lambda_1^p \dots \lambda_n^p} + \text{tr} \left\{ \begin{bmatrix} \frac{1}{d_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{d_n} \end{bmatrix} P \right\} n \right) \quad (3.8)$$

$$KL(P, Q) = \frac{1}{2} \left(\ln \frac{d_1 \dots d_n}{\lambda_1^p \dots \lambda_n^p} + \sum_{k=1}^n \frac{p_{kk}}{d_k} - n \right) \quad (3.9)$$

$$\Rightarrow KL(P, Q) = \frac{1}{2} \sum_{k=1}^n \left(\ln \frac{d_k}{\lambda_k^p} + \frac{p_{kk}}{d_k} - 1 \right) \quad (3.10)$$

where p_{kk} is the $(k, k)^{th}$ element of P .

Note that the objective function can be written as the sum of n functions, each is a function of one variable, i.e., d_k . Define:

$$f_k(d_k) = \frac{1}{2} \left(\ln \frac{d_k}{\lambda_k^p} + \frac{p_{kk}}{d_k} - 1 \right) \quad (3.11)$$

So,

$$KL(P, Q) = \sum_{k=1}^n f_k(d_k) \quad (3.12)$$

Equating 3.11 has exactly the same structure as AR with α replaced by d_k and $\sum_{k=1}^n \lambda_k^p$ by p_{kk} . So that the optimal solution of 3.11 is $d_k = p_{kk}$ and optimal solution of DMR is $\{d_k = p_{kk}\}_{k=1}^n$ by previous lemma. This result is summarized below as lemma 3.2.

Lemma 3.2 *The minimum value of the objective function given in (3.10) above is $KL^*(P, Q) = \ln \frac{p_{11} \dots p_{nn}}{\lambda_1^p \dots \lambda_n^p}$ and the optimum solution is $d_k^* = p_{kk}$ where p_{kk} is the diagonal entry of the matrix P .*

Solutions of these two approaches give some information about the general solution. From now on, we will consider that Q is an approximation of P that may not be diagonal but Q belongs to a set probably containing finite number of elements. Q can be decomposed as $Q = VDVT^T$ where V is an orthonormal matrix (eigenvector matrix) and D is the diagonal matrix that contains eigenvalues of Q . In other words now we assume that eigenvectors and eigenvalues of P are represented as elements of a set that has finite elements. Under these conditions, the aim is to select the best V and D that minimizes the Kullback-Liebler distance among these probability densities. Such a selection can be considered as decoding stage of a coding algorithm. Following section will exploit this idea.

3.1.3 Eigenvalue - Eigenvector Representation (EER)

In this section, we assume that we apply eigenvalue decomposition for P and quantize eigenvalues and eigenvectors of P to finite values and finite vectors. So, we have $P = TD_pT^T$ and $Q = VD_qV^T$ where $D_p = \text{diag}(\lambda_1^p, \dots, \lambda_n^p)$ and $D_q = \text{diag}(\lambda_1^q, \dots, \lambda_n^q)$. λ_i^p and λ_i^q are eigenvalues of P and Q , respectively. $T = [e_1^p \dots e_n^p]$ where e_i^p is the i^{th} normalized eigenvector of P and e_i^p is orthogonal to e_j^p if $i \neq j$. Similarly, $V = [e_1^q \dots e_n^q]$ where e_i^q is the i^{th} normalized eigenvector of Q and e_i^q is orthogonal to e_j^q ($i \neq j$). Writing the KL distance (Eq. 3.1) in terms of eigenvectors and eigenvalues and using the fact that $T^T = T^{-1}$ and $V^T = V^{-1}$ and noting that the determinant of a matrix equals to multiplication of eigenvalues:

$$KL(P, Q) = \frac{1}{2} \left(\sum_{i=1}^n \ln \frac{\lambda_i^q}{\lambda_i^p} + \text{tr} \left\{ VD_q^{-1}V^{-1}TD_pT^{-1} \right\} - n \right) \quad (3.13)$$

Since $\text{tr}\{AB\} = \text{tr}\{BA\}$, we can write (3.13) as:

$$KL(P, Q) = \frac{1}{2} \left(\sum_{i=1}^n \ln \frac{\lambda_i^q}{\lambda_i^p} + \text{tr} \left\{ D_q^{-1}V^{-1}TD_pT^{-1}V \right\} - n \right) \quad (3.14)$$

Let $K = T^{-1}V = T^TV$, then,

$$KL(P, Q) = \frac{1}{2} \left(\sum_{i=1}^n \ln \frac{\lambda_i^q}{\lambda_i^p} + \text{tr} \left\{ D_q^{-1} K^{-1} D_p K \right\} - n \right) \quad (3.15)$$

Writing more explicitly, we obtain the distance equation in terms of eigenvalues and eigenvectors as:

$$KL(P, Q) = \frac{1}{2} \left(\sum_{i=1}^n \ln \frac{\lambda_i^q}{\lambda_i^p} + \sum_{j=1}^n \sum_{i=1}^n \left\{ \frac{\lambda_j^p}{\lambda_i^q} k_{ij}^2 \right\} - n \right) \quad (3.16)$$

where k_{ij} represents $(i, j)^{th}$ element of the matrix K.

3.1.3.1 EER Analysis for nxn Matrices

In this section, general solution for EER will be given. By representing eigenvalues and eigenvalues of the actual covariance, we will obtain a representation of the covariance matrix. Now, we will assume that eigenvalues of P are represented quite well, i.e., $\lambda_i^p \sim \lambda_i^q$ and investigate the eigenvector representation of P . So, distance equation, which can also be taken as the cost function, turn out to be the following one.

$$KL(P, Q) \propto D(P, Q) = \sum_{j=1}^n \sum_{i=1}^n \left\{ \frac{\lambda_i^p}{\lambda_i^q} k_{ij}^2 \right\} \quad (3.17)$$

Since $K = T^TV$, $k_{ij} = \langle e_i^p, e_j^q \rangle$. So, $D(P, Q)$ is obtained as:

$$D(P, Q) = \sum_{j=1}^n \sum_{i=1}^n \left\{ \frac{\lambda_j^p}{\lambda_i^q} \langle e_i^p, e_j^q \rangle^2 \right\} \quad (3.18)$$

This formulation shows that even if $\lambda_i^p = \lambda_i^q$, they still have a role on the selection of V matrix. Furthermore analysis and intuition shows that one must pay more attention on the selection of $\{e_i^q, e_j^q\}$ pair if $\{\lambda_i^q, \lambda_j^q\}$ are far from each other.

In the following subsection, we will give an algorithm that selects the best matrix among a set of matrices Ω that minimizes the D function. For this purpose, assume that $\{\lambda_i^p\}_{i=1}^n$ and $\{\lambda_i^q\}_{i=1}^n$ are ordered sets that are decreasing. Rewriting the minimization problem:

$$\min_{e^q} D(P, Q) \quad s.t. \quad \langle e_i^q, e_j^q \rangle = \delta_{i,j} \quad (3.19)$$

The above defined optimization is not easy to solve due to the complex interaction between the vectors that must be selected. We will give an algorithm that creates a suboptimal solution. The algorithm is based on using multiple code books, $\{\Omega_i\}_{i=1}^{n-1}$, each corresponding to a different e_i^p vector. Define $c_{j,i} = \frac{\lambda_j^p}{\lambda_i^q}$. Note that since the sets $\{\lambda_i^p\}_{i=1}^n$ and $\{\lambda_i^q\}_{i=1}^n$ have decreasing elements, we can say that $c_{j,i}$'s are decreasing with respect to the first argument, i.e., $c_{j,i} \geq c_{j+1,i}$ for $j = 1, \dots, n-1$ and $i = 1, \dots, n$. Similarly, $c_{j,i}$'s are increasing with respect to the second argument, i.e., $c_{j,i} \geq c_{j,i-1}$ for $j = 1, \dots, n$ and $i = 1, \dots, n-1$. We propose an algorithm based on these observations. the algorithm first solves the following minimization problem by simply computing the cost for all possible $e_{1,t}^q \in \Omega_1$ and choosing the one that gives the minimum cost. According to the definitions given, subcost function is defined as:

$$C_1^t = \sum_{i=1}^n \frac{\lambda_1^p}{\lambda_i^q} \langle e_i^p, e_{1,t}^q \rangle^2 = \sum_{i=1}^n c_{1,i} \langle e_i^p, e_{1,t}^q \rangle^2 \quad (3.20)$$

that fixes the value of $e_{1,t}^q$ as $e_{1,*}^q$.

In the second step, we first use an orthonormal transformation T_1^* that transforms $e_{1,*}$ to $e_1 = [1 \ 0 \dots 0]^T$, i.e., $T_1^* e_{1,*}^q = [1 \ 0 \dots 0]^T$. The transformation matrix T_1^* can be obtained by applying Gramm-Schmidt orthogonalization to a properly selected matrix. It can obviously be done in an offline fashion and can be stored as attached to each vector $e_{1,t}^q \in \Omega_1$. The next step of the algorithm transfers all e_i^p vectors to an $n-1$ dimensional space by using T_1^* , and eliminating the first element. The first step is repeated for the new problem in the reduced dimensional space. Once the optimal vector of Step 2 is found it is augmented by adding a zero to the first component and transformed back by using $(T_1^*)^{-1}$ that gives $e_{2,*}^q$. Algorithm continues in this way for $n-1$ steps. The n^{th} vector is the one that is orthogonal to all of the previously computed vectors.

Lemma 3.3 $\langle e_{1,*}^q, e_{2,*}^q \rangle = 0$

Proof $\langle e_{1,*}^q, e_{2,*}^q \rangle = (e_{1,*}^q)^T e_{2,*}^q = (T_1^* e_{1,*}^q)^T T_1^* e_{2,*}^q = e_1^T (T_1^* e_{2,*}^q) = 0$ since the

first component of $(T_1^* e_{2,*}^q)$ is zero.

Corollary 3 $\langle e_{i,*}^q, e_{j,*}^q \rangle = 0$ for $i \neq j$

Proof This is a simple generalization of the Lemma (3.3).

Note that the above algorithm can only generate a suboptimal solution since each sub minimization problem uses the result of the previous minimization problem that may produce a non-optimal solution for the complete problem. It is possible to overcome this difficulty partially and obtain a better solution to the expense of computation. This can be done by delaying the decisions to a later step. Such a strategy obviously increases the computational load exponentially. A procedure that has two steps will be given in the next algorithm. The number of the steps of the algorithm is $n - 1$ and 'Step k' is defined as a selection procedure of e_k^q from the corresponding codebook. The states at each step are associated with the elements of the related codebook, for example the number of the states of the first step is cardinality of the codebook Ω_1 . Computational load comes mainly from the transformation matrices. The algorithm is given below:

Algorithm 1 1. Find all partial costs of Step 1, i.e., compute the cost

$$C_1^t = \sum_{i=1}^n c_{1,i} \langle e_i^p, e_{1,t}^q \rangle^2 \text{ for each } e_{1,t}^q \in \Omega_1.$$

2. Apply transformation matrices and find total costs corresponding to Step 1 and Step 2. For each state of Step 2, find the minimum total cost of first two states and fix the value of $e_{1,*}^q$.

3. Repeat the above procedure for all steps.

3.1.3.2 EER Analysis for 2x2 Matrices

For two dimensional vectors, the distance is obtained as:

$$KL(P, Q) = \frac{1}{2} \left(\ln \frac{\lambda_1^q}{\lambda_1^p} + \ln \frac{\lambda_2^q}{\lambda_2^p} + \sum_{j=1}^2 \sum_{i=1}^2 \left(\frac{\lambda_i^p}{\lambda_i^q} k_{ij}^2 \right) \right) - 1 \quad (3.21)$$

For 2x2 covariance matrices, form of eigenvector matrices may be taken as in the following form.

$$M = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad (3.22)$$

where θ represents the angle between the first eigenvector and the vector $[1 \ 0]'$. Using this fact, and knowing $K = T^{-1}V$, the matrix K can be written as:

$$K = \begin{bmatrix} \cos(\theta_f) & -\sin(\theta_f) \\ \sin(\theta_f) & \cos(\theta_f) \end{bmatrix} \quad (3.23)$$

where $\theta_f = \theta_q - \theta_p$, θ_q and θ_p are the angles between the first eigenvector of P and Q , and the vector $[1 \ 0]'$.

Angle concept can be clarified by considering the Figure (3.2).

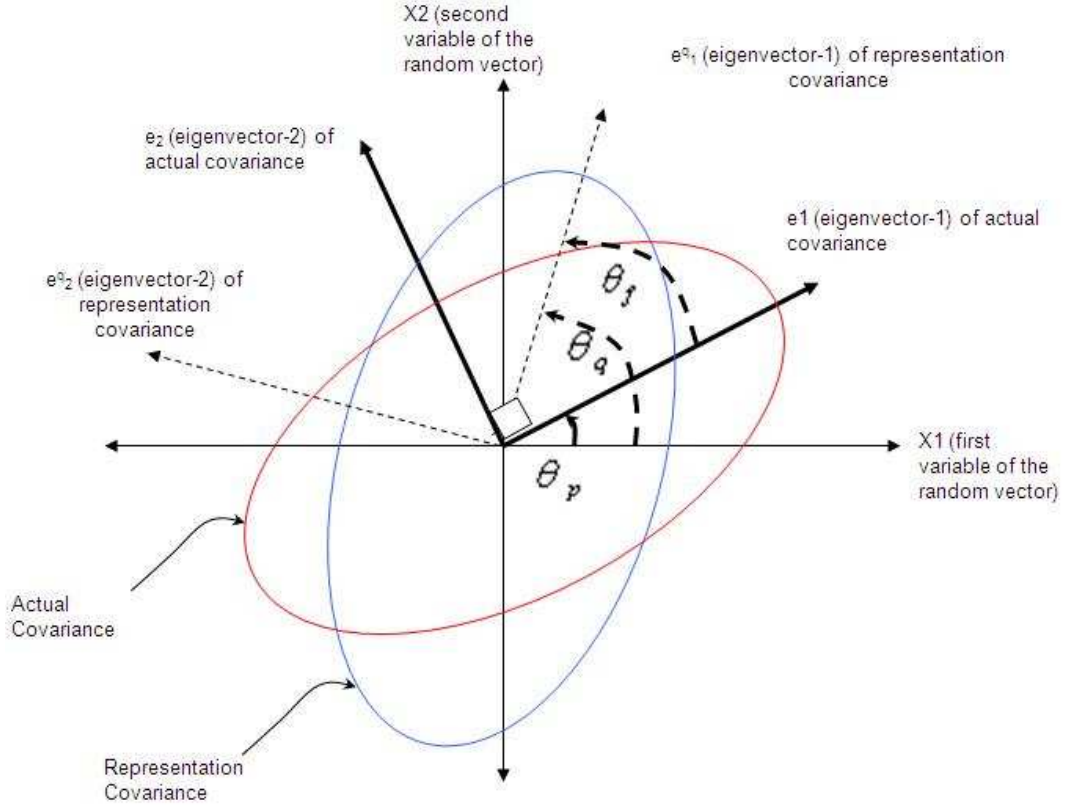


Figure 3.2: Schematic representation of the angle concept

Inserting K given in (3.23) into the equation 3.21 to obtain the distance in terms of eigenvalues and θ_f :

$$KL(P, Q) = \frac{1}{2} \left(\ln \frac{\lambda_1^q}{\lambda_1^p} + \ln \frac{\lambda_2^q}{\lambda_2^p} + \left(\frac{\lambda_1^p}{\lambda_1^q} + \frac{\lambda_2^p}{\lambda_2^q} \right) \cos^2 \theta_f + \left(\frac{\lambda_1^p}{\lambda_2^q} + \frac{\lambda_2^p}{\lambda_1^q} \right) \sin^2 \theta_f \right) - 1 \quad (3.24)$$

Now, assume that we define the errors between eigenvalues of P and Q as in the equations (3.25):

$$\lambda_{er1} = \lambda_1^p - \lambda_1^q \implies \lambda_1^q = \lambda_1^p - \lambda_{er1} \quad (3.25)$$

$$\lambda_{er2} = \lambda_2^p - \lambda_2^q \implies \lambda_2^q = \lambda_2^p - \lambda_{er2} \quad (3.26)$$

Inserting λ_1^q and λ_2^q into (3.24), we obtain:

$$\begin{aligned}
KL(\lambda_{er1}, \lambda_{er2}, \theta_f) = & \frac{1}{2} \left(\ln \frac{\lambda_1^p - \lambda_{er1}}{\lambda_1^p} + \ln \frac{\lambda_2^p - \lambda_{er2}}{\lambda_2^p} + \right. \\
& \left(\frac{\lambda_1^p}{\lambda_1^p - \lambda_{er1}} + \frac{\lambda_2^p}{\lambda_2^p - \lambda_{er2}} \right) \cos^2 \theta_f + \\
& \left. \left(\frac{\lambda_1^p}{\lambda_2^p - \lambda_{er2}} + \frac{\lambda_2^p}{\lambda_1^p - \lambda_{er1}} \right) \sin^2 \theta_f \right) - 1 \quad (3.27)
\end{aligned}$$

Distance KL reaches its minimum value at $[\lambda_{er1} \lambda_{er2} \theta_f]^T = \vec{0}$. Let us find second order Taylor series expansion around $[\lambda_{er1} \lambda_{er2} \theta_f]^T = \vec{0}$ and investigate the behavior of the distance regarding the eigenvalues, the angle difference and the eigenvalue errors.

$$\begin{aligned}
KL(\lambda_{er1}, \lambda_{er2}, \theta_f) = & \underbrace{KL(0, 0, 0)}_0 + \underbrace{\nabla KL(\lambda_{er1}, \lambda_{er2}, \theta_f)^T}_{=0 \text{ since } \vec{0} \text{ is optimal point}} \bigg|_{\vec{0}} \begin{bmatrix} \lambda_{er1} \\ \lambda_{er1} \\ \theta_f \end{bmatrix} + \\
& \begin{bmatrix} \lambda_{er1} \\ \lambda_{er1} \\ \theta_f \end{bmatrix}^T H(KL(\lambda_{er1}, \lambda_{er2}, \theta_f)) \bigg|_{\vec{0}} \begin{bmatrix} \lambda_{er1} \\ \lambda_{er1} \\ \theta_f \end{bmatrix} + \\
& h.o.t. \quad (3.28)
\end{aligned}$$

Ignoring h.o.t., approximate distance in terms of λ_{er1} , λ_{er2} , θ_f and eigenvalues is:

$$KL(\lambda_{er1}, \lambda_{er2}, \theta_f) \cong \begin{bmatrix} \lambda_{er1} \\ \lambda_{er1} \\ \theta_f \end{bmatrix}^T \begin{bmatrix} \frac{1}{(\lambda_1^p)^2} & 0 & 0 \\ 0 & \frac{1}{(\lambda_1^p)^2} & 0 \\ 0 & 0 & 2\left(\frac{\lambda_1^p}{\lambda_2^p} + \frac{\lambda_2^p}{\lambda_1^p} - 2\right)\theta_f^2 \end{bmatrix} \begin{bmatrix} \lambda_{er1} \\ \lambda_{er1} \\ \theta_f \end{bmatrix} \quad (3.29)$$

which can also be written as:

$$KL(\lambda_{er1}, \lambda_{er2}, \theta_f) \cong \frac{1}{(\lambda_1^p)^2} \lambda_{er1} + \frac{1}{(\lambda_2^p)^2} \lambda_{er2} + 2\left(\frac{\lambda_1^p}{\lambda_2^p} + \frac{\lambda_2^p}{\lambda_1^p} - 2\right)\theta_f^2 \quad (3.30)$$

Now, we can derive some important conclusions by focusing on the equation (3.29):

1. For small eigenvalues, λ_{er1} and λ_{er2} has high weighting because they are multiplied by the inverse of square of eigenvalues. So, small eigenvalues should be quantized more precisely so that distance can be reduced.
2. In the equation, θ_f has a multiplication factor which is a function of the ratio of eigenvalues and but not the individual eigenvalues. For this reason, while quantizing the angle θ_p , eigenvalue ratio should have an important role. For example, for $\frac{\lambda_1^p}{\lambda_2^p} \approx 1$, θ_f has not any effect on the distance so it may have any value. However, when $\frac{\lambda_1^p}{\lambda_2^p}$ takes high values, we should be careful while quantizing θ_p so that we may reduce the θ_f value as much as possible. Actually, this idea above sounds sensible, because of the fact that when the ratio is around one, the cross-section of pdf of the random vector takes the circle-like shape. So, whatever θ_p is, the shape does not change and for this reason θ_p has no importance.

3.2 Other Approximations of Covariance Matrices

Surely, we can offer some other representation matrices for the covariance matrix where these representations may be in any structure. However the KL analysis may not be easy for some of them.

Some of these representations may be given as below:

3.2.1 Whole Matrix Representation (WMR)

In this representation, covariance matrices are taken with no modification, i.e., \mathbf{Q} is a 6x6 matrix which is a represented version of the actual covariance. By inspection, optimal q_{ij} should be equal to p_{ij} , i.e., $q_{ij}^* = p_{ij} \quad \forall i, j \in [1, N]$. Still we can use KL to analyze the quantized covariance matrix. However, it is not easy to find an analytical procedure of computation of a good codebook for it.

3.2.2 Block Diagonal Matrix Representation (BDMR)

This representation is based on representing the covariance matrix by 2x2 block diagonal matrices. Only 2x2 diagonal blocks of P are taken and quantization is applied to them, i.e., modified covariance matrix to be represented is $\hat{P} = \text{blkdiag}(P_1, \dots, P_n)$ and the represented version of it is $Q = \text{diag}(B_1, \dots, B_n)$ where P_i 's and B_i 's are 2x2 matrices and P_i 's are diagonal blocks of the actual covariance.

Representing P by means of its 2x2 diagonal blocks is proposed in this study because of the simple relationship between the eigenvalues and the eigenvectors of 2x2 symmetric matrix. This relationship is elaborated in the following sections. Our study shows that selection of diagonal block is extremely important since "0" (zero) of the sub-blocks assumes uncorrelatedness between the corresponding variables. Whenever this assumption is not satisfied by a given problem, the quality of the representation (i.e., 2x2 diagonal blocks) degrades dramatically.

Surely, for $n=1$, our BDMR matrix gets same with the DMR matrix and with the WMR matrix for $n=N$.

Also note that we assume the dimension of the covariance N is an even number.

CHAPTER 4

EXPERIMENTS AND ANALYSIS

To analyze and compare the representation methods proposed in the previous chapter, a simulation environment has been generated. This environment is almost the same as the one given in [2] and [3]. Similar to these references, we will use a track fusion system that fuses track information generated by two radars. Covariance matrices generated by the filters at the radars are quantized and sent to the fusion node. At the fusion node, these quantized matrices are used for track fusion. Yet, quantization of state estimate is not the main concern of this study. So, state estimate is assumed to be sent without quantization.

The quantization of the covariance matrix is made for different representations of it which are given in the previous chapter with those additionally proposed in this chapter. For these representations, some performance criteria is generated to compare different representations.

4.1 Simulation Model

As a model, we have chosen a fusion system that has two monopulse radars sending track information to a center. State estimates and the corresponding error covariance matrices are taken as the track information. We assume that the state estimate of the target is sent without any quantization process. However, the covariance matrix should be quantized since we have limited bandwidth for sending the covariance. The experimental setup of the track

fusion system should use the covariance matrices of the tracks found by the radar filters to make a proper fusion at the fusion center. Furthermore, because of the nonlinear nature of the radar measurements, these covariances can not converge to a fixed value. This system model is very appropriate for our purpose of analyzing the effect of covariance approximation on track fusion systems since this system generates time-varying covariance matrix information in the system that should be interchanged between the nodes of the system for optimal sensor fusion. Configuration of the experimental setup is given in Figure (4.1).

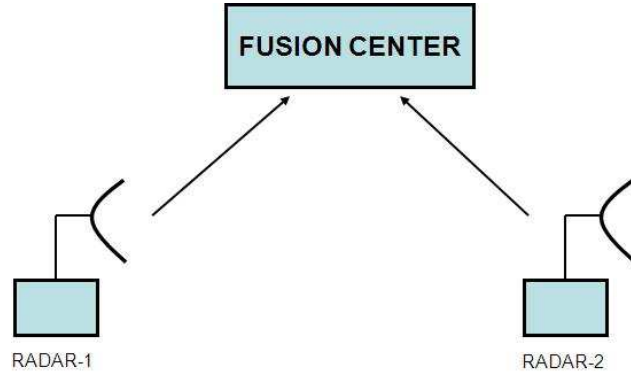


Figure 4.1: Simulation System Model

In the model, two monopulse radars detect targets, produce measurements from them and generate tracks from these measurements. They generate the state estimation and the corresponding filter output error covariance matrix for the target and send this track information to the fusion center. We as-

sume that only digital communication is possible between the radars and the fusion center. We also assume bandwidth of the communication channel is limited. So, the filter output covariance information is sent to the fusion center after a quantization process. In order to investigate the effect of covariance representation, we will assume that the quantization is only applied to the generated covariance matrix, but not to the estimated state. We assume that these radars work synchronously and data association at the fusion center is made perfectly.

In the next section, more detailed information on this system will be given. Target motion and measurement models, radar tracker filter, fusion scheme and the criteria for performance comparison will be discussed as main topics.

4.1.1 Target Motion and Measurement Models

Target motion is modeled as a constant velocity motion in 3D space. Target motion equations can be written as:

$$X_{k+1} = AX_k + v_k \quad (4.1)$$

$$Y_k = h(X_k) + w_k \quad (4.2)$$

where X_k represents the true state of the target at time k and Y_k represents the corresponding sensor measurement. The matrix A and the function $h(\cdot)$ are known, and assumed to represent an observable system. Both the process noise v_k and the measurement noise w_k are assumed to be white, zero-mean and Gaussian and they are assumed to be independent from each other. Their covariances are taken as $E\{v_k v_k^T\} = Q_k$ and $E\{w_k w_k^T\} = R_k$.

Assume that the time interval is T and $X_k = [x_k \ v_{x,k} \ y_k \ v_{y,k} \ z_k \ v_{z,k}]'$ where x_k, y_k, z_k denote the position and $v_{x,k}, v_{y,k}$ and $v_{z,k}$ denote the corresponding velocities, respectively. Y_k^i equals to $[r_k^i \ b_k^i \ e_k^i]'$ where r, b and e represents range, bearing and elevation angles for the i^{th} radar, respectively.

The matrix A is taken as the following one:

$$A = blkdiag\left(\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}\right) \quad (4.3)$$

and $Q_k = Q$ is taken as:

$$Q = blkdiag\left(\begin{bmatrix} \frac{T^3}{3}q_x & \frac{T^2}{2}q_x \\ \frac{T^2}{2}q_x & Tq_x \end{bmatrix}, \begin{bmatrix} \frac{T^3}{3}q_y & \frac{T^2}{2}q_y \\ \frac{T^2}{2}q_y & Tq_y \end{bmatrix}, \begin{bmatrix} \frac{T^3}{3}q_z & \frac{T^2}{2}q_z \\ \frac{T^2}{2}q_z & Tq_z \end{bmatrix}\right) \quad (4.4)$$

where q_x , q_y , and q_z represent process noise variances in x, y, and z axis, respectively. Our measurement model is:

$$r_k^i = \sqrt{(x_k - x_r^i)^2 + (y_k - y_r^i)^2 + (z_k - z_r^i)^2} + w_r(k) \quad (4.5)$$

$$b_k^i = \tan^{-1}\left(\frac{y_k - y_r^i}{x_k - x_r^i}\right) + w_b(k) \quad (4.6)$$

$$e_k^i = \tan^{-1}\left(\frac{z_k - z_r^i}{\sqrt{(x_k - x_r^i)^2 + (y_k - y_r^i)^2}}\right) + w_e(k) \quad (4.7)$$

where $[x_r^i \ y_r^i \ z_r^i]'$ denotes the position of the i^{th} radar in cartesian coordinates.

In a monopulse radar, calculation of the cross-covariance matrix of $w_b(k)$ and $w_e(k)$ is given in [18] and [19] as the following equation:

$$Cov\left\{\begin{matrix} w_b(k) \\ w_e(k) \end{matrix} \middle| \mathfrak{R}_0\right\} = \frac{1}{\mathfrak{R}_0} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\mathfrak{R}}{\mathfrak{R} + 1} \begin{pmatrix} \eta_b \\ \eta_e \end{pmatrix} \begin{pmatrix} \eta_b \\ \eta_e \end{pmatrix}' \right) \quad (4.8)$$

where \mathfrak{R} and \mathfrak{R}_0 are expected and observed sum channel signal-to-noise ratios, respectively. η_b and η_e are the electronic off-beam angles (normalized to beam-width) in bearing and elevation.

According to the target motion model given at the equation (4.1), true trajectory of the target is generated randomly. Using the true state, the measurements of the radars for this target are obtained using the measurement model given above. The algorithm for true target trajectory and corresponding radar measurement generation is given below:

1. Firstly, choose an initial random state for the target in a 6 dimensional cube. We assume that the initial state is uniformly distributed within this cube for any target, so that initial state of the target $X_0 = [x_0 \ v_{x,0} \ y_0 \ v_{y,0} \ z_0 \ v_{z,0}]'$ is chosen as: $x_0, y_0 \in \text{unif}[-40000, 40000]$, $v_{x,0}, v_{y,0} \in \text{unif}[-200, 200]$, $z_0 \in \text{unif}[0, 5000]$ and $v_{z,0} \in \text{unif}[0, 5]$. Then, start generating the true state by the equation (4.1) where v_k is taken by:

$$v_k = \text{sqr}t(Q) * t_k \quad (4.9)$$

where t_k is 6x1 gaussian random vector with covariance I_6 . q_x, q_y and q_z are taken as 900, 900, 1 (m^2), respectively. Note that these values are selected in order to generate a simulator that is relatively realistic. T is taken as 4 seconds.

2. After obtaining the true trajectory, range, bearing and elevation measurements for both of the radars are calculated by the equations (4.5),(4.6) and (4.7), respectively. Sensor locations are taken as $[x_r^1 \ y_r^1 \ z_r^1]' = [0 \ 0 \ 0]'$ and $[x_r^2 \ y_r^2 \ z_r^2]' = [40000 \ 0 \ 0]'$. $w_r(k)$ is taken as Gaussian with variance 100. Bearing and elevation noises have been found by the following equation:

$$\begin{bmatrix} w_b(k) \\ w_e(k) \end{bmatrix} = \text{sqr}t\left(\text{Cov}\{w_b(k), w_e(k)\}\right) * t_k \quad (4.10)$$

where t_k is a Gaussian random variable with variance 1 and $\text{Cov}\{w_b(k), w_e(k)\}$ is calculated by the equation (4.8). $\mathfrak{R}, \mathfrak{R}_0, \eta_b$ and η_e have taken the values 34, 3000, 2 and 2, respectively.

At the end, we have the true state of the target and the corresponding radar measurements.

4.1.2 Radar Tracker

Previously, we have obtained the true state of the target and the corresponding radar measurements. At this point, we need local tracker to obtain the state

of the target from the measurements. Since we have nonlinear radar measurement equation in our model, we used Extended-Kalman Filtering (EKF) algorithm for the state estimation. Flowchart of EKF algorithm is given in Figure (4.2):

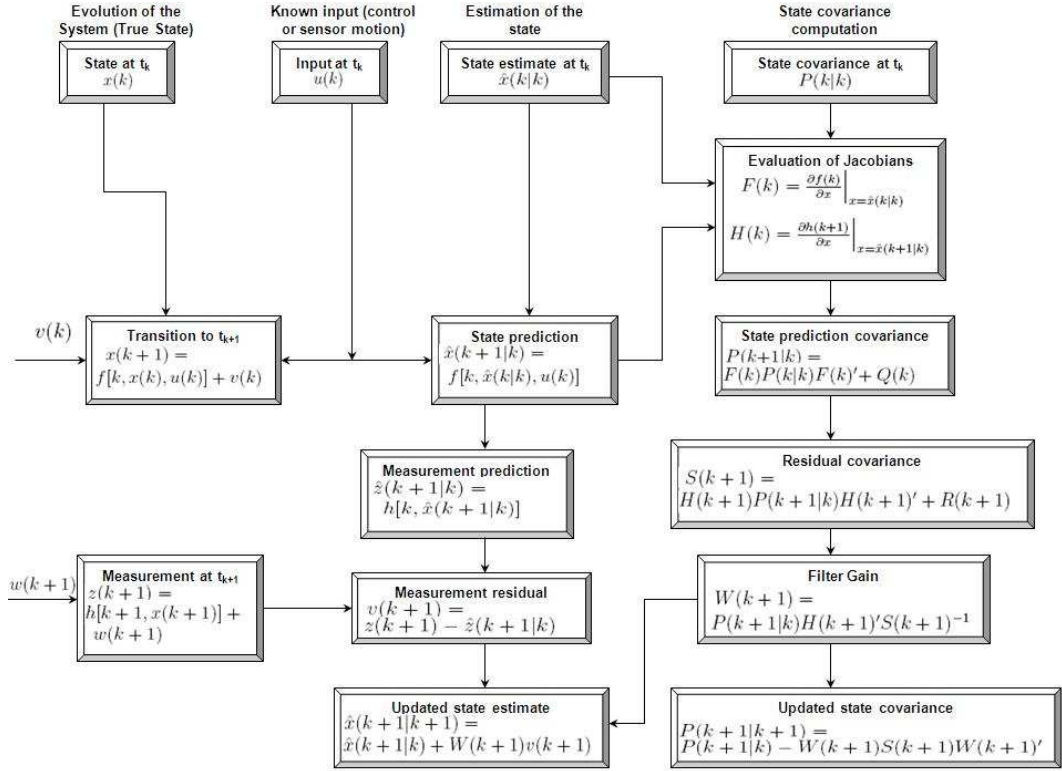


Figure 4.2: Flowchart of EKF Algorithm (One Cycle),[1]

Obviously, for our model, i.e., $u(k) = 0$ and the process is linear. We have evaluated the Jacobian matrix for $h(X_k)$ and used this Jacobian in our

algorithm. All the parameters like the process noise covariance or the time interval are chosen as strictly the same as with those of the target trajectory and measurement generation process.

4.1.3 Track-to-Track Fusion

In the literature, there are several fusion methods used at multisensor target tracking systems. Distributed and centralized fusion are some basic methods that are preferred by common fusion systems. In this thesis, we prefer to use centralized fusion architecture in which sensor track information is sent to a center node where the information is fused. There are several methods to solve track fusion problem such as weighted covariance fusion (WCF)[1], information matrix fusion [20], best linear unbiased estimate fusion [21], and tracklet fusion [22]. In our study, we prefer to use WCF method for the experiments in a Sensor-to-Sensor fusion scheme [23]. Schematic representation of Sensor-to-Sensor Track fusion structure is given in the Figure (4.3).

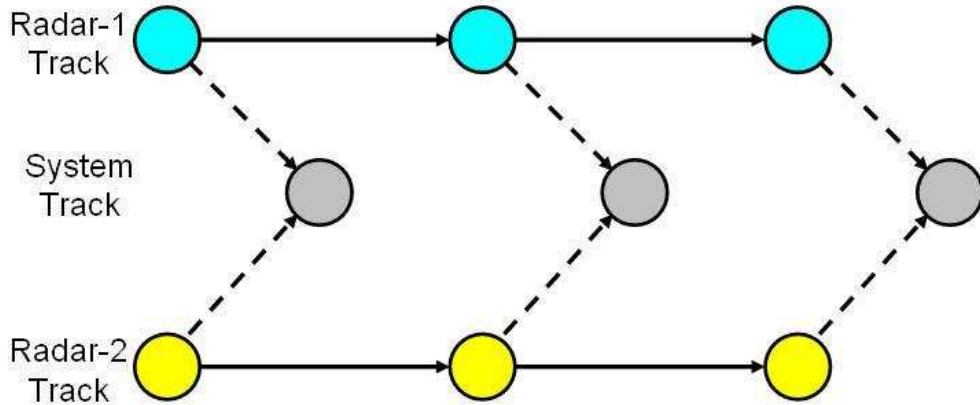


Figure 4.3: Sensor-to-Sensor track Fusion Scheme

In fusion systems, the state estimates of different sensors have the common process noise which should be removed for optimal track fusion. WCF method considers this dependency by the usage of the cross-covariance of the corresponding estimation errors. So, the state fusion equation of WCF is given as:

$$\hat{x}^f = \hat{x}^i + (P^i - P^{ij})(P^i + P^j - P^{ij} - P^{ji})^{-1}(\hat{x}^i - \hat{x}^j) \quad (4.11)$$

and the corresponding covariance is:

$$P^f = P^i - (P^i - P^{ij})(P^i + P^j - P^{ij} - P^{ji})^{-1}(P^i - P^{ji}) \quad (4.12)$$

where \hat{x}^i is the state estimate of sensor i and P^i is the corresponding state estimate error covariance. P^{ij} is the cross-covariance between the errors of sensors i and j . There are some techniques [1] to calculate the cross-covariances in the literature. However, in our study, we will use none of them because we ignore the effect of correlated process noise and take $P^{ij} = P^{ji} = 0$.

4.2 Performance Evaluation

After mentioning about fusion method, the performance criteria for comparing proposed representations of the covariance matrices will be discussed in this section. For a fused state of a target which has used true covariance matrices - P^1 and P^2 -, true average mean-square error (MSE) is calculated first. This MSE is considered as "**true error**" of the fusion process. Secondly, for the fusion system output which uses the approximated version (Q) of the actual covariance (P), the average MSE is recalculated. This MSE is called as "**representation error**". Normalizing this error by the true error, we obtain a ratio which gives an idea about the degradation of the usage of the approximated covariance matrices.

True MSE calculation for a track with frame length N is performed only for the last state estimate, i.e.,

$$MSE_{true} = \|\hat{x}(true)_N^f - x_N\| \quad (4.13)$$

where x_N is the true state at time N and $\hat{x}(true)_N^f$ is the fused state which is calculated by the equation (4.11) assuming there is no effect of correlated process noise.

$$\hat{x}(true)_k^f = \hat{x}_k^1 + P_k^1(P_k^1 + P_k^2)^{-1}(\hat{x}_k^1 - \hat{x}_k^2) \quad (4.14)$$

In the equation above, $\hat{x}(true)_k^f$ is the fused state, \hat{x}_k^1 is the state estimate of radar-1 and P_k^1 is the corresponding true covariance matrix, \hat{x}_k^2 is the state estimate of radar-2 and P_k^2 is the corresponding true covariance matrix and k denotes time. MSE calculation for true covariance matrices is summarized in the Figure (4.4). The reason for selecting only one element, i.e., the last element of each track, is to eliminate the bias in the statistics. In the opposite end, one can use only one track to get the desired statistics, but this may be biased due to some special position of the track with respect to the radar.

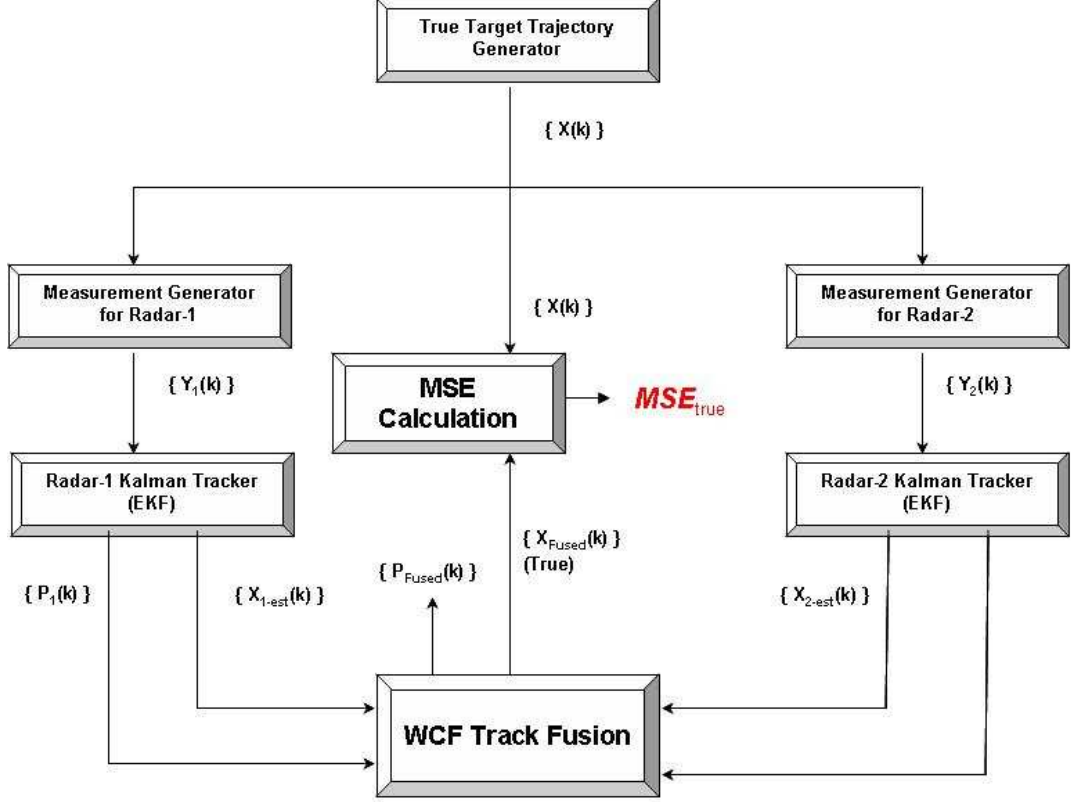


Figure 4.4: MSE calculation process for the true covariance matrices

Similarly, calculation of MSE which belongs to the represented covariance matrices is given by the following equations:

$$MSE_{repr} = \|\hat{x}(repr)_N^f - x_N\| \quad (4.15)$$

where x_N is the true state at time N and $\hat{x}(repr)_k^f$ is calculated by the equation (4.11).

$$\hat{x}(repr)_k^f = \hat{x}_k^1 + Q_k^1(Q_k^1 + Q_k^2)^{-1}(\hat{x}_k^1 - \hat{x}_k^2) \quad (4.16)$$

In the equation (4.16), $\hat{x}(repr)_k^f$ is the fused state, \hat{x}_k^1 is the state estimate of radar-1 and Q_k^1 is the representation of the corresponding true covariance

matrix, \hat{x}_k^2 is the state estimate of radar-2 and Q_k^2 is the representation of the corresponding true covariance matrix. The Figure (4.5) presents the process of the MSE calculation for the representation matrices.

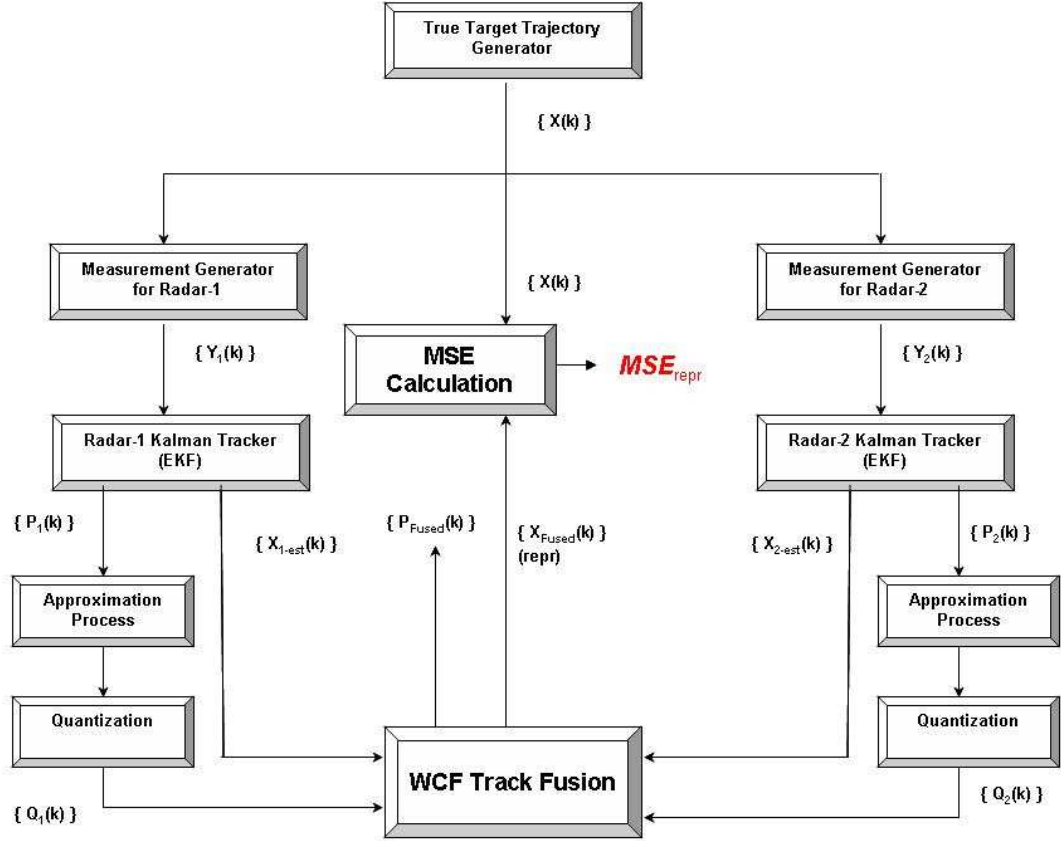


Figure 4.5: MSE calculation process for the represented covariance matrices

It is important to note that the measurements and the generated tracks are the same for both of the MSE calculation procedures. It is also important to mention that we perform the MSE calculation of a track for only its last

estimate because we assume the covariance information of frames do not change rapidly, which may give biased MSE's.

Finally, defining normalized MSE (MSE_{norm}), we reach a measure of performance of the representation matrix for one track.

$$MSE_{norm} = \frac{MSE_{repr}}{MSE_{true}} \quad (4.17)$$

We will obtain MSE_{norm} for the different representations of covariance matrices so that we will be able to compare them from this point of view.

4.3 Representations of Covariances Specific to Target Tracking and Their Quantization

In this section, we will first introduce two possible forms of block diagonal representations of covariance matrices, then we will describe at least one method to quantize all of the representations. All quantizations use Kullback-Leibler as a distance measure between two covariance matrices. Quantization levels are changed from 1 bit to 10 bits.

4.3.1 Representation Methods

Previously we have defined some approximation methods for a covariance matrix in Chapter (3). In this chapter, we will again discuss about these approximation methods as well as their possible quantizations. These quantization processes are realized for both of the radars in the system.

4.3.1.1 Quantization Alfa Representation (AR)

Previously, we have defined AR in Chapter (3) and found that if we approximate P as αI , optimal α is found as the average of the eigenvalues of the covariance matrix P . So, the approximation is obtained by calculating average of the eigenvalues of the covariance matrix. Then, this average is quantized

by uniform quantization method before sending to the fusion center. The procedure is briefly described in the Figure (4.6).

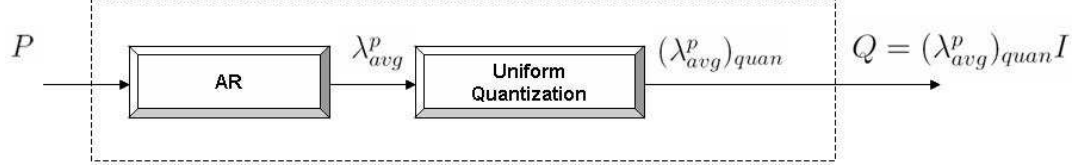


Figure 4.6: AR approximation process

To determine the distribution of the eigenvalue average, we obtained the histogram of this variable using training data which we will discuss later. Figure (4.7) presents the corresponding histogram and quantization levels for 5 bits are shown by "x"'s on the same figure.

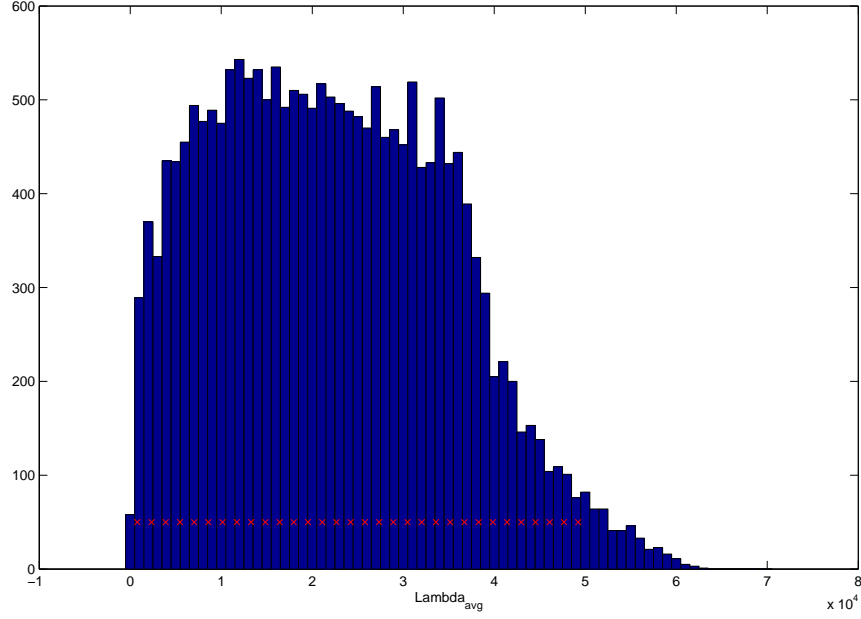


Figure 4.7: Histogram of eigenvalue average and uniform quantization levels for 5 bits.

As the histogram in Figure (4.7) shows, uniform quantization in the range (5000-35000) is quite reasonable. Some more effort can be spend for a better quantization outside this range. However, since the main subject of the thesis is to compare representations rather than getting the optimal quantizer, a simple uniform quantization is selected. Considering the histogram, we have taken the quantization interval as $[0, 50000]$ since the probability of λ_{avg} being grater than 50000 is shown to be small.

4.3.1.2 Diagonal Matrix Representation (DMR) and Its Quantization

Similarly, DMR method is also analyzed in Chapter (3). DMR is defined as representing the covariance matrix P by a diagonal matrix Q such that

$Q = \text{diag}(d_1, \dots, d_n)$. Optimal d_k 's are found as the $(k, k)^{th}$ element of P which is defined as p_{kk} . So, the approximation process is performed on the diagonal entries of the covariance matrix, then K-means algorithm is used for the quantization. KL K-means quantization method is performed to the vectors F where $F = [\{p_{ii}\}_{i=1}^6]$.

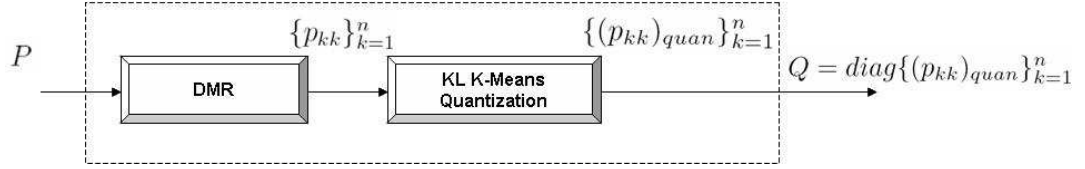


Figure 4.8: DMR approximation process

4.3.1.3 Block Diagonal Matrix Representation (BDMR) and Its Quantization

This representation is based on representing the covariance matrix by 2x2 block diagonal matrices. The 2x2 blocks themselves are either represented by a 3 dimensional vector that contains the elements of the block matrix which is Whole Matrix Representation (WMR) or by $(\lambda_1, \lambda_2, \theta)$ triple named as Eigenvalue-Eigenvector Representation for 2x2 matrices (EER-2x2).

So, there are some representations that might be suitable for 2x2 matrix approximations. The first one is Whole Matrix Representation (WMR), which will also be discussed for 6x6 dimensional matrix case, and the other is Eigenvalue-Eigenvector Representation for 2x2 matrices (EER-2x2).

Note that our aim for this proposal is investigation of the performance of BDMR and comparing this method with the others.

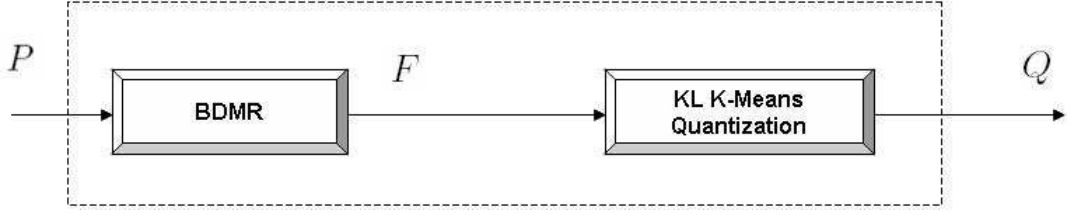


Figure 4.9: BDMR approximation process

1. BDMR-WMR method is the representation of all of the elements of block diagonal matrices and constructing a vector containing these elements. Define the vector $F = [\{p_{ij}\}_{i,j=1}^2 \quad \{p_{ij}\}_{i,j=3}^4 \quad \{p_{ij}\}_{i,j=5}^6]$ for $i \leq j$ whose dimension is 1x9. Here, we take the advantage of the symmetry property of the covariance matrix. Then apply vector quantization for this vector to generate the matrix Q . Note that positive definiteness of all the matrices in the codebook must be checked.
2. BDMR-EER-2x2 In Chapter (3), we have emphasized that a 2x2 matrix can be represented by its eigenvalues and its eigenvector angle. So define $B_1 = P(1 : 2, 1 : 2)$, $B_2 = P(3 : 4, 3 : 4)$ and $B_3 = P(5 : 6, 5 : 6)$. Find eigenvector angles and eigenvalues of these matrices and generate the vector to be quantized. Define this vector as

$F = [\{\lambda_1^{B_i}\}_{i=1}^3 \quad \{\lambda_2^{B_i}\}_{i=1}^3 \quad \{\theta^{B_i}\}_{i=1}^3]$ where θ^{B_i} denotes the eigenvector angle of the matrix B_i and $\lambda_1^{B_i}$ and $\lambda_2^{B_i}$ corresponds to the first and second eigenvector of the matrix B_i , respectively. Then apply KL K-means quantization method to find the codebook.

The state of the system can be described in various ways, i.e., the order of the variables within the state can be changed. So that, we can have the different covariances matrices containing the same elements in different locations. We investigate this to discover importance of the the correlation information between different vector variables. For BDMR, we define two types of the state vector. The first one is the traditional one- $X_k = [x_k \quad v_{x,k} \quad y_k \quad v_{y,k} \quad z_k \quad v_{z,k}]'$ (**type-I**) and the second one is $X_k = [x_k \quad y_k \quad v_{x,k} \quad v_{y,k} \quad z_k \quad v_{z,k}]'$ (**type-II**).

Notice that, the covariance matrix for these different states contains the same elements but in different locations, e.g., $p_{3,3}$ in type-I is the same as the element $p_{2,2}$ in type-II or $p_{2,1}$ in type-I is same as $p_{3,1}$ etc.

4.3.1.4 Whole Matrix Representation (WMR)

In this representation, covariance matrices are taken with no modification and quantized using K-means algorithm and KL as distance measure.

4.3.2 KL K-Means Clustering Algorithm

We use KL K-means clustering technique to generate a codebook of the covariance matrices. As indicated in [3], K-means classification method is based on partitioning the training data into K regions. The centroids of these clusters are then selected as their representatives in the codebook. Then, coding of a data is done by selecting the closest code vector to that data. This quantization technique is composed of two steps: "codebook generation" and "quantization process".

4.3.2.1 Codebook Generation

Codebook generation requires a large amount of "typical" data. We have obtained this data by making Monte-Carlo runs by using our models. So, random track generation procedure is done for Radar-1 as given in "Target and Measurement Model" section (4.1.1). 32768 tracks have been generated and the covariance of the 50th frame of these tracks is stored to avoid transient effects of Kalman filtering. These covariances are stored in the training set which is required to generate the codebooks. In the Figure (4.10), some of the generated tracks and cross-sections of density functions at 50th frame is illustrated.

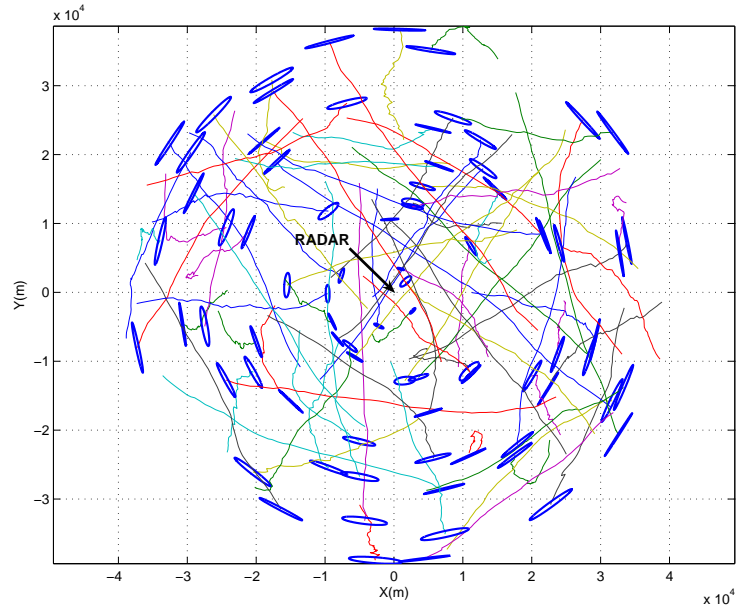


Figure 4.10: Generated Tracks for Training and Cross-Sections of the Corresponding Covariances

In our work, we followed a basic K-means clustering method to generate the codebooks for each representation methods except AR. Instead of usage of traditional L2 norm distance, we preferred using KL distance measure while codebook generation. Before explaining the codebook generation algorithm, some explanations on notation should be made:

TDS : Training data set which includes all the covariance matrices generated.

N : number of covariance matrices in TDS.

CB : codebook set

n : number of elements in the codebook.

W : weighting vector which provides flexibility in selectivity of vector elements.

The algorithm for generating "n clusters" is as follows:

1. Select n random initial vectors, $C_n \in CB$ and put them into the codebook. (Figure (4.11))
2. for all $P_{i=1}^N \in TDS$
 - (a) find the closest cluster in the codebook in KL sense,i.e., Find C_m , $m \in 1 : n$ which minimizes $KL(P_i, C_m)$ where $i = 1, \dots, N$.
 - (b) update C_m as: $C_m^{new} = C_m + W \cdot (P_i - C_m)$ where W is the weighting vector and KL represents the Kullback-Leibler distance between the covariances.

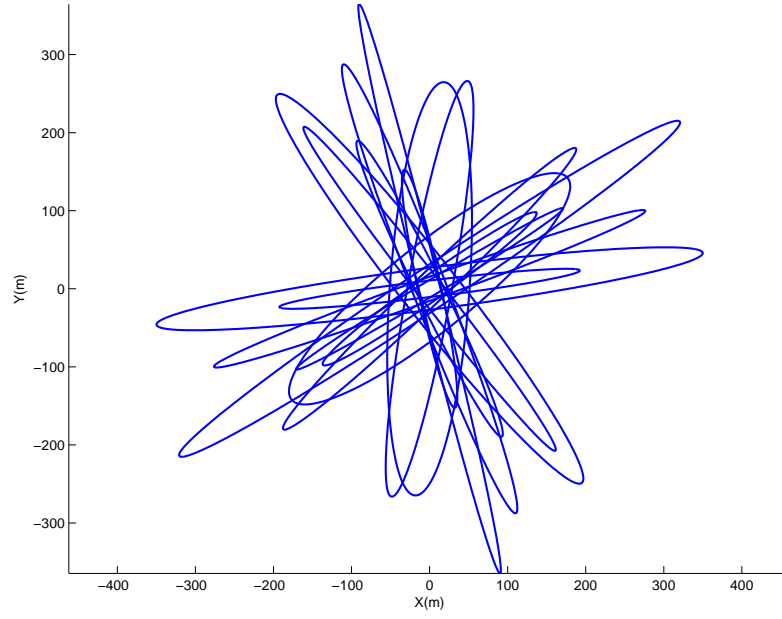


Figure 4.11: Cross-Sections of initial selected covariances for WMR method

After the selection of the initial covariances, 4-bit KL K-means quantizer generates the covariances as given in Figure 4.12. Then, these covariances are added to the corresponding codebook.

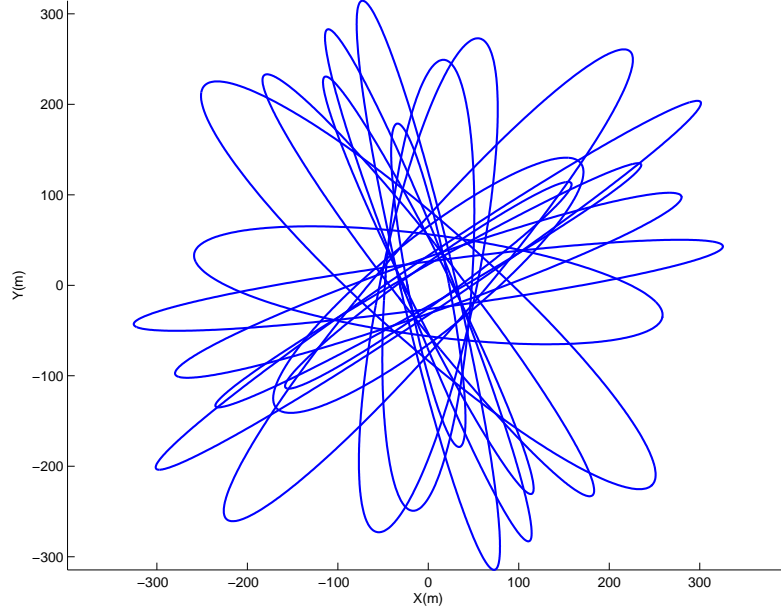


Figure 4.12: Cross-Sections of final codebook covariances of WMR method for 4 bits.

This codebook generation algorithm is performed for each representation method and for different number of bits.

4.3.2.2 Quantization

After giving the explanation about codebook generation process, quantization of a given covariance matrix will be explained in this section. In this process, after getting the covariance matrix required to be quantized, we need to decide the best codebook element that can represent it. For this purpose, we again benefit KL as a distance measure and produced the following algorithm:

1. take the output covariance (P) of the radar tracker

2. find the closest codebook element in KL sense.,i.e., find C_m , $m \in 1 : n$ which minimizes $KL(P, C_m)$ and obtain the representation matrix.

At both of the radar sites, after generating the covariance matrix, the best matrix within the codebook is chosen by the criteria of KL distance minimization that gives out the matrix Q . So, we obtain an approximation to our actual covariance matrix. After determining the approximations for both of the radar, we calculate average MSE_{norm} at the fusion center where the approximate matrices has been used.

4.3.3 Results

Before giving the MSE_{norm} for each representation methods for various number of bits, an important experimental result should be presented. This experiment is based on using the approximate covariances **without any quantization**. However, this experiment is only applicable for AR, DMR, BDMR1 and BDMR2 and **it reveals the correlation between the fusion performance and the KL distance**. For 1000 tracks, MSE_{norm} and the average KL distances between the actual covariances and the represented ones at both radar sites are calculated.

Table 4.1: Performance of Representation Matrices without Quantization.

Method	Average MSE_{norm}	Average KL at Radar-1	Average KL at Radar-2
AR	2.91	10.13	10.14
DMR	2.63	3.24	3.28
BDMR-I	2.63	2.46	2.49
BDMR-II	1.17	1.16	1.16

Table 4.1 shows us there exists a correlation between the fusion system performance and Kullback Leibler criteria. This correlation gives the intuition that one can get better performance if the representation matrices are close to the actual covariances in the KL distance sense.

Finally, **using the representation matrices after quantization**, the performance of each representation method has been obtained. The Figure

(4.13) shows their performances for various number of bits.

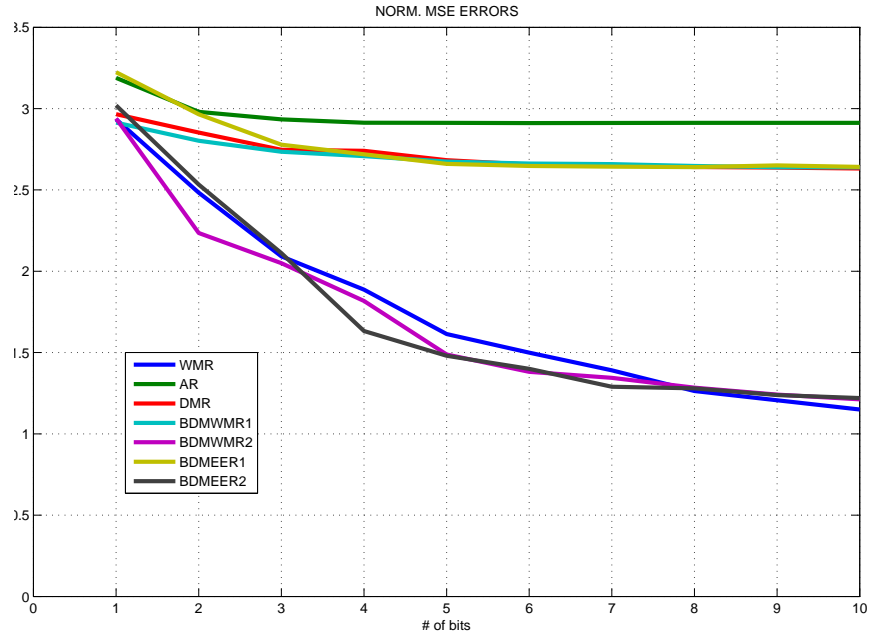


Figure 4.13: Performances of Representation Matrices

As can be seen in the Figure (4.13), DMR, BDMWMR1 and BDMEER1 has almost the same performances for each number of available bits and AR has slightly higher MSE_{norm} values than these methods. In addition, all these methods show almost constant performance for higher bit rates. The good ones WMR, BDMWMR2 and BDMEER2 have decreasing MSE_{norm} characteristics with increasing number of available bits. Also note that these three

representations have almost the same fusion performances. This shows us the importance of cross-correlation terms between X and Y components and proves that the cross-correlation term should not be underestimated during quantization process.

For better understanding of performances of different representations, cross-section of both actual and quantized covariances are demonstrated in the Figures (4.14)-(4.19).

All the Figures (4.14)-(4.19) have the following common properties:

1. Quantization level is 4-bits.
2. There are two plots belonging to two different targets and both targets have two tracks that is generated by two radars.
3. All quantities related with radar-1 are in blue color while they are red for radar-2.
4. Ellipses represent covariance matrices and dashed-line ellipses represent their quantized versions.
5. "X"'s represent the estimated states of the radars.
6. Green square with the letter "P" shows the true fused state with true covariance and cyan square with the letter "Q" shows the fused state with quantized covariance matrix.

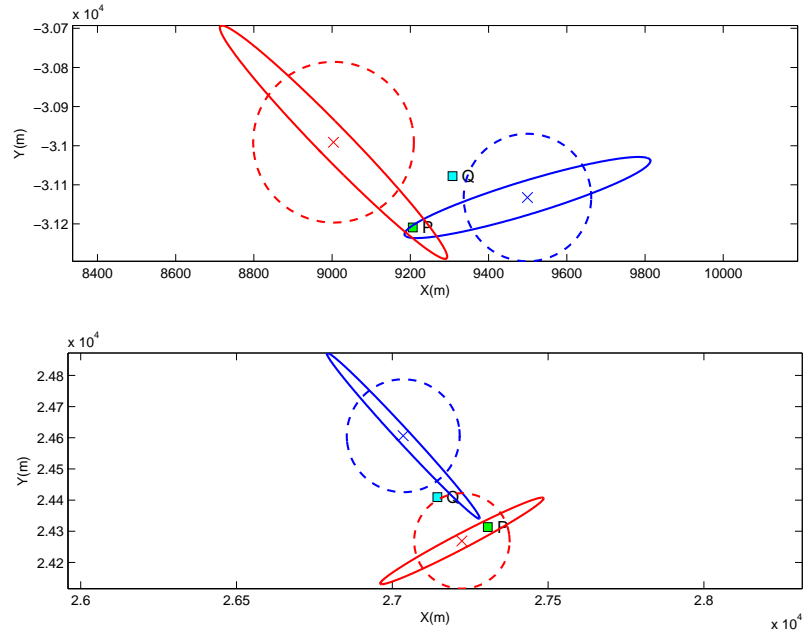


Figure 4.14: AR method examples

Figure (4.14) shows four snapshots of AR-uniform quantization result. In all of them, the representation of the covariance matrix is a circle since it is a diagonal matrix having the same value on its diagonals.

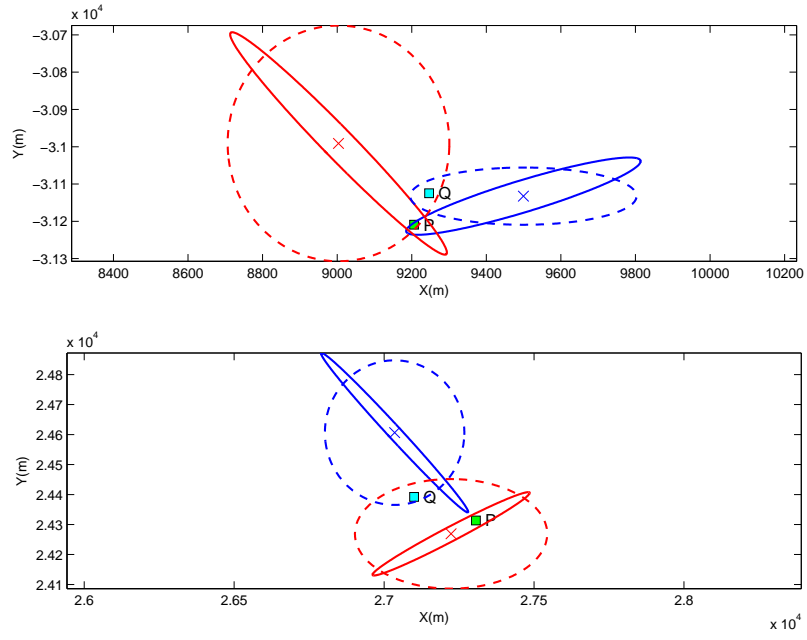


Figure 4.15: DMR method examples

In the DMR model, covariance matrix is diagonal but with different diagonal elements. As a result, covariance matrices are represented by ellipses instead of circles. However, since eigenvectors of these matrix are along x and y axes, its representation is not oblique. This is obviously a better representation than AR but still not good enough as high MSE_{norm} shows.

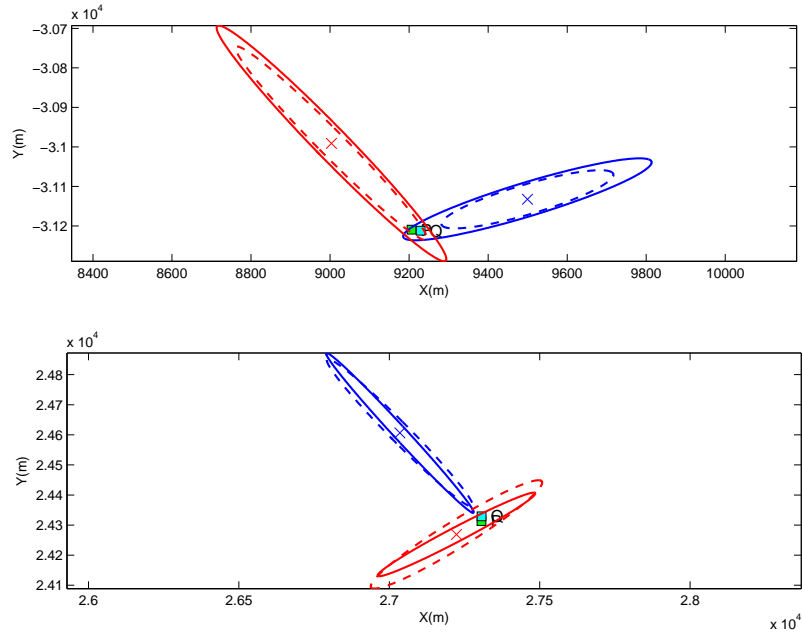


Figure 4.16: BDMWMR2 method examples

In this method, the 6x6 covariance matrix is approximated by 2x2 diagonal blocks. The first block corresponds to the position (x,y). The actual covariance matrices show high correlation between x and y components and the quantization scheme is well adopted to this scheme as Figure (4.16) shows. The good performance of the method indicates that the correlation between the blocks is small.

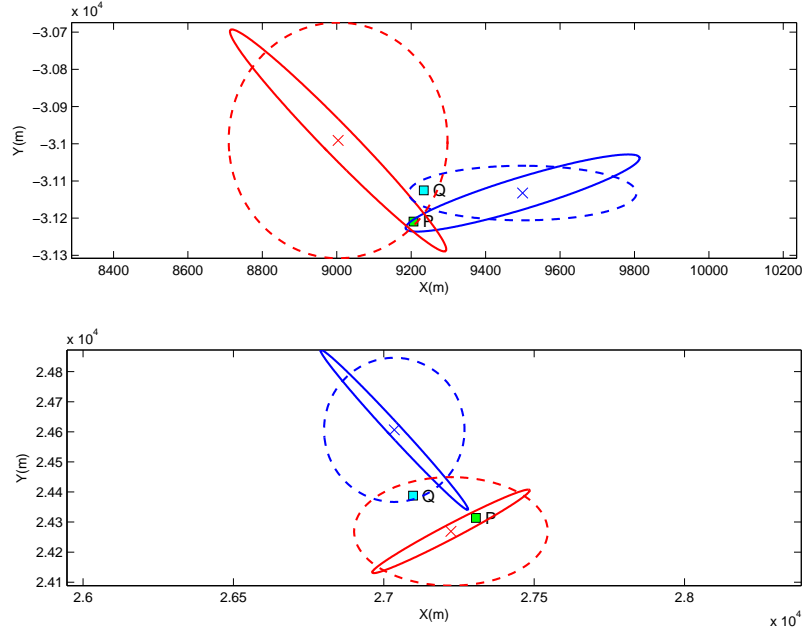


Figure 4.17: BDMWMR1 method examples

BDMWMR1 and BDMWMR2 are same in using quantization and representation methods except that variables are interchanged and non-zero correlations between (x, V_x) and/or (y, V_y) and/or (z, V_z) are allowed while forcing the vectors to be uncorrelated by selecting the corresponding appropriate block diagonalization. Figure (4.17) indicates that this assumption is not satisfactory due to the huge difference between the actual covariances and their representations. Furthermore, it can be clearly seen from the MSE_{norm} results, there is almost no correlation between the elements of these vectors, i.e., $(x$ and $V_x)$ and $(y$ and $V_y)$. Briefly, comparison of performances of BDMWMR1 and BDMWMR2 clearly shows that there is a strong correlation between the position elements $(x$ and $y)$ while there is nearly no correlation between the positions and velocities $(x$ and $V_x)$ and $(y$ and $V_y)$.

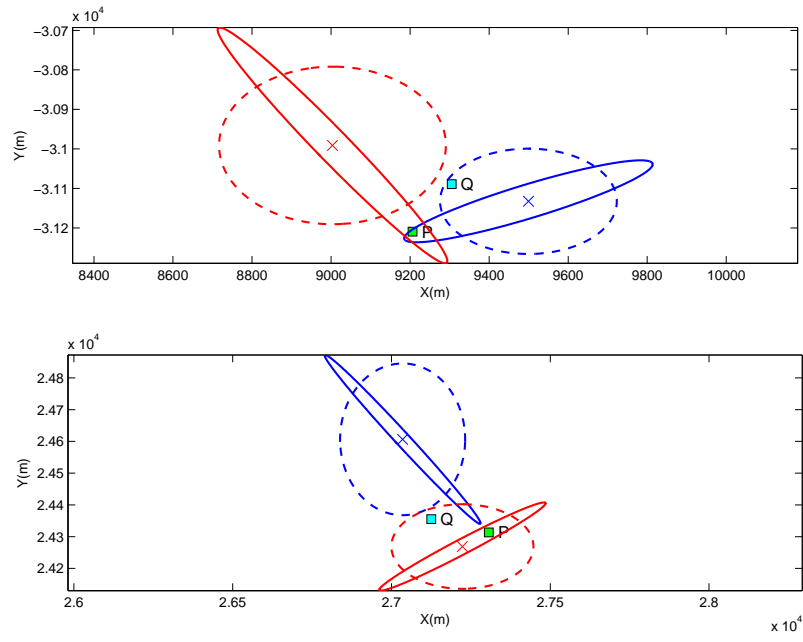


Figure 4.18: BDMEER1 method examples

This case shows very similar performance to BDMWMR1 case and the conclusions are again the same on the correlation between the elements of the state vector.

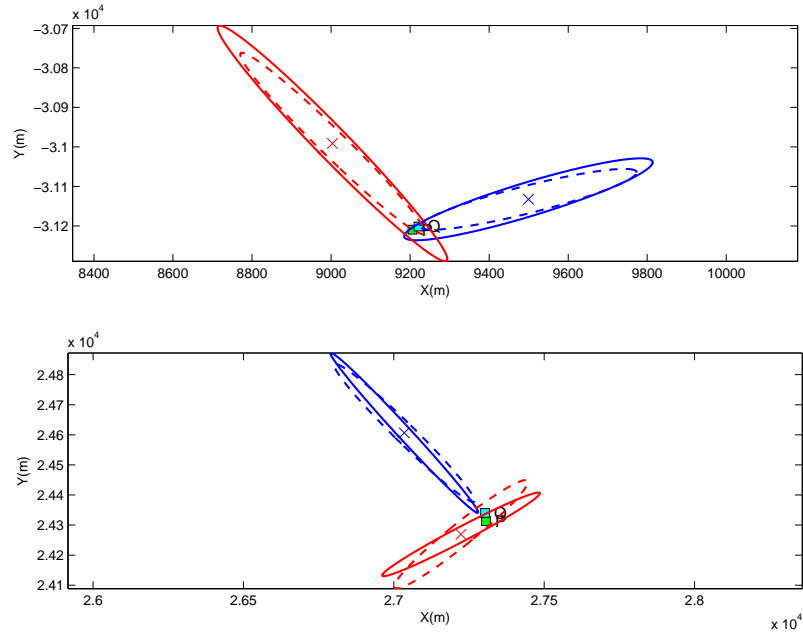


Figure 4.19: BDMEER2 method examples

BDMEER2 method shows almost same performance with BDMWMR2 method and the conclusions on the correlation between the elements of the state vector are again the same.

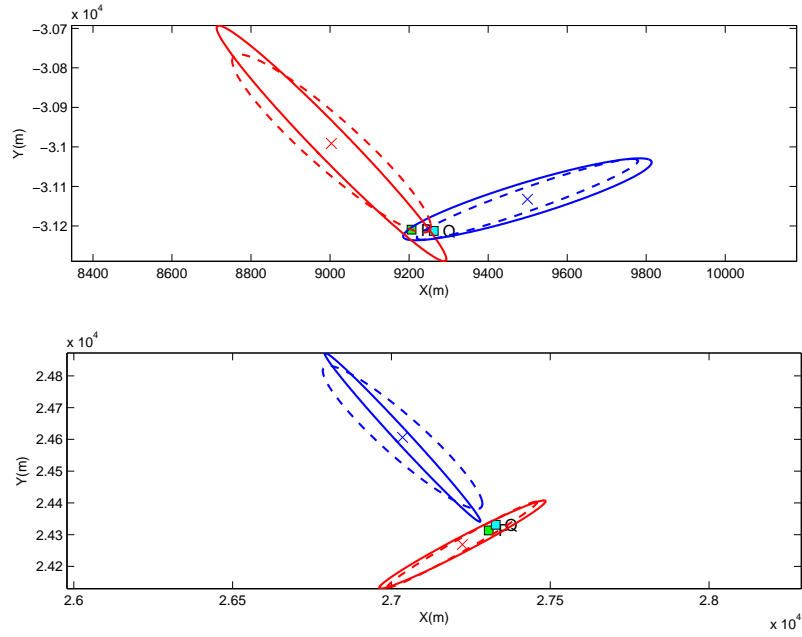


Figure 4.20: WMR method examples

Finally, Figure (4.20) shows the result of quantizing the matrices with no assumption on diagonality.

The overall conclusion that can be drawn from these observations are the following:

1. Proper modeling of correlation is extremely important for good approximation.
2. For very simple models like AR or DMR, it is unnecessary to use large number of bits.
3. Performances of the most complete representations, i.e., WMR, does not deviate much from simplified but appropriate models like BDMWMR2 or

BDMEER2 even for large number of bits. This is due to the fact that the assumptions made on zero blocks of covariance matrices actually holds.

These results reveal the performance of all the representation methods proposed in this thesis. One can choose any of these methods according to available bandwidth in the system. Also note that the quantization technique exploited for these representation methods has significant effects on the system performance and quantization should be made efficiently and intelligently.

CHAPTER 5

SUMMARY, CONCLUSION AND FUTURE WORK

This thesis is mainly concerned with sharing problem of covariance matrix between the nodes of a track fusion system. To begin with, in Chapter 1 some introduction on covariance matrices is given such as the importance of it in a track fusion system. In Chapter 2, we proposed a Gaussian model and derived a reduced distance measure between the actual covariance matrix and the representative one. This distance is generated using Kullback-Leibler measure which is a well-known statistical distance between two probability density functions. In Chapter 3, utilizing this reduced distance we have found optimal matrices that can best represent the covariance matrix under some structural simplifications. These structures are called "representation matrices" and are named as AR, DMR etc. Moreover, we assumed that the covariance matrix can be represented in terms of eigenvalues and eigenvectors and we derived Kullback-Leibler distance between the actual covariance and the representative one for 2x2 matrix in terms of these parameters. Finally, in Chapter 4, to compare the previously proposed representation matrices, we set up an experimental simulation environment which simulates a two-radar track fusion system. In addition to the previously defined representations in Chapter 3, block diagonal representation is proposed in two possible forms by selecting different sub-blocks where the selection is system specific. And finally, covariance matrix itself is taken as as whole.

The fusion system is composed of two-radars using Extended Kalman Filter (EKF) as the tracker and gives out the time-varying covariance information which should be sent to the fusion center for optimal sensor fusion. Weighted Covariance Fusion (WCF) method provided us a performance measure to compare various representation methods. In the model, we also need to make quantization of the covariance matrices. K-means clustering algorithm is exploited to quantize the representations. In our study, K-means clustering algorithm uses Kullback-Leibler distance measure instead using the traditional L2 norm distance. The experiments were performed for bits from 1 to 10 to understand the performances of each representation for different number of bits.

At the end of Chapter 4, we presented the results of the experiments and gave some cross-section examples of the representation covariance matrices. According to the results, we proved that for track fusion systems, the correlation between target positions (i.e. x and y) is much more important than that of the target position and velocity (i.e. (x, V_x) and (y, V_y)). We also see that it is unnecessary to use high number of bits for AR, DMR, BDMWMR1 and BDMEER1 methods. Additionally, methods DMR, BDMWMR1 and BDMEER1, more or less, present the same performance for high bit rates. On the other hand, the performances of WMR, BDMWMR2 and BDMEER2 methods increase with increasing number of bits and these methods show similar performance characteristics for various number of bits.

Innovations of this thesis study:

1. A reduced Kullback-Leibler distance measure is obtained between two Gaussian random variables. This form is only a function of the means and the covariances of the densities which can be considered very useful.
2. Different matrix representation methods are proposed and compared. In the literature there is no study that provides the opportunity to see such a comparison.
3. This study has searched for whether it is possible to make the represen-

tation by the eigenvectors and eigenvalues of the covariance matrix or not.

4. K-means clustering algorithm uses Kullback-Leibler distance criteria. This is a new approach for covariances belonging to Gaussian densities since there is not much literature on covariance quantization and the existing ones like [2] and [3] use K-means vector quantization technique that is based on Euclidean distance.

In the thesis, we have also made some analysis on eigenvalue and eigenvector representation of $n \times n$ covariance matrices. This work aimed to analyze whether utilizing eigenvalues and eigenvectors provides advantage for the representation or not. So, eigenvalue/eigenvector representation method is a future work waiting to be handled.

REFERENCES

- [1] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. YBS Publishing, 1995.
- [2] Y. Ruan and P. Willett, “A quantized track fusion strategy,” *Proceedings of the 43rd IEEE Conference on Decision and Control*, Dec. 2004.
- [3] ———, “A quantization architecture for track fusion,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 2, pp. 738–754, Apr. 2005.
- [4] B. Mielczarek and W. Krzymie, “Flexible channel feedback quantization in multiple antenna systems,” *Proceedings of 2005 IEEE Vehicular Technology Conference*, vol. 1, pp. 620–624, May 2005.
- [5] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood: Artech House, 1999.
- [6] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *J. Basic Eng.*, vol. 82, no. 1, pp. 35–46, Mar. 1960.
- [7] M. Liggins, C. Chong, I. Kadar, M. Alford, V. Vannicola, and S. Thomopoulos, “Distributed fusion architectures and algorithms for target tracking,” *Proceedings of the IEEE*, vol. 85, no. 1, pp. 95–107, Jan. 1997.
- [8] C. Chong, *Multisensor Multitarget Tracking: Advanced Applications*. Norwood: Artech House, 1990.
- [9] O. Drummond, “Feedback in track fusion without process noise,” *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, vol. 2561, pp. 369–383, July 1995.
- [10] G. Frenkel, “Multisensor tracking of ballistic targets,” *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, vol. 2561, pp. 337–346, July 1995.
- [11] P. Mahalanobis, “On the generalised distance in statistics,” *Proceedings of the National Institute of Science of India*, 1936.
- [12] K. C. Chang, R. K. Saha, and Y. Bar-Shalom, “On optimal track-to-track fusion,” *IEEE Transactions on Aerospace and Electronics Systems*, vol. 33, no. 4, pp. 1271–1276, Oct. 1997.

- [13] S. Kullback, *Information Theory and Statistics*. New York: Dover Publications, 1968.
- [14] J. Bucklew, *Large Deviation Techniques in Decision, Simulation and Estimation*. New York: Wiley, 1990.
- [15] R. M. Gray and T. Linder, "Mismatch in high-rate entropy-constrained vector quantization," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1204–1217, May 2003.
- [16] P. Vanheeghe, E. Duflos, P. Dumont, and V. Nimier, "Sensor management with respect to danger level of targets," *Proceedings of the 40th IEEE Conference on Decision and Control*, pp. 4439–4444, Dec. 2001.
- [17] M. N. Do, "Fast approximation of kullback-leibler distance for dependence trees and hidden markov models," Submitted to *IEEE Signal Processing*, pp. 115–118, Apr. 2003.
- [18] W. Blair and M. Brandt-Pierce, "Statistical description of monopulse parameters for tracking rayleigh targets," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 597–611, Apr. 1998.
- [19] P. Willett, W. Blair, and Y. Bar-Shalom, "On the correlation between horizontal and vertical monopulse measurements," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 533–549, Apr. 2003.
- [20] C. Y. Chong, K. C. Chang, and S. Mori, "Distributed tracking in distributed sensor networks." In *Proceedings of the American Control Conference*, 1986.
- [21] X. R. Li and K. Zhang, "Optimality and efficiency of optimal distributed fusion." In *Proceedings of the Workshop on Estimation, Tracking and Data Fusion: A Tribute to Yaakov Bar-Shalom*, pp. 37–50, May 2001.
- [22] O. Drummond, "Track and tracklet fusion using data from distributed sensors." In *Proceedings of the Workshop on Estimation, Tracking and Data Fusion: A Tribute to Yaakov Bar-Shalom*, May 2001.
- [23] C.-Y. Chong, S. Mori, B. Barker, and K.-C. Chang, "Architectures and algorithms for track association and fusion," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 15, no. 1, pp. 5–13, Jan. 2000.