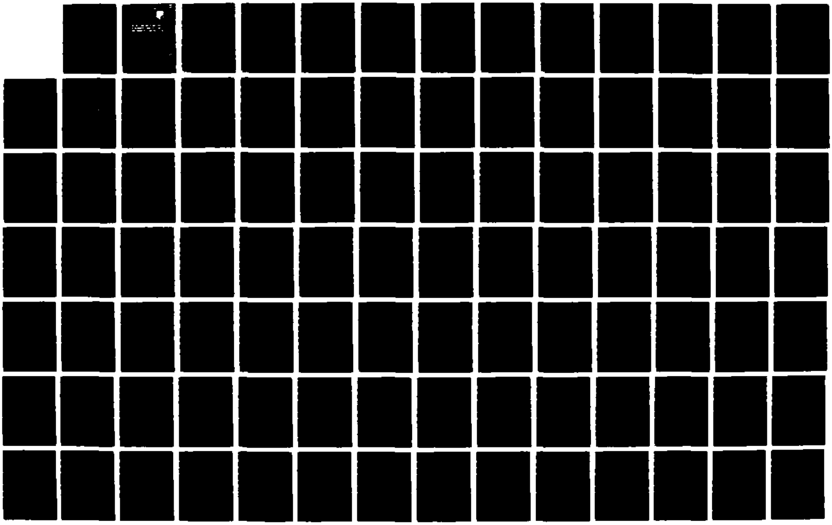
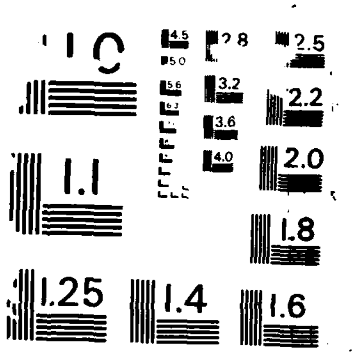
AD-A193 934 ON THE DESIGN AND OPTIMIZATION OF DISTRIBUTED SIGNAL 1/3
DETECTION AND PARAME..(U) SYRACUSE UNIV NY DEPT OF
ELECTRICAL AND COMPUTER ENGINEERING..
UNCLASSIFIED I Y HOBALLAH ET AL. SEP 87 RADC-TR-87-130 F/G 12/9 NL





(2)

DTIC FILE COPY

RADC-TR-87-130
Final Technical Report
September 1987

AD-A193 934

ON THE DESIGN AND OPTIMIZATION OF DISTRIBUTED SIGNAL DETECTION AND PARAMETER ESTIMATION SYSTEMS

Syracuse University

Imad Youssef Hoballah and Pramod K. Varshney

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700

This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-87-130 has been reviewed and is approved for publication.

APPROVED: Vincent C. Vannicola

VINCENT C. VANNICOLA
Project Engineer

APPROVED: Frank J. Rehm

FRANK J. REHM
Technical Director
Directorate of Surveillance

FOR THE COMMANDER: John A. Ritz

JOHN A. RITZ
Directorate of Plans & Programs

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (OCTS) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notice on a specific document requires that it be returned.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS N/A
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) N/A	5. MONITORING ORGANIZATION REPORT NUMBER(S) RADC-TR-87-130
6a. NAME OF PERFORMING ORGANIZATION Syracuse University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Rome Air Development Center (OCTS)
6c. ADDRESS (City, State, and ZIP Code) Department of Electrical and Computer Engr Link Hall Syracuse NY 13244-1240	7b. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Rome Air Development Center 8b. OFFICE SYMBOL (If applicable) OCTS	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F30602-81-C-0169
8c. ADDRESS (City, State, and ZIP Code) Griffiss AFB NY 13441-5700	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
61102F	2305	J8	PC

11. TITLE (Include Security Classification)
ON THE DESIGN AND OPTIMIZATION OF DISTRIBUTED SIGNAL DETECTION AND PARAMETER ESTIMATION SYSTEMS

12. PERSONAL AUTHOR(S)
Imad Youssef Hoballah, Pramod K. Varshney

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM Jun 84 TO Dec 86	September 1987	216

16. SUPPLEMENTARY NOTATION
N/A

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP		
09	03		Fusion	Detection
17	04		Surveillance	Estimation
			Remote Receivers	

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

In this report, the problems of hypothesis testing and parameter estimation in a distributed framework are considered. First, hypothesis testing in distributed systems with data fusion is treated. The approach can easily be applied to decentralized systems without data fusion. Optimal decision rules at the detectors and optimal fusion rules are derived for the distributed hypothesis testing problems using the Neyman-Pearson criterion, the general Bayesian criterion and the minimum equivocation criterion. Correspondence between information theory and detection theory is established. Decentralized postdetection integration problems are also considered and optimum fusion rules, as well as optimum decision rules at the individual detectors are obtained for two proposed schemes. Next, decentralized Bayesian parameter estimation is considered and optimum estimation rules at the local estimators and optimum combining rules are obtained for the minimum mean square error criterion, the absolute error criterion and the uniform cost function criterion.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Vincent Vannicola	22b TELEPHONE (Include Area Code) (315) 330-4437	22c OFFICE SYMBOL RADC (OCTS)

DD Form 1473, JUN 86 Previous editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Notations

DD problem ≡ Distributed detection system with local inference

DDF problem ≡ Distributed detection system with data fusion

DPEF problem ≡ Distributed parameter estimation system with fusion

MED ≡ Minimum equivocation detection

$Y = (y_1^T, y_2^T, \ldots, y_N^T)^T \equiv$ Total observation vector

$y_i = (y_{i1}, y_{i2}, \ldots, y_{in})^T \equiv$ Observation vector at sensor i, $i = 1,2,\ldots,N$

$x = (u_1, u_2, \ldots, u_N)^T \equiv$ Decision vector

$Y^\mu = (y_1^T, y_2^T, \ldots, y_{\mu-1}^T, y_{\mu+1}^T, \ldots, y_N^T)^T$

$x^\mu = (u_1, u_2, \ldots, u_{\mu-1}, u_{\mu+1}, \ldots, u_N)^T$

$x_\mu^j = (u_1, u_2, \ldots, u_\mu = j, \ldots, u_N)^T, \quad j = 0,1,\ldots,M-1$

$P(Y|H_j)$ = Conditional density function of the observation vector Y given H_j, $j = 0,1,\ldots,M-1$

$P_D \equiv$ Probability of detection of the overall system

$P_F \equiv$ Probability of false alarm of the overall system

$P_M \equiv$ Probability of miss of the overall system

$P_{Di} \equiv$ Probability of detection at detector i, $i = 1,2,\ldots,N$

$P_{Fi} \equiv$ Probability of false alarm at detector i, $i = 1,2,\ldots,N$

$P_{Mi} \equiv$ Probability of miss at detector i, $i = 1,2,\ldots,N$

$M_x = D_x = P(x|H_1)$

$F_x = P(x|H_0)$

$P_{k\underline{x}} = P(u=k|\underline{x})$, $k = 0,1,\ldots,M-1$

$P_{\underline{x}j} = P(\underline{x}|H_j)$, $j = 0,1,\ldots,M-1$

$S_0 \equiv$ Set of all j such that u_j is an element of $\underline{x}$ and $u_j = 0$

$S_0 \equiv \{u_j \mid j = 1,2,\ldots,N \text{ and, } u_j = 0\}$

$S_1 \equiv \{u_k \mid k = 1,2,\ldots,N \text{ and, } u_k = 1\}$

$M_{\underline{x}}^{\mu} = D_{\underline{x}}^{\mu} = P(x^{\mu}|H_1)$

$F_{\underline{x}}^{\mu} = P(x^{\mu}|H_0)$

$P_{k\underline{x}}^{\mu j} = P_{kx}\Big|_{u_\mu=j}$ $k,j = 0,1,\ldots,M-1$

$S_0^{\mu} = \{u_j \mid j = 1,2,\ldots,\mu-1,\mu+1,\ldots,N \text{ and, } u_j = 0\}$

$S_1^{\mu} = \{u_k \mid k = 1,2,\ldots,\mu-1,\mu+1,\ldots,N \text{ and, } u_k = 1\}$

$\sum_{\underline{x}} \equiv$ Summation over all possible values of $\underline{x}$

$\sum_{\underline{x}-\underline{x}^*} \equiv$ Summation over all possible values of $\underline{x}$ except $\underline{x}^*$

$C_{ij} \equiv$ Cost of deciding $u = i$ when H_j is present

$i,j = 0,1,\ldots,M-1$

$A_{\underline{x}^{\mu}} \equiv P(u=1|\overset{1}{x_{\mu}}) - P(u=1|\overset{0}{x_{\mu}})$

$\delta(.)/\delta x \equiv$ Partial derivative with respect to the varible x

$R_{ms} \equiv$ Minimum mean-square-error function

$\hat{a}_{ms} \equiv$ Minimum mean-square estimate

$R_{abs} \equiv$ Minimum absolute error function

$\hat{a}_{abs} \equiv$ Minimum absolute error estimate

$R_{unif} \equiv$ Minimum uniform cost function

$\hat{a}_{unif} \equiv$ Minimum uniform estimate

Table of Contents

I. Introduction

1.1 Introduction

Classical signal detection and estimation involve centralized signal processing. Traditionally, a single sensor is employed for observations and the data is processed at a central processor. More recently, surveillance systems are employing multiple sensors for observations to improve system performance parameters such as reliability and speed, and also to increase the coverage and the number of targets under consideration. If there is no constraint on communication channel and processor bandwidths, complete observations may be brought to a central processor for data processing. In this case, the signal processing is still centralized in nature as shown in Figure 1.1. The theory of centralized signal detection and estimation is very well understood and the solutions to problems such as optimum Bayesian detection, Neyman-Pearson detection, minimum equivocation detection and many estimation problems are available in standard textbooks [1].

The goal in this report is to consider some signal detection and parameter estimation problems when the signal processing is not centralized in nature. Multiple sensors which may be spatially distributed or located at one

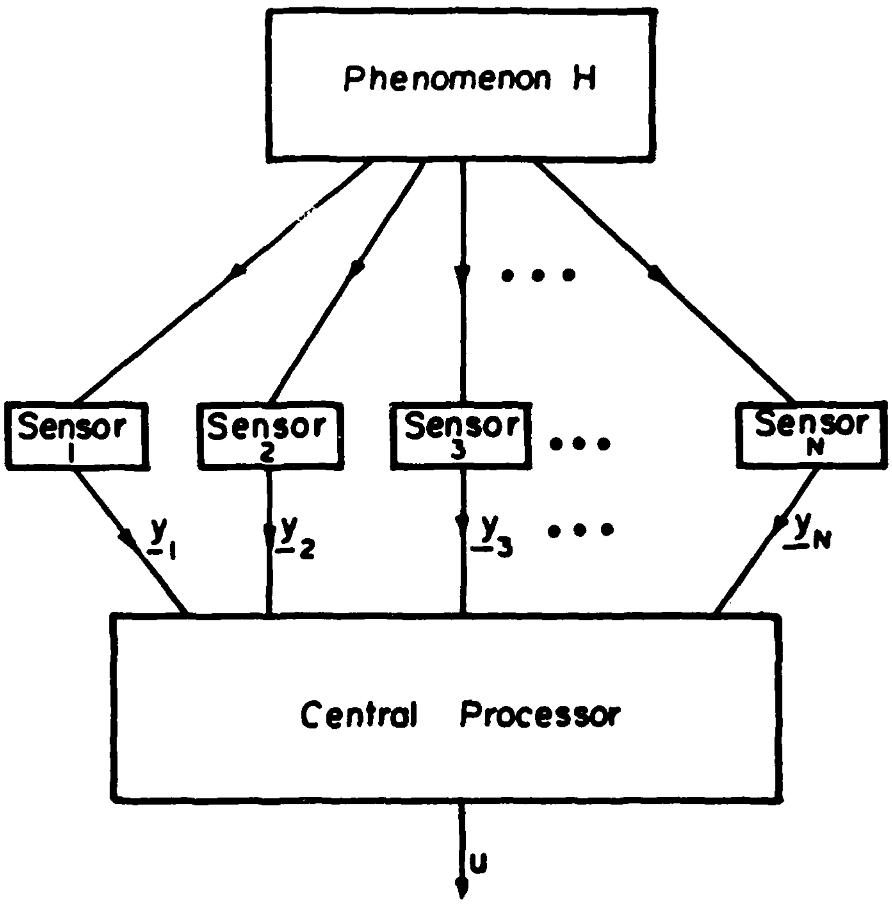


Figure 1.1 Distributed Sensor System with Central Computation

location are employed for making observations. The sensors have signal processing capabilities and some or all of the processing is done at the sensor itself. Therefore, such systems perform distributed computation as opposed to centralized computation. In the fully distributed system shown in Figure 1.2, all the signal processing is performed at the sensors and the inference is available locally. In some other distributed sensor systems, where a global inference is desired, partial results are transmitted to a data fusion center where they are appropriately combined to yield the global inference as shown in Figure 1.3. Some other distributed sensor network topologies which involve hierarchical structures have also been studied in the literature [6]. It should be pointed out that there are many practical reasons for deploying multiple sensor surveillance systems with distributed computation. These include cost, reliability, survivability and limitations on communication bandwidth. As pointed out previously, the classical approach to signal detection and estimation has dealt with centralized problems and analytical solutions are readily available. The classical theory needs to be extended to be able to solve distributed hypothesis testing and estimation problems. The goal of the work, reported in this report, is to consider and obtain analytical solutions to several distributed detection and parameter estimation problems. In the next two sections, we introduce the notation and terminology and set up the distributed

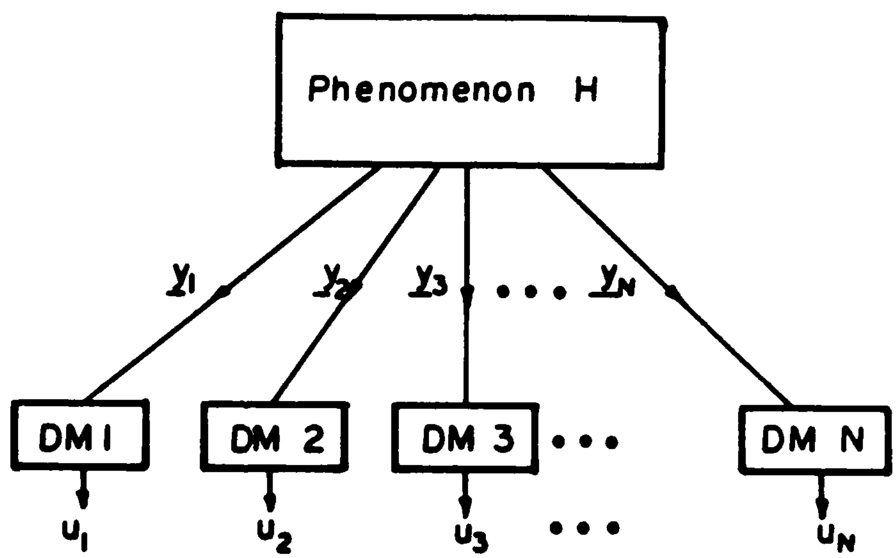


Figure 1.2 Distributed Sensor System with Local Inference

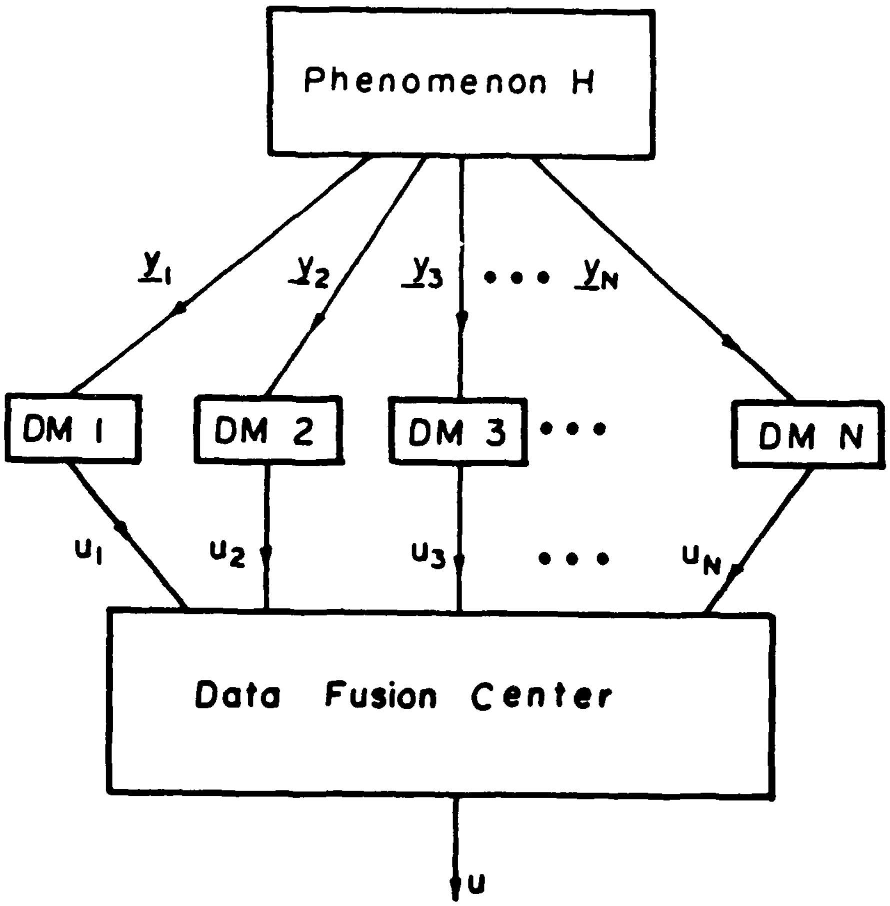


Figure 1.3 Distributed Sensor System with Data Fusion

detection and distributed parameter estimation problems respectively. We also briefly discuss the relevant literature on the subject.

1.2 Distributed Hypothesis Testing

1.2.1 System Description and Terminology

In this report, we shall consider two different distributed hypothesis testing configurations. The first one is the fully distributed system shown in Figure 1.2. In this system, all the signal processing is performed at the sensors and the inference is generated locally. We will refer to this distributed detection configuration as the DD problem throughout this report. In the second system of Figure 1.3, partial results from the sensors are transmitted over bandlimited channels to the data fusion center where they are combined according to a fusion rule to yield the global inference. This distributed detection with fusion configuration will be referred to as the DDF problem throughout this report. It should be pointed out that the DDF problem can be reduced to the DD problem by a suitable choice of the fusion rule and cost functions.

In Figure 1.4, we present the basic block diagram of a distributed detection system. This block diagram includes both the DD problem and the DDF problem. The basic

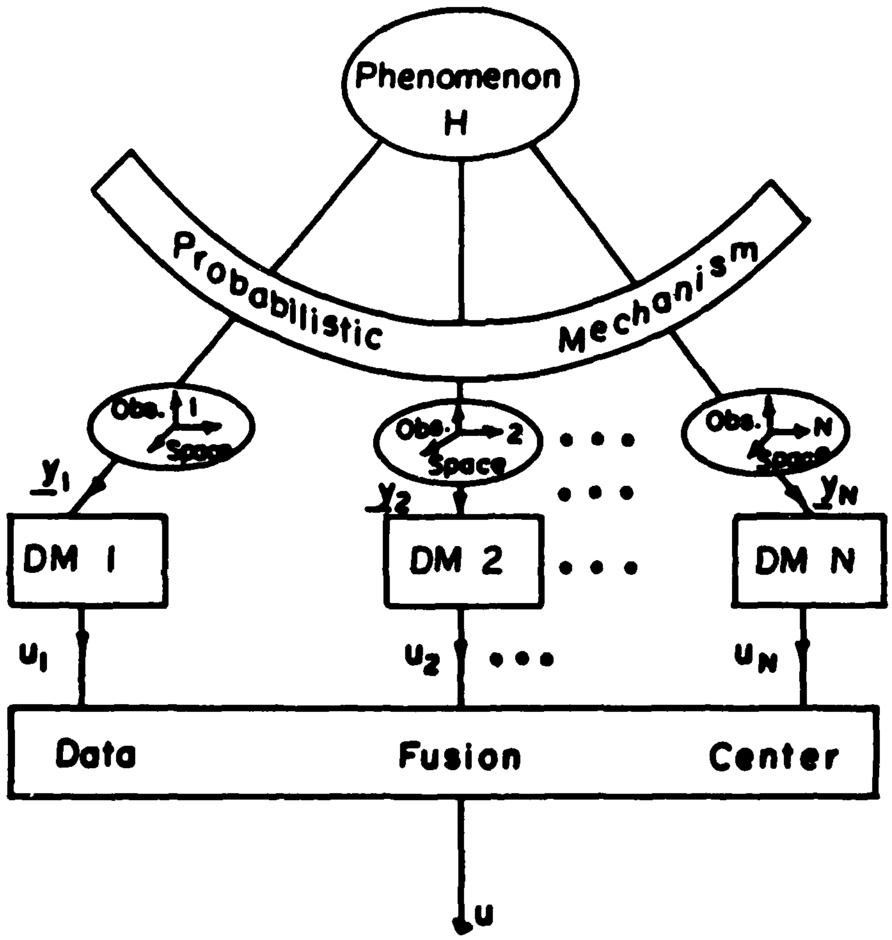


Figure 1.4 Distributed Sensor System with Data Fusion

components of the problem are as follows:

1. A source which generates an output. The output could be one of M possible choices. These choices are referred to as hypotheses, and are denoted by $H_0, H_1, \ldots, H_{M-1}$. The a priori probabilities of the hypotheses are denoted by $P_j = P(H_j)$, $j = 0,1,\ldots,M-1$.

2. A probabilistic transition mechanism that can be viewed as a mechanism which, based on the knowledge of the true hypothesis and some probabilistic law, chooses a point in an observation space.

3. An observation space, consisting of points in an n x N dimensional space. Each point is represented by an observation vector $\underline{Y}$,

$$\underline{Y} = (\underline{y}_1^T, \underline{y}_2^T, \ldots, \underline{y}_N^T)^T$$

where $\underline{y}_i$ is the observation vector at the detector i, $i = 1,2,\ldots,N$. Each $\underline{y}_i$ in turn is given by

$$\underline{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})^T \qquad i = 1,2,\ldots,N$$

where n is the number of observations at each detector. Without loss of generality, we have assumed that the number of observations is the same at each of the detectors. The observation vector has a known nN-dimensional conditional density function $p(\underline{Y}|Hj)$, $j = 0,1,\ldots,M-1$.

4. A decision rule, $u_i = g(\underline{y}_i)$, for each detector i, $i = 1,2,\ldots,N$. After observing the outcome in the observation space, each detector uses this rule to guess as to which hypothesis is present. Each decision maker will assign the points in its own observation space to one of the hypotheses. The vector consisting of the local decisions will be denoted by $\underline{x}$, i.e., $\underline{x} = (u_1,u_2,\ldots,u_N)^T$.

5. A data fusion rule $u = f(\underline{x})$ which is used by the data fusion center to declare as to which hypothesis is present on the basis of the vector $\underline{x}$. Note that the data fusion rule does not exist in the DD problem.

The block diagram and the notation introduced above will be used in Chapters 2, 3, 4 and 5 where we will solve some distributed detection problems. Next, we review some of the reported work on distributed detection.

1.2.2 Previous Work

Some recent work on the detection problem with multiple sensors has been reported in the literature (e.g.[3-18]). Tenney and Sandell [3] extended the classical Bayesian decision theory to the case where they considered the DD problem . This extension does not follow from the classical theory in a straightforward manner because the decision rules at the individual detectors are coupled. Sadjadi [4], treated the problem of optimum detection with N decentralized sensors selecting among M possible hypotheses, with no data fusion. Further work along these lines has been performed in [5,6]. Lauer and Sandell [5], considered the Bayesian detection of signal waveforms in the presence of noise. Ekchian and Tenney [6], formulated the detection problem for various distributed sensor network topologies. Kushner and Pacut [7] conducted a simulation study of a specific distributed detection problem. Teneketzis [8,9] has also solved some interesting decentralized detection problems, namely a version of the Wald problem and the quickest detection problem. Tsitsiklis and Athans [10], have considered the computational complexity of decentralized decision problems. Their results point to the inherent computational difficulty of the problem and suggest that optimality may be an elusive goal. Conte, D'Addio, Farina and Longo [11], have considered the design and performance evaluation of optimum and suboptimum multistatic radar receivers. Their suboptimum structure is a special case of the DDF problem that we shall be considering in this report. Stearns [12] considered different combining schemes in order to determine as to which achieves the best possible receiver operating characteristic (ROC). Some related work which has been reported from a control standpoint has been reported. Sandell and Athans [13], and Radner [14], considered the decentralized static linear quadratic Guassian (LQG) problem and derived appropriate decision rules. Other information structural problems were studied by Ho [15,16,17]. Sandell, Varaiya, Athans and Safanov

[18], have surveyed the decentralized control methods for large scale systems.

In the next section, we introduce the notation and terminology for the distributed parameter estimation problem. We also discuss the relevant literature on the subject.

1.3 Distributed Parameter Estimation

1.3.1 System Description and Terminology

A detailed block diagram of a distributed parameter estimation system with fusion rule is shown in Figure 1.5. Local parameter estimates are obtained at the individual sensors and are transmitted to the fusion center where they are combined to yield the global estimate. We will refer to this distributed parameter estimation with fusion structure as the DPEF problem. The basic components of the problem are as follows

1. A parameter space, consisting of points which correspond to the random parameter, a, to be estimated. For the single parameter case, it corresponds to segments of the line $-\infty < a < +\infty$. The probability density function of the random parameter, a, is denoted by p(a).

2. A probabilistic transformation which maps the parameters to an observation space.

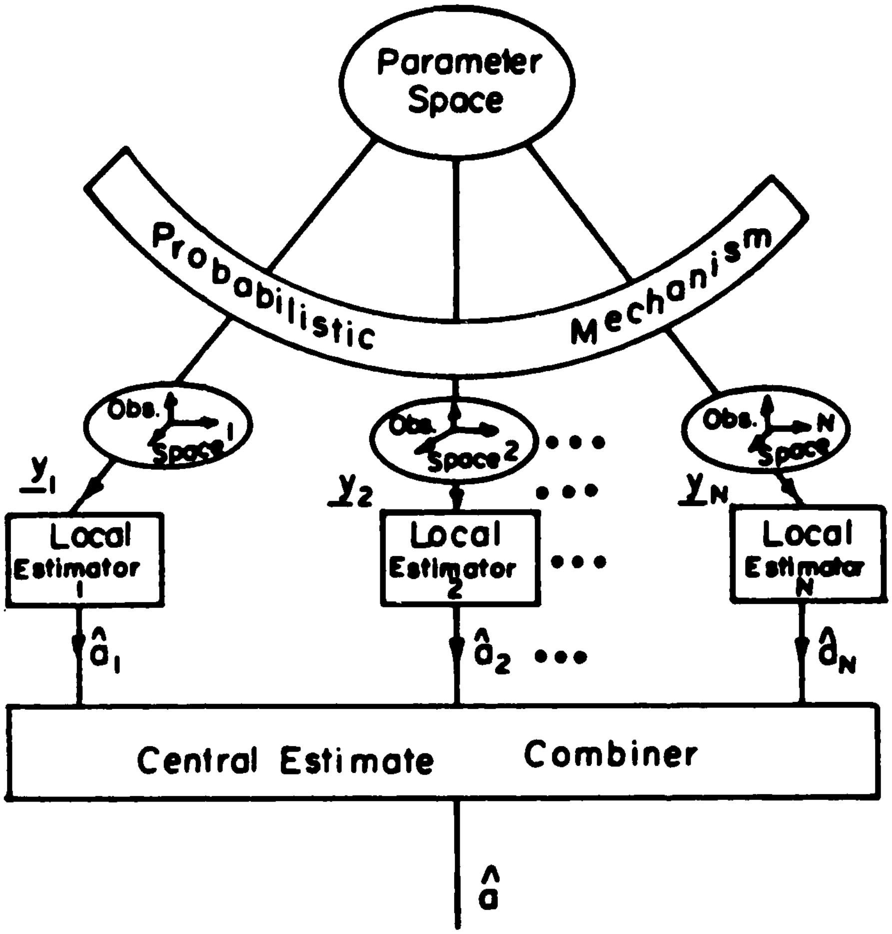


Figure 1.5 Distributed Estimation System with Estimate Combining

3. An observation space consisting of points in an $n \times N$ dimensional space. Each point is represented by the observation vector, $\underline{Y}$, which is

$$\underline{Y} = (\underline{y}_1^T, \underline{y}_2^T, \ldots, \underline{y}_N^T)^T$$

where $\underline{y}_i$ is the observation vector at the estimator i. Each $\underline{y}_i$, in turn is given by $\underline{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})^T$, where n is the number of observations collected by each estimator. Without loss of generality, we have assumed that the number of observations is the same at each of the estimators.

The observation vector has a known conditional density function

$$p(Y|a) = p(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n|a)$$

4. An estimation rule $\hat{a}_i = h_i(\underline{y}_i)$ at each estimator i, $i = 1,2,\ldots,N$. Each estimator uses its own estimation rule to map its observation to an estimate. We will assume that estimator i, $i = 1,2,\ldots N$, does not have knowledge of the observations at other estimators j, $j \neq i$, $j = 1,2,\ldots,N$. The vector consisting of the local estimates will be denoted by

$$\underline{h} = (h_1(\underline{y}_1), h_2(\underline{y}_2), \ldots, h_N(\underline{y}_N))^T.$$

5. A combining rule $\hat{a} = f_c(\underline{h}) = f_c(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_N)$ which, based on the values $\hat{a}_1$, $\hat{a}_2, \ldots,$ and $\hat{a}_N$, gives the global estimate of the parameter a.

The block diagram shown in Fig. 1.5 and the notation introduced in this subsection will be used in Chapter 6 where we will solve some distributed parameter estimation problems. Next, we briefly review some reported work on distributed estimation.

1.3.2 Previous Work

Most of the work reported in the literature has dealt with decentralized state estimation problems. Borkar and Varaiya [19], considered asymptotic agreement in distributed estimation problems. In their case, each of several agents updates its estimate and transmits it to a randomly chosen set of the other agents. They showed that the common limit which ring members agree upon depend upon the order in which estimates are transmitted. Teneketzis and Varaiya [20], studied the consensus problem in distributed estimation with inconsistent beliefs. They considered the case when two people's estimates of the same random variable are available, and discussed convergence of the estimates to the same value. In [21], Chang and Tabaczynski considered multisensor state estimation with applications to the target tracking problem. Willsky, Bello, Castanon, Levy and Verghese [22], considered combining and updating local estimates and regional maps. The estimates may be generated at different locations or at different times. They emphasized the conceptual similarity between many problems in decentralized control and in the analysis of random fields. In [23], they developed a framework for the study of centralized estimation problems where the decentralized estimation problem is imbeded into an equivalent scattering problem. In [24], they derived algorithms for different mapping problems in a unified framework. Castanon and Teneketzis [25], obtained a distributed processing algorithm which recovers exactly the centralized conditional distribution when only the sufficient statistics are communicated. Washburn and Teneketzis in [26] analyzed the performance for hybrid state estimation problems. Varshney and Varshney [27], have considered recursive estimation with uncertain observations in a multisensor environment.

In the next section, we present the report organization.

1.4 Report Organization

In this report we consider some distributed detection problems as well as distributed estimation of random parameters. Optimal decision rules and fusion rules are derived for the distributed detection problems considered. Similarly, for the distributed parameter estimation problems, optimal local estimators and combining rules are obtained.

In Chapter 2, we present the distributed Neyman-Pearson detection problem where we design an optimum decision system. First, when the fusion rule is known, the decision rules at the individual detectors are derived so as to minimize the probability of miss (or to maximize the probability of detection) under a constraint on the probability of false alarm. Next, when the decision rules at the individual detectors are known, we derive the optimum fusion rule using the same criterion. We also discuss the overall solution where we simultaneously obtain the optimum fusion and decision rules.

In Chapter 3, we treat the problem of distributed Bayesian detection with data fusion. The optimum decision rules at the sensors and, the optimum fusion rule are derived. Several special cases such as the independent observation case and the identical detector case are discussed in detail.

In Chapter 4, we present two schemes to be used for distributed postdetection decision making.

In Chapter 5, we solve the minimum equivocation detection problem for the DD and the DDF systems. Optimal decision rules at the individual detectors for both systems and, optimal data fusion rule for the DDF problem are derived. Our criterion is to maximize the mutual information (or equivalently to minimize the equivocation) between the input and the output.

In Chapter 6, we consider the distributed parameter estimation problem when multiple estimators are used and a combining rule is employed to obtain the global estimate. Optimal combining and estimation rules are derived.

In Chapter 7, we summarize the results and also present some suggestions for future research.

II. Neyman-Pearson Detection Using Multiple Sensors

2.1 Introduction

The theory of signal detection using a single radar is very well understood [1,2]. The Bayesian approach to the optimum detection problem requires the knowledge of the a priori probabilities and the costs. An optimum detection rule is then obtained which minimizes the average cost of detection. For most radar detection problems, the Bayesian approach is inappropriate because the required information, i.e., a priori probabilities and costs, may not be available. For this reason, the Neyman-Pearson criterion, which does not require the above knowledge, is employed extensively in radar detection systems. Under this criterion, a constraint is placed on the probability of false alarm and, the probability of detection is maximized (or the probability of miss is minimized). The detection rule thus obtained is used for signal detection.

In this chapter, we develop the Neyman-Pearson decision theory for signal detection using multiple radars. We assume the structure shown in Figure 1.3, i.e, individual decisions from the radars are fed to a data fusion center which yields the global decision. A constraint on the probability of false alarm of the overall system (global decision) is placed and the probability of miss of the overall system is minimized.

In Section 2.2, we formulate the problem and define the notation and terminology. In Section 2.3, we develop the Neyman-Pearson decision theory when multiple sensors are used for surveillance. We consider the problem of binary hypothesis testing using multiple detectors. First, when the fusion rule is known, decision rules at individual detectors are obtained. These rules are functions of the data fusion scheme being employed and are, in general, coupled. Secondly, when the decision rules at the detectors are given, the optimum fusion rule is derived. The overall solution to the problem is also presented. Special cases of "AND" and "OR" data fusion rules are considered. A specific example is presented in Section 2.4. Finally, the results are discussed in Section 2.5.

2.2 Problem Statement

We consider a binary hypothesis testing problem with the following two hypotheses:

$$\begin{aligned} H_0 &: \quad \text{Target is absent, and} \\ H_1 &: \quad \text{Target is present.} \end{aligned} \qquad (2\text{-}1)$$

We consider the system structure shown in Figure 1.3 where a data fusion center is used along with the distributed sensors.

The observations at the i^{th} detector are denoted by y_i, $i=1,2,\ldots,N$. We further assume that the joint conditional probability density function $p(\underline{y}_1,\underline{y}_2,\ldots,\underline{y}_N|H_j)$, $j=0,1$, is known. Each detector employs a decision rule $g_i(\underline{y}_i)$ to make a local decision u_i, $i=1,2,\ldots,N$ where

$$u_i = \begin{cases} 0 & \text{if detector } i \text{ decides } H_0 \\ 1 & \text{if detector } i \text{ decides } H_1 \end{cases} \qquad (2\text{-}2)$$

$$i = 1,2,\ldots,N$$

The data fusion center determines the overall or global decision for the system, u, based on individual decisions, i.e.,

$$u = f(u_1,u_2,\ldots,u_N).$$

As an example of data fusion rules, we present the "AND" and "OR" data fusion rules for the special case of two detectors (N=2) in Table 2.1. The global decision u is simply a Boolean AND or OR of the Boolean variables u_1 and u_2. Other data fusion rules involving other Boolean operations on variables (u_i), can be formulated, e.g., majority logic. Later in this chapter we will consider "AND" and "OR" fusion rules as examples. It should be pointed out that these

		"AND"	"OR"
u_1	u_2	u	u
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

Table 2.1. "AND" and "OR" Fusion Rules

fusion rules are not necessarily optimum. The derivation of optimum data fusion rules for multiple sensor detection systems is also considered in this chapter.

The goal of this chapter is to develop the Neyman-Pearson decision theory for detection systems with multiple sensors. For this, we define the probability of false alarm P_F, the probability of miss P_M, and the probability of detection P_D of the overall system

$$
\begin{aligned}
P_F &= \mathrm{Prob}(u=1 \mid H_0) \\
P_M &= \mathrm{Prob}(u=0 \mid H_1) \qquad\qquad (2\text{-}3) \\
P_D &= \mathrm{Prob}(u=1 \mid H_1).
\end{aligned}
$$

The probability of miss and the probability of false alarm for individual detectors can be defined in a similar manner and are denoted by P_{Mi} and P_{Fi}, $i=1,2,\ldots,N$, respectively.

Two problems are considered here. In the first one, assuming the fusion rule is known, we find the decision rules at the individual detectors which minimize the probability of miss P_M, under the constraint that the probability of false alarm satisfies $P_F \le \alpha$. In the second one, when the decision rules of the detectors are given, we derive the fusion rule to satisfy the same condition on P_F and minimize P_M. In the next section, we employ the Lagrange multiplier method for the solution of the problems.

2.3 Distributed Neyman-Pearson Detection

2.3.1 Optimum Decision Rules

We consider the binary-hypothesis Neyman-Pearson detection problem with N sensors. In this subsection, we first assume that the observations at the individual detectors are statistically independent and, that the conditional probability densities $p(\underline{y}_i|H_j)$, $i=1,2,\ldots,N$, $j=0,1$, and the fusion rule $f(u_1,u_2,\ldots,u_N)$ are known. We wish to maximize P_D (or equivalently minimize P_M) under the constraint that P_F satisfies the inequality $P_F \leq \alpha$. Following the approach taken in classical Neyman-Pearson analysis, we form the function

$$\Gamma = P_M + L\,[P_F - \alpha] \tag{2-4}$$

where L is the Lagrange multiplier.

We assume that the fusion rule is not necessarily deterministic. In order to be able to express P_F and P_M in terms of the probability of false alarm and the probability of miss of the individual detectors, i.e. P_{Fi}'s and P_{Mi}'s, we define the following quantities

$\underline{x} = (u_1,u_2,\ldots,u_N)^T$, a vector whose elements take values zero or one, representing the decisions of the individual detectors.

$$M_x = \prod_{S_0} P_{Mj} \prod_{S_1} (1 - P_{Mk}) = P(\underline{x} \mid H_1); \qquad (2\text{-}5\text{-}a)$$

$$F_x = \prod_{S_0} (1 - P_{Fj}) \prod_{S_1} P_{Fk} = P(\underline{x} \mid H_0); \qquad (2\text{-}5\text{-}b)$$

and

$$P_{kx} = \mathrm{Prob}(\, u=k \mid \underline{x}) \qquad k = 0,1; \qquad (2\text{-}5\text{-}c)$$

where,

$$S_0 = \text{set of all } j,\ 1 \le j \le N \text{ and } u_j = 0 \qquad (2\text{-}6\text{-}a)$$

$$S_1 = \text{set of all } k,\ 1 \le k \le N \text{ and } u_k = 1 \qquad (2\text{-}6\text{-}b)$$

Then, we may express P_M and P_F as follows

$$P_M = \sum_{\underline{x}} P_{0x} M_x \qquad (2\text{-}7)$$

and,

$$P_F = \sum_{\underline{x}} P_{1x} F_x \qquad (2\text{-}8)$$

where

$\sum_{\underline{x}} \equiv$ summation over all possible vectors $\underline{x}$.

Substituting P_F and P_M in (2-4), Γ can be expressed as

$$\Gamma = \sum_{\underline{x}} P_{0x} M_x + L\,(\sum_{\underline{x}} P_{1x} F_x - \alpha) \qquad (2\text{-}9)$$

Expanding (2-9) in terms of $P_{M\mu}$ and $P_{F\mu}$, the probability of miss and the probability of false alarm of the μ^{th} detector respectively (μ = 1,2,...,N), Γ becomes

$$\Gamma = P_{M\mu} \sum_{\underline{x}} P_{0\underline{x}}^{\mu 0} M_{\underline{x}}^{\mu} + (1 - P_{M\mu}) \sum_{\underline{x}} P_{0\underline{x}}^{\mu 1} M_{\underline{x}}^{\mu}$$

$$+ L \left((1 - P_{F\mu}) \sum_{\underline{x}} P_{1\underline{x}}^{\mu 0} F_{\underline{x}}^{\mu} + P_{F\mu} \sum_{\underline{x}} P_{1\underline{x}}^{\mu 1} F_{\underline{x}}^{\mu} - \alpha\right) \quad (2\text{-}10)$$

where

$$M_{\underline{x}}^{\mu} = \prod_{S_0^{\mu}} P_{Mj} \prod_{S_1^{\mu}} (1 - P_{Mk}) = P(\underline{x}^{\mu}|H_1) \qquad (2\text{-}11\text{-a})$$

$$F_{\underline{x}}^{\mu} = \prod_{S_0^{\mu}} (1 - P_{Fj}) \prod_{S_1^{\mu}} P_{Fk} = P(\underline{x}^{\mu}|H_0) \qquad (2\text{-}11\text{-b})$$

$$P_{k\underline{x}}^{\mu j} = P_{k\underline{x}}\Big|_{u_\mu = j} \qquad j = 0,1 \qquad (2\text{-}11\text{-c})$$

$$\underline{x}^{\mu} = (u_1, u_2, \ldots, u_{\mu-1}, u_{\mu+1}, \ldots, u_N)^T \qquad (2\text{-}12\text{-a})$$

$$S_0^{\mu} = S_0 \text{ where } u_\mu \text{ is excluded} \qquad (2\text{-}12\text{-b})$$

and,

$$S_1^{\mu} = S_1 \text{ where } u_\mu \text{ is excluded.} \qquad (2\text{-}12\text{-c})$$

We can write (2-10) as

$$\Gamma = P_{M\mu} \sum_{\underline{x}} (P_{0x}^{\mu 0} - P_{0x}^{\mu 1}) M_x^{\mu} + \sum_{\underline{x}} P_{0x}^{\mu 1} M_x^{\mu}$$

$$+ L \, (P_{F\mu} \sum_{\underline{x}} (P_{1x}^{\mu 1} - P_{1x}^{\mu 0}) F_x^{\mu} - \alpha + \sum_{\underline{x}} P_{1x}^{\mu 0} F_x^{\mu}) \qquad (2-13)$$

Then, Γ can be expressed as

$$\Gamma = C_{M\mu} P_{M\mu} + K_{1\mu} + L \, [C_{F\mu} P_{F\mu} - \alpha + K_{2\mu}] \qquad (2-14)$$

where

$$C_{M\mu} = \sum_{\underline{x}} M_x^{\mu}(P_{0x}^{\mu 0} - P_{0x}^{\mu 1}) \qquad (2-15-a)$$

$$K_{1\mu} = \sum_{\underline{x}} M_x^{\mu} P_{0x}^{\mu 1} \qquad (2-15-b)$$

$$C_{F\mu} = \sum_{\underline{x}} F_x^{\mu} (P_{1x}^{\mu 1} - P_{1x}^{\mu 0}) \qquad (2-15-c)$$

and,

$$K_{2\mu} = \sum_{\underline{x}} F_x^{\mu} P_{1x}^{\mu 0} \qquad (2-15-d)$$

It should be noted that in the above formulation system-wide performance is being optimized rather than the performance of each individual detector. The decision rules obtained in this manner will not, in general, be the same as the ones obtained when the detectors are treated independently of each other. In fact, the decision rules at the individual detectors and their computation will be coupled. Now we proceed with the solution of the problem.

While deriving the decision rule at one detector, it would be assumed that the decision rules at all other detectors have already been obtained. The decision rule at this detector will be obtaind in terms of the decision rules at the other detectors. A simultaneous solution of the N equations obtained in this manner will yield the desired decision rules.

Assuming that $C_{M\mu}$ and $C_{F\mu}$ are not zero, we may rewrite (2-14) as

$$\frac{\Gamma}{C_{M\mu}} = P_{M\mu} + \frac{L\, C_{F\mu}}{C_{M\mu}}\left[P_{F\mu} - \frac{\alpha - K_{\alpha\mu}}{C_{F\mu}}\right] + \frac{K_{1\mu}}{C_{M\mu}} \qquad \mu=1,2,\ldots,N$$

(2-16)

Note that $C_{M\mu}$, $C_{F\mu}$ and $K_{1\mu}/C_{M\mu}$ are all functions of the fusion rule and the other detectors and are thus independent of the μ^{th} detector. Since $K_{1\mu}/C_{M\mu}$ is independent of detector μ, we can ignore it during minimization. Also, we let,

$$\frac{\Gamma}{C_{M\mu}} = \Gamma_\mu \tag{2-17-a}$$

$$\frac{L}{C_{M\mu}}\, C_{F\mu} = L_\mu \tag{2-17-b}$$

and,

$$\alpha_\mu = \frac{\alpha - K_{\alpha\mu}}{C_{F\mu}} \tag{2-17-c}$$

We then have

$$\Gamma_\mu = P_{M\mu} + L_\mu \, [P_{F\mu} - \alpha_\mu] \qquad \mu = 1,2,\ldots,N \qquad (2\text{-}18)$$

Minimization of Γ_μ yields the following likelihood ratio test (LRT) at the μ^{th} detector

$$\Omega_\mu(y_\mu) = \frac{p(y_\mu|H_1)}{p(y_\mu|H_0)} \mathop{\gtrless}_{H_0}^{H_1} L_\mu = \frac{L\, C_{F\mu}}{C_{M\mu}} = t_\mu \qquad (2\text{-}19)$$

where t_μ is the solution of

$$P_{F\mu} = \int_{t_\mu}^{\infty} p(\Omega_\mu|H_0)\, d\Omega_\mu = \alpha_\mu \qquad \mu = 1,2,\ldots,N \qquad (2\text{-}20)$$

Observe that the threshold of the μ^{th} detector is a function of the fusion rule and also the probability of false alarm of the other detectors and thus, the other thresholds.

Repeating this procedure for all the detectors, we will get N nonlinear equations in N unknowns. A simultaneous solution of these equations yields the set of thresholds which minimize the probability of miss of the overall system under a constraint on the probability of false alarm of the overall system. It should be noted that in the special case when the minimal solution is at the end point of the observation interval, the Lagrange multiplier method fails. In such a situation the above procedure must be modified.

For illustration purposes, we present the quantities defined previously for the case of two detectors and, we obtain specific results for the fusion rules "AND" and "OR".

<u>Two-Detector Neyman-Pearson Detection</u>

$$P_{kij} = \mathrm{Prob}(u=k \mid u_1=i,\ u_2=j) \qquad i,j,k=0,1$$

$$M_{00} = P_{M1}\ P_{M2}$$

$$M_{01} = P_{M1}(1 - P_{M2})$$

$$M_{10} = P_{M2}(1 - P_{M1})$$

$$M_{11} = (1 - P_{M2})(1 - P_{M1})$$

$$F_{00} = (1 - P_{F1})(1 - P_{F2})$$

$$F_{01} = (1 - P_{F1})P_{F2}$$

$$F_{10} = P_{F1}(1 - P_{F2})$$

$$F_{11} = P_{F1}\ P_{F2}$$

$$P_M = P_{000}\ P_{M1}\ P_{M2} + P_{001}P_{M1}(1 - P_{M2}) + P_{010}(1 - P_{M1})P_{M2}$$
$$+ P_{011}(1 - P_{M1})(1 - P_{M2})$$

$$P_F = P_{100}(1 - P_{F1})(1 - P_{F2}) + P_{101}(1 - P_{F1})P_{F2}$$

$$+ P_{111} P_{F1} P_{F2} + P_{110} P_{F1}(1 - P_{F2})$$

$$C_{M1} = P_{M2}(P_{000} - P_{001} - P_{010} + P_{011}) + (P_{001} - P_{011})$$

$$C_{F1} = P_{F2}(P_{100} - P_{101} - P_{110} + P_{111}) + (P_{110} - P_{100})$$

$$K_{11} = P_{010} P_{M2} + P_{011}(1 - P_{M2})$$

$$K_{21} = (P_{101} - P_{100})P_{F2} + P_{100}$$

$$C_{M2} = P_{M1}(P_{000} - P_{001} - P_{010} + P_{011}) + (P_{010} - P_{011})$$

$$C_{F2} = P_{F1}(P_{100} - P_{101} - P_{110} + P_{111}) + (P_{101} - P_{100})$$

$$K_{22} = P_{100}(1-P_{F1}) + P_{110}P_{F1} = (P_{110} - P_{100})P_{F1} + P_{100}$$

"AND" Fusion Rule

This rule requires that

$P_{000} = 1$ $\qquad$ $P_{010} = 1$

$P_{100} = 0$ $\qquad$ $P_{110} = 0$

$P_{001} = 1$ $\qquad$ $P_{011} = 0$

$P_{101} = 0$ $\qquad$ $P_{111} = 1.$

Therefore,

$$C_{Mi} = 1-P_{Mj}$$

$$C_{Fi} = P_{Fj} \qquad i \neq j \quad i,j=1,2$$

$K_{ei} = 0$

The two LRT's and the corresponding equations for the thresholds are

$$\Omega_i(y_i) \underset{H_0}{\overset{H_1}{\gtrless}} L \frac{P_{Fj}}{1-P_{Mj}} = t_i \qquad (2\text{-}21\text{-}a)$$

and,

$$\int_{t_i}^{\infty} p\ (\Omega_i | H_0)\ d\Omega_i = \frac{\alpha}{P_{Fj}} \qquad i \neq j \quad i,j=1,2 \qquad (2\text{-}21\text{-}b)$$

<u>"OR" Fusion Rule</u>

In this case,

$P_{000} = 1$ $\qquad$ $P_{010} = 0$

$P_{100} = 0$ $\qquad$ $P_{110} = 1$

$P_{001} = 0$ $\qquad$ $P_{011} = 0$

$P_{101} = 1$ $\qquad$ $P_{111} = 1.$

Therefore,

$C_{Mi} = P_{Mj}$

$C_{Fi} = 1-P_{Fj} \qquad i \neq j \qquad i,j=1,2$

$K_{ei} = P_{Fj}$

The two LRT's and the corresponding equations for the thresholds for the "OR" fusion rule are

$$\Omega_i(y_i) \underset{H_0}{\overset{H_1}{\gtrless}} L \frac{1 - P_{Fj}}{P_{Mj}} = t_i \qquad (2\text{-}22\text{-}a)$$

and,

$$\int_{t_i}^{\infty} p\ (\Omega_i | H_0)\ d\Omega_i = \frac{\alpha - P_{Fj}}{1 - P_{Fj}} \qquad (2\text{-}22\text{-}b)$$

2.3.2 Optimum Fusion Rule

In the previous subsection, we derived optimum decision rules at the individual detectors when the fusion rule is initially specified. The optimization criterion was the minimization of the overall probability of miss under a probability of false alarm constraint. In this subsection, we obtain the optimum fusion rule when the individual detectors (decision rules) are specified, using the same optimization criterion. We consider the binary hypothesis problem as stated before.

We consider the system structure of Figure 1.3. We assume that the conditional joint probability density function $p(y_1,y_2,y_3,\ldots,y_N|H_j)$, $j = 0,1$ is known. Each detector employs a decision rule $g_i(y_i)$ to make a local decision u_i, $i=1,2,\ldots,N$, where

$$u_i = \begin{cases} 0 & \text{if detector } i \text{ decides } H_0 \\ 1 & \text{if detector } i \text{ decides } H_1 \end{cases} \qquad (2\text{-}23)$$

The data fusion center determines the overall or global decision for the system, u, based on individual decisions i.e.,

$$u = f(u_1, u_2, \ldots, u_N) \qquad (2\text{-}24)$$

Assume that the decision rules for the individual detectors are known. As before, the probability of miss, the probability of detection and, the probability of false alarm of the overall system may be written as

$$P_M = \sum_{\underline{x}} P_{0\underline{x}} P_{\underline{x}1} \qquad (2\text{-}25\text{-a})$$

$$P_D = \sum_{\underline{x}} P_{1\underline{x}} P_{\underline{x}1} \qquad (2\text{-}25\text{-b})$$

and,

$$P_F = \sum_{\underline{x}} P_{1\underline{x}} P_{\underline{x}0} \qquad (2\text{-}25\text{-c})$$

where

$$P_{k\underline{x}} = P(u=k\,|\,\underline{x}) \qquad k = 0,1 \qquad (2\text{-}26\text{-a})$$

and

$$P_{\underline{x}j} = P(\underline{x}\,|\,H_j) \qquad j = 0,1 \qquad (2\text{-}26\text{-b})$$

As before, the function to be minimized is

$$\Gamma = P_M + L [P_F - \alpha] \qquad (2\text{-}27)$$

or equivalently, the function to be maximized is

$$F = P_D - L [P_F - \alpha] \qquad (2\text{-}28)$$

where L is the Lagrange multiplier. Substituting for P_D and P_F in (2-28), we get

$$F = \sum_{\underline{x}} P_{1\underline{x}} P_{\underline{x}1} - L [\sum_{\underline{x}} P_{1\underline{x}} P_{\underline{x}0} - \alpha] \qquad (2\text{-}29)$$

Now, expanding F as a function of a specific value of $\underline{x}$, $\underline{x}^* = (u_1^*, u_2^*, \ldots, u_N^*)$, we may rewrite (2-29) as

$$F = P_{1\underline{x}^*} P_{\underline{x}^*1} - L [P_{1\underline{x}^*} P_{\underline{x}^*0} - \alpha^{*\prime}] + K(\underline{x}^*) \qquad (2\text{-}30)$$

where

$$\alpha^{*\prime} = \alpha - \sum_{\underline{x} \neq \underline{x}^*} P_{1\underline{x}} P_{\underline{x}0} \qquad (2\text{-}31\text{-a})$$

$$K(\underline{x}^*) = \sum_{\underline{x} \neq \underline{x}^*} P_{1\underline{x}} P_{\underline{x}1} \qquad (2\text{-}31\text{-b})$$

and,

$\sum_{\underline{x} \neq \underline{x}^*}$ = summation over all possible values of $\underline{x}$ except $\underline{x}^*$.

We wish to maximize F by varying $P_{1\underline{x}^*}$ when the probability of false alarm satisfies $P_F \le \alpha$. This problem is equivalent to maximizing $P_{1\underline{x}^*} P_{\underline{x}^*1}$ under the constraint $P_{1\underline{x}^*} P_{\underline{x}^*0} \le \alpha^{*\prime}$.

Dividing (2-30) by the constant P_{x^*1}, we have

$$F^* = P_{1x^*} - L^* [P_{1x^*} - \alpha^*] + K'(\underline{x}^*) \qquad (2\text{-}32)$$

where

$$F^* = F / P_{x^*1} \qquad (2\text{-}32\text{-}a)$$

$$L^* = L \, P_{x^*o} / P_{x^*1} \qquad (2\text{-}32\text{-}b)$$

$$\alpha^* = \alpha^{*\prime} / P_{x^*o}. \qquad (2\text{-}32\text{-}c)$$

and,

$$K'(\underline{x}^*) = K(\underline{x}^*) / P_{x^*1} \qquad (2\text{-}32\text{-}d)$$

$K'(\underline{x}^*)$ is a constant in P_{1x^*}. Equation (2-32) represents the equation of a straight line in P_{1x^*}. The maximum is achieved by setting P_{1x^*} as follows

$$\begin{array}{ccc} & P_{1x^*} = \min(\alpha^*, 1) & \\ & > & \\ 1 & & L^* \\ & < & \\ & P_{1x^*} = 0 & \end{array} \qquad (2\text{-}33\text{-}a)$$

or

$$\begin{array}{ccc} & P_{1x^*} = \min(\alpha^*, 1) & \\ & > & \\ \dfrac{P_{x^*1}}{P_{x^*o}} & & L \\ & < & \\ & P_{1x^*} = 0 & \end{array} \qquad (2\text{-}33\text{-}b)$$

For each value of $\underline{x}^m$ we get one equation. Therefore, we have a total of 2^N equations. This set of equations specifies the optimum fusion rule so as to minimize the probability of miss under the probability of false alarm constraint $P_F \leq \alpha$.

Now, we present an efficient search procedure to implement the above result.

<u>Search Procedure</u>

1- Compute the quantities $P_{\underline{x}1} / P_{\underline{x}0}$ for all possible values of $\underline{x}$. Arrange these quantities in a decreasing order. Denote this ordered sequence by $\{\beta_m\}$, $m = 1,2,\ldots,2^N$ where $\beta_m = P_{\underline{x}^m 1} / P_{\underline{x}^m 0}$. Define a sequence $\{\sigma_m\}$ corresponding to $\{\beta_m\}$, where $\sigma_m = P_{\underline{x}^m 1} / \beta_m$.

2- Set $k = 1$.

3- Form the sum $S_k = \sum_{m=1}^{k} \sigma_m$.

4- Compare S_k to α.

If $S_k < \alpha$, set $k = k + 1$ and go back to step 3.

5- Set $P_{1\underline{x}^m} = 1$ for all $m \leq k-1$,

$P_{1\underline{x}^k} = \alpha_k$ and,

$P_{1\underline{x}^m} = 0$ for all $m > k$

where

$$\alpha_k = (\alpha - \text{[illegible]}_{k-1}) / P_{X_0}^{k}.$$

This procedure is illustrated with an example in the next section.

Now, if we assume that the observations at the individual detectors are conditionally independent, the solution to this problem is given by

$$P_{1X}^{*} = \min(\alpha^{*}, 1)$$

$$\frac{M_X^{*}}{F_X^{*}} \; \begin{matrix} > \\ < \end{matrix} \; L \qquad (2\text{-}34)$$

$$P_{1X}^{*} = 0$$

In the next section we present two examples, where the case of two detectors is considered. In the first one, we assume that the fusion rule is known and we solve for the optimum decision rules at the two detectors. In the second one, we solve for the best fusion rule when the detectors have already been designed.

2.4 Examples

Example 2.1

Let us assume that the observations at both detectors under the two hypotheses are exponentially distributed, i.e.,

$$p(y_i|H_0) = \begin{cases} \exp(-y_i) & y_i \geq 0 \quad i=1,2 \\ 0 & \text{otherwise} \end{cases} \tag{2-35-a}$$

$$p(y_i|H_1) = \begin{cases} (1/\Theta_i)\exp(-y_i/\Theta_i) & y_i \geq 0 \quad i=1,2 \\ 0 & \text{otherwise} \end{cases} \tag{2-35-b}$$

where the signal to noise ratio (SNR_i) $i=1,2$ is given by

$$SNR_i = \Theta_i - 1 \qquad i=1,2$$

and,

$$\Theta_i > 1$$

We should note that the above is an approximate model for several radar problems. e.g., if a square law envelope detector is used for narrowband Gaussian signal and noise, the output follows the exponential law [28, P.15-10]. The likelihood ratios at the detectors are given by

$$\Omega_i(y_i) = \frac{P(y_i/H_1)}{P(y_i/H_0)} = \frac{(1/\Theta_i)\exp(-y_i/\Theta_i)}{\exp(-y_i)}$$

$$= (1/\Theta_i)\exp[y_i(1-(1/\Theta_i))] \qquad i = 1,2 \tag{2-36-a}$$

The LRT is

$$(1/\Theta_i)\exp(y_i(1-(1/\Theta_i))) \underset{H_0}{\overset{H_1}{\gtrless}} t_i$$

or

$$y_i[1 - (1/\theta_i)] \underset{H_0}{\overset{H_1}{\gtrless}} \ln(t_i\theta_i)$$

or

$$y_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\theta_i}{\theta_i - 1} \ln(t_i\theta_i) = t'_i \qquad i=1,2. \qquad (2\text{-}36\text{-}b)$$

where t_i can be expressed as a function of t'_i as

$$t_i = (1/\theta_i) \exp(t'_i(1 - (1/\theta_i))) \qquad i=1,2 \qquad (2\text{-}36\text{-}c)$$

The probability of false alarm and the probability of miss for the two detectors are given by

$$P_{Fi} = \int_{t'_i}^{\infty} e^{-y_i} \, dy_i = e^{-t'_i}$$

$$= \exp\left[- \frac{\theta_i}{\theta_i - 1} \ln(\theta_i \ t_i)\right] = (\theta_i \ t_i)^{(\theta_i/(1-\theta_i))}$$

$$i=1,2 \qquad (2\text{-}37\text{-}a)$$

and,

$$P_{Mi} = \int_0^{t'_i} (1/\theta_i) \exp(-y_i/\theta_i)\, dy_i = 1 - \exp(-t'_i/\theta_i)$$

$$= 1-\exp\left[-\frac{1}{\theta_i - 1} \ln(\theta_i\ t_i) \right] = 1 - (\theta_i t_i)^{(1/(1-\theta_i))}$$

$$i=1,2 \qquad (2\text{-}37\text{-}b)$$

From (2-17), the two thresholds also satisfy the following sets of equations

$$P_{Fi} = \frac{\alpha - K_{Bi}}{C_{Fi}} \qquad i=1,2 \qquad (2\text{-}38)$$

and,

$$t_i = \frac{L\ C_{Fi}}{C_{Mi}} \qquad i=1,2 \qquad (2\text{-}39)$$

A simultaneous solution of the above equations yields the desired thresholds. The solution requires the knowledge of the fusion rule. Next, we consider the "OR" and "AND" fusion rules.

"OR" fusion rule

In this case,

$$C_{Fi} = 1 - P_{Fj} = 1 - \exp(-t'_j)$$

$$C_{Mi} = P_{Mj} = 1 - \exp(-t'_j/\theta_j) \qquad i \neq j \quad i,j = 1,2$$

$$K_{Bi} = P_{Fj} = \exp(-t'_j)$$

We substitute the above into (2-38), we get

$$P_{Fi} = \frac{\alpha - K_{2i}}{C_{Fi}} \qquad i=1,2$$

$$P_{Fi} = \frac{\alpha - P_{Fj}}{1 - P_{Fj}} \qquad i \neq j \quad i,j=1,2 \tag{2-40-a}$$

Also

$$t_i = \frac{L\,(1 - P_{Fj})}{P_{Mj}} \qquad i \neq j \quad i,j=1,2 \tag{2-40-b}$$

Solving (2-40-b) for L (for i =1,2) and equating the results we get

$$t_2 \frac{1 - P_{F2}}{P_{M2}} = t_1 \frac{1 - P_{F1}}{P_{M1}} \tag{2-41}$$

Substituting for t_i, P_{Fi}, and P_{Mi}, i=1,2 in (2-41) and (2-40-a), we get

$$\frac{1}{\Theta_2} \exp[t'_2(1-(1/\Theta_2))] \frac{1 - \exp(-t'_2)}{1 - \exp(-t'_2/\Theta_2)}$$

$$= \frac{1}{\Theta_1} \exp[t'_1(1-(1/\Theta_1))] \frac{1 - \exp(-t'_1)}{1 - \exp(-t'_1/\Theta_1)} \tag{2-42-a}$$

and,

$$\exp(-t'_1) = \frac{\alpha - \exp(-t'_2)}{1 - \exp(-t'_2)} \qquad (2\text{-}42\text{-}b)$$

Next, let $\exp(-t'_1) = A_1$ and $\exp(-t'_2) = A_2$, and substitute in (2-42-a) and (2-42-b), we have

$$\frac{1 - A_2}{\theta_2 (A_2^{(1-1/\theta_2)} - A_2)} = \frac{1 - A_1}{\theta_1 (A_1^{(1-1/\theta_1)} - A_1)} \qquad (2\text{-}43)$$

and,

$$A_1 = \frac{\alpha - A_2}{1 - A_2} \qquad (2\text{-}44)$$

As a numerical example, let $\theta_1=2$, $\theta_2=4$, and $\alpha = .25$ and solve for A_2. We get the value of A_2 that corresponds to the optimum solution such that

$$A_2 = 0.175$$

and, the value of A_1 is

$$A_1 = 0.0909090$$

which lead to the following values of t'_1 and t'_2

$$t'_1 = 1.742969$$

$$t'_2 = 2.3978953$$

with a value of $P_M = 0.24671711$. We show the ROC for this problem in Figure 2.1. Next, the "AND" case is presented. In this case we have a situation where the threshold is at one of the end points.

"AND" Fusion Rule

In this case, we have

$$C_{Fi} = P_{Fj} = \exp(-t'_j) \qquad i \neq j \quad i,j=1,2 \tag{2-45-a}$$

$$C_{Mi} = 1 - P_{Mj} = \exp(-t'_j/\theta_j) \qquad i \neq j \quad i,j=1,2 \tag{2-45-b}$$

From (2-39), we have

$$\frac{C_{M1}\ t_1}{C_{F1}} = \frac{C_{M2}\ t_2}{C_{F2}}$$

or

$$\frac{(1 - P_{M2})}{P_{F2}\ t_2} = \frac{(1 - P_{M1})}{P_{F1}\ t_1} \tag{2-46}$$

Substituting (2-36-c) and (2-45) into (2-46), we obtain

$$\frac{\exp(-t'_2/\theta_2)}{\exp(-t'_2)}\ \frac{\theta_2}{\exp[t'_2(1-1/\theta_2)]} = \frac{\exp(-t'_1/\theta_1)}{\exp(-t'_1)}\ \frac{\theta_1}{\exp[t'_1(11/\theta_1)]} \tag{2-47}$$

which leads to the condition $\theta_1=\theta_2$.

Since in general, $\theta_1 \neq \theta_2$, we conclude that the minimum corresponds to the end points. Solving this problem we find

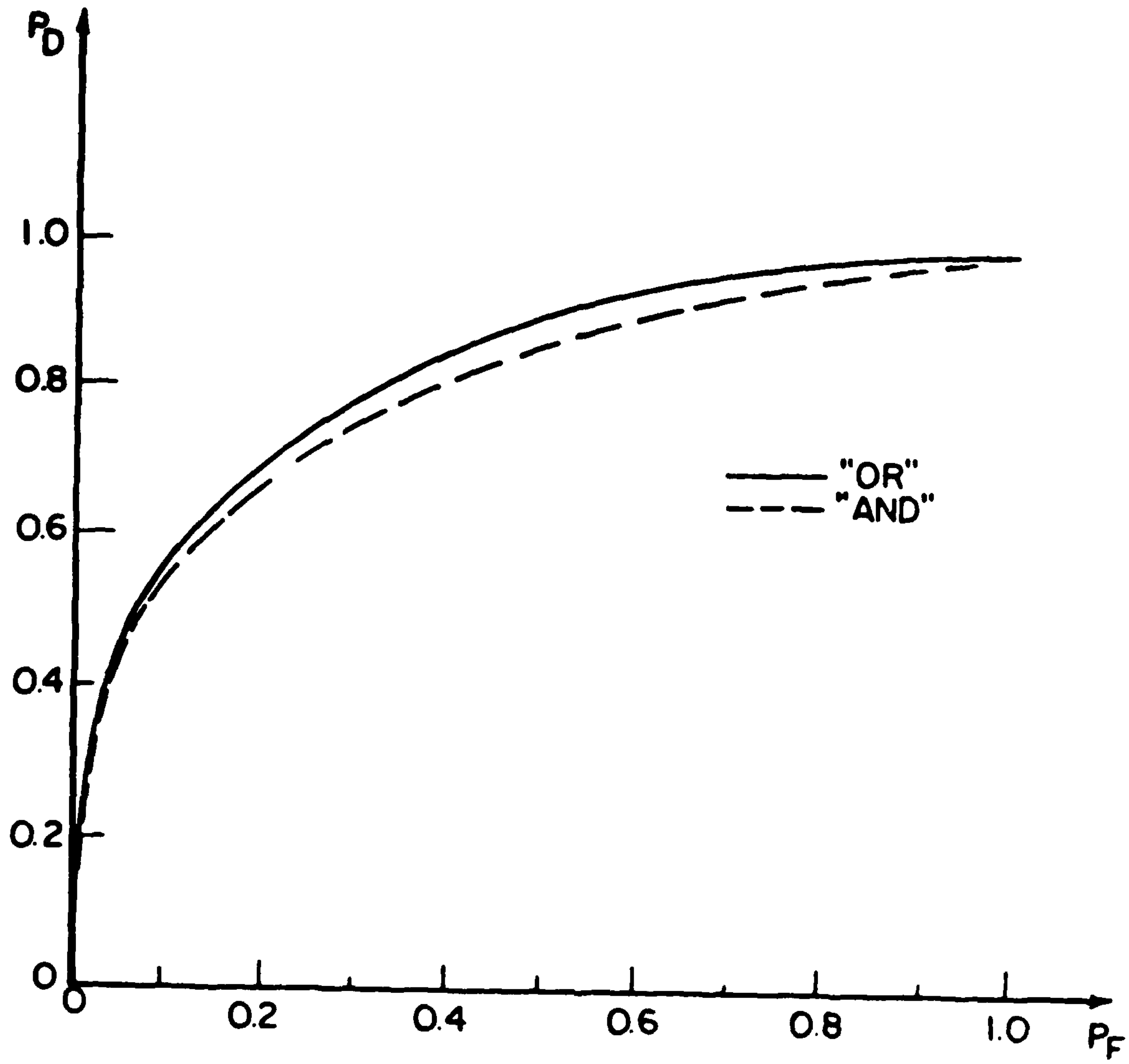


Figure 2.1 Receiver Operating Characteristic for Example 2.1

that $P_M = P_{M2}$ and, $P_F = P_{F2}$ from where we can solve for the value of t'_1 (and t'_2) . We show the ROC for the case when $\theta_1=2$ and $\theta_2=4$ in Figure 2.1.

EXAMPLE 2.2

Assume that the probabilities of false alarm and miss of the two detectors with independent observations, which have already been designed, are given by

$$P_{F1} = 0.3,\ P_{F2} = 0.4,\ P_{M1} = 0.2 \ \text{and},\ P_{M2} = 0.25.$$

We wish to find the optimum fusion rule which minimizes the probability of miss of the overall system when the system's probability of false alarm satisfies $P_F \le 0.21$.

For this example, using F_{00}, F_{01}, F_{10}, F_{11}, M_{00}, M_{01}, M_{10} and M_{11} (the quantities defined previously), the ordered sequence (β_m) and the corresponding $\underline{x}^m$ are tabulated as follows

$$\beta_1 = \frac{.8 \times .75}{.3 \times .4} = \frac{0.6}{0.12} = 5, \qquad \underline{x}^1 = (11)^T,$$

$$\beta_2 = \frac{.8 \times .25}{.3 \times .6} = \frac{0.2}{0.18} = 1.111, \qquad \underline{x}^2 = (10)^T,$$

$$\beta_3 = \frac{.2 \times .75}{.7 \times .4} = \frac{0.15}{0.28} = 0.535, \qquad \underline{x}^3 = (01)^T,$$

$$\beta_4 = \frac{.2 \times .25}{.7 \times .6} = \frac{0.05}{0.42} = 0.119 \qquad \text{and,} \qquad x^4 = (00)^T,$$

where $x^m = (i_1{}^m, i_2{}^m)^T$. We then have

$\Phi_1 = 0.12,$

$\Phi_2 = 0.18$ and,

$\Phi_3 = 0.28.$

Since the value of P_F that we want is 0.21, from the procedure outlined in the previous section we conclude that

$$P_{111} = 1,\ P_{110} = 1,\ P_{100} = 0 \ \text{ and, } P_{101} = \frac{0.21 - 0.12}{0.18} = 0.5.$$

This fusion rule yields the value of P_M to be 0.3.

2.5 Discussion

In this chapter, we have considered the signal detection problem when multiple sensors are used for surveillance and a global decision is desired. Local decisions are fed to a data fusion center where a global decision is obtained based on a given fusion rule. The Neyman-Pearson criterion for signal detection is used for system optimization. A constraint is placed on the probability of false alarm and the probability of miss of the overall system is minimized. When the fusion rule is given, the decision rules at the individual detectors were derived. The decision rules and their computation at individual detectors are coupled. We have considered the special cases of "AND" and "OR" fusion rules. We have also presented two examples for illustration. While computing the decision rules, one may obtain multiple solutions. Only the feasible solutions are to be kept. The optimum fusion rule, when the decision rules of the individual detectors are known, was also derived. It should be pointed out that solving equations (2-19) and (2-34) simultaneously, in the case of independunt observations, solves the problem of finding both, the optimum fusion rule and the optimum decision rules at the individual detectors. A solution of similar equations, in the case of dependent observations, yields the overall solution to the problem with dependent observations.

III. Distributed Bayesian Hypothesis Testing

3.1 Introduction

Bayesian hypothesis testing problem for centralized systems has been dealt with extensively in the literature [1]. Given the a priori probabilities and conditional densities of the observations for each hypothesis, fixed costs are assigned to each possible course of action. Then, optimum decision rules are derived so as to minimize the average cost. The resulting decision rule is a likelihood ratio test.

There has been some recent effort to extend the Bayesian hypothesis testing formulation to the case of distributed sensors. Tenney and Sandell [3], have solved the binary DD problem for the case of two sensors, i.e., they have treated the distributed detection problem without a fusion center. The cost assignment may reflect the effect of fusion but the design of a fusion rule has not been considered. The work has been extended by Lauer and Sandell to detection of signal waveforms in noise in [5]. They have also briefly considered some more general situations such as the dependent observations case. Sadjadi [4], extended Tenney and Sandell's work, [3], to include multiple hypotheses and more than two sensors for the DD problem.

The contribution of the work reported in this chapter is to present a generalized Bayesian formulation of the distributed detection problem. The formulation may be used to obtain solutions to both the DD problem as well as the DDF problem. Thus, previous work becomes a special case of the work reported here. In addition, we consider some special cases such as the case of independent observations and identical detectors in detail.

In Section 3.2, we formulate and solve the binary Bayesian hypothesis testing problem for the DDF system. In Section 3.3, we extend the study to the case of M hypotheses. In Section 3.4, we present some special cases. In particular, we present the solution for the case of independent observations. Special attention is paid to the case of binary hypothesis and identical detectors. Examples and some numerical results are presented in Section 3.5. Section 3.6, contains a discussion of the results obtained in this chapter.

3.2 Distributed Binary Hypothesis Testing with Data Fusion

3.2.1 Problem Statement

In this section, we consider the binary Bayesian hypothesis testing problem where we have two hypotheses H_0 and H_1. We consider the system shown in Figure 1.3. which is also reproduced as Figure 3.1. Each detector i, based on its observation vector $\underline{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})^T$ makes a

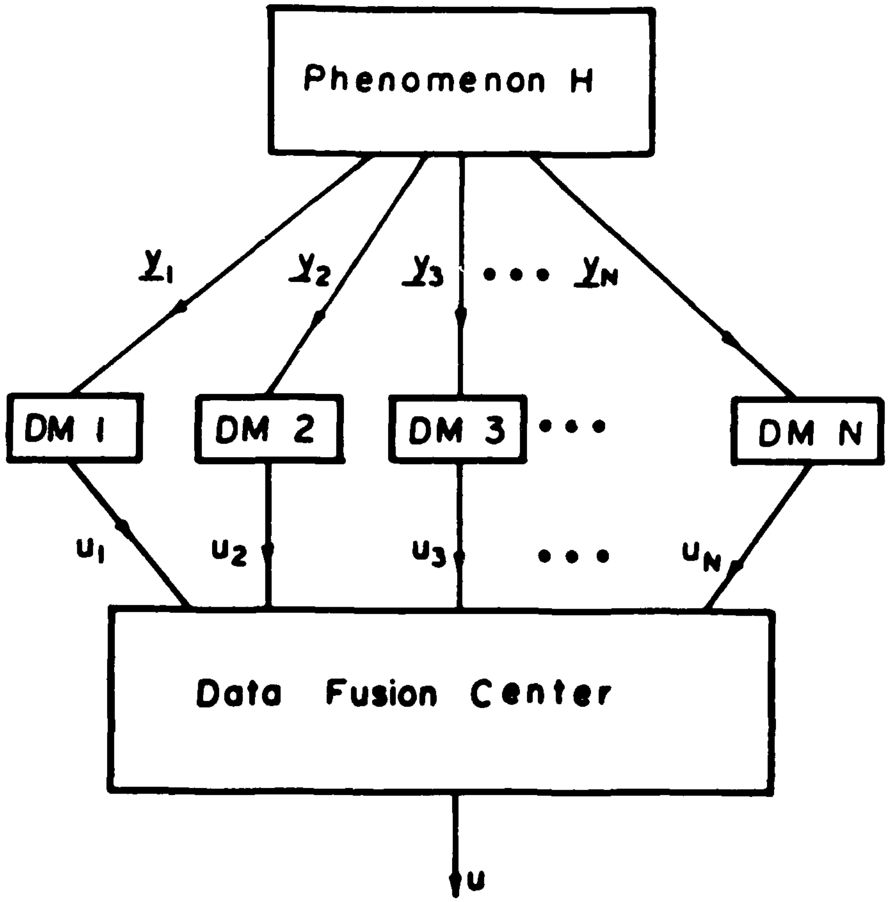


Figure 3.1 Distributed Sensor System with Data Fusion

decision u_i, $i = 1,2,\ldots,N$. Each decision, u_i, may take the value 0 or 1, depending on whether the detector i decides H_0 or H_1.

The probability of false alarm and the probability of detection of detector i are denoted by P_{Fi} and P_{Di} respectively. The overall probability of false alarm and the probability of detection are denoted by P_F and P_D respectively. The global decision, u, is made, knowing the decision vector containing the individual decisions, i.e., $\underline{x} = (u_1,u_2,\ldots,u_N)^T$. We assume that the global decision, u, depends only on the decision vector $\underline{x}$, and does not depend on the observations at the individual detectors.

The goal of this section is to develop the theory of Bayesian hypothesis testing for the DDF system, i.e., design the optimal system (obtain both the fusion rule and the decision rules at the individual decision makers) so as to minimize the average cost. The Bayes risk function that we wish to minimize can be written as

$$R_B = \sum_{i=0}^{1} \sum_{j=0}^{1} C_{ij}\, P_j\, P_r(\text{decide } H_i | H_j \text{ is present}) \qquad (3\text{-}1)$$

where C_{ij} is the cost of deciding H_i when H_j is present and P_j is the a priori probability of hypothesis H_j, $i,j = 0,1$. Throughout this chapter, it will be assumed that C_{ij} and P_j

are known. We can rewrite (3-1) as

$$R_B = C_{00}\, P_0\, Pr(u=0|H_0 \text{ is present})$$

$$+ C_{10}\, P_0\, Pr(u=1|H_0 \text{ is present})$$

$$+ C_{01}\, P_1\, Pr(u=0|H_1 \text{ is present})$$

$$+ C_{11}\, P_1\, Pr(u=1|H_1 \text{ is present}), \qquad (3\text{-}2)$$

or,

$$R_B = P_0(C_{00}(1 - P_F) + C_{10}P_F) + P_1(C_{01}(1 - P_D) + C_{11}P_D) \qquad (3\text{-}3)$$

Substituting $1-P_0$ for P_1 and rearranging (3-3), we have

$$R_B = P_F\ (\ P_0(C_{10} - C_{00}))\ + P_D\ ((1 - P_0)(C_{11} - C_{01}))$$

$$+ C_{01}(1 - P_0) + C_{00}P_0 \qquad (3\text{-}4)$$

or,

$$R_B = C_F\ P_F - C_D\ P_D + C \qquad (3\text{-}5)$$

where

$$C_F = P_0\ (\ C_{10} - C_{00}) \qquad (3\text{-}6\text{-}a)$$

$$C_D = (1 - P_0)(C_{01} - C_{11}) \qquad (3\text{-}6\text{-}b)$$

and,

$$C = C_{01}\ (1 - P_0) + C_{00}\ P_0. \qquad (3\text{-}6\text{-}c)$$

Throughout the discussion, we will assume that making a wrong decision is more costly than making a correct decision i.e.,

$$C_{10} > C_{00}$$

and, (3-7)

$$C_{01} > C_{11}$$

which imply that $C_F > 0$ and $C_D > 0$. Before we continue, we need to define the following two sets of conditional probabilities,

$$P(\underline{x}|H_j) = P(u_1,u_2,\ldots,u_N|H_j) \tag{3-8-a}$$

which is the probability of deciding u_1 at the first detector, u_2 at the second detector,..., and u_N at the N^{th} detector when H_j is present, $j,u_1,u_2,\ldots,u_N = 0,1$.

$$P(u=i|\underline{x}) = P(u=i|u_1,u_2,\ldots,u_N) \tag{3-8-b}$$

which is the probability of making a global decision i, when the individual detector decisions are $u_1,u_2,\ldots,u_N$, where $i,u_1,u_2,\ldots,u_N = 0,1$.

Thus, the probability of false alarm can be expressed as

$$P_F = \sum_{\underline{x}} P(u=1|\underline{x})\, P(\underline{x}|H_0) \tag{3-9}$$

and the probability of detection is

$$P_D = \sum_{\underline{x}} P(u=1|\underline{x})\ P(\underline{x}|H_1) \tag{3-10}$$

Substituting (3-9) and (3-10) into (3-5), we have

$$R_B = C_F \sum_{\underline{x}} P(u=1|\underline{x})\ P(\underline{x}|H_0)$$

$$- C_D \sum_{\underline{x}} P(u=1|\underline{x})\ P(\underline{x}|H_1) + C \tag{3-11}$$

Next, we proceed with the solution to the DDF problem.

3.2.2 Optimum System

First, we assume that the detectors have been designed and we obtain the fusion rule which minimizes R_B. The result is presented in Theorem 3.1. Then, we assume that the fusion rule is known and we derive decision rules at the individual detectors which again minimize R_B. This result is presented in Theorem 3.2.

Theorem 3.1

Given the decision rules at the individual detectors, the following fusion rule minimizes the risk function for the binary DDF problem

$$\frac{P(\underline{x}|H_1)}{P(\underline{x}|H_0)} \underset{P(u=1|\underline{x})=0}{\overset{P(u=1|\underline{x})=1}{\gtrless}} \frac{C_F}{C_D} \quad \text{for all } \underline{x} \tag{3-12}$$

The minimum risk, $R_{s,min}$, is

$$R_{s,min} = C - \sum_{S} [\, C_D\, P(\underline{x}|H_1) - C_F\, P(\underline{x}|H_0)] \qquad (3\text{-}13)$$

where

$$S = \{\, \underline{x} : [C_D\, P(\underline{x}|H_1) - C_F\, P(\underline{x}|H_0)] > 0 \,\} \qquad (3\text{-}14)$$

and C, C_D and C_F are as defined before.

Proof

We assume that the detectors in Figure 3.1 have already been designed, i.e., for a given observation vector $\underline{Y}$,

$$\underline{Y} = (\underline{y}_1^T, \underline{y}_2^T, \ldots, \underline{y}_N^T)^T,$$

the decision vector $\underline{x} = (u_1, u_2, \ldots, u_N)^T$, and the conditional probabilitiy densities $P(\underline{x}|H_j)$, j = 0,1, are known. Let $\underline{x}^* = (u^*_1, u^*_2, \ldots, u^*_N)^T$ be one out of the 2^N possible decision vectors. Then, R_s from equation (3-11) can be expressed as

$$R_s = P(u=1|\underline{x}^*)\, [C_F\, P(\underline{x}^*|H_0) - C_D\, P(\underline{x}^*|H_1)] + K(\underline{x}^*) \qquad (3\text{-}17)$$

where

$$K(\underline{x}^*) = \sum_{\underline{x} \neq \underline{x}^*} P(u=1|\underline{x})\, [C_F\, P(\underline{x}|H_0) - C_D\, P(\underline{x}|H_1)] + C$$

(3-18)

(3-17) is the equation of a straight line in $P(u=1|\underline{x}^*)$ where the slope is $(C_F\, P(\underline{x}^*|H_0) - C_D\, P(\underline{x}^*|H_0))$ and, $K(\underline{x}^*)$ is a constant in $P(u=1|\underline{x}^*)$. Since we have assumed that the costs C_{ij} are preassigned and the detectors have already been designed, the quantities C_F, $P(\underline{x}^*|H_0)$, C_D and $P(\underline{x}^*|H_1)$are known. $P(u=1|\underline{x}^*)$ is a probability, and it takes values in [0,1]. In order to minimize R_B, we must have

$P(u=1|\underline{x}^*) = 0$ if the slope is positive and,

$P(u=1|\underline{x}^*) = 1$ if the slope is negative.

Or,

$$C_F\, P(\underline{x}^*|H_0) - C_D\, P(\underline{x}^*|H_1) \underset{P(u=1|\underline{x}^*)=1}{\overset{P(u=1|\underline{x}^*)=0}{\gtrless}} 0 \qquad (3\text{-}19)$$

Using the cost assumption given in (3-7), (3-19) is equivalent to

$$\Omega(\underline{x}^*) = \frac{P(\underline{x}^*|H_1)}{P(\underline{x}^*|H_0)} \underset{P(u=1|\underline{x}^*)=0}{\overset{P(u=1|\underline{x}^*)=1}{\gtrless}} \frac{C_F}{C_D} \qquad (3\text{-}20\text{-a})$$

or,

$$\Omega(\underline{x}^*) \underset{u=0}{\overset{u=1}{\gtrless}} \frac{C_F}{C_D} \qquad (3\text{-}20\text{-}b)$$

which represents the fusion rule for any decision vector $\underline{x}$.

Now, we obtain the value of the minimum risk, R_{Bmin}. Substituting from (3-20) into (3-11), the risk function can be written as

$$R_{Bmin} = C + \sum_{S} (C_F\, P(\underline{x}|H_0) - C_D\, P(\underline{x}|H_1)) \qquad (3\text{-}21\text{-}a)$$

or,

$$R_{Bmin} = C - \sum_{S} (C_D\, P(\underline{x}|H_1) - C_F\, P(\underline{x}|H_0)) \qquad (3\text{-}21\text{-}b)$$

where

$$S = \{ \underline{x} \; : \; \Omega(\underline{x}) > C_F/C_D \}.$$

We remark that for all $\underline{x}$ in S $(C_D P(\underline{x}|H_1) - C_F\, P(\underline{x}|H_0))$ is nonnegative and therefore, $R_{Bmin} \le C$.

Q.E.D.

Theorem 3.2

Given the fusion rule $u = f(u_1,u_2,\ldots,u_N)$ and the conditional densities $p(\underline{y}_1,\underline{y}_2,\ldots,\underline{y}_N|H_j)$, $j = 0,1$, the decision rule at the μ^{th} detector, $\mu = 1,2,\ldots,N$, which minimizes R_m is given by

$$p(\underline{y}_\mu|H_1) \sum_{\underline{X}^\mu} \int_{\underline{Y}^\mu} d\underline{Y}^\mu \, A_{\underline{X}}^\mu \, P(\underline{X}^\mu|\underline{Y}^\mu) \, C_D \, p(\underline{Y}^\mu|\underline{y}_\mu,H_1) \underset{u_\mu=0}{\overset{u_\mu=1}{\gtrless}}$$

$$p(\underline{y}_\mu|H_0) \sum_{\underline{X}^\mu} \int_{\underline{Y}^\mu} d\underline{Y}^\mu \, A_{\underline{X}}^\mu \, P(\underline{X}^\mu|\underline{Y}^\mu) \, C_F \, p(\underline{Y}^\mu|\underline{y}_\mu,H_1)$$

(3-22-a)

where

$$\underline{Y}^\mu = (\underline{y}_1{}^T, \underline{y}_2{}^T, \ldots, \underline{y}_{\mu-1}{}^T, \underline{y}_{\mu+1}{}^T, \ldots, \underline{y}_N{}^T)^T, \qquad (3\text{-}22\text{-}b)$$

$$A_{\underline{X}}^\mu = P(u=1|\overset{1}{\underline{x}}_\mu) - P(u=1|\overset{0}{\underline{x}}_\mu), \qquad (3\text{-}22\text{-}c)$$

and,

$$\overset{j}{\underline{x}}_\mu = (u_1, u_2, \ldots, u_\mu=j, \ldots, u_N)^T. \qquad (3\text{-}22\text{-}d)$$

<u>Proof</u>

We can expand R_m given in (3-11) in terms of u_μ, the decision of a specific detector μ, μ = 1,2,...,N. We then have

$$R_m = C + \sum_{\substack{u_1 \ldots u_N \\ \text{except } u_\mu}} \!\!\cdots\!\! \sum \{(P(u=1|u_1,\ldots,u_\mu=1,\ldots,u_N) \times$$

$$[C_F P(u_1,\ldots,u_\mu=1,\ldots,u_N|H_0) - C_D P(u_1,\ldots,u_\mu=1,\ldots,u_N|H_1)]$$

$$+(P(u=1|u_1,\ldots,u_\mu=0,\ldots,u_N) \times$$

$$[C_F P(u_1,\ldots,u_\mu=0,\ldots,u_N|H_0) - C_D P(u_1,\ldots,u_\mu=0,\ldots,u_N|H_1)]\})$$

(3-23)

For notational convenience, we define the following

$$\underline{x}_\mu^{j} = (u_1,\ldots,u_\mu=j,\ldots,u_N)^T = \underline{x}\Big|_{u_\mu=j} \qquad j = 0,1 \qquad \text{(3-23-a)}$$

$$\underline{x}^\mu = (u_1,\ldots,u_{\mu-1},u_{\mu+1},\ldots,u_N)^T, \qquad \text{(3-23-b)}$$

and,

$\sum\limits_{\underline{x}^\mu}$ ≡ summation over all possible values of u_i such that $i \neq \mu$.

(3-23-c)

Then, (3-23) becomes

$$R_m = C + \sum_{\underline{x}^\mu}\Big(P(u=1|\underline{x}_\mu^1)\,[C_F\, P(\underline{x}_\mu^1|H_0) - C_D\, P(\underline{x}_\mu^1|H_1)]$$

$$+ P(u=1|\underline{x}_\mu^0)\,[C_F\, P(\underline{x}_\mu^0|H_0) - C_D\, P(\underline{x}_\mu^0|H_1)]\Big)$$

(3-24)

Using the relationship $P(\underline{x}_\mu^0|H_j) = P(\underline{x}^\mu|H_j) - P(\underline{x}_\mu^1|H_j)$, in (3-24), we rewrite R_m as

$$R_m = C + \sum_{\underline{x}^\mu}\Big(P(u=1|\underline{x}_\mu^0)[C_F\, P(\underline{x}^\mu|H_0) - C_D\, P(\underline{x}^\mu|H_1)]$$

$$+[P(u=1|\underline{x}_\mu^1) - P(u=1|\underline{x}_\mu^0)][C_F\, P(\underline{x}_\mu^1|H_0) - C_D P(\underline{x}_\mu^1|H_1)]\Big)$$

(3-25)

Let

$$A_{\underline{x}^M} = P(u=1|\overset{1}{\underline{x}_M}) - P(u=1|\overset{0}{\underline{x}_M}) \tag{3-26-a}$$

and,

$$C_M = \sum_{\underline{x}^M} P(u=1|\overset{0}{\underline{x}_M}) \, [C_F \, P(\underline{x}^M|H_0) - C_D \, P(\underline{x}^M|H_1)] + C \tag{3-26-b}$$

Then, R_m may be written as,

$$R_m = \sum_{\underline{x}^M} A_{\underline{x}^M}[C_F \, P(\overset{1}{\underline{x}_M}|H_0) - C_D \, P(\overset{1}{\underline{x}_M}|H_1)] + C_M \tag{3-27}$$

But

$$P(\underline{x}|H_j) = \int_{\underline{Y}} P(\underline{x}|\underline{Y}) . \, p(\underline{Y}|H_j) \, d\underline{Y} \tag{3-28}$$

where

$$\int_{\underline{Y}} \quad \text{is the integral over all possible elements of } \underline{Y} \tag{3-29}$$

Recall our earlier assumption that the decision of each detector depends only on its own observation. Therefore, we may write

$$P(\underline{x}|\underline{Y}) = \prod_{i=1}^{N} P(u_i|y_i) \tag{3-30-a}$$

and,

$$P(x_\mu^j|Y) = P(u_\mu=j|y_\mu) \cdot P(\underline{x}^\mu|Y^\mu) \qquad j = 0,1 \qquad \text{(3-30-b)}$$

Substituting (3-28), (3-29), and (3-30) into (3-27), we have

$$R_m = \int_{\underline{y}_\mu} dy_\mu \; P(u_\mu=1|y_\mu) \; \{ \sum_{\underline{x}^\mu} \int_{\underline{Y}^\mu} dY^\mu \; A_{\underline{x}^\mu} \; P(\underline{x}^\mu|Y^\mu)$$

$$\times \; [C_F \; p(\underline{Y}|H_0) - C_D \; p(\underline{Y}|H_1)]\}$$

(3-31)

At this point, we assume that all the detectors except the μ^{th} detector have been designed and obtain the decision rule at the μ^{th} detector so as to minimize R_m. This procedure will be repeated for all of the N detectors. A simultaneous solution of the resulting N nonlinear equations will yield the desired result. From (3-31), we obtain the decision rule at the μ^{th} detector as

$$\sum_{\underline{x}^\mu} \int_{\underline{Y}^\mu} A_{\underline{x}^\mu} \; P(\underline{x}^\mu|\underline{Y}^\mu)[C_F \; p(\underline{Y}|H_0) - C_D \; p(\underline{Y}|H_1)] \; dY^\mu \; \underset{P(u_\mu=1|\underline{y}_\mu)=1}{\overset{P(u_\mu=1|\underline{y}_\mu)=0}{\gtrless}} \; 0$$

(3-32)

But

$$p(Y|Hj) = p(\underline{Y}^\mu|y_\mu,H_j) \cdot p(y_\mu|H_j) \qquad \text{(3-33)}$$

Substituting from (3-33) into (3-32), we get the following decision rule

$$p(y_M|H_1) \sum_{\underline{X}^M} \int_{\underline{Y}^M} A_{X^M} \; C_D \; P(\underline{X}^M|\underline{Y}^M) \; p(\underline{Y}^M|y_M,H_1) \; d\underline{Y}^M \;\; \overset{u_M=1}{\underset{u_M=0}{\gtrless}}$$

$$p(y_M|H_0) \sum_{\underline{X}^M} \int_{\underline{Y}^M} A_{X^M} \; C_F \; P(\underline{X}^M|\underline{Y}^M) \; p(\underline{Y}^M|y_M,H_0) \; d\underline{Y}^M \qquad (3\text{-}34)$$

It is clear from (3-34) that $P(u_M|y_M)$ can be either 0 or 1. Thus, we have a non randomized decision rule for each detector.

Q.E.D.

Corollary 3.1

The overall solution to the binary Bayesian DDF problem, i.e., obtaining the decision rules at the individual detectors and the fusion rule which jointly minimize R_B, is obtained by solving N equations of the form (3-34) and 2^N equations of the form (3-12) simultaneously. We remark that the above equations are coupled and nonlinear. It is a difficult task to solve them when N becomes large. The case of independent observations is discussed later in Section 3.4. In this case, the equations are easier to deal with and the decision rules at the individual detectors become classical likelihood ratio tests.

In the next section, we generalize the results to the case of N detectors and M hypotheses.

3.3 Generalization to M-ary Hypothesis Testing

3.3.1 Problem Statement

In this section, we consider the Bayesian DDF problem illustrated in Figure 3.1, with N decision makers and M hypotheses, $H_0, H_1, \ldots, H_{M-1}$. Each decision maker i, uses its observation vector $\underline{y}_i$ to make a decision u_i, $i=1,2,\ldots,N$. Depending on whether the i^{th} decision maker decides H_0, $H_1, \ldots,$ or H_{M-1}, u_i may take the values $0,1,\ldots,$ or $M-1$ respectively. Knowing the decision vector $\underline{x}$, which contains the decisions of the individual decision makers, i.e., $\underline{x} = (u_1, u_2, \ldots, u_N)^T$, the global decision, u, is made. This global decision may again take the values $0,1,2,\ldots,$ or $M-1$. It does not depend on the observations at the individual decision makers. Here, we develop the theory of Bayesian hypothesis testing for the generalized DDF system, i.e., obtain both, the fusion rule and the decision rules at the individual decision makers so as to minimize the average cost R_B. Now, we state a lemma that we will use for the solution to our problem in the next subsection.

Lemma 3.1

Let $F = \sum_{i=0}^{M-1} P_i C_i$ where C_i are known positive constants and P_i satisfy $\sum_{i=0}^{M-1} P_i = 1$ and, $P_i \geq 0$. The minimum value of F, F_{min}, is equal to C_{min}, where

$$C_{min} = \mathrm{Min}\,(C_0, C_1, \ldots, C_{M-1})$$

This minimum F_{min} is achieved by setting

$$P_i = \begin{cases} 1 & \text{for } C_i = C_{min} \\ 0 & \text{otherwise} \end{cases}$$

3.3.2 Optimum System

The generalized result is stated in the following theorem

Theorem 3.3

Assume that the costs C_{ij}, $i,j = 0,1,\ldots,M-1$ and the a priori probabilities are known. For the Bayesian DDF system of Figure 3.1, the optimum fusion rule and the optimum decision rules at the detectors are obtained by solving the following sets of equations simultaneously

$$P(u=m|\underline{x}^*) = \begin{cases} 1 & \text{if } C_m^* < C_i^* \text{ for all } i,\ i=0,1,\ldots,M-1, \\ & \text{and } i \neq m. \\ 0 & \text{otherwise.} \end{cases}$$

for all $\underline{x}^*$, $\underline{x}^* = (u_1^*, u_2^*, \ldots, u_N^*)^T$, $u_k^* = 0,1,\ldots,M-1$

and,

$$P(u_\mu = h \mid \underline{y}_\mu) = \begin{cases} 1 & \text{if } I_\mu(h) \le I_\mu(m) \text{ for all } m, \text{ such} \\ & \text{that } h \ne m \text{ and } h,m = 0,1,\ldots,M-1 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu = 1,2,\ldots,N$$

where

$$C_i^* = \sum_{j=0}^{M-1} C_{ij}\, P(\underline{x}^* \mid H_j)\, P(H_j)$$

and, $I_\mu(h)$ is given by

$$I_\mu(h) = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} p(\underline{y}_\mu \mid H_j) \int_{\underline{Y}^\mu} d\underline{Y}^\mu\, P_j C_{ij}$$

$$\times \sum_{\underline{x}^\mu} [P(u=i \mid \overset{j}{\underline{x}_\mu}) - P(u=i \mid \overset{h}{\underline{x}_\mu})]\, P(\underline{x}^\mu \mid \underline{Y}^\mu)\, p(\underline{Y}^\mu \mid \underline{y}_\mu, H_j)$$

Proof

In this case, we write the Bayesian risk that we wish to minimize as

$$R_B = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} C_{ij}\, P_j\, \Pr(\text{decide } H_i \mid H_j \text{ is present})$$

or

$$R_B = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P(u=i \mid H_j)\, P_j\, C_{ij} \tag{3-35}$$

where again, C_{ij} is the cost of deciding H_i when H_j is present and, P_j is the a priori probability of hypothesis H_j, $i,j = 0,1,\ldots,M-1$. Recall our assumption that the global decision, u, does not depend on the observations. It only depends on $\underline{x}$. Therefore, we may write

$$P(u=i|H_j) = \sum_{\underline{x}} Pr(u=i|\underline{x})\, Pr(\underline{x}|H_j) \quad . \tag{3-36}$$

Substituting (3-36) into (3-35) , we have

$$R_m = \sum_{\underline{x}} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P(u=i|\underline{x})\, P(\underline{x}|H_j)\, P(H_j)\, C_{ij} \tag{3-37}$$

Next, we proceed with the solution to our problem. First, obtain the optimum fusion rule which minimizes R_m.

Optimum Fusion Rule

While deriving the fusion rule, we assume that the decision makers have already been designed. Therefore, $P(\underline{x}|H_j)$ is known for all possible decision vectors $\underline{x}$, $\underline{x} = (u_1,u_2,\ldots,u_N)^T$. Let $x^* = (u_1^*,u_2^*,\ldots,u_N^*)^T$ be one of the M^N possible decision vectors. Separating the terms which depend on $\underline{x}^*$ in (3-37), R_m may be expressed as

$$R_m = \sum_{i=0}^{M-1} P(u=i|\underline{x}^*) \sum_{j=0}^{M-1} C(u=i,H_j)\, P(\underline{x}^*|H_j)\, P(H_j)$$

$$+ \sum_{i=0}^{M-1} \sum_{\underline{x}\neq\underline{x}^*} P(u=i|\underline{x}) \sum_{j=0}^{M-1} C_{ij}\, P(\underline{x}|H_j)\, P(H_j) \tag{3-38}$$

In equation (3-38), C_{ij}, $P(H_j)$ and $P(\underline{x}|H_j)$ are known for all i,j such that $i,j = 0,1,\ldots,M-1$ and all vectors $\underline{x}$. Therefore, the second term does not depend on $P(u=i|\underline{x}^*)$ and, minimizing R_m^* with respect to $P(u=\ |\underline{x}^*)$ minimizes R_m, where

$$R_m^* = \sum_{i=0}^{M-1} P(u=i|\underline{x}^*) \sum_{j=0}^{M-1} C_{ij}\, P(\underline{x}^*|H_j)\, P(H_j) \qquad (3\text{-}39)$$

Now, applying the result of Lemma 3.1 to our problem, where

$$C_i = \sum_{j}^{M-1} C_{ij}\, P(\underline{x}^*|H_j)\, P(H_j) \qquad (3\text{-}40)$$

$R_{m\,min}$ is achieved by setting $P(u=m|x^*) = 1$ for the value m of u for which C_i^* is minimum. In this case the value of $R_{m\,min}$ is equal to C_m^*. Thus, the optimal fusion rule is

$$P(u=m|\underline{x}^*) = \begin{cases} 1 & \text{if } C_m^* \le C_i^* \text{ for all } i \text{ such that} \\ & i = 0,1,\ldots,M-1 \text{ and } i \neq m \\ 0 & \text{otherwise} \end{cases} \qquad (3\text{-}41)$$

This last equation states that, for any decision vector $\underline{x}^* = (u_1^*, u_2^*, \ldots, u_N^*)$, the fusion center decides that H_m is present where $C_m^* = \min(C_0^*, C_1^*, \ldots, C_{M-1}^*)$.

Next, we obtain the decision strategies at the individual decision makers. We obtain the decision strategy at the μ^{th} detector, $\mu = 1,2,\ldots,N$. While deriving the decision strategy at the μ^{th} detector, $\mu = 1,2,\ldots,N$, we assume that all the other detectors have already been designed.

Optimum Detectors

Recall that R_B is given by (3-37). We reproduce it here

$$R_B = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P_j C_{ij} \sum_{\underline{x}} [P(u=i|\underline{x}) P(\underline{x}|H_j)] \tag{3-37}$$

In terms of μ, the μ^{th} detector, R_B may be rewritten as

$$R_B = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P_j C_{ij} \sum_{k=0}^{M-1} \sum_{\overset{k}{\underline{x}}_\mu} [P(u=i|\overset{k}{\underline{x}}_\mu) P(\overset{k}{\underline{x}}_\mu|H_j)] \tag{3-42}$$

where

$$\underline{x}^\mu = (u_1, u_2, \ldots, u_{\mu-1}, u_{\mu+1}, \ldots, u_N)^T \tag{3-43-a}$$

$$\overset{k}{\underline{x}}_\mu = (u_1, u_2, \ldots, u_\mu = k, \ldots, u_N)^T \tag{3-43-b}$$

and,

$$\sum_{\overset{k}{\underline{x}}_\mu} \equiv \text{summation over all possible values of } \overset{k}{\underline{x}}_\mu. \tag{3-43-c}$$

Recall the assumption that, each detector's decision depends only on its observation and that it does not depend on the hypothesis present. Therefore, we write

$$P(\underline{x}|H_j) = \int_{\underline{Y}} P(\underline{x}|\underline{Y})\, p(\underline{Y}|H_j)\, d\underline{Y} \tag{3-44}$$

which, using our earlier assumption that decisions at different decision makers are independent, becomes

$$P(\underline{x}|H_j) = \int_{\underline{Y}} \prod_{m=0}^{M-1} P(u_m|y_m)\; p(\underline{Y}|H_j)\; d\underline{Y} \qquad (3\text{-}45)$$

similarly,

$$P(\overset{k}{\underline{x}_M}|H_j) = \int_{\underline{y}_M} dy_M \; P(u_M{=}k\,|\,\underline{y}_M)p(\underline{y}_M|H_j)$$

$$= \int_{\underline{Y}^M} d\underline{Y}^M \; P(\underline{x}^M|\underline{Y}^M)\; p(\underline{Y}^M|\underline{y}_M,H_j) \qquad (3\text{-}46)$$

where Y^M and, $\int_{\underline{Y}^M}$ were defined in (3-2) and (3-29) respectively and,

$$P(\underline{x}^M|\underline{Y}^M) = \prod_{\substack{m=0\\ m\neq M}}^{M-1} P(u_m|\underline{y}_m) \qquad (3\text{-}47)$$

Substituting from (3-46) into (3-42) and interchanging summation and integration over y_M, we may rewrite R_m as

$$R_m = \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \int_{\underline{y}_M} P(u_M=k|\underline{y}_M)\; p(\underline{y}_M|H_j)$$

$$\times \sum_{i=0}^{M-1} P_j\, C_{ij} \sum_{\underline{x}^M} P(u=i|\underline{x}_M)$$

$$\times \int_{\underline{Y}^M} P(\underline{x}^M|\underline{Y}^M)\; p(\underline{Y}^M|\underline{y}_M,H_j)\; dY \tag{3-48}$$

We have the following relationship,

$$\int_{\underline{y}_M} d\underline{y}_M\; P(u_M=k|\underline{y}_M)\; p(\underline{y}_M|H_j) \int_{\underline{Y}^M} dY^M\; P(\underline{x}^M|Y^M)\; p(\underline{Y}^M|\underline{y}_M,H_j)$$

$$= \begin{cases} \displaystyle\int_{\underline{y}_M} P(u_M=k|\underline{y}_M)\; p(\underline{y}_M|H_j) \int_{\underline{Y}^M} P(\underline{x}^M|\underline{Y}^M)\; p(\underline{Y}^M|\underline{y}_M,H_j)\; dY & \text{if } k \neq j \\[2ex] \displaystyle\int_{\underline{y}_M} \Big(1 - \sum_{\substack{h=0 \\ h\neq j}}^{M-1} P(u_M=h|\underline{y}_M)\Big)\; p(\underline{y}_M|H_j) \int_{\underline{Y}^M} P(\underline{x}^M|\underline{Y}^M) p(\underline{Y}^M|\underline{y}_M,H_j)\; dY & \text{if } k = j \end{cases} \tag{3-49}$$

Using the above equation, R_m can be expressed as

$$R_m = K(h) + R_m(h) \tag{3-50}$$

where

$$K(h) = \sum_{j=0}^{M-1} \int_{\underline{y}_M} p(\underline{y}_M|H_j) \sum_{i=0}^{M-1} P_j\, C_{ij} \sum_{\underline{x}^M} P(u=i|\underline{x}_M^{j})$$

$$\times \int_{\underline{Y}^M} d\underline{Y}\; P(\underline{x}^M|\underline{Y}^M)\; p(\underline{Y}^M|y_M,H_j) \qquad (3\text{-}51\text{-}a)$$

and,

$$R_m(h) = \sum_{j=0}^{M-1} \sum_{\substack{h=0 \\ h\neq j}}^{M-1} \int_{y_M} P(u_M=h|y_M)\; p(\underline{y}_M|H_j) \sum_{i=0}^{M-1} P_j\, C_{ij}$$

$$\times \left(\sum_{x^M} [P(u=i|x_M^{j})-P(u=i|\underline{x}_M^{h})] \right) \int_{Y^M} d\underline{Y}\; P(\underline{x}^M|\underline{Y}^M)\; p(\underline{Y}^M|\underline{y}_M,H_j)$$

$$(3\text{-}51\text{-}b)$$

We remark that, $K(h)$ is a constant independent of $P(u_M=h|\underline{y}_M)$. Then, minimizing R_m with respect to $P(u_M=h|\underline{y}_M)$ is equivalent to minimizing $R_m(h)$. we define $I_M(h)$ as

$$I_M(h) = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} p(\underline{y}_M|H_j) \int_{\underline{Y}^M} d\underline{Y}^M\; P_j C_{ij}$$

$$\times \sum_{\underline{x}^M} [P(u=i|\underline{x}_M^{j})-P(u=i|\underline{x}_M^{h})]\; P(\underline{x}^M|\underline{Y}^M)\; p(\underline{Y}^M|\underline{y}_M,H_j)$$

$$(3\text{-}52)$$

Substituting from (3-52) into (3-51), we have

$$R_m(h) = \int_{y_\mu} dy_\mu \sum_{\substack{h=0 \\ h \neq j}}^{M-1} P(u_\mu = h \mid y_\mu)\, I_\mu(h) \tag{3-53}$$

Therefore, it is possible to minimize $R_m(h)$, and consequently R_m, by choosing the following decision strategy for the decision maker μ, $\mu=1,2,\ldots,N$,

$$P(u_\mu = h \mid y_\mu) = \begin{cases} 1 & \text{if } I_\mu(h) \leq I_\mu(m) \text{ for all } m \text{ such that} \\ & h \neq m \text{ and } h,m = 0,1,\ldots,M-1 \\ 0 & \text{otherwise} \end{cases} \tag{3-54}$$

Repeating this procedure for all of the N decision makers, a total of N equations is obtained. Simultaneous solution of these equations yield the decision strategies at all decision makers which minimize R_m.

Again, note that these decision strategies are nonrandomized. The equations are coupled and highly nonlinear. Therefore, the solution is not straightforward. This problem becomes easier if we assume that observations at all decision makers are independent of each other. This latter case is discussed in the next section.

Overall Solution

The overall solution to the M-hypothesis Bayesian DDF problem with N decision makers, i.e., obtaining the decision rules at the individual decision makers and the fusion rule which minimize R_B, is obtained by solving N equations of the form (3-54) and M^N equations of the form (3-41) simultaneously.

Q.E.D.

In the next section, we consider the Bayesian DDF problem when the observations at the decision makers are independent of each other, and some other special cases.

3.4 Special Cases

In the previous section, we obtained the solution to the DDF problem with M hypotheses and N decision makers so as to minimize the Bayesian risk. The solution consisted of M^N + N equations which must be solved simultaneously to yield the decision rules at the individual decision makers and the fusion rule. We noted the computational difficulty in obtaining the solution due to the coupling and nonlinearity of the equations. This computational difficulty is reduced considerably by invoking the independence assumption on the observations i.e.,

$$p(\underline{Y}|H_j) = \prod_{i=1}^{N} p(\underline{y}_i|H_j) \qquad (3\text{-}55)$$

Next, we present the solution to the Bayesian DDF problem with independent observation.

3.4.1 Independent Observations

If the observations at the decision makers are independent, the detectors become threshold detectors and the solution to the problem requires solving the following equations

$$P(u=m|x^*) = \begin{cases} 1 & \text{if } C_m^* < C_i^* \text{ for all } i \text{ such that } i \neq m \text{ and } m,i = 0,1,\ldots,M-1 \\ 0 & \text{otherwise} \end{cases} \quad (3\text{-}56\text{-}a)$$

for all $\underline{x}^*$, $\underline{x}^* = (u_1^*, u_2^*, \ldots, u_N^*)^T$, $\quad u_k^* = 0,1,\ldots,M-1$

and,

$$P(u_\mu = h|\underline{y}_\mu) = \begin{cases} 1 & \text{if } I_\mu(h) \leq I_\mu(m) \text{ for all } m \text{ such that } h \neq m \text{ and } h,m = 0,1,\ldots,M-1 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu = 1,2,\ldots,N \quad (3\text{-}56\text{-}b)$$

where

$$C_i^* = \sum_{j=0}^{M-1} C_{ij} \prod_{k=1}^{N} P(u_k^*|H_j)\, P(H_j) \quad (3\text{-}57\text{-}a)$$

and,

$$I_\mu(h) = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} p(y_\mu|H_j)\, P_j\, C_{ij}$$

$$* \sum_{x^\mu} [P(u=i|x_\mu^j) - P(u=i|x_\mu^h)] \prod_{\substack{m=1 \\ m \neq \mu}}^{N} \int_{y_\mu} P(u_m|y_m)\, p(y_m|H_j)\, dy_m \tag{3-57-b}$$

<u>Binary Case</u>

In the case of binary hypothesis testing, we may use the procedure developed in Chapter 2 and get, the following likelihood ratio tests

$$\Omega_\mu(y_\mu) = \frac{p(y_\mu|H_1)}{p(y_\mu|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} L_\mu = \frac{C_{F\mu}}{C_{M\mu}} = t_\mu$$

$$\mu = 1,2,\ldots,N \tag{3-58}$$

and,

$$\frac{D_{\underline{x}}}{F_{\underline{x}}} \underset{P(u=1|\underline{x})=0}{\overset{P(u=1|\underline{x})=1}{\gtrless}} \frac{C_F}{C_D} \qquad \text{for all } \underline{x} \tag{3-59}$$

where

$$C_{F\mu} = P_0\, (C_{10} - C_{00}) \sum_{\underline{x}} F_{\underline{x}}^\mu\, (P_{1\underline{x}}^{\mu 1} - P_{1\underline{x}}^{\mu 0}) \tag{3-60-a}$$

$$C_{M\mu} = P_1\, (C_{01} - C_{11}) \sum_{\underline{x}} M_{\underline{x}}^\mu\, (P_{0\underline{x}}^{\mu 0} - P_{0\underline{x}}^{\mu 1}) \tag{3-60-b}$$

$$F_x^{\mu} = \prod_{S_0^{\mu}} (1 - P_{Fj}) \prod_{S_1^{\mu}} P_{Fk} \qquad (3\text{-}60\text{-}c)$$

$$M_x^{\mu} = \prod_{S_0^{\mu}} P_{Mj} \prod_{S_1^{\mu}} (1 - P_{Mk}) \qquad (3\text{-}60\text{-}d)$$

$$D_x = \prod_{S_0} (1 - P_{Dj}) \prod_{S_1} P_{DK} \qquad (3\text{-}60\text{-}e)$$

$$F_x^{*} = \prod_{S_0} (1 - P_{Fj}) \prod_{S_1} P_{Fk} \qquad (3\text{-}60\text{-}f)$$

$P_{jx}^{\mu k}$ = probability of deciding j when the decision of detector μ is k and $\underline{x}$ is given, k,j = 0,1 and, μ = 1,2,...,N (3-60-g)

and,

S_0, S_1, S_0^{μ} and S_1^{μ} are given by (2-6) and (2-12)

Note that the solution of the binary Bayesian DDF problem requires the solution of $2^N + N$ equations. There are 2^{2^N} possible fusion rules and we must select one out of these. It is clear that the number of fusion rules grows very rapidly with N. In order to make the problem tractable, we next obtain the solution for the case of identical detectors with independent observation i.e.,

$$p(Y|H_j) = \prod_{i=1}^{N} p(y_i|H_j) \qquad (3\text{-}61\text{-}a)$$

and,

$$p(y_k|H_j) = p(y_h|H_j) \qquad \text{for all } h,k;\ h,k = 1,2,\ldots,N$$
$$j = 0,1$$
$$(3\text{-}61\text{-}b)$$

3.4.2 Identical Detectors

In this subsection, we continue with the binary hypothesis testing problem. We assume that the detectors are identical with independent observations (3-61-a) and (3-61-b) and obtain the optimum decision rules at the detectors and the optimum fusion rule so as to minimize R_B.

Since the detectors are identical, it is obvious that all of the detectors should have the same thresholds and, consequently, the same values of probability of detection and probability of false alarm. i.e., $P_{Di} = P_{Dj}$ and $P_{Fi} = P_{Fj}$ for all i,j such that i,j = 1,2,...,N. Thus,

$$P_{1\underline{x}_i} = P_{1\underline{x}_j} \qquad \text{for all } \underline{x}_i \text{ and } \underline{x}_j \text{ such that } \underline{x}_i \text{ and } \underline{x}_j \text{ have the same number of ones.} \qquad (3\text{-}62)$$

For notational convenience, we define

$$Q_j = \text{Prob(decide } u=1|\underline{x} \text{ contains } j \text{ ones)} \quad j = 0,1,\ldots,N \qquad (3\text{-}63)$$

Therefore, for a given value of the threshold at the detectors, the fusion rule is given by the following set of equations,

$$[(1-p_D)/(1-p_F)]^N \underset{Q_0=0}{\overset{Q_0=1}{\gtrless}} L \qquad \text{(stage 0)}$$

$$(p_D/p_F)\,[(1-p_D)/(1-p_F)]^{(N-1)} \underset{Q_1=0}{\overset{Q_1=1}{\gtrless}} L \qquad \text{(stage 1)}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$(p_D/p_F)^i\,[(1-p_D)/(1-p_F)]^{(N-i)} \underset{Qi=0}{\overset{Q_i=1}{\gtrless}} L \qquad \text{(stage N-i)}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$(p_D/p_F)^N \underset{Q_N=0}{\overset{Q_N=1}{\gtrless}} L \qquad \text{(stage N)}$$

(3-64)

where p_D and p_F are the probability of detection and the probability of false alarm of the individual decision makers respectively, and $L = (C_F/C_D)$.

Since we assumed that the threshold is known, p_D and p_F are also known. We claim that the optimum fusion rule in this special case reduces to a "K out of N" fusion rule. Next, we justify this choice of the fusion rule and obtain the optimum value of K.

"K out of N" fusion rule

It is well known that the receiver operating characteristic curves (ROC's) of optimal detection systems are convex [12]. The ROC's of the individual detectors in an optimal system are also convex. Moving down in the set of equations describing the fusion rule, (3-64) i.e., from stage i to stage $i+1$, $i = 0,1,\ldots,N-1$, we notice that we are multiplying the left hand side term by $(p_D/p_F)(1-p_F)/(1-p_D)$. From Figure 3.2-a, we see that $(p_D/p_F) \geq 1$, and from Figure 3.2-b, we observe that $(1-p_D)/(1-p_F) \leq 1$. Therefore, $(p_D/p_F)[(1-p_F)/(1-p_D)]$ is always greater than or equal to unity. Thus, if the left hand side term becomes greater than the right hand side term at any stage i, $i=0,1,2,\ldots,N-1$, the left hand side term will remain greater than the right hand side term for all further stages, i.e., for all j such that $j \geq i$. Therefore,

$$Q_i = Q_{i+1} = Q_{i+2} = \ldots = Q_N = 1 \qquad (3\text{-}65)$$

which is clearly a "K out of N" fusion rule.

Next, we obtain the optimum value of K, K_{opt}, which minimizes R_B for the "K out of N" fusion rule.

Optimum Value of K

Using a "K out of N" fusion rule, the probability of

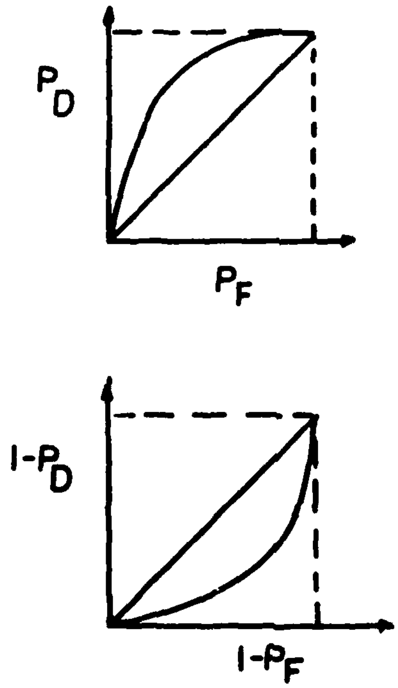


Figure 3.2 Receiver Operating Charateristic (ROC)

(a) Optimum ROC (b) Inverted ROC

detection P_D is given by

$$P_D = \sum_{i=K}^{N} \binom{N}{i} (p_D)^{i} (1 - p_D)^{N-i} \tag{3-66-a}$$

and the probability of false alarm P_F

$$P_F = \sum_{i=K}^{N} \binom{N}{i} (p_F)^{i} (1 - p_F)^{N-i} \tag{3-66-b}$$

where

$$\binom{N}{i} = \frac{N!}{i!\,(N-i)!} \tag{3-66-c}$$

and,

$$n! = n(n-1)(n-2)\ldots(2)(1) \tag{3-66-d}$$

Substituting (3-66-a) and (3-66-b) into (3-5), for any value K' of K the risk function, denoted by $R_B(K')$ is given by

$$R_B(K') = \sum_{i=K'}^{N} \binom{N}{i} \left[C_F(p_F)^{i} (1 - p_F)^{N-i} - C_D(p_D)^{i} (1 - p_D)^{N-i}\right] + C \tag{3-67-a}$$

Similarly, $R_B(K'+1)$ can be expressed as

$$R_B(K'+1) = \sum_{i=K'+1}^{N} \binom{N}{i} \left[C_F(p_F)^{i} (1- p_F)^{N-i} - C_D(p_D)^{i} (1- p_D)^{N-i}\right] + C \tag{3-67-b}$$

First, we will show that $R_B(.)$ has a single minimum. If it is a decreasing function or an increasing function, then clearly it has a single minimum at one of its end points. Otherwise, it remains to be shown that it has exactly one decreasing part and one increasing part implying that it has a single minimum. To show this, we examine the sign of $(R_B(K'+1) - R_B(K'))$, or equivalently, the sign of $F(K')$, where

$$R_B(K'+1) - R_B(K') = \binom{N}{K'} [C_D(p_D)^{K'}(1-p_D)^{N-K'} - C_F(p_F)^{K'}(1-p_F)^{N-K'}] \tag{3-68-a}$$

and,

$$F(K') = [C_D(p_D)^{K'}(1-p_D)^{N-K'} - C_F(p_F)^{K'}(1-p_F)^{N-K'}] \tag{3-68-b}$$

The function $F(.)$ can be used to express the nature of the function $R_B(.)$ as follows

$$F(K') \begin{array}{c} R_B(K') < R_B(K'+1) \\ > \\ \\ < \\ R_B(K') > R_B(K'+1) \end{array} 0 \tag{3-69}$$

Now, let

$$\Omega(K') = (p_D/p_F)^{K'} [(1-p_D)/(1-p_F)]^{N-K'} \tag{3-69-a}$$

Then, (3-69) is equivalent to

$$\Omega(K') \quad \underset{R_B(K') > R_B(K'+1)}{\overset{R_B(K') < R_B(K'+1)}{\gtrless}} \quad (C_F/C_D) = L \qquad (3\text{-}70)$$

Let K" be the first value of K (K increasing from 0 to N-1), where $\Omega(.) \geq L$. Then, using arguments similar to those used previously, we conclude that $\Omega(K) > L$ for all K > K". This means that $R_B(.)$ has a single minimum and K" is the value of K which minimizes $R_B(.)$. Let us denote this value of K by K_{opt}.

Now, we find the value of K_{opt}. We compare $\Omega(K')$ to L. We rewrite (3-70)

$$\Omega(K') = (p_D/p_F)^{K'} [(1-p_D)/(1-p_F)]^{N-K'} \quad \underset{R_B(K') > R_B(K'+1)}{\overset{R_B(K') < R_B(K'+1)}{\gtrless}} \quad L \qquad (3\text{-}71)$$

Taking the log of both sides, (3-70) becomes

$$\log\{(p_D/p_F)^{K'} [(1-p_D)/(1-p_F)]^{N-K'}\} \quad \underset{R_B(K') > R_B(K'+1)}{\overset{R_B(K') < R_B(K'+1)}{\gtrless}} \quad \log(L) \qquad (3\text{-}72)$$

Rearranging (3-72), we have

$$K'\log((p_D/p_F)(1-p_F)/(1-p_D)) + N\log((1-p_D)/(1-p_F)) \underset{R_B(K') > R_B(K'+1)}{\overset{R_B(K') < R_B(K'+1)}{\gtrless}} \log(L) \tag{3-73}$$

which may be written as

$$K'\log((p_D/p_F)(1-p_F)/(1-p_D)) \underset{R_B(K') > R_B(K'+1)}{\overset{R_B(K') < R_B(K'+1)}{\gtrless}} \log(L[(1-p_F)/(1-p_D)]^N) \tag{3-74}$$

Since $p_D \geq p_F$ and $1-p_D \leq 1-p_F$, $\log((p_D/p_F)(1-p_F)/(1-p_D))$ is nonnegative and therefore, we can divide both sides by this term without affecting the sign of the inequality. Then, (3-74) becomes

$$K' \underset{R_B(K') > R_B(K'+1)}{\overset{R_B(K') < R_B(K'+1)}{\gtrless}} \frac{\log(L[(1-p_F)/(1-p_D)]^N)}{\log((p_D/p_F)(1-p_F)/(1-p_D))} \tag{3-75}$$

Let

$$N^* = \frac{\log(L[(1-p_F)/(1-p_D)]^N)}{\log((p_D/p_F)(1-p_F)/(1-p_D))} \tag{3-76}$$

Thus, K_{opt} is given by

$$K_{opt} = \begin{cases} \lceil N^* \rceil & \text{if } N^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3-77}$$

where $\lceil \ \rceil$ is the ceiling function.

Overall Solution

The solution to the overall problem i.e., optimum decision rules at the detectors and optimum fusion rule, can be obtained by solving equations (3-58) and (3-77) simultaneously.

In the next section, we present an example.

3.5 Example

In this section, we consider a binary Bayesian hypothesis testing problem using two detectors for illustration. As in Example 2.1, we assume that the observations at both detectors are independent and exponentially distributed, i.e.,

$$p(y_i|H_0) = \begin{cases} \exp(-y_i) & \text{if } \ y_i \geq 0 \\ & i=1,2 \\ 0 & \text{otherwise} \end{cases} \tag{3-88-a}$$

$$p(y_i|H_1) = \begin{cases} (1/\theta_i)\exp(-y_i/\theta_i) & \text{if } y_i \ge 0 \\ 0 & \text{otherwise} \end{cases} \quad i=1,2 \qquad (3\text{-}88\text{-}b)$$

Here, we will design the DDF system, i.e., obtain the optimum fusion rule and the optimum decision rules at the detectors, so as to minimize the Bayesian risk R_B. Recall from (2-36-a) that, the likelihood ratios at the individual detectors are

$$\Omega_i(y_i) = \frac{p(y_i|H_1)}{p(y_i|H_0)} = \frac{1}{\theta_i} \exp\,(y_i(1-(1/\theta_i))) \qquad i=1,2$$

(3-89)

and the likelihood ratio tests (LRT) are

$$y_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\theta_i}{1-\theta_i} \ln(t_i\ \theta_i) = t_i' \qquad i = 1,2$$

(3-90)

The probability of miss and the probability of false alarm at the individual detectors are given by,

$$P_{Mi} = 1 - (\theta_i\ t_i)^{(1/(1-\theta_i))} \qquad i = 1,2 \qquad (3\text{-}91\text{-}a)$$

and,

$$P_{Fi} = (\theta_i\ t_i)^{(\theta_i/(1-\theta_i))} \qquad i = 1,2 \qquad (3\text{-}91\text{-}b)$$

The thresholds and, the inequalities describing the fusion rule are,

$$t_1 = L \; \frac{P_{F2}(P_{1\underline{00}} - P_{1\underline{01}} - P_{1\underline{10}} + P_{1\underline{11}}) + (P_{1\underline{10}} - P_{1\underline{00}})}{P_{M2}(P_{0\underline{00}} - P_{0\underline{01}} - P_{0\underline{10}} + P_{0\underline{11}}) + (P_{0\underline{01}} - P_{0\underline{11}})} \tag{3-92-a}$$

$$t_2 = L \; \frac{P_{F1}(P_{1\underline{00}} - P_{1\underline{01}} - P_{1\underline{10}} + P_{1\underline{11}}) + (P_{1\underline{01}} - P_{1\underline{00}})}{P_{M1}(P_{0\underline{00}} - P_{0\underline{01}} - P_{0\underline{10}} + P_{0\underline{11}}) + (P_{0\underline{10}} - P_{0\underline{11}})} \tag{3-92-b}$$

$$\frac{1 - P_{D1}}{1 - P_{F1}} \; \frac{1 - P_{D2}}{1 - P_{F2}} \quad \underset{P_{1\underline{00}}=0}{\overset{P_{1\underline{00}}=1}{\gtrless}} \quad L \tag{3-93-a}$$

$$\frac{1 - P_{D1}}{1 - P_{F1}} \; \frac{P_{D2}}{P_{F2}} \quad \underset{P_{1\underline{01}}=0}{\overset{P_{1\underline{01}}=1}{\gtrless}} \quad L \tag{3-93-b}$$

$$\frac{P_{D1}}{P_{F1}} \; \frac{1 - P_{D2}}{1 - P_{F2}} \quad \underset{P_{1\underline{10}}=0}{\overset{P_{1\underline{10}}=1}{\gtrless}} \quad L \tag{3-93-c}$$

$$\frac{P_{D1}}{P_{F1}} \; \frac{P_{D2}}{P_{F2}} \quad \underset{P_{1\underline{11}}=0}{\overset{P_{1\underline{11}}=1}{\gtrless}} \quad L \tag{3-93-d}$$

where L is given by

$$L = \frac{C_F}{C_D} = \frac{P_0 \; (C_{10} - C_{00})}{P_1 \; (C_{01} - C_{11})} \tag{3-94}$$

In Figure 3.3, we present the ROC curves for this example with $\Theta_1 = 2$ and $\Theta_2 = 4$ for the two known fusion rules "AND" and "OR". Note that the ROC curves are the same ones as shown in Figure 2.1. It is clear from these curves that the "OR" fusion rule is better. For the same example as above, Figure 3.4 shows the ROC curve for the optimum DDF system. Each point is obtained by deriving the optimum decision rules at the detectors and optimum fusion rule. The optimum ROC is the same as the one obtained bu using the "OR". Then, it is clear that, for this example,"OR" is the optimum fusion rule. Figure 3.5 shows the ROC curve for the optimum DDF system with $\Theta_1=6$ and $\Theta_2=8$.

It should be noted that, in general, the optimum fusion rule is not necessarily the "OR" fusion rule. It can be any nonrandomized fusion rule, e.g., for the case when $\Theta_1 = 2$, $\Theta_2 = 40$ and high value of P_0, the optimum fusion rule is the decision of the second detector i.e., $u = u_2$. The same is true for the case when $(\Theta_1,\Theta_2) = (2,98)$.

Next, we obtain the optimum value of K for the "K out of N" fusion rule which minimizes R_B. In this example, 20 detectors are used (N=20). These detectors have independent identical exponential distributions. For different values of Θ, we plotted R_B' ($R_B'= R_B-C$) as a function of K and obtained the value of K at which R_B is minimum. Then, we compared it with the value of K obtained using equation (3-77). In all cases, the two results agreed with each

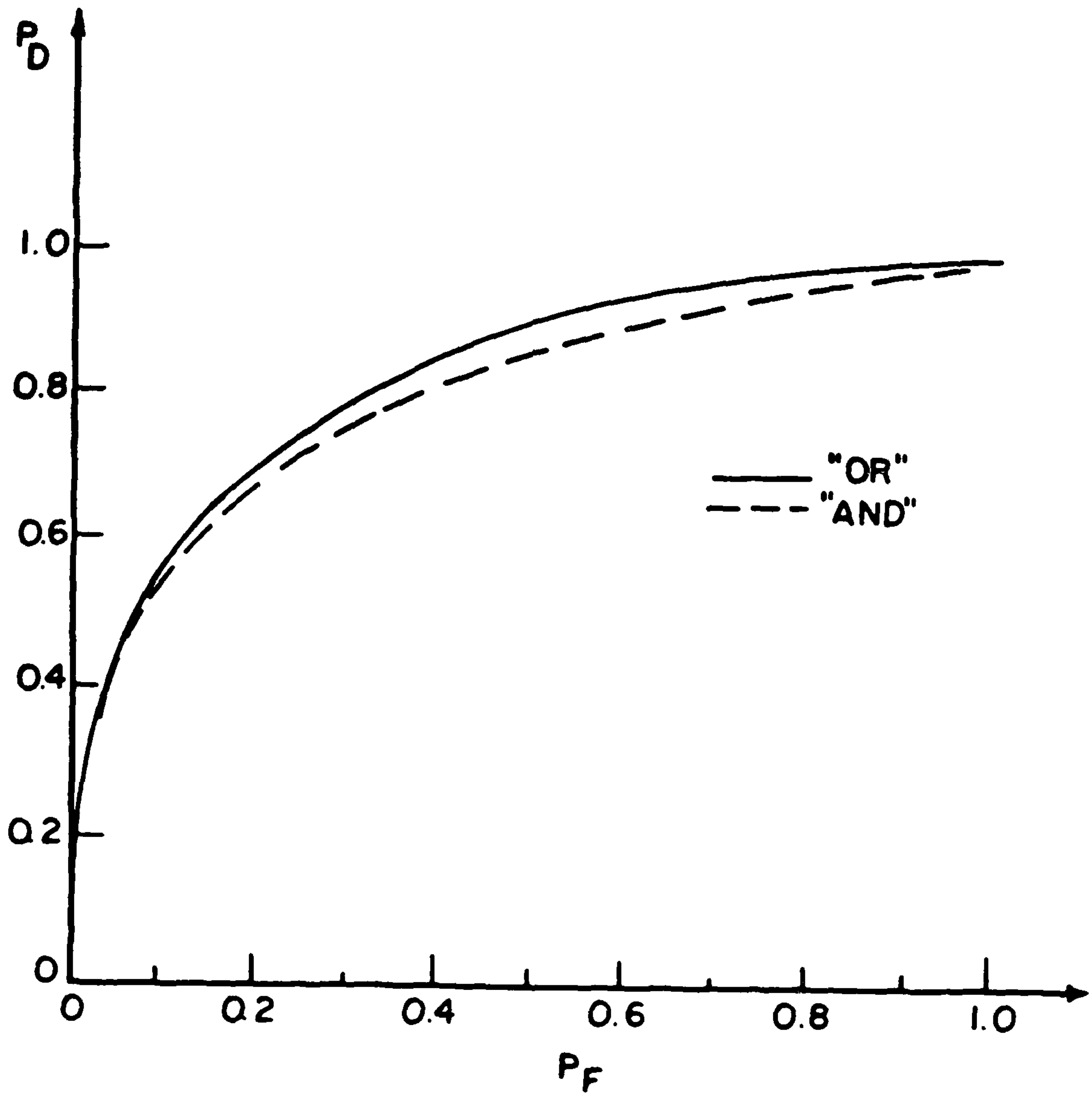
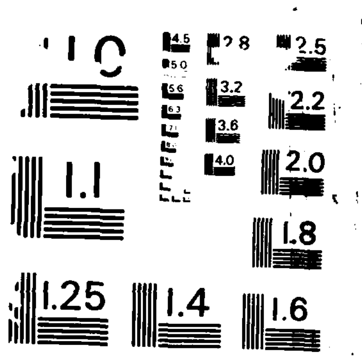


Figure 3.3 Receiver Operating Characteristic for Example 2.1 and the "AND" and "OR" Fusion Rules

AD-A193 934 ON THE DESIGN AND OPTIMIZATION OF DISTRIBUTED SIGNAL DETECTION AND PARAME..(U) SYRACUSE UNIV NY DEPT OF ELECTRICAL AND COMPUTER ENGINEERING.. 2/3

UNCLASSIFIED I Y HOBALLAH ET AL. SEP 87 RADC-TR-87-130 F/G 12/9 NL



COPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

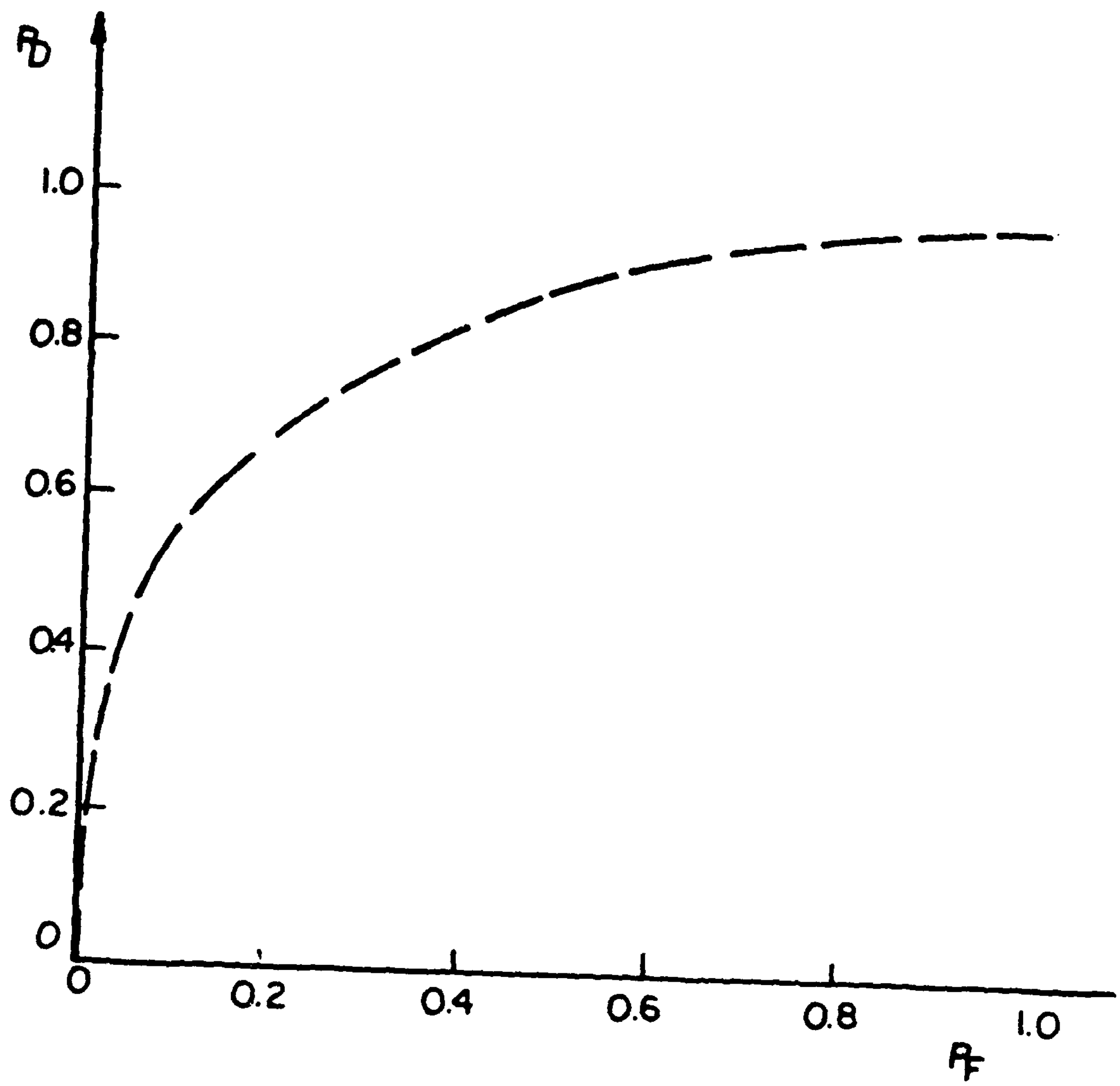


Figure 3.4 Optimum Receiver Operating Characteristic for Example 2.1

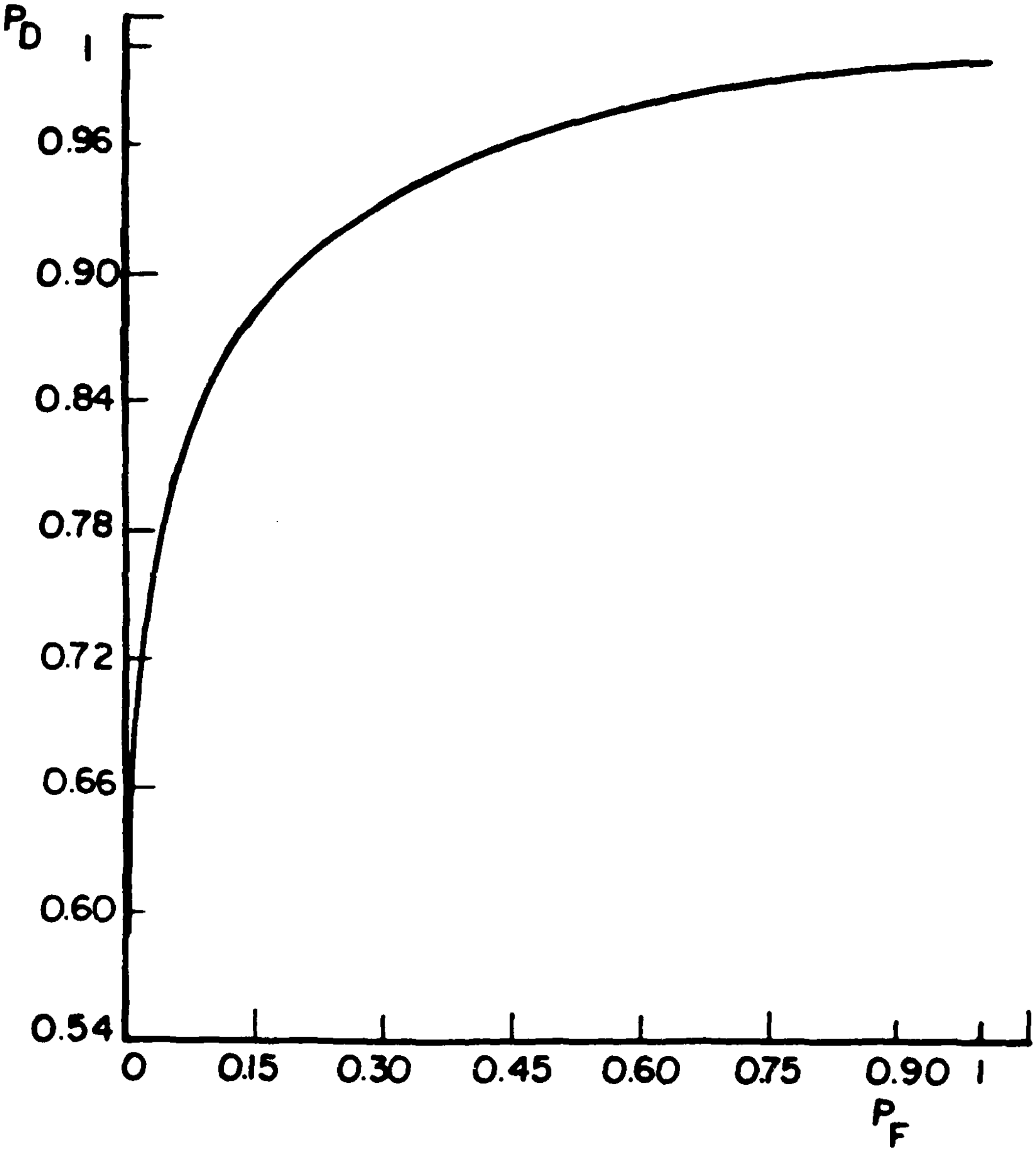


Figure 3.5 Optimum Receiver Operating Characteristic for Example 2.1 and ([illegible]$_1$, [illegible]$_2$) = (6,8)

other. For example, in Figures 3.6, 3.7, 3.8, and 3.9 we show the R_B' vs. K curves for P_0 = 0.6 and $(t', \theta) = (2,2)$, (6,2), (7,4) and (10,4) respectively. In Figure 3.6, $K_{opt} = 4$, in Figure 3.7, $K_{opt} = 1$, in Figure 3.8, $K_{opt} = 3$ and, in Figure 3.9, $K_{opt} = 2$.

3.6 Discussion

In this chapter, we have presented a generalized approach to the problem of distributed Bayesian hypothesis testing with data fusion. It should be noted that with a proper cost assignment and a proper fusion rule, the DDF system can be reduced to the DD system which is thus a special case of the problem considered in this chapter. Minimum probability of error is widely used as an optimization criterion in the design of optimum detection systems. For this criterion, we can obtain the results by simply setting the costs appropriately. It is to be noted that, in general, the fusion rule does not have to be "AND" or "OR". The fusion rule may be any nonrandomized rule.

In the next chapter, we apply the results obtained in this chapter to an interesting problem in radar signal detection namely, double threshold detection.

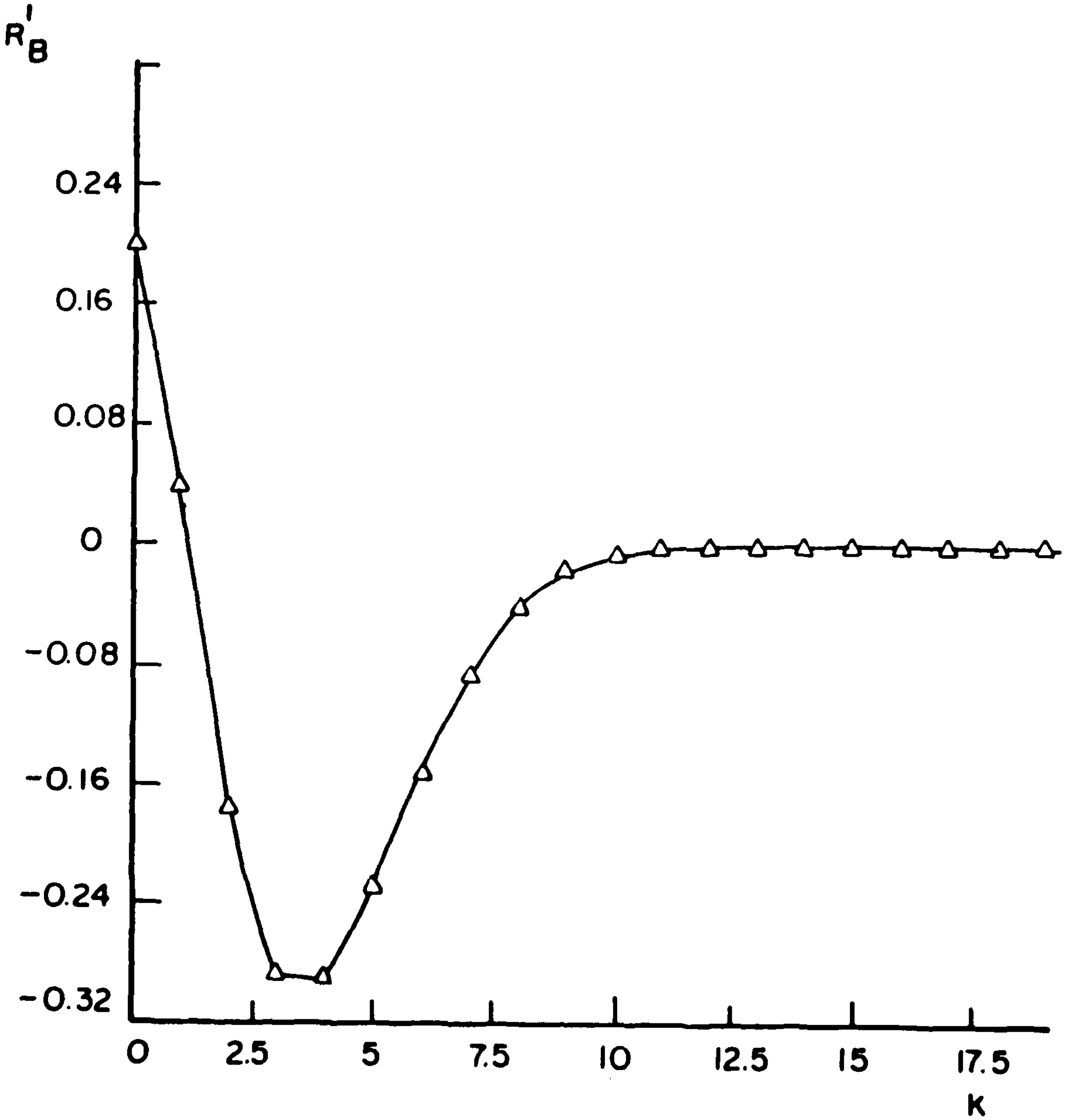


Figure 3.6 Risk Versus K for Example 2.1 and t' = θ = 2 and P_0 = 0.6

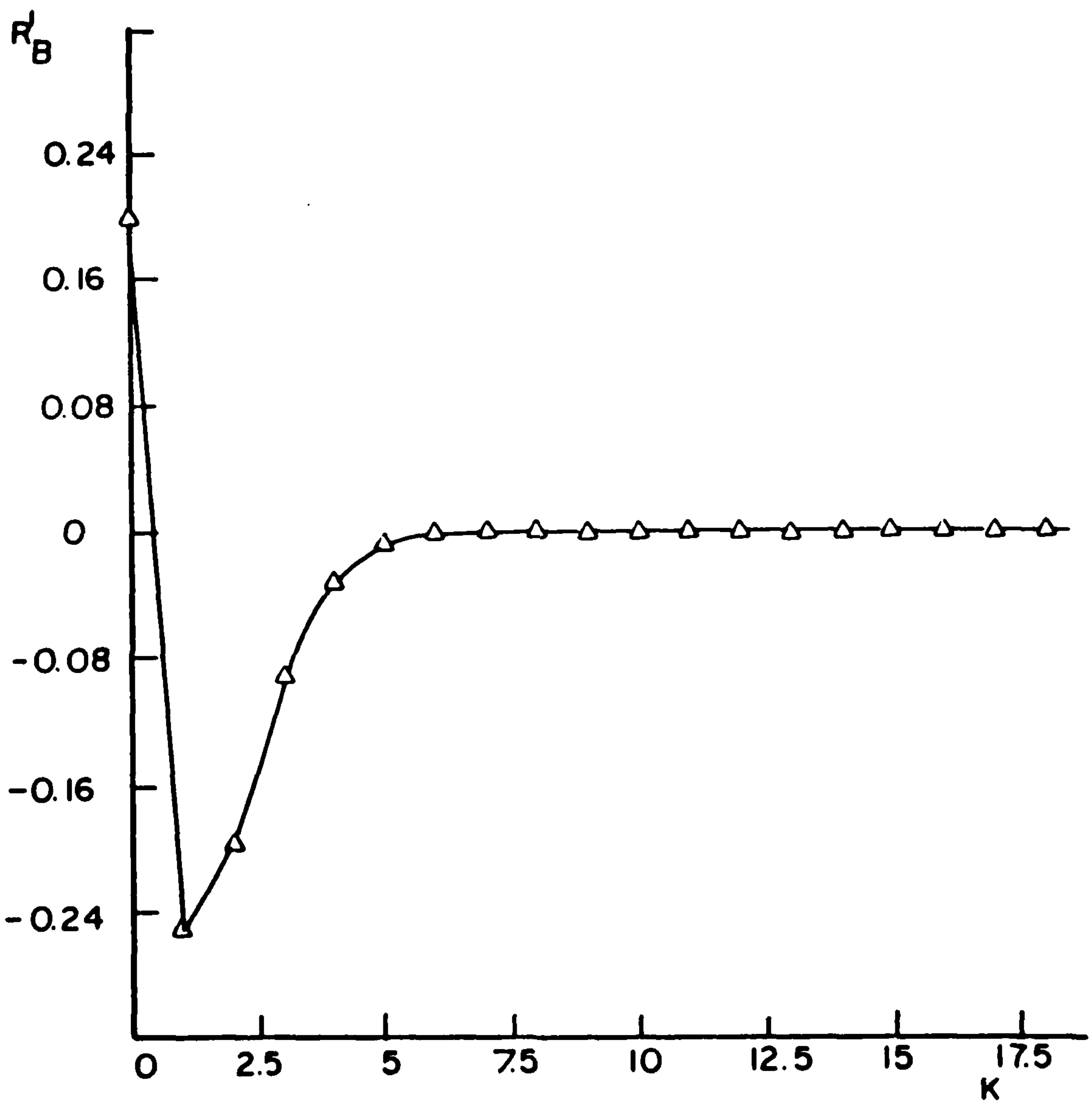


Figure 3.7 Risk Versus K for Example 2.1, $(t',n) = (6,2)$ and $P_0 = 0.6$

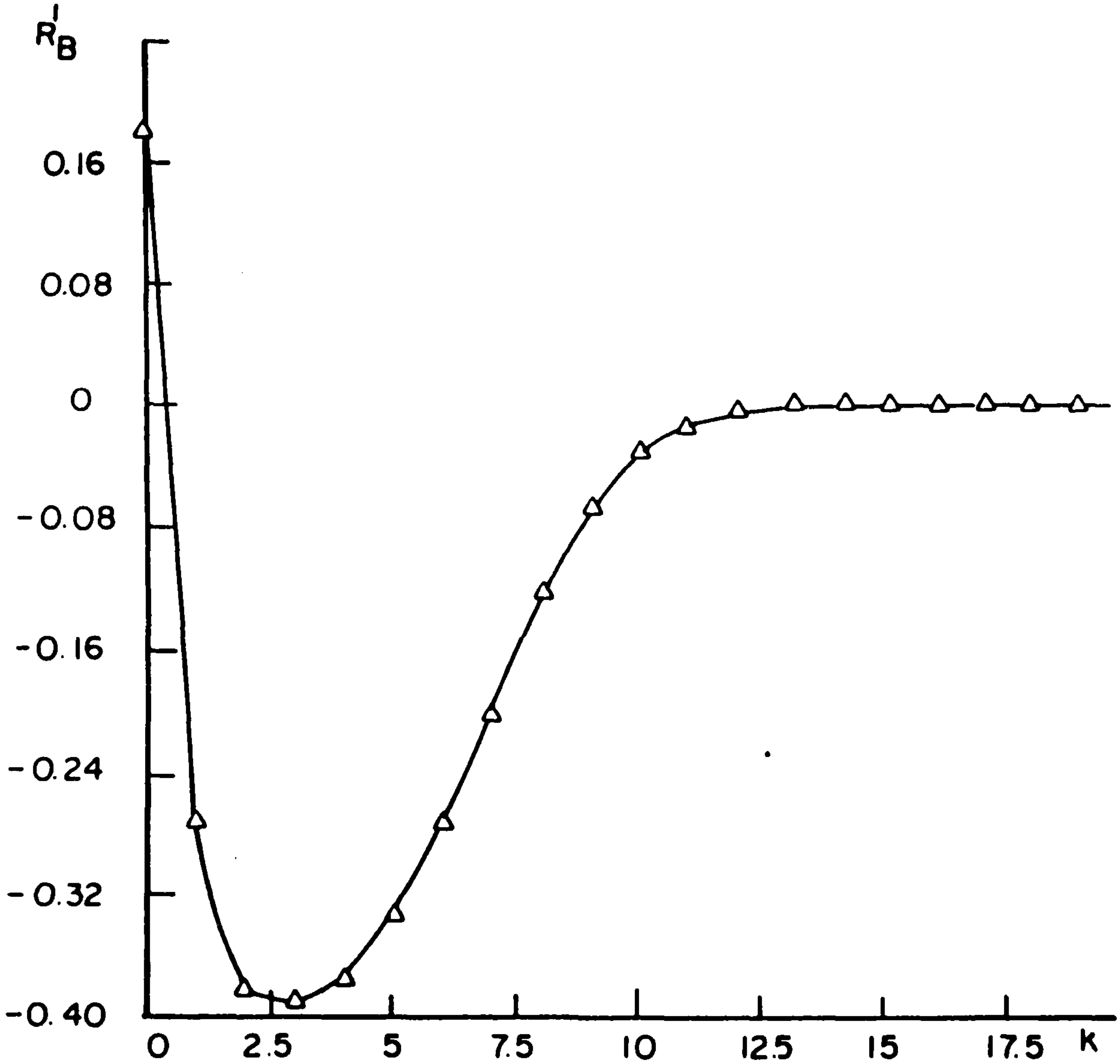


Figure 3.8 Risk Versus K for Example 2.1, (t',θ) = (7,4) and P_0 = 0.6

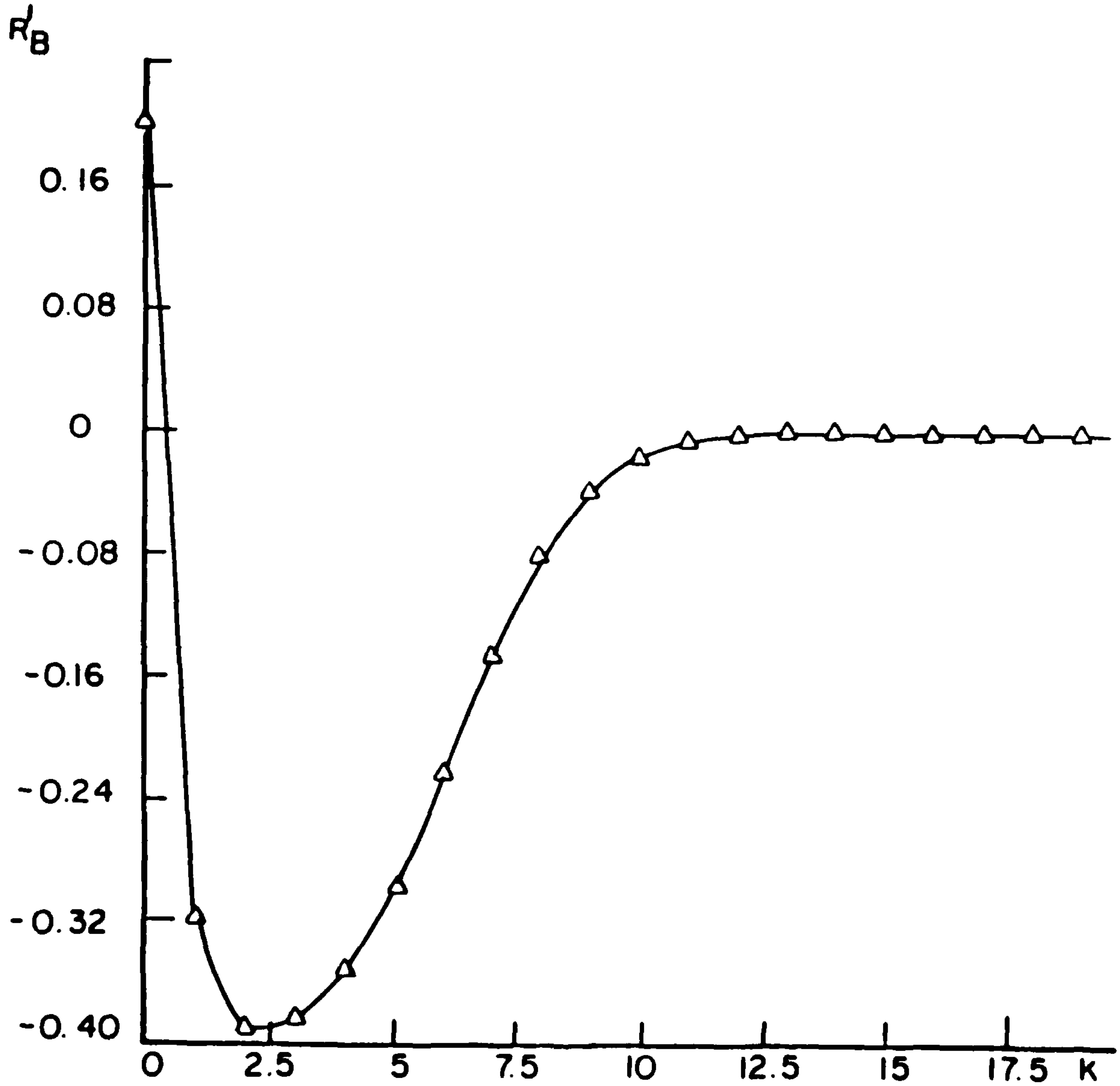


Figure 3.9 Risk Versus K for Example 2.1, (t',n) = (10,4) and P_0 = 0.6

IV. Double Threshold Detection With Distributed Sensors

4.1 Introduction

A postdetection integration technique known as double threshold detection is often employed for target detection in radar systems. Greater detectability can be achieved by using predetection integration. But due to simplicity and cost, postdetection integration is often implemented even though its performance is slightly poorer.

A block diagram of the double threshold detection system is shown in Figure 4.1, [2]. Out of N waveform envelope samples coming out of the matched filter corresponding to an N pulse train, the number of signals that exceed a first threshold, T', at the sampling instant is counted . If this number is equal to or greater than a second threshold , K, the target is declared present. For single sensors, this technique has been analyzed in [2, 31-36]. The system optimization criterion used is the Neyman-Pearson criterion. It should be observed that the double threshold detection system and the binary DDF system with identical detectors considered in Section 3.4.2 are analogous and the results obtained in Section 3.4.2 can be employed to obtain the results for the double threshold detection. From our results in Chapter 3, the values of the two thresholds, namely, K and T (T is related to T' through the likelihood ratio), are given below

P.M.F = Pulse Matched Filter

Env. Det. = Envelope Detector

S.R.G = Sampler Range Gate

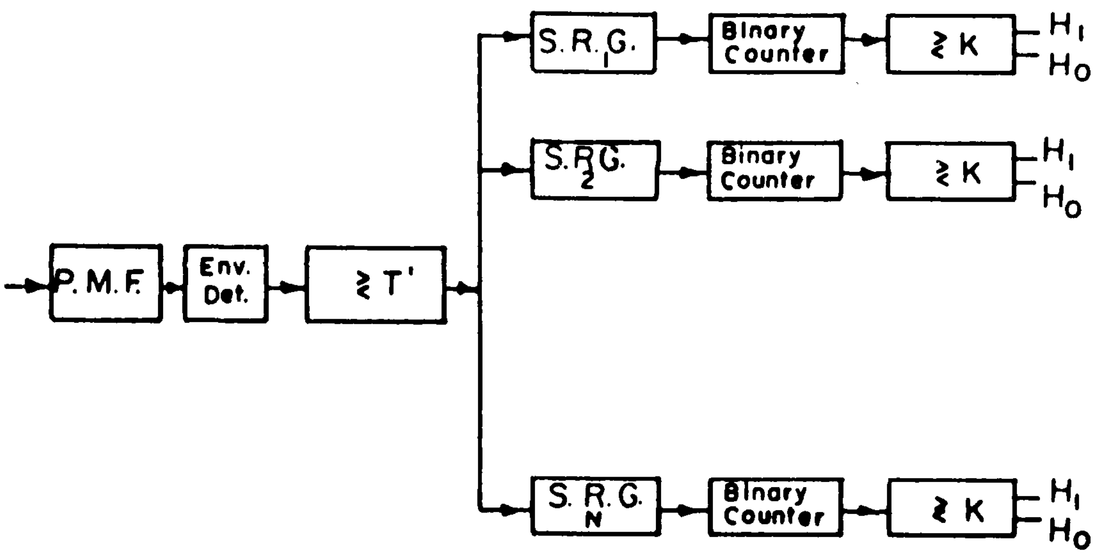


Figure 4.1 A Range Gated Binary Integration for Pulse Train Waveform

Let

$$N^{*} = \frac{\log\left[L \left(\frac{1-p_F}{1-p_D}\right)^{N} \right]}{\log\left[\left(\frac{p_D}{p_F}\right) \frac{(1-p_F)}{(1-p_D)} \right]} \qquad (4\text{-}1)$$

where

p_F = the probability of false alarm of an element of the observation vector

p_D = the probability of detection of an element of the observation vector

and,

L is the Lagrange multiplier.

Then, the optimum value of K is given by,

$$K = \begin{cases} \lceil N^{*} \rceil & \text{if } N^{*} \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (4\text{-}2)$$

The value of the threshold, T, in the likelihood ratio test for any sample is obtained by solving the equation obtained after taking the derivative of equation (2-28), setting $(\delta P_D/\delta P_F) = T$, and setting the derivative equal to zero. Then, T is given by

$$T = L \frac{c_F}{c_M} \tag{4-3}$$

where

$$c_F = \sum_{i=K}^{N} \binom{N}{i} \left[p_F^{(i-1)} \, (1 - p_F)^{(N-i-1)} \, (i - N\, p_F) \right] \tag{4-4-a}$$

$$c_D = \sum_{i=K}^{N} \binom{N}{i} \left[p_D^{(i-1)} \, (1 - p_D)^{(N-i-1)} \, (i - N\, p_D) \right] \tag{4-4-b}$$

The probability of false alarm for the detection system is equal to α, i.e.,

$$P_F = \alpha. \tag{4-5}$$

Recall that the probability of detection and the probability of false alarm of the detector were given in (3-66).

Next, we extend the above result to the case when multiple sensors are employed along with a data fusion center. Two schemes are proposed and system parameters are obtained for each of the schemes.

4.2 Postdetection Integration With Distributed Sensors

In this section, we propose two alternate schemes for post detection integration when distributed sensors are employed. These schemes differ in the bandwidths required for channels connecting the sensors and the fusion center.

Scheme 1

The block diagram of the first scheme is shown in Figure 4.2. Each detector i, receives an observation vector $\underline{y}_i$, $i = 1,2,\ldots,N$. Each observation vector consists of n components. We assume that each detector compares each element of its observation vector, y_{ij}, $i = 1,2,\ldots,N$, $j = 1,2,\ldots,n$, to a threshold T_i'. If the j^{th} element of the observation vector at the i^{th} detector exceeds the first threshold , we set the the j^{th} element of the decision vector of the i^{th} detector equal to 1, i.e., $u_{ij} = 1$, $i = 1,2,\ldots,N$, $j = 1,2,\ldots,n$. Otherwise, we set $u_{ij} = 0$. Thus, each detector i, $i = 1,2,\ldots,N$ has an n-dimensional decision vector $\underline{u}_i$, $\underline{u}_i = (u_{i1},u_{i2},\ldots,u_{in})^T$, associated with it. Note that without loss of generality , we have assumed that each observation vector y_i consists of the same number of elements. Each individual detector sends its decision vector, $\underline{u}_i$, to a data fusion center where a global decision, u, is made using a fusion rule. The global decision is based on a decision vector $\underline{x}$ which is obtained by concatenating the decision vectors $\underline{u}_i$, $i = 1,2,\ldots,N$, i.e.,

$$\underline{x} = (u_{11},u_{12},\ldots,u_{1n},u_{21},u_{22},\ldots,u_{2n},\ldots,u_{N1},u_{N2},\ldots,u_{Nn})^T \qquad (4\text{-}6)$$

Next, we present the thresholds at the individual detectors and the fusion rule. The results are obtained directly from the results of Chapters 2 and 3. We treat the

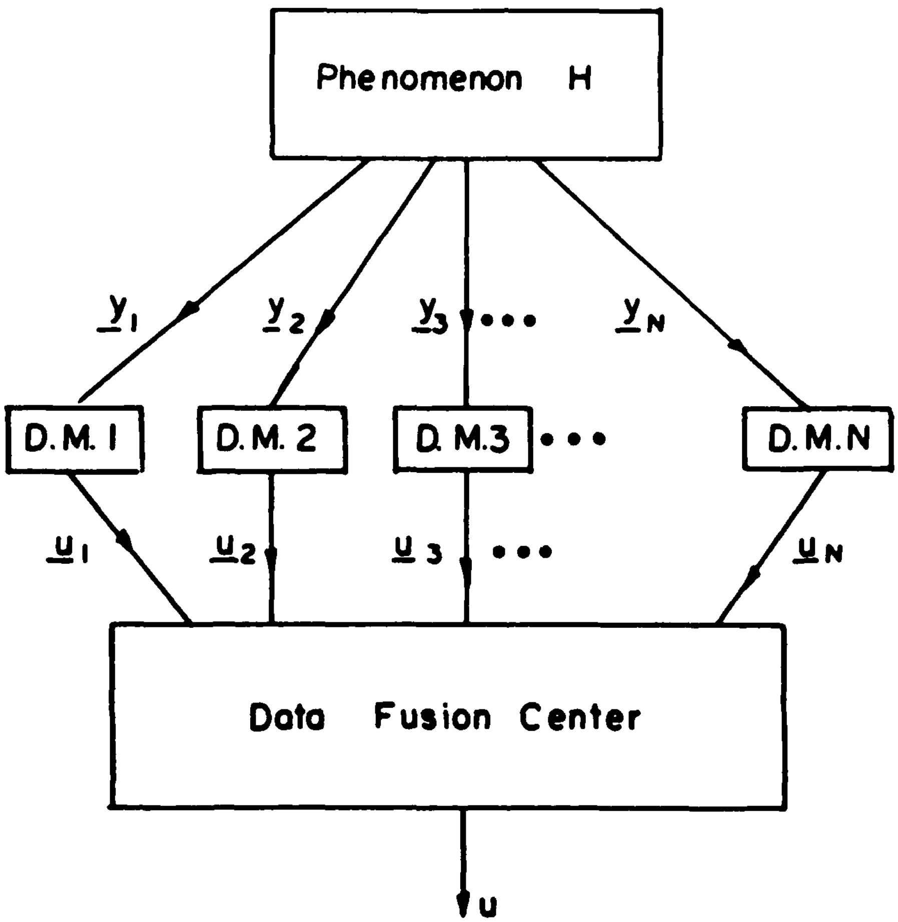


Figure 4.2 Postdetection Integration Using Distributed Sensors, Scheme 1

problem as a system with nxN detectors. The goal here is to design a system which minimizes the probability of miss, when the probability of false alarm is constrained by $P_F = \alpha \le \alpha'$. In this case, the optimum fusion rule is given by

$$1 \begin{array}{c} P_{1X^*} = \min(\alpha^*, 1) \\ > \\ < \\ P_{1X^*} = 0 \end{array} L^* \tag{4-7}$$

and, the threshold in the likelihood ratio test is

$$T = L \frac{C_{F\mu j}}{C_{M\mu j}} \tag{4-8}$$

where

$$\alpha^* = \alpha - \sum_{X \neq X^*} P_{1X}\, P_{X0} \tag{4-8-a}$$

and, P_{1X}, P_{X0} are given in (2-5). $C_{M\mu j}$ and $C_{F\mu j}$ are given in (2-15) except that the subscript μj corresponds to the j^{th} element of the decision vector of detector μ, $\mu = 1,\ldots,N$ and, $j = 1,\ldots,n$.

<u>Scheme 2</u>

The block diagram of the second scheme is shown in Figure 4.3. In this scheme, each detector acts as a double threshold detector. Based on its own decision vector

D.F. = Data Fusion Center

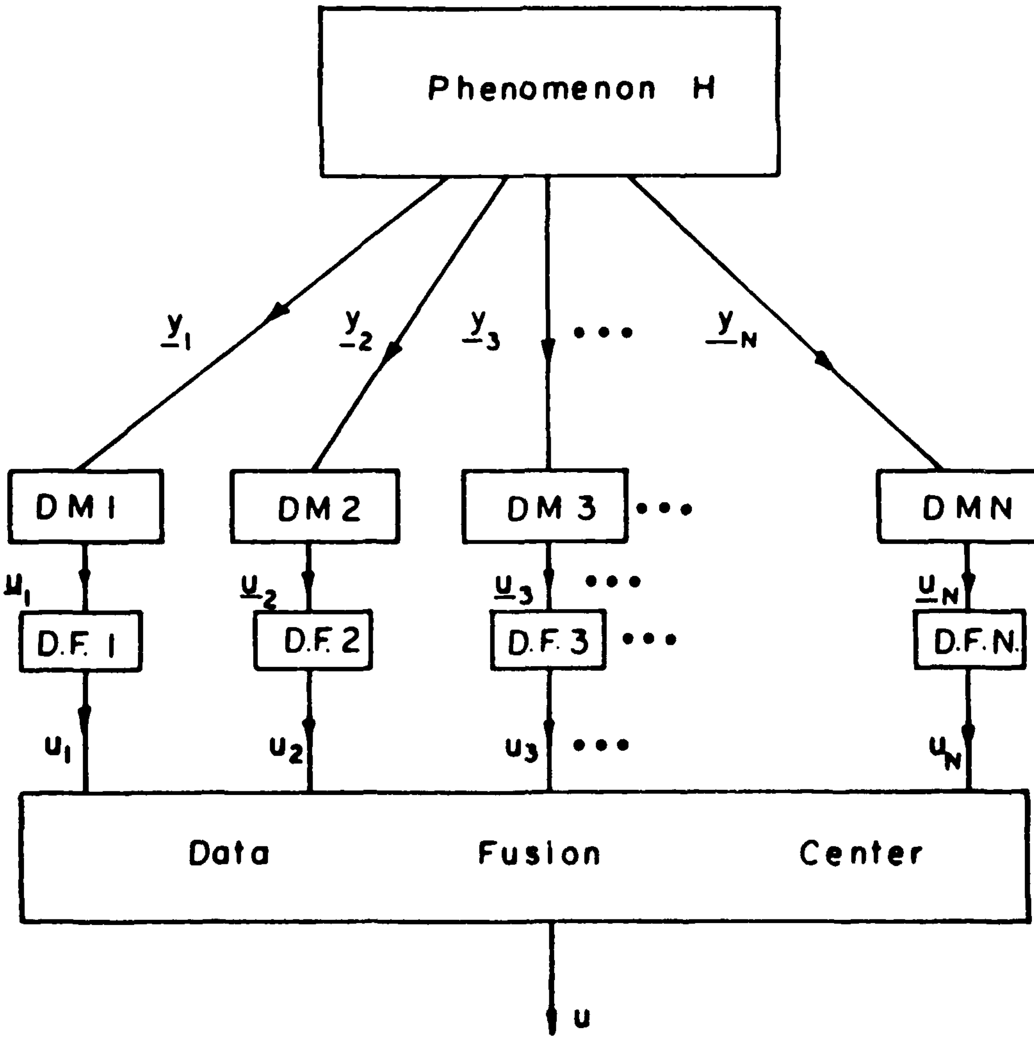


Figure 4.3 Postdetection Integration Using Distributed Sensors, Scheme 2

u_i, $\underline{u}_i = (u_{i1}, u_{i2}, \ldots, u_{in})^T$, it uses a second threshold to make a decision u_i, $i = 1,2,\ldots,N$, which is sent to the data fusion center where a global decision, u, is made. The function to be maximized in this case is the following

$$\Gamma = P_D - L\,[P_F - \alpha] \tag{4-9}$$

or equivalently,

$$\Gamma = \sum_{\underline{x}} P_{1\underline{x}}\, D_{\underline{x}} - L\,[\sum_{\underline{x}} P_{1\underline{x}}\, F_{\underline{x}} - \alpha] \tag{4-10}$$

where P_D and P_F are the global probabilities of detection and false alarm, respectively. $P_{1\underline{x}}$, $D_{\underline{x}}$, and $F_{\underline{x}}$ were defined earlier in (2-5). For any specific decision vector $\underline{x}^*$, the fusion rule is given by

$$\begin{array}{ccc} & P_{1\underline{x}^*} = \min(\alpha^*, 1) & \\ & > & \\ 1 & & L^* \\ & < & \\ & P_{1\underline{x}^*} = 0 & \end{array} \tag{4-11}$$

where

$$\alpha^* = \alpha - \sum_{\underline{x} \neq \underline{x}^*} P_{\underline{x}0}\, P_{1\underline{x}} \tag{4-11-a}$$

and

$$L^* = L\,\frac{P_{\underline{x}^*0}}{P_{\underline{x}^*1}} \tag{4-11-b}$$

The second threshold K_μ for detector μ, $\mu = 1,2,\ldots,N$ is given by

$$
K_\mu = \begin{cases} \lceil N^* \rceil & \text{if } N^* \geq 0 \\ 0 & \text{Otherwise} \end{cases} \qquad (4\text{-}12)
$$

where

$$
N^* = \frac{\log\left[L_\mu \left[\dfrac{1-p_{F\mu}}{1-p_{D\mu}} \right]^n \right]}{\log\left[\dfrac{p_{D\mu}}{1-p_{F\mu}} \quad \dfrac{1-p_{F\mu}}{1-p_{D\mu}} \right]} \qquad (4\text{-}12\text{-}a)
$$

$p_{F\mu}$ = probability of false alarm of any element of the observation vector of detector μ

$p_{D\mu}$ = probability of detection of any element of the observation vector of detector μ

$P_{F\mu}$ = probability of false alarm of detector μ

$$
P_{F\mu} = \sum_{i=K_\mu}^{n} \binom{n}{i} p_{F\mu}^{(i)} (1-p_{F\mu})^{(n-i)}
$$

$$
= \alpha_\mu = \frac{\alpha - K_{B\mu}}{C_{F\mu}} \qquad \mu = 1,2,\ldots,N \qquad (4\text{-}13)
$$

$P_{D\mu}$ = probability of detection of detector μ

$$= \sum_{i=K_\mu}^{n} \binom{n}{i} p_{D\mu}^{(i)} (1-p_{D\mu})^{(n-i)} \qquad (4\text{-}14)$$

$$L_\mu = L \frac{C_{F\mu}}{C_{D\mu}} \qquad (4\text{-}15)$$

$$C_{F\mu} = \sum_{\underline{x}} F_{\underline{x}}^{\mu} (P_{1\underline{x}}^{\mu 1} - P_{1\underline{x}}^{\mu 0}) \qquad (4\text{-}16)$$

$$C_{D\mu} = \sum_{\underline{x}} D_{\underline{x}}^{\mu} (P_{1\underline{x}}^{\mu 1} - P_{1\underline{x}}^{\mu 0}) \qquad (4\text{-}17)$$

$$K_{B\mu} = \sum_{\underline{x}} F_{\underline{x}}^{\mu} P_{1\underline{x}}^{\mu 0} \qquad (4\text{-}18)$$

$$D_{\underline{x}}^{\mu} = \prod_{S_0^\mu} (1 - P_{Dj}) \prod_{S_1^\mu} P_{Dk}) \qquad (4\text{-}19)$$

$$F_{\underline{x}}^{\mu} = \prod_{S_0^\mu} (1 - P_{Fj}) \prod_{S_1^\mu} P_{Fk} \qquad (4\text{-}20)$$

S_0 = set of all j, such that u_j is an element of $\underline{x}$ and $u_j=0$ (4-21)

S_1 = set of all k, such that u_k is an element of $\underline{x}$ and $u_k=1$ (4-22)

$S_0^\mu = S_0$ where u_μ is excluded (4-23)

$S_1^\mu = S_1$ where u_μ is excluded. (4-24)

The first threshold, t_μ, of detector μ is given by

$$t_\mu = L_\mu \frac{C'_{F\mu}}{C'_{D\mu}} \tag{4-25}$$

where

$$C'_{F\mu} = \sum_{i=K_\mu}^{n} \binom{n}{i} [p_{F\mu}^{(i-1)} (1 - p_{F\mu})^{(n-i-1)} (i - n\, p_{F\mu})] \tag{4-26-a}$$

$$C'_{D\mu} = \sum_{i=K_\mu}^{n} \binom{n}{i} [p_{D\mu}^{(i-1)} (1 - p_{D\mu})^{(n-i-1)} (i - n\, p_{D\mu})] \tag{4-26-b}$$

and, $p_{F\mu}$ and $p_{D\mu}$ are the probability of false alarm and the probability of detection of an element of the observation vector of detector μ.

The overall solution to the problem is obtained by solving (4-11), (4-12) and (4-25) simultaneously. In the special case where the detectors are identical, we could view it as a triple threshold detection system, where the equations become easier to solve.

V. Distributed Minimum Equivocation Detection

5.1 Introduction

In statistical decision theory, a variety of criteria are used for the optimization of detectors. For example, in the Bayesian formulation a fixed cost is assigned to each possible course of action and then, the average cost is minimized. In applications where such costs are available and are meaningful, Bayesian cost formulation provides an excellent choice for system optimization. However, this may not be the case in all applications. In some applications, our interest may be the amount of information that we are able to transfer, e.g., in telephone channels we are concerned with the amount of information transmitted rather than the nature of the information itself. In such situations, cost may be a variable and entropy-based cost functions may be more meaningful.

In this chapter, we shall use entropy based cost functions and derive optimum multisensor detection systems. For single sensor detection problems, Middleton [29], used such a criterion for the design of an optimum decision system where he minimized the equivocation between the input and the output. In the next section, we will review some basic information theory definitions. We will also show the correspondence between the classical binary detection system and, the binary communication channel. Then, we present the results for the single sensor case. In Section 5.3, we solve the minimum equivocation detection problem for the DDF system. In Section 5.4, we present the solution to the minimum equivocation detection problem for the DD system. In Section 5.5, we present a numerical example.

5.2 Preliminaries

5.2.1 Basic Information Theory Definitions

In this section, first we briefly present some definitions from information theory [41, 42], which will be used in the rest of this chapter. Let x and y be two discrete random variables taking values from the sets $(x_1, x_2, \ldots, x_M)$ and $(y_1, y_2, \ldots, y_N)$ respectively. Let $P(x_m)$ and $P(y_n)$ denote the associated probability measures and $P(x_m|y_n)$ and $P(x_m, y_n)$ denote the conditional and joint distributions respectively.

The entropty $h(x)$, which measures the uncertainty about x is defined by

$$h(x) = E\{\log[1/P(x)]\} = \sum_{x} P(x) \log\{1/P(x)\} \tag{5-1}$$

The conditional entropy (equivocation) $h(x|y)$, which measures the uncertainty about x given y, is given by

$$h(x|y) = E\{\log[1/P(x|y)]\} = \sum_{x}\sum_{y} P(x,y) \log[1/P(x|y)] \tag{5-2}$$

The mutual information, $I(x;y)$, is the amount of information provided about x by y and is defined as

$$I(x;y) = h(x) - h(x|y) \tag{5-3}$$

Also, we have the following relationships

$$\begin{aligned} I(x;y) &= I(y;x) \\ &= h(y) - h(y|x) \\ &= h(x) + h(y) - h(x,y) \end{aligned} \tag{5-4}$$

where $h(x,y)$ is the entropy of x and y, defined as

$$h(x,y) = E\{\log [1/P(x,y)]\} = \sum_x \sum_y P(x,y) \log [1/P(x,y)] \tag{5-5}$$

5.2.2 Correspondence Between Detection Theory and Information Theory

Block diagrams of a classical binary detection system and a binary communication channel are shown in Figure 5.1. We observe the correspondence between the two problems as follows. The source in the detection problem can be viewed as the information source in the information transmission problem. The boxed part of the detection problem corresponds to the channel of the information transmission system. The decisions in Figure 5.1.(a) may be looked at as the output of the channel in Figure 5.1.(b) The probability of detection, P_D, the probability of miss, P_M, and the probability of false alarm, P_F, in the detection problem are

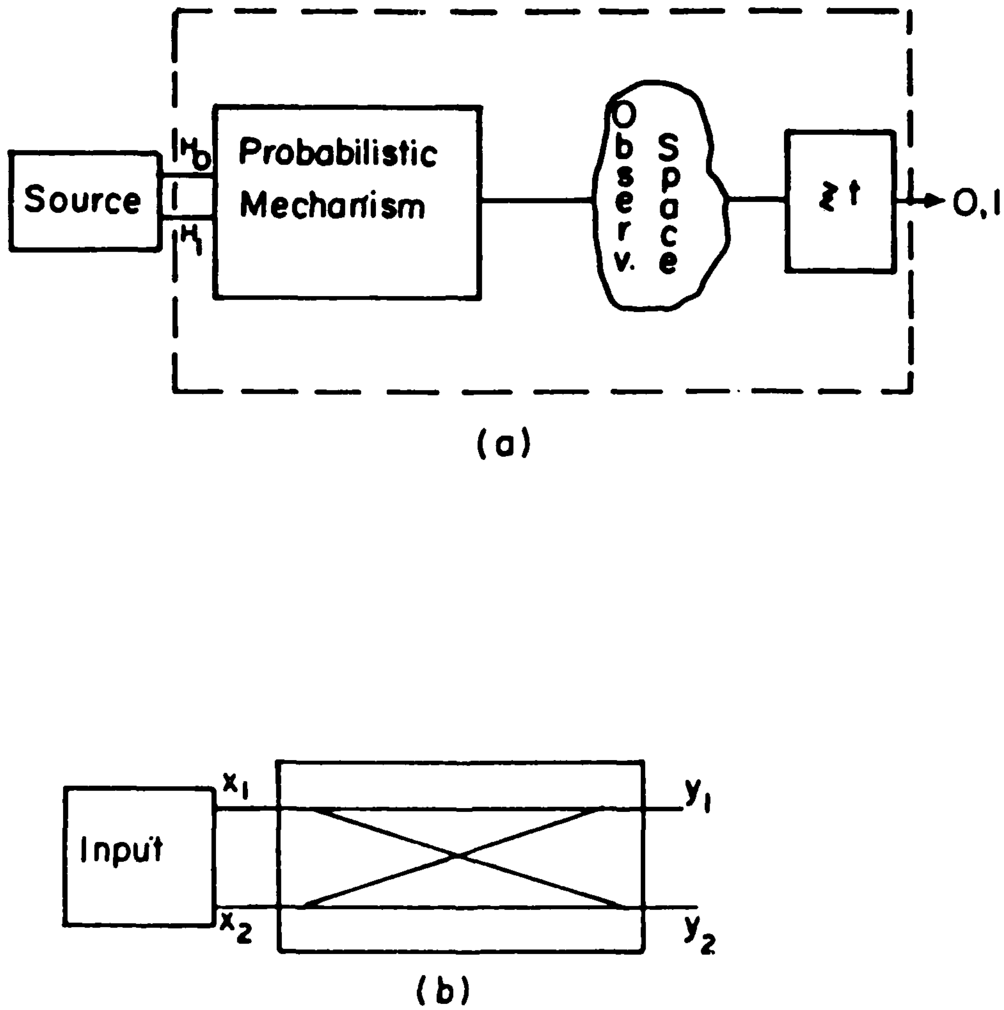


Figure 5.1 Correspondence Between Detection and Information Theories

(a) A Detection System (b) A Channel

equivalent to the transition probabilities for the information transmission problem as indicated in Figure 5.2. In this case, the input is a random variable H which may assume one of the two values 0 or 1, (H = i corresponds to the hypothesis H_i being present i = 0,1). The output is the decision random variable u which may again assume the value of 0 or 1. Therefore, the transition probabilities are given as follows.

The probability of deciding "0" when a "0" is sent, $P(u=0|H=0)$, is

$$P(u=0|H=0) = 1-P_F \qquad (5\text{-}6\text{-}a)$$

Similarly,

$$P(u=1|H=0) = P_F \qquad (5\text{-}6\text{-}b)$$

$$P(u=0|H=1) = P_M = 1 - P_D \qquad (5\text{-}6\text{-}c)$$

and,

$$P(u=1|H=1) = P_D \qquad (5\text{-}6\text{-}d)$$

As indicated earlier, in this chapter we will consider detection problems where we minimize the equivocation between the input and the output. Throughout this chapter, we will denote this problem by the MED problem. In this case, we are interested in minimizing an average cost, where the cost is not a constant but, is a function of the a

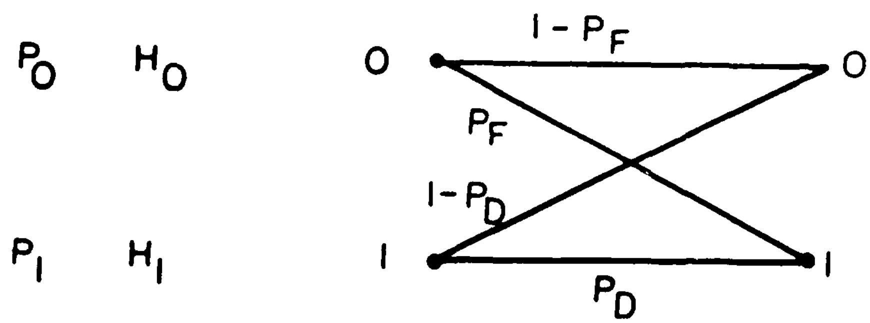


Figure 5.2 Transmission Channel

posteriori probability of H given u. This cost function is

$$C(u,H) = \frac{1}{P(H|u)} \tag{5-7}$$

Recall that, the average Bayesian risk is defined as

$$R_B = E(C(u,H)) \tag{5-8}$$

where C(u,H) is a constant for a given pair u and H. For the MED problem, the average cost is the conditional entropy of H given u, i.e.,

$$h(H|u) = E\ (\ \log [\ 1\ /\ P(H|u)]) \tag{5-9}$$

Knowing P_0 and P_1, our problem is to obtain decision rules so as to minimize the average cost h(H|u) which is equivalent to the problem of maximizing the mutual information I(H;u). This is obvious from the expression which relates I(H;u) and h(H|u), i.e.,

$$I(H;u) = h(H) - h(H|u) \tag{5-10}$$

and, realizing that h(H) is constant when the a priori probabilities P_0 and P_1 are known. Thus, the minimization of equivocation is equivalent to the maximization of I(H;u).

Later on in this subsection, we conclude that the point (P_D,P_F) corresponding to the detection system which maximizes the mutual information, lies on the receiver operating characteritic curve, (ROC), of an optimum Bayesian detector. In order to prove this result, we first need to prove the following theorem.

Theorem 5.1

Given the a priori probabilities P_0 and P_1, then for each value of the probability of false alarm P_F (or probability of detection P_D), the minimum mutual information $I_{min}(H;u)$ is achieved at the point where $P_D = P_F$.

Proof

For the channel model shown in Figure 5.2, the mutual information $I(H;u)$ is given by

$$I(H;u) = \sum_H \sum_u P(H,u) \log \left(\frac{P(H|u)}{P(H)} \right) \qquad (5\text{-}11)$$

and the a posteriori probabilities are the following

$$P(u=0) = P_0 (1 - P_F) + (1 - P_0) (1 - P_D)$$

$$= \alpha_0 \qquad (5\text{-}12\text{-}a)$$

and,

$$P(u=1) = P_0 P_F + (1 - P_0) P_D$$

$$= \alpha_1 \qquad (5\text{-}12\text{-}b)$$

where P_0 and P_1 are the a priori probabilities, $P(H_0)$ and $P(H_1)$ respectively. In Figure 5.3, we show a typical sketch of the mutual information as a function of P_D and P_F for

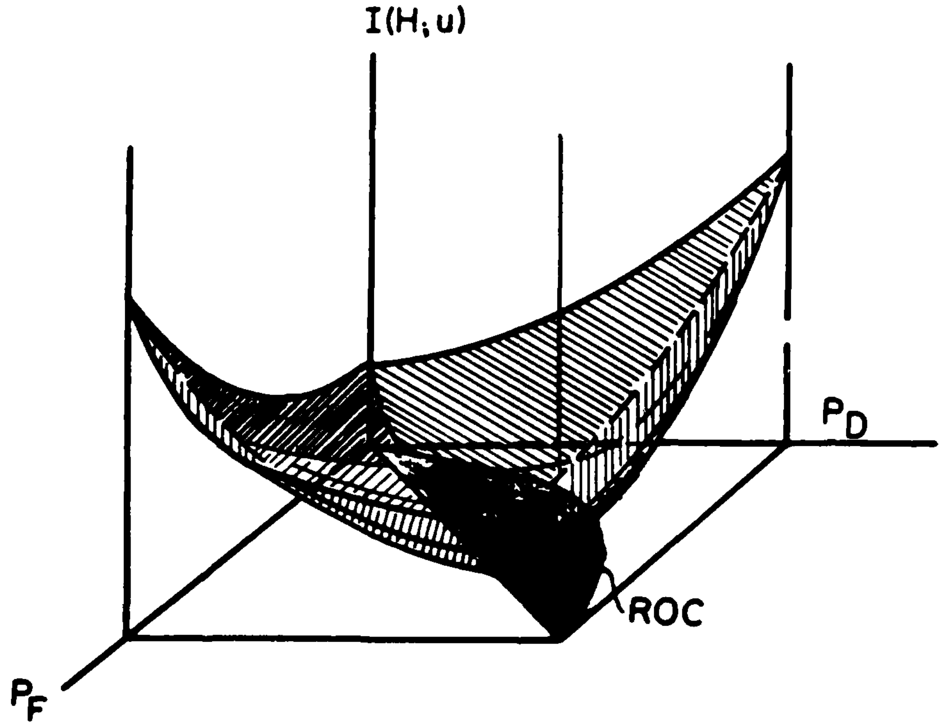


Figure 5.3 Typical Mutual Information Curve Versus the Transition Probabilities

given a priori probabilities. Substituting from (5-12) into (5-11) and cancelling terms, we get

$$I(H;u) = P_O\,(1-P_F)\ \log \frac{1-P_F}{P_O\,(1-P_F)+(1-P_O)(1-P_D)}$$
$$+\ P_O\ P_F\ \log \frac{P_F}{P_O\,P_F+(1-P_O)\,P_D}$$
$$+\ (1-P_O)(1-P_D)\ \log \frac{1-P_D}{P_O\,(1-P_F)+(1-P_O)(1-P_D)}$$
$$+\ (1-P_O)\,P_D\ \log \frac{P_D}{P_O\,P_F+(1-P_O)\,P_D} \qquad (5\text{-}13)$$

Depending on the value of P_O, It is known that the maximum of $I(H;u)$ is achieved by setting both P_M (P_M = $1-P_D$) and P_F, to be either "0" or "1", i.e., $P_F = P_M = 0$ or $P_F = P_M = 1$. $I(H;u)$ is a concave upward function in the transition probabilities $P(u|H)$ (refer to Theorem 1.7, [41], for details). Thus, for a fixed value of P_F (or P_D), $I(H;u)$ is concave upward function in P_D (or P_F), as shown in Figure 5.4. Therefore, for a given probability of false alarm P_{F1} of P_F (or a given value of probability of detection P_{D1} of P_D), if a minimum exists and is interior to the interval (0,1), the value of P_D which minimizes the mutual information $I(P_{F1},P_D)$ is obtained by setting the derivative of I(.) with respect to P_D equal to zero. Taking the derivative of I(.) with respect to P_D, we have

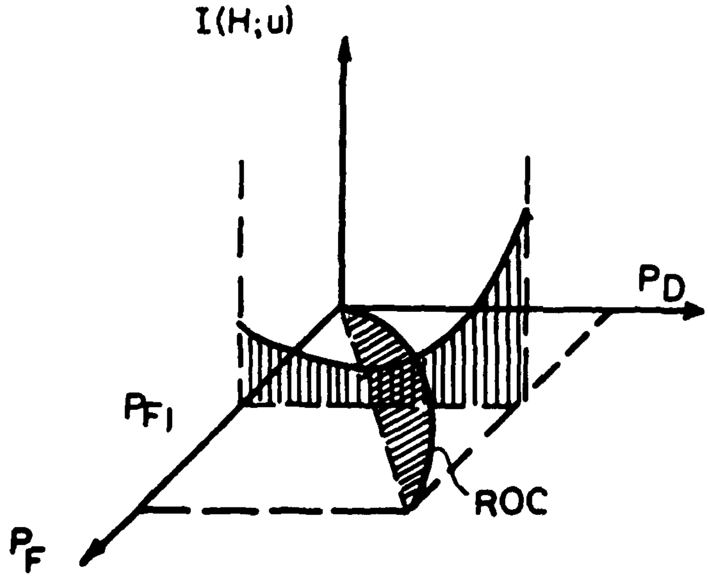


Figure 5.4 Mutual Information Versus the Probability of Detection for a Given Value P_{F1} of P_F

$$\frac{\delta I(H;u)}{\delta P_D} = + \frac{P_o\,(1 - P_{F1})(1 - P_o)}{P_o\,(1-P_{F1}) + (1-P_o)(1-P_D)} - \frac{P_o\,P_{F1}\,(1 - P_o)}{P_o\,P_{F1} + P_D(1-P_o)}$$

$$- (1-P_o) \log \frac{1 - P_D}{P_o\,(1-P_{F1}) + (1-P_o)(1-P_D)}$$

$$+ \frac{(1 - P_o)P_D}{P_D} - \frac{(1 - P_o)(1 - P_D)}{(1 - P_D)}$$

$$+ \frac{(1 - P_o)^2\,(1-P_D)}{P_o\,(1-P_{F1}) + (1-P_o)(1-P_D)}$$

$$+ (1-P_o) \log \frac{P_D}{P_o\,P_{F1} + (1-P_o)\,P_D}$$

$$- \frac{(1 - P_o)^2\,P_D}{P_o\,P_{F1} + (1-P_o)\,P_D} \tag{5-14}$$

Rearranging (5-14), we obtain

$$\frac{\delta I(H;u)}{\delta P_D} = (1 - P_o) \log \left(\frac{P_D\,[P_oP_D - P_oP_{F1} + (1 - P_D)]}{(1 - P_D)\,[P_D - P_oP_D + P_oP_{F1}]}\right) \tag{5-15}$$

Assuming P_o, $P_1 \neq 0$, the right hand side term is equal to zero if the argument of the logarithmic function is equal to 1, i.e.,

$$\frac{(1 - P_D)\,[P_D - P_oP_D + P_o\,P_{F1}]}{P_D\,[P_oP_D - P_oP_{F1} + (1 - P_D)]} = 1 \tag{5-16}$$

which yields

$$P_0 \, [P_D - P_{F1}] = 0 \qquad (5\text{-}16\text{-}a)$$

Since $P_0 \neq 0$, we have

$$P_D = P_{F1} \qquad (5\text{-}17)$$

Q.E.D

Now we show that the ROC of the optimum detector which maximizes $I(H;u)$ is in the shaded area of Figure 5.5, which is the region between the ROC corresponding to the optimum Bayesian detection rule and the $P_D = P_F$ line.

We use the following properties of the receiver operating characteristic curves (ROC) of optimum detectors

1. The ROC curves for optimum Bayesian detection systems, lie in the area above the line $P_D = P_F$, and

2. These ROC curves are always convex cap.

In Figure 5.6, we show some typical ROC curves associated with different decision rules for a given detection system. Curve (1) represents the ROC corresponding to the optimum detection rule (using the Bayesian formulation). Curve (2) is the inverted ROC [12], which corresponds to the worst detection rule. Therefore, the ROC corresponding to any detection rule for the system has to be within the region encircled by curves (1) and (2).

For all detection purposes, for a given value of P_F we would like to maximize the optimum probability of detection

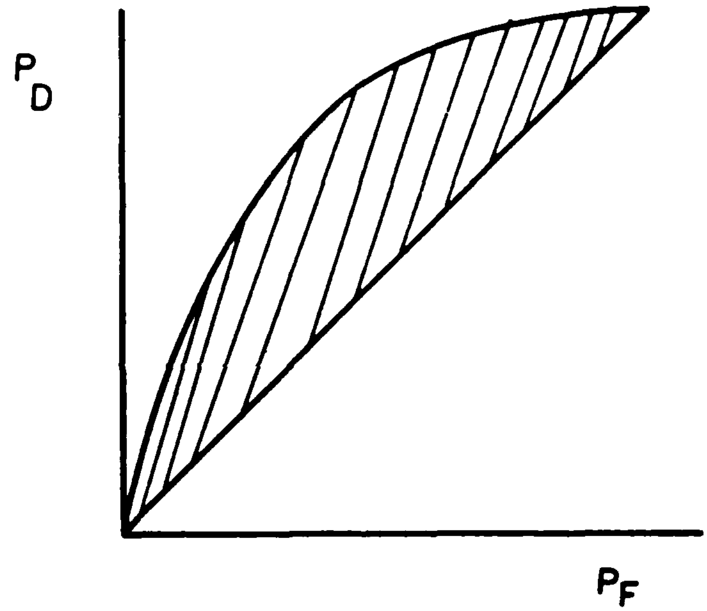


Figure 5.5 Receiver Operating Characteristic

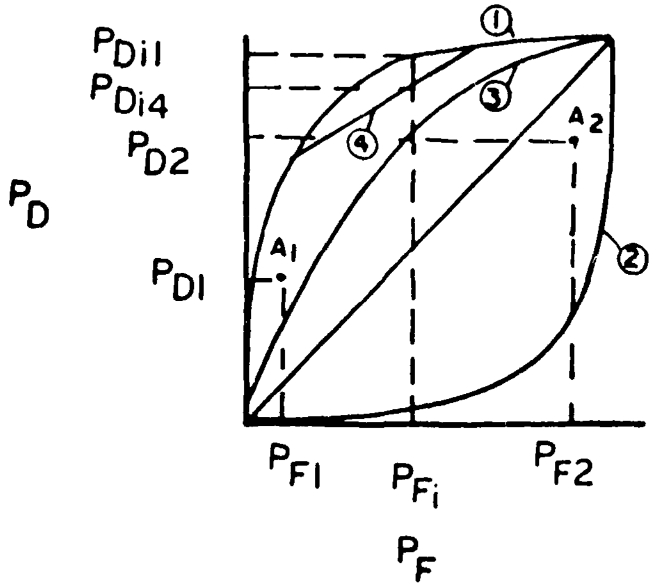


Figure 5.6 Optimum and Nonoptimum ROC

P_D. Mutual information is a symmetric function in its transition probabilities, i.e., the value of I(H|u) at any point A_1 with coordinates (P_{F1},P_{D1}), such that A_1 is above the $P_D = P_F$ line, is also achieved at a point A_2 with coordinates $(1-P_{F1},1-P_{D1})$ below the $P_D = P_F$ line, i.e., $I(P_F,P_D) = I(1-P_F,1-P_D)$. Therefore, we restrict the study to ROC curves above the $P_D = P_F$ line, i.e., the ROC curves lie in the shaded region of Figure 5.5.

ROC's corresponding to nonoptimum decision rules will have all or a part of the curve below the ROC corresponding to the optimum decision rule as shown by curves (3) and (4). Therefore, for any given probability of false alarm P_{F1}, such that $P_{F1} \in (0,1)$, the corresponding probability of detection on curve (1), P_{D11} (on the optimum ROC) and the one on curve (4), P_{D14} (on a nonoptimum ROC) satisfy the following

$$\text{for all } P_{F1} \text{ such that } P_{F1} \in (0,1),\ P_{D11} \geq P_{D14}$$

Finally, since I(H|u) is a concave upward function in P_D for a given value P_{F1} of P_F, and the minimum value of I(H|u) is achieved by choosing $P_D = P_{F1}$, then for values of P_D such that $P_D \geq P_{F1}$, I(H|u) is an increasing function in P_D. Therefore, if $P_{DJ} \geq P_{DK}$ then $I(P_{F1},P_{DJ}) \geq I(P_{F1},P_{DK})$ i.e., for a fixed value P_{F1}, the corresponding value of P_D obtained using the Neyman-Pearson criterion maximizes I(H|u). Therefore, the pair (P_D,P_F) which maximizes I(H|u) lies on the ROC of the optimum Bayesian detection system.

Next, we present the solution to the detection problem using one sensor for the minimum equivocation criterion.

5.2.3 Single Sensor Minimum Equivocation Detection

As pointed out earlier, the classical detection problem, utilizing a single sensor which minimizes equivocation has been solved in [29]. To validate our approach, in this subsection, we apply it to the single detector problem and verify that we obtain the same result. Later on in this chapter, we will extend this work to distributed multisensor situations.

In this problem, given the a priori probabilities P_0 and P_1, and the conditional densities $p(y|H_j)$ $j=0,1$, the objective is to find the optimum decision rule which maximizes the mutual information $I(H;u)$ (or minimize the equivocation $h(H|u)$).

We consider the system (channel) shown in Figure 5.2. We recall that the optimum decision rule which maximizes the mutual information $I(H,u)$ has its operating point (P_D,P_F) on the ROC of the optimum Bayesian detector which is a threshold detector. Thus, the optimum MED detector can be implemented as a threshold detector with the following likelihood ratio test

$$\Omega(y) = \frac{p(\underline{y}|H_1)}{p(\underline{y}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} t \qquad (5\text{-}18)$$

Also P_D and P_F are given by

$$P_D = P(u=1|H_1) = \int_t^{+\infty} p(\Omega|H_1)\, d\Omega \qquad (5\text{-}19\text{-}a)$$

and,

$$P_F = P(u=1|H_0) = \int_t^{+\infty} p(\Omega|H_0)\, d\Omega \qquad (5\text{-}19\text{-}b)$$

The mutual information is given by (5-13), or equivalently

$$\begin{aligned} I(H;u) = {} & P_0\,(1 - P_F)\,\log\,(1 - P_F) - P_0\,(1 - P_F)\,\log\,\alpha_0 \\ & - (1 - P_0)\,(1 - P_D)\,\log\,\alpha_0 + P_0\,P_F\,\log\,P_F \\ & - (1 - P_0)\,P_D\,\log\,\alpha_1 + (1 - P_0)\,P_D\,\log\,P_D \\ & + (1 - P_0)\,(1 - P_D)\,\log\,(1 - P_D) - P_0\,P_F\,\log\alpha_1 \end{aligned} \qquad (5\text{-}20)$$

where α_0 and α_1 are the a posteriori probabilities, given in (5-12). Taking the derivative of $I(H;u)$ with respect to P_F, and setting

$$\frac{\delta P_D}{\delta P_F} = t \tag{5-21}$$

and rearranging the result, we get

$$\frac{\delta I(H|u)}{\delta P_F} = - P_0 \log \frac{1 - P_F}{P_F} + [\, P_0 + t\ (1 - P_0)]\log \alpha_0$$

$$- t\ (1 - P_0) \log \frac{1 - P_D}{P_D} - [P_0 + t\ (1 - P_0)]\log \alpha_1 \tag{5-22}$$

Setting this derivative equal to zero and, solving for t, we get

$$t = \frac{- P_0\ [\ \log(\alpha_0/\alpha_1) - \log((1-P_F)/P_F)]}{(1 - P_0)\ [\ \log(\alpha_0/\alpha_1) - \log((1-P_D)/P_D)]} \tag{5-23}$$

Since P_D and P_F depend on t, the right hand side is also a function of t. After solving for t, P_D and P_F can be obtained. It can be shown that the value of t in (5-23) is the same as obtained by Middleton [29] and Gabrielle [37]. Next, we solve the MED problem for the DDF system.

5.3 Distributed Minimum Equivocation Detection with Data Fusion

In this section, we consider the binary hypothesis testing system shown in Figure 1.3. We assume that the observations at different detectors are independent. We find the optimum fusion rule and the optimum decision rules at the individual detectors which maximize the mutual information $I(H;u)$.

Again we assume that the global decision, u, depends only on the present decision vector $\underline{x}$ and that the decision u_i of detector i depends only on its own observation $\underline{y}_i$ $i=1,2,\ldots,N$. We will use the result obtained in Section 5.2, i.e., P_D and P_F for the system which maximizes the mutual information correspond to a point on the ROC of an optimum system with minimum Bayesian risk .

Using the results of Section 3.4 for independent observations, we conclude that the detectors used are threshold detectors. First, we assume that $I(H;u)$ is a function of a variable v, belonging to an interval (a,b) and, we find an equation in v which can be solved to yield the maximum of $I(H;u)$ with respect to v. Then, we assume that the detectors have already been designed and we obtain the optimum fusion rule which maximizes $I(H;u)$. After that, we assume that the fusion rule is known and we obtain the optimum decision rules at the individual detectors so as to maximize $I(H;u)$. Finally, we obtain the overall solution, i.e., we obtain the optimum fusion rule and the optimum decision rules at the detectors which maximize $I(H;u)$.

As we mentioned earlier, before we proceed with the solution to the problem, we present the following theorem.

Theorem 5.2

Let $I(H;u)$ be the mutual information between the random variables H and u. Let v be a variable such that $v \in (a,b)$. and $I(H;u) = f(v)$, (the mutual information is a function of v). The extremum (maximum or minimum) of $I(H;u)$ is obtained by solving

$$P_1(C_{11} - C_{01}) \frac{\delta P_D}{\delta v} - P_0(C_{00} - C_{10}) \frac{\delta P_F}{\delta v} \begin{array}{c} > \\ = \\ < \end{array} 0 \quad \begin{array}{l} \text{I is increasing in v} \\ \text{I has an extremum} \\ \text{I is decreasing in v} \end{array} \tag{5-24}$$

where

$$C_{ij} = \log \left(\frac{P(u_i, H_j)}{P(u_i)\, P(H_j)} \right) \qquad i,j = 0,1 \tag{5-25}$$

Proof

The mutual information $I(H;u)$ is given by (5-4). The logarithmic cost function is given in (5-25). Rearranging $I(H;u)$ and substituting C_{ij}, $i,j = 0,1$, into (5-4), we get

$$I(H;u) = P_0\,(C_{10} - C_{00})\,P_F - P_1\,(C_{01} - C_{11})\,P_D + P_0\,C_{00} + P_1 C_{01} \tag{5-26}$$

Let v be a variable such that v $\in$ (a,b), and I(H;u) is a function of v. The value v_{ext} of v, where the mutual information has an extremum (if the extremum exists and $\in$(a,b)), is obtained by setting the derivative of I(H;u) with respect to v equal to zero. Taking the partial derivative of I(H;u) with respect to v and rearranging the result, we get

$$\frac{\delta I(H;u)}{\delta v} = F(\delta v) + P_0 \, (C_{10} - C_{00}) \frac{\delta P_F}{\delta v} - P_1 \, (C_{01} - C_{11}) \frac{\delta P_D}{\delta v} \tag{5-27}$$

where

$$F(\delta v) = P_0 \left(\frac{\delta C_{10}}{\delta v} - \frac{\delta C_{00}}{\delta v} \right) P_F - P_1 \left(\frac{\delta C_{01}}{\delta v} - \frac{\delta C_{11}}{\delta v} \right) P_D$$

$$+ P_0 \frac{\delta C_{00}}{\delta v} + P_1 \frac{\delta C_{01}}{\delta v} \tag{5-28}$$

or equivalently

$$F(\delta v) = \frac{\delta C_{00}}{\delta v} [\, P_0 \, (1 - P_F)] + \frac{\delta C_{01}}{\delta v} [\, P_1 \, (1 - P_D)]$$

$$+ \frac{\delta C_{10}}{\delta v} P_0 \, P_F + \frac{\delta C_{11}}{\delta v} P_1 \, P_D \tag{5-29}$$

But

$$\frac{\delta C_{00}}{\delta v} = \frac{-\dfrac{\delta P_F}{\delta v}}{(1 - P_F)} - \frac{-\dfrac{\delta P_F}{\delta v} P_0 - \dfrac{\delta P_D}{\delta v} P_1}{P_0 (1 - P_F) + P_1 (1 - P_D)} \tag{5-30-a}$$

$$\frac{\delta C_{10}}{\delta v} = \frac{\dfrac{\delta P_F}{\delta v}}{P_F} - \frac{\dfrac{\delta P_F}{\delta v} P_0 + \dfrac{\delta P_D}{\delta v} P_1}{P_0 P_F + P_1 P_D} \tag{5-30-b}$$

$$\frac{\delta C_{01}}{\delta v} = \frac{-\dfrac{\delta P_D}{\delta v}}{(1 - P_D)} - \frac{-\dfrac{\delta P_F}{\delta v} P_0 - \dfrac{\delta P_D}{\delta v} P_1}{P_0 (1 - P_F) + P_1 (1 - P_D)} \tag{5-30-c}$$

and,

$$\frac{\delta C_{11}}{\delta v} = \frac{\dfrac{\delta P_D}{\delta v}}{P_D} - \frac{\dfrac{\delta P_F}{\delta v} P_0 + \dfrac{\delta P_D}{\delta v} P_1}{P_0 P_F + P_1 P_D} \tag{5-30-d}$$

Substituting (5-30) into (5-29) and rearranging, we have

$$F(\delta v) = P_0 \left[\frac{\delta P_F}{\delta v} - \frac{\delta P_F}{\delta v} \right] + P_1 \left[\frac{\delta P_D}{\delta v} - \frac{\delta P_D}{\delta v} \right]$$

$$+ \left[\frac{\delta P_F}{\delta v} P_0 + \frac{\delta P_D}{\delta v} P_1 \right] \left(\left[\frac{P_0 (1 - P_F) + P_1 (1 - P_D)}{P_0 (1 - P_F) + P_1 (1 - P_D)} - \frac{P_0 P_F + P_1 P_D}{P_0 P_F + P_1 P_D} \right] \right) \tag{5-31}$$

which is equal to zero. Then, $\frac{\delta I(H|u)}{\delta v}$ becomes

$$\frac{\delta I(H|u)}{\delta v} = P_0\,(C_{10} - C_{00})\,\frac{\delta P_F}{\delta v} - P_1\,(\,C_{01} - C_{11})\,\frac{\delta P_D}{\delta v} \tag{5-32-a}$$

or

$$\frac{\delta I(H|u)}{\delta v} = P_1\,(C_{11} - C_{01})\,\frac{\delta P_D}{\delta v} - P_0\,(\,C_{00} - C_{10})\,\frac{\delta P_F}{\delta v} \tag{5-32-b}$$

But, the following is always true

$$\frac{\delta I(H|u)}{\delta v} \begin{matrix} > \\ = \\ < \end{matrix} 0 \qquad \begin{matrix} I(H|u) \text{ is an increasing function in } v \\ I(H|u) \text{ has an extremum} \\ I(H|u) \text{ is a decreasing function in } v \end{matrix} \tag{5-33}$$

and thus, we have the desired result.

Q.E.D.

Now, we proceed with the solution to the problem. First, we solve for the optimum fusion rule.

<u>Optimum Fusion Rule</u>

Figure 1.3 shows the system under consideration. We assume that the detectors have already been designed. We

further assume that the final decision, u, depends on the decision vector $\underline{x}$ and not on the observations. The objective is to obtain the fusion rule, i.e., find $P_{1\underline{x}^*}$, the probability of deciding u = 1, when the present decision vector is $\underline{x}^* = (u_1^*, u_2^*, \ldots, u_N^*)^T$, a specific value of $\underline{x}$. Recall that the probability of detection P_D, and the probability of false alarm P_F are given by

$$P_D = \sum_{\underline{x}} P_{1\underline{x}} P_{\underline{x}1} \quad \text{and} \quad P_F = \sum_{\underline{x}} P_{1\underline{x}} P_{\underline{x}0} \tag{5-34}$$

Let the variable v, defined earlier, be the probability $P_{1\underline{x}^*}$, and (a,b) be the interval (0,1). Then, (5-32) becomes

$$\frac{\delta I(H|u)}{\delta P_{1\underline{x}^*}} = P_1 \, (C_{11} - C_{01}) \frac{\delta P_D}{\delta P_{1\underline{x}^*}} - P_0 \, (C_{00} - C_{10}) \frac{\delta P_F}{\delta P_{1\underline{x}^*}} \tag{5-35}$$

Substituting from (5-34) into (5-35), we have

$$\frac{\delta I(H|u)}{\delta P_{1\underline{x}^*}} = P_1 \, (C_{11} - C_{01}) \, P_{\underline{x}^*1} - P_0 \, (C_{00} - C_{10}) \, P_{\underline{x}^*0} \tag{5-36}$$

which is in general non zero. Thus, we have the following decision rule

$$P_1 \, (C_{11} - C_{01}) \, P_{\underline{x}^*1} - P_0 \, (C_{00} - C_{10}) \, P_{\underline{x}^*0} \quad \begin{array}{c} \text{Set } P_{1\underline{x}^*} = 1 \\ > \\ < \\ \text{Set } P_{1\underline{x}^*} = 0 \end{array} \quad 0 \tag{5-37}$$

We assume that

$$C_{00} \geq C_{10} \tag{5-38-a}$$

and,

$$C_{11} \geq C_{01} \tag{5-38-b}$$

which is equivalent to

$$\log \frac{P(u=i|H_i)}{P(u=i|H_j)} \geq 0 \tag{5-39-a}$$

or $\qquad i \neq j,\ i, j = 0, 1.$

$$\frac{P(u=i|H_i)}{P(u=i|H_j)} \geq 1 \tag{5-39-b}$$

Therefore,

$$\begin{array}{c} P(u=1|\underline{x}^*) = 1 \\ \dfrac{P_{\underline{x}^*1}}{P_{\underline{x}^*0}} \begin{array}{c} > \\ < \end{array} \dfrac{P_0\ (C_{00} - C_{10})}{P_1\ (C_{11} - C_{01})} \\ P(u=1|\underline{x}^*) = 0 \end{array} \tag{5-40}$$

Next, we find the optimum decision rules at the detectors.

<u>Optimum Decision Rules at The Detectors</u>

We again consider the system shown in Figure 1.3. We assume that the fusion rule is known and we find the decision rules at the individual detectors. While deriving the decision rule at detector μ, $\mu = 1,2,\ldots,N$, we assume that other detectors have already been designed. Using the same approach used in finding the fusion rule and, setting $v = P_{F\mu}$ and $(a,b) = (0,1)$, the derivative of $I(H;u)$ with respect to $P_{F\mu}$ is given by

$$\frac{\delta I(H;u)}{\delta P_{F\mu}} = P_1\ (C_{11} - C_{01})\ \frac{\delta P_D}{\delta P_{F\mu}} - P_0\ (C_{00} - C_{10})\ \frac{\delta P_F}{\delta P_{F\mu}} \tag{5-41}$$

Using the chain rule, we have

$$\frac{\delta P_D}{\delta P_{F\mu}} = \frac{\delta P_D}{\delta P_{D\mu}}\ t_\mu \tag{5-42}$$

where we have used the result obtained previously, i.e., the detectors are threshold detectors and the fact that

$$t_\mu = \delta P_{D\mu}/\delta P_{F\mu}$$

Substituting (5-42) into (5-41) , we have

$$\frac{\delta I(H;u)}{\delta P_{F\mu}} = P_1\ (C_{11} - C_{01})\ \frac{\delta P_D}{\delta P_{D\mu}}\ t_\mu - P_0\ (C_{00} - C_{10})\ \frac{\delta P_F}{\delta P_{F\mu}} \tag{5-43}$$

Setting (5-43) equal to zero and solving for t_μ, we get

$$t_\mu = \frac{P_0\ (C_{10} - C_{00})\ \dfrac{\delta P_F}{\delta P_{F\mu}}}{P_1\ (C_{01} - C_{11})\ \dfrac{\delta P_D}{\delta P_{D\mu}}} \tag{5-44}$$

As expected, the expression of t_k obtained here has the same form as the one we found in the Bayesian case, except that here C_{ij} $i,j = 0,1$ are not constants. We get N coupled, nonlinear equations, which sould be solved simultaneously to yield the desired result.

The overall solution to the problem, i.e., finding the optimum decision rules at the individual detectors and the optimum fusion rule which maximize $I(H;u)$ (or minimize $h(H|u)$) can be obtained by solving 2^N equations of the form (5-40) and N equations of the form (5-44), simultaneously. Many solutions may result. Only the ones which correspond to the absolute maximum of $I(H;u)$ (or minimum of $h(H|u)$) are to be kept.

5.4 Distributed Minimum Equivocation Detection

5.4.1 Introduction

In this section, we find the optimum decision rules at the individual detectors which maximize the mutual information, $I(H;\underline{x})$, for the system shown in Figure 5.7. The binary hypothesis testing problem is considered and, the observations at the individual detectors are assumed to be independent i.e.,

$$p(Y|H_j) = \prod_{i=1}^{N} p(y_i|H_j) \qquad (5\text{-}45)$$

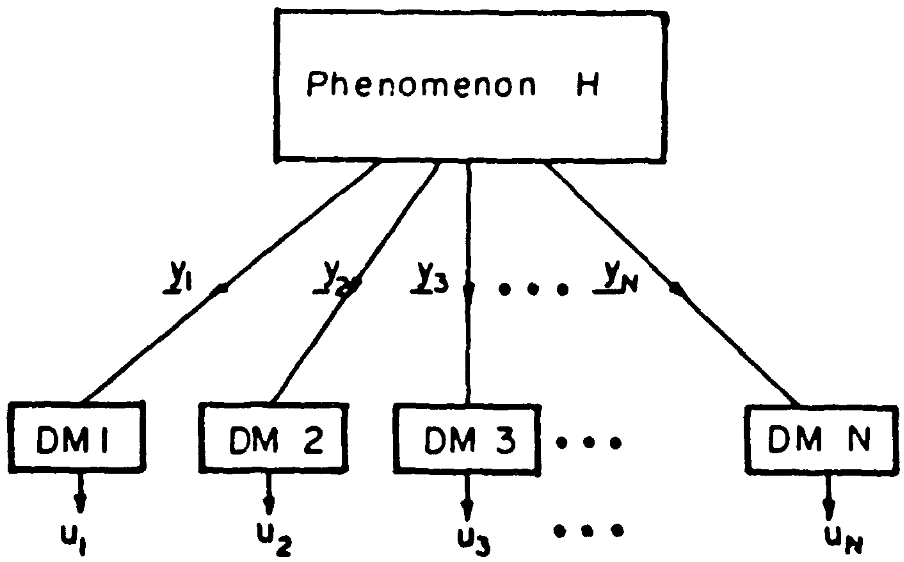


Figure 5.7 Distributed Sensor System with Local Inference

We recall that in this case, x is the decision vector and is given by $\underline{x} = (u_1, u_2, \ldots, u_N)^T$ where u_i is the decision of detector i, $i=1,2,\ldots,N$, and the decision of each detector only depends on its own observation.

First, we will show that, each detector should use a likelihood ratio test (threshold detector). Then, we will derive the expressions which yield the optimum threshold at the individual detector. Later on in this section, we will present a suboptimum solution to the problem considered in Section 5.3. In the suboptimum solution, we will solve the problem of optimizing the mutual information in two separate stages, i.e., in the first stage, we maximize $I(H;\underline{x})$ and then in the second stage, we maximize $I(\underline{x};u)$.

Next, we will show that for the detection system shown in Figure 5.7 with multiple sensors and minimum equivocation criterion, the detectors are threshold detectors. Before we consider the general case of N detectors ($N > 2$), we consider the case of two detectors.

Two-Detector Case

In this case, we use an approach similar to the one used in the previous section. First, we show that if P_{F1}, P_{F2} and P_{D2} are given, the value of P_{D1} which minimizes $I(H;\underline{x})$ is equal to P_{F1}. Then, we use the convexity property of $I(H;\underline{x})$ to show that the pair (P_{Fi}, P_{Di}), $i = 1,2$, which maximizes $I(H;\underline{x})$ is a point on the receiver operating characteristic curve (ROC) of the optimum threshold detector, using the Bayesian formulation. Then, we find the thresholds which maximize $I(H,\underline{x})$.

Consider the channel model shown in Figure 5.8. The transition probabilities are

$$P(u_1=0,u_2=0|H_0) = (1 - P_{F1})\ (1 - P_{F2}) \tag{5-46-a}$$

$$P(u_1=0,u_2=0|H_1) = (1 - P_{D1})\ (1 - P_{D2}) \tag{5-46-b}$$

$$P(u_1=0,u_2=1|H_0) = (1 - P_{F1})\ P_{F2} \tag{5-46-c}$$

$$P(u_1=0,u_2=1|H_1) = (1 - P_{D1})\ P_{D2} \tag{5-46-d}$$

$$P(u_1=1,u_2=0|H_0) = P_{F1}\ (1 - P_{F2}) \tag{5-46-e}$$

$$P(u_1=1,u_2=0|H_1) = P_{D1}\ (1 - P_{D2}) \tag{5-46-f}$$

$$P(u_1=1,u_2=1|H_0) = P_{F1}\ P_{F2} \tag{5-46-g}$$

and,

$$P(u_1=1,u_2=1|H_1) = P_{D1}\ P_{D2} \tag{5-46-h}$$

The output probabilities $P(\underline{x} = 00)$, $P(\underline{x} = 01)$, $P(\underline{x} = 10)$ and $P(\underline{x} = 11)$ are denoted by α_{00}, α_{01}, α_{10} and α_{11} respectively. They are given by

$$\alpha_{00} = P_0\ (1- P_{F1})(1 - P_{F2}) + (1 - P_0)(1 - P_{D1})(1 - P_{D2}) \tag{5-47-a}$$

$$\alpha_{01} = P_0\ (1 - P_{F1})\ P_{F2} + (1 - P_0)\ (1 - P_{D1})\ P_{D2} \tag{5-47-b}$$

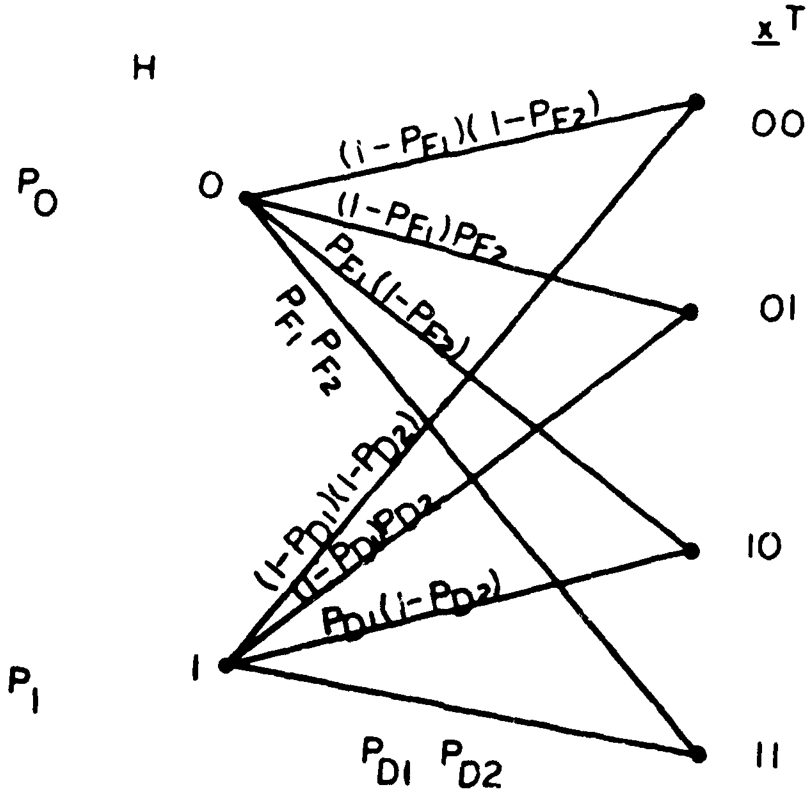


Figure 5.8 Single-Input Double-Output Channel

$$\alpha_{10} = P_o P_{F1} (1 - P_{F2}) + (1 - P_o) P_{D1} (1 - P_{D2}) \tag{5-47-c}$$

and,

$$\alpha_{11} = P_o P_{F1} P_{F2} + (1 - P_o) P_{D1} P_{D2} \tag{5-47-d}$$

The joint probabilities are

$$P(u_1=0,u_2=0,H_o) = P_o(1 - P_{F1}) (1 - P_{F2}) \tag{5-48-a}$$

$$P(u_1=0,u_2=0,H_1) = (1 - P_o)(1 - P_{D1}) (1 - P_{D2}) \tag{5-48-b}$$

$$P(u_1=0,u_2=1,H_o) = P_o (1 - P_{F1}) P_{F2} \tag{5-48-c}$$

$$P(u_1=0,u_2=1,H_1) = (1 - P_o) (1 - P_{D1}) P_{D2} \tag{5-48-d}$$

$$P(u_1=1,u_2=0,H_o) = P_o P_{F1} (1 - P_{F2}) \tag{5-48-e}$$

$$P(u_1=1,u_2=0,H_1) = (1 - P_o) P_{D1} (1 - P_{D2}) \tag{5-48-f}$$

$$P(u_1=1,u_2=1,H_o) = P_o P_{F1} P_{F2} \tag{5-48-g}$$

and,

$$P(u_1=1,u_2=1,H_1) = (1 -P_o) P_{D1} P_{D2} \tag{5-48-h}$$

where again P_o and P_1 are the a priori probabilities, and P_{Fi} and P_{Di} are the probability of false alarm and the probability of detection of detector i, i = 1,2. Substituting (5-46), (5-47) and (5-48) into (5-4), $I(H|\underline{x})$

becomes

$$
\begin{aligned}
I(H;x) = -P_o \,[\, & \log(P_o) + P_{F1}P_{F2}\log(P_oP_{F1}P_{F2}) \\
& + (1-P_{F1})(1-P_{F2})\log(P_o(1-P_{F1})(1-P_{F2})) \\
& + (1-P_{F1})P_{F2}\log(P_o(1-P_{F1})P_{F2}) \\
& + P_{F1}(1-P_{F2})\log(P_oP_{F1}(1-P_{F2}))\,] \\
& +(1-P_o)\,[\,(1-P_{D1})(1-P_{D2})\log((1-P_o)(1-P_{D1})(1-P_{D2})) \\
& \quad + (1-P_{D1})P_{D2}\log((1-P_o)(1-P_{D1})P_{D2}) \\
& \quad + P_{D1}(1-P_{D2})\log((1-P_o)P_{D1}(1-P_{D2})) \\
& \quad + P_{D1}P_{D2}\log((1-P_o)P_{D1}P_{D2})-\log((1-P_o))] \\
& - [\alpha_{oo}\log(\alpha_{oo}) + \alpha_{o1}\log(\alpha_{o1}) + \alpha_{1o}\log(\alpha_{1o}) \\
& \quad + \alpha_{11}\log(\alpha_{11})\,]
\end{aligned}
\tag{5-49}
$$

Taking the derivative of $I(H;x)$ with respect to P_{F1}, we get

$$
\frac{\delta I(H;x)}{\delta P_{F1}} = -P_o \log\left(\frac{1-P_{F1}}{P_{F1}}\right) + P_o\,(1-P_{F2})\log\frac{\alpha_{oo}}{\alpha_{o1}} + P_o\,P_{F2}\log\frac{\alpha_{1o}}{\alpha_{11}} \tag{5-50}
$$

Setting $P_{F1} = P_{D1}$ and cancelling terms, we have

$$\left.\frac{\delta I(H;\underline{x})}{\delta P_{F1}}\right|_{P_{F1} = P_{D1}} = -P_o \log\left(\frac{1 - P_{D1}}{P_{D1}}\right) + P_o P_{FE} \log\left(\frac{1 - P_{D1}}{P_{D1}}\right)$$

$$+ P_o (1 - P_{FE}) \log\left(\frac{1 - P_{D1}}{P_{D1}}\right)$$

$$= \left[\log\left(\frac{1 - P_{D1}}{P_{D1}}\right)\right] [- P_o + P_o - P_o P_{FE} + P_o P_{FE}]$$

$$= 0 \tag{5-51}$$

We conclude that the point $P_{F1} = P_{D1}$ corresponds to at least a lo_al minimum for given values of P_{F1}, P_{FE} an P_{DE}. Since $I(H;\underline{x})$ is convex in P_{D1} (Theorem 1.7, [41]), this minimum is the absolute minimum of $I(H;\underline{x})$ for given P_{F1}, P_{FE} and P_{DE}. Similar argument is valid for the second detector. Using this result and an argument similar to the one in the previous section, it can be shown that the maximum $I(H;\underline{x})$ is achieved by setting the (P_{Fi} and P_{Di}) i=1,2 as points on the ROC of the optimum threshold detectors of detectors 1 and 2 with Bayesian criterion.

<u>N-Detector Case</u>

The same argument may be used for the general case of N detectors, N $\geq$ 2. We conclude this subsection by saying that the detection system which maximizes $I(H;\underline{x})$ has its operating point (P_D,P_F) on the ROC of the optimum Bayesian detector and in the case of independent observations the detectors would be threshold detectors.

5.4.2 Optimum DD System

In this subsection, we find the thresholds which maximize $I(H;\underline{x})$ for the DD system shown in Figure 5.8.

We recall that for the threshold detectors, the probability of detection P_{Di} and the probability of false alarm P_{Fi} $i=1,2,\dots,N$ are respectively given by

$$P_{Di} = \int_{t_i}^{\infty} p(\Omega_i|H_1)\, d\Omega_i \qquad i = 1,2,\dots,N \qquad (5\text{-}52\text{-}a)$$

and,

$$P_{Fi} = \int_{t_i}^{\infty} p(\Omega_i|H_0)\, d\Omega_i \qquad i = 1,2,\dots,N \qquad (5\text{-}52\text{-}b)$$

where t_i is the threshold of detector i and Ω_i is the likelihood ratio defined as

$$\Omega_i(\underline{y}_i) = \frac{p(\underline{y}_i|H_1)}{p(\underline{y}_i|H_0)} \qquad i = 1,2,\dots,N \qquad (5\text{-}53)$$

For each detector the likelihood ratio test is the following

$$\Omega_i(\underline{y}_i) \underset{H_1}{\overset{H_0}{\gtrless}} t_i \qquad i = 1,2,\ldots,N \qquad (5\text{-}54)$$

and,

$$t_i = \frac{\delta P_{Di}}{\delta P_{Fi}} \qquad i = 1,2,\ldots,N \qquad (5\text{-}55)$$

<u>Two Detector Case</u>

In this case, $I(H|\underline{x})$ is given by (5-49). Taking the derivative of $I(H|\underline{x})$ with respect to P_{Fi}, we have

$$\frac{\delta I(H|\underline{x})}{\delta P_{Fi}} = -P_o \log\left(\frac{1 - P_{Fi}}{P_{Fi}}\right) - (1 - P_o)\frac{\delta P_{Di}}{\delta P_{Fi}} \log\left(\frac{1 - P_{Di}}{P_{Di}}\right)$$

$$- \frac{\delta}{\delta P_{Fi}}[\, \alpha_{00} \log \alpha_{00} + \alpha_{10} \log \alpha_{10} + \alpha_{01} \log \alpha_{01} + \alpha_{11} \log \alpha_{11} \,] \qquad (5\text{-}56)$$

Rearranging (5-56), setting the result equal to zero and solving for t_i, we get

$$t_1 = - \frac{P_0 \left[\log\left(\frac{\alpha_{00}}{\alpha_{10}}\right) + P_{F2}\log\left(\frac{\alpha_{01}\,\alpha_{10}}{\alpha_{00}\,\alpha_{11}}\right) - \log\left(\frac{1 - P_{F1}}{P_{F1}}\right)\right]}{(1-P_0)\left[\log\left(\frac{\alpha_{00}}{\alpha_{10}}\right) + P_{D2}\log\left(\frac{\alpha_{01}\,\alpha_{10}}{\alpha_{00}\,\alpha_{11}}\right) - \log\left(\frac{1 - P_{D1}}{P_{D1}}\right)\right]} \qquad \text{(5-57-a)}$$

Similarly,

$$t_2 = - \frac{P_0 \left[\log\left(\frac{\alpha_{00}}{\alpha_{01}}\right) + P_{F1}\log\left(\frac{\alpha_{01}\,\alpha_{10}}{\alpha_{00}\,\alpha_{11}}\right) - \log\left(\frac{1 - P_{F2}}{P_{F2}}\right)\right]}{(1-P_0)\left[\log\left(\frac{\alpha_{00}}{\alpha_{01}}\right) + P_{D1}\log\left(\frac{\alpha_{01}\,\alpha_{10}}{\alpha_{00}\,\alpha_{11}}\right) - \log\left(\frac{1 - P_{D2}}{P_{D2}}\right)\right]} \qquad \text{(5-57-b)}$$

Solving (5-57-a) and (5-57-b) simultaneously yields the optimum thresholds which maximize the mutual information.

<u>Three Detector Case</u>

In the case of 3 detectors, we may use the same procedure. Defining $\alpha_{\underline{x}} = \alpha_{u_1u_2u_3}$ as the probability of the output vector $\underline{x} = (u_1, u_2, u_3)^T$, the expression of t_1 is then,

$$t_1 = - \left(\frac{P_0}{1 - P_0} \right) \left(\frac{K_1}{K_2} \right) \qquad \text{(5-58-a)}$$

where

$$K_1 = \log \left(\frac{1 - P_{F1}}{P_{F1}}\right) - (1 - P_{F2})(1 - P_{F3}) \log \frac{\alpha_{000}}{\alpha_{100}}$$

$$- (1 - P_{F2})\, P_{F3} \log \frac{\alpha_{001}}{\alpha_{101}} + P_{F2}\, (1 - P_{F3}) \log \frac{\alpha_{010}}{\alpha_{110}}$$

$$- P_{F2}\, P_{F3} \log \frac{\alpha_{011}}{\alpha_{111}} \qquad \text{(5-58-b)}$$

and,

$$K_2 = \log \left(\frac{1 - P_{D1}}{P_{D1}}\right) - (1 - P_{D2})(1 - P_{D3}) \log \frac{\alpha_{000}}{\alpha_{100}}$$

$$- (1 - P_{D2})\, P_{D3} \log \frac{\alpha_{001}}{\alpha_{101}} + P_{D2}\, (1 - P_{D3}) \log \frac{\alpha_{010}}{\alpha_{110}}$$

$$- P_{D2}\, P_{D3} \log \frac{\alpha_{011}}{\alpha_{111}} \qquad \text{(5-58-c)}$$

Similar expressions are available for t_2 and t_3. Solving the resulting three equations simultaneously yields the optimum solution.

General case

In a similar manner, for the case of N detectors shown in Figure 5.7, we get the following expression for the threshold of the μ^{th} detector, $\mu = 1,2,\ldots,N$,

$$t_\mu = \frac{-\; P_0 \left[\log \left(\dfrac{1 - P_{F\mu}}{P_{F\mu}} \right) - \sum\limits_{\underline{x}} F_x^\mu \log \left(\alpha_x^{\mu 0} / \alpha_x^{\mu 1} \right) \right]}{(1-P_0) \left[\log \left(\dfrac{1 - P_{D\mu}}{P_{D\mu}} \right) - \sum\limits_{\underline{x}} D_x^\mu \log \left(\alpha_x^{\mu 0} / \alpha_x^{\mu 1} \right) \right]} \tag{5-59}$$

where F_x^μ and D_x^μ are defined in (2-11) and $\sum\limits_x$ is the summation over all possible decision vectors $\underline{x}$. Therefore, we have N nonlinear coupled equations. Solving these equations simultaneously yields the optimum solution, i.e., optimum decision rules at the detectors to maximize $I(H;\underline{x})$ (or minimize the equivocation $h(H|\underline{x})$. We may get many solutions, only the feasible solutions are to be kept.

5.4.3 Suboptimum Solution for the DDF system

When we considered the DDF system with N sensors in Section 5.3, we concluded that a simultaneous solution of $N + 2^N$ coupled nonlinear equations is required. This is a difficult task especially when N becomes large. In this section we consider a suboptimum solution where we optimize the system in two stages , i.e., first we maximize $I(H;\underline{x})$ and then, maximize $I(\underline{x};u)$.

Maximization of $I(H;\underline{x})$ has already been considered in Section 5.4.2. While maximizing $I(\underline{x};u)$, we assume that the detectors have already been designed, i.e., the probability of $\underline{x} = (u_1, u_2, \ldots, u_N)^T$ is known for all $\underline{x}$ which makes $h(\underline{x})$ a constant. We wish to maximize $I(\underline{x},u)$ (or minimize $h(\underline{x}|u)$) for the system shown in Figure 5.9. We use the following expressions of $I(\underline{x};u)$ and $h(\underline{x}|u)$

$$I(\underline{x};u) = h(\underline{x}) + h(u) - h(\underline{x},u) \tag{5-60-a}$$

and,

$$h(\underline{x}|u) = h(\underline{x},u) - h(u) \tag{5-60-b}$$

Minimizing the equivocation is equivalent to simultaneously minimizing $h(x,u)$ and maximizing $h(u)$, if possible. We recall that $h(x,u)$ is a convex downward (cap) function in $P(\underline{x},u)$. Thus, for a given value of $P(\underline{x})$ the absolute minimum for $h(\underline{x},u)$ is achieved by setting $P(u|\underline{x})$ to be either "0" or "1". A corresponding maximum of $h(u)$ is possible by choosing the values of $P(u|x)$ which set $P(u=0)$ and $P(u=1)$ as close as possible to 1/2, because the absolute maximum of $h(u)$ is 1 for $P(u=0) = P(u=1) = 1/2$. Note that it is desirable to set $u = 0$ when $x = (0,\ldots,0)^T$ and $u = 1$ when $\underline{x} = (1,\ldots,1)^T$.

Let us consider the question if it is possible that the maximum value of $I(\underline{x};u)$ (or minimum value of $h(\underline{x}|u)$) is achieved at points where $P(u|x)$ are not set to be 0's or 1's, e.g. by setting $P(u)$ such that $P(u=0) = P(u=1) = 1/2$ which maximizes $h(u)$, and set $P(u|x)$ accordingly, or any other possible combination. In answering this, we refer to the fact that since $h(\underline{x}|u)$ is convex downward function in

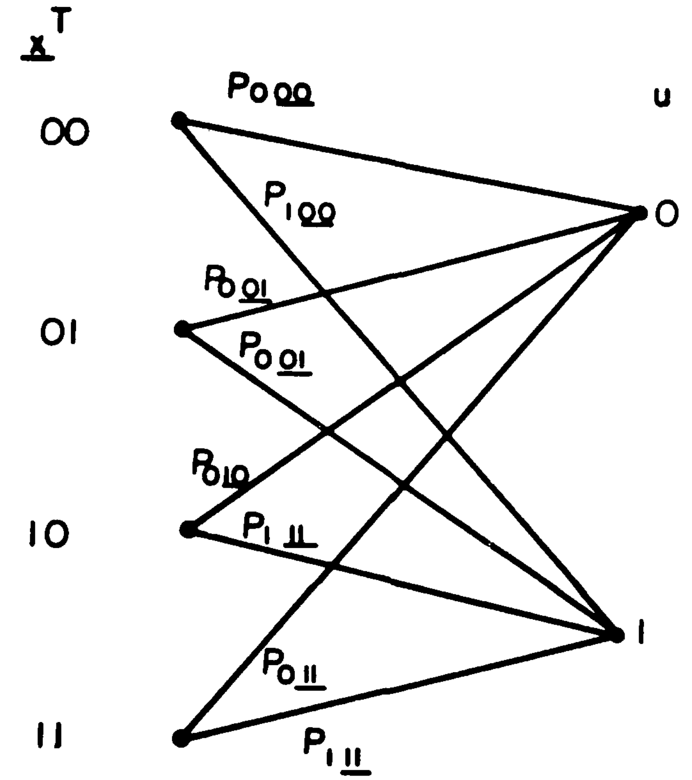


Figure 5.9 Double-Input Single-Output Channel

the transition probabilities, it has to achieve its maximum at a point where the transition probabilities form a combination of "0's" and "1's". Therefore, $P(u|\underline{x})$ can take the values zero or one only, (for all possible combinations, (0,1) $h(\underline{x},u)$ is the same).

5.5 Numerical Examples

Example 5.1

In this example, we consider the same system as considered in the example of Section 3.5. The system has two detectors with independent observations given by (3-88). We will design the system (both DD and DDF) so as to minimze the equivocation between the input and the output.

Recall that in this case, the detectors are threshold detectors. In Figure 5.10, we present the ROC curve for this example (DDF system) when $\theta_1 = 2$ and $\theta_2 = 4$. Again, in this case the "OR" fusion rule is superior to the "AND" fusion rule. We also show the curve of $I(H;u)$ versus P_0 in Figure 5.11. Figure 5.12 shows the curves of t_1' and t_2' versus P_0, where t_i', $i = 1,2$, is the value of the threshold at detector i, i= 1,2 as defined (3-90).

For the DD configuration, when $(\theta_1,\theta_2) = (2,4)$, the curve of $I(H;\underline{x})$ versus P_0 is shown in Figure 5.13. The curves of t_1' and t_2' are shown in Figure 5.14. When $(\theta_1,\theta_2) = (3,6)$, the curves of t_1' and t_2' are shown in

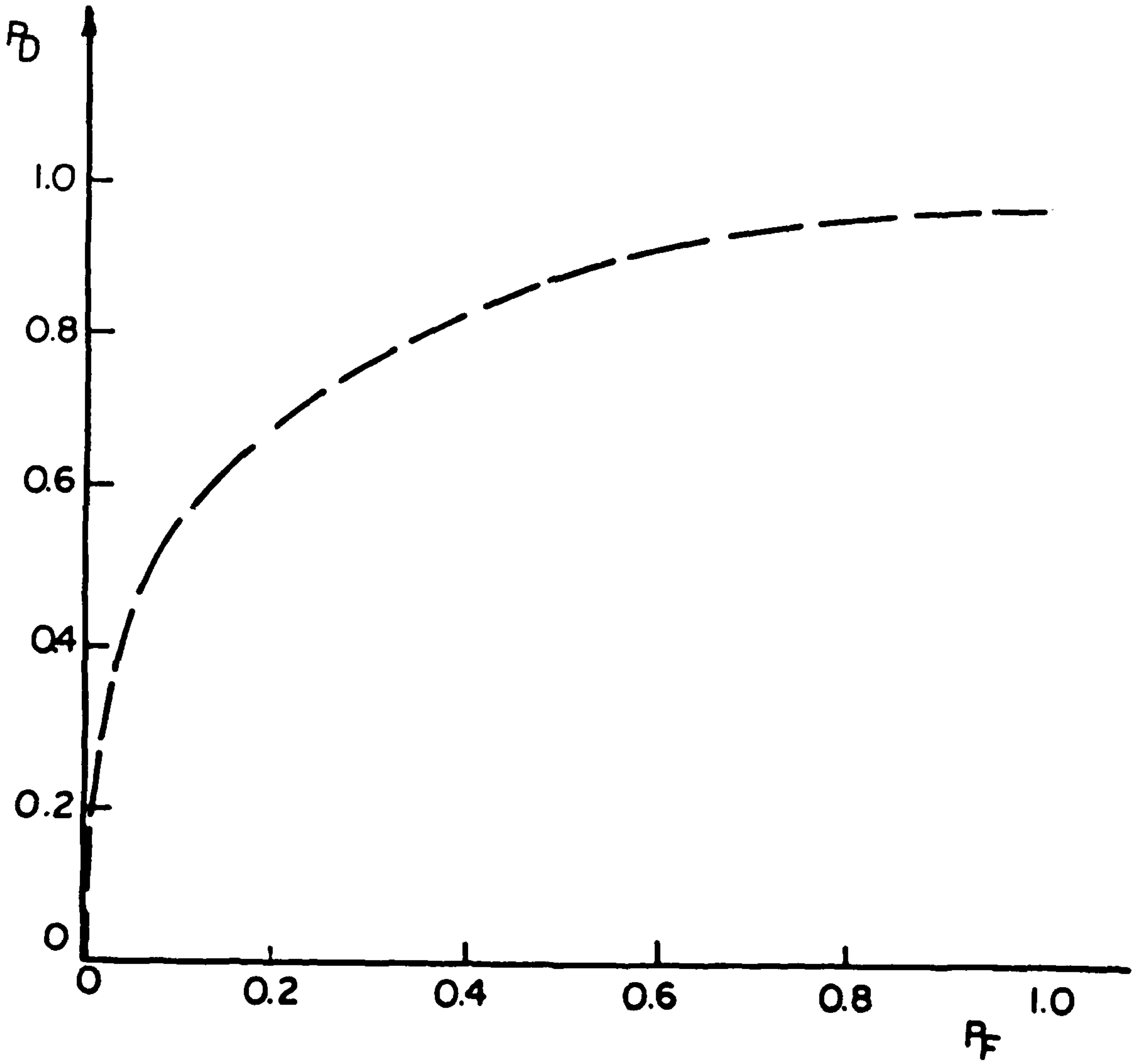


Figure 5.10 Receiver Operating Characteristic for Example 5.1 and $(\theta_1, \theta_2) = (2,4)$. (DDF Configuration)

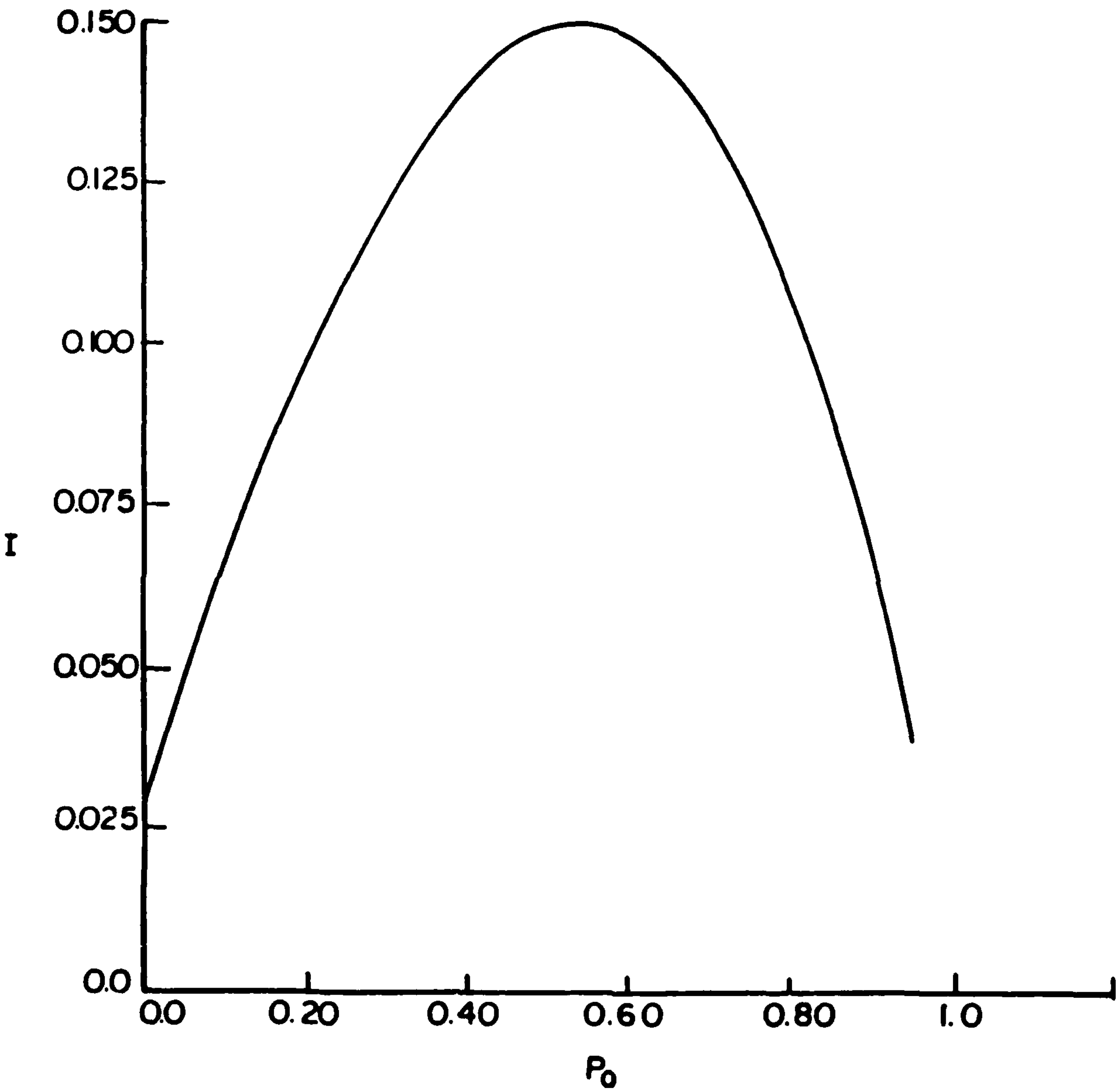


Figure 5.11 $I(H;u)$ Versus P_0 for Example 5.1 and $(\theta_1,\theta_2) = (2,4)$. (DDF Configuration)

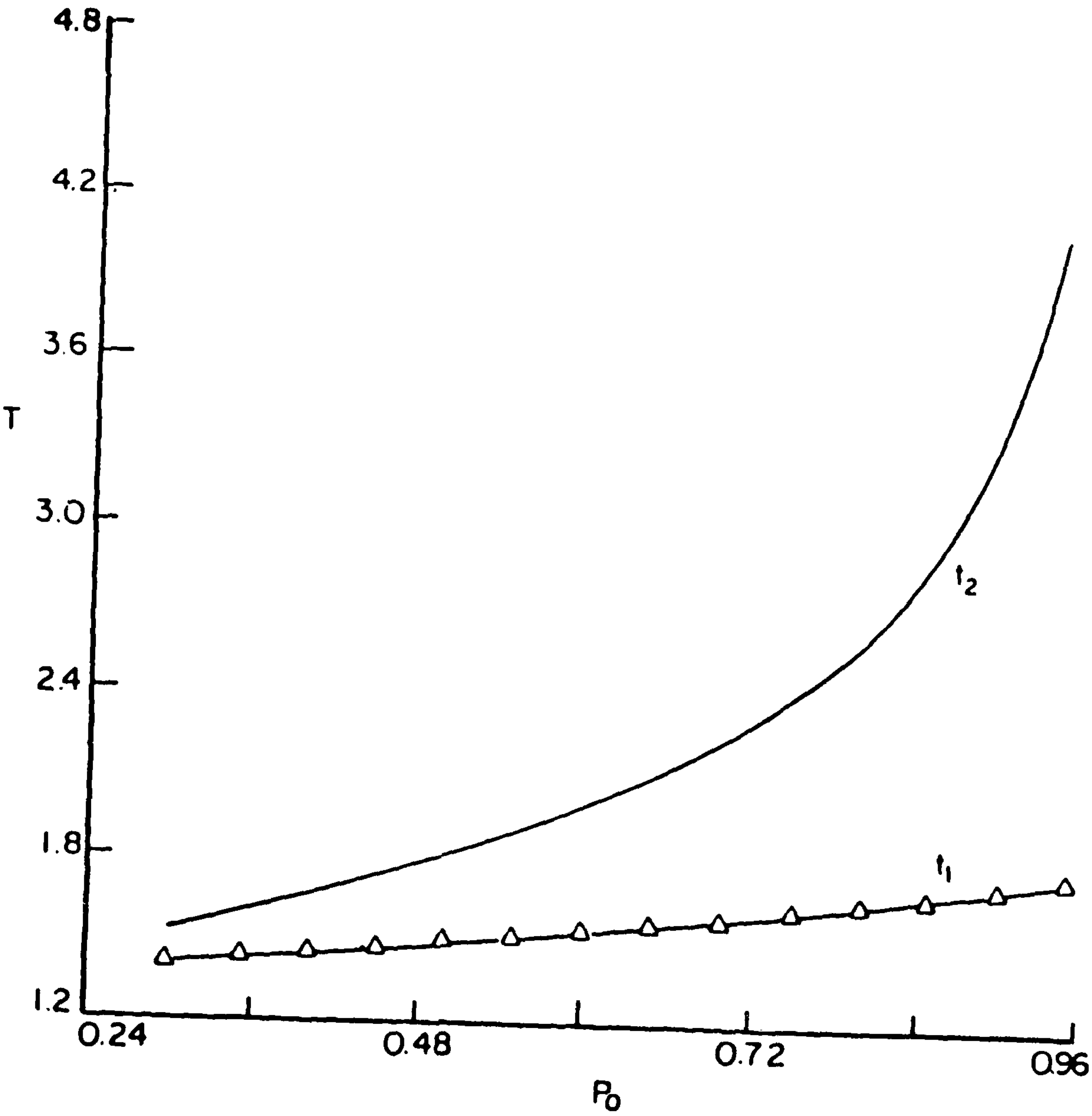


Figure 5.12 t'_1 and t'_2 Versus P_0 for Example 5.1 and $(\theta_1, \theta_2) = (2,4)$ (DDF Configuration)

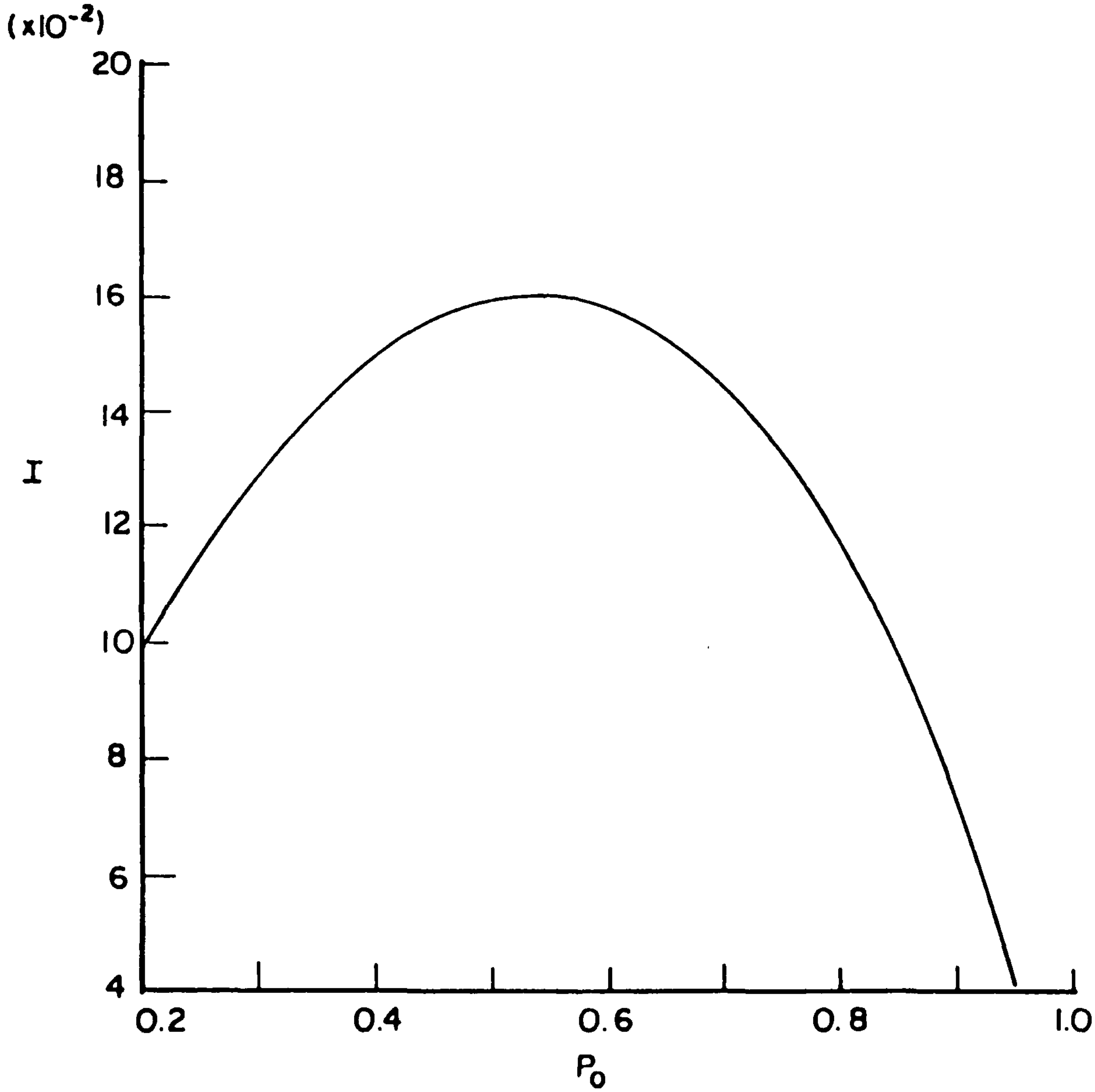


Figure 5.13 $I(H;\mathbf{y})$ Versus P_0 for Example 5.1 and $(\theta_1, \theta_2) = (2,4)$ (DD Configuration)

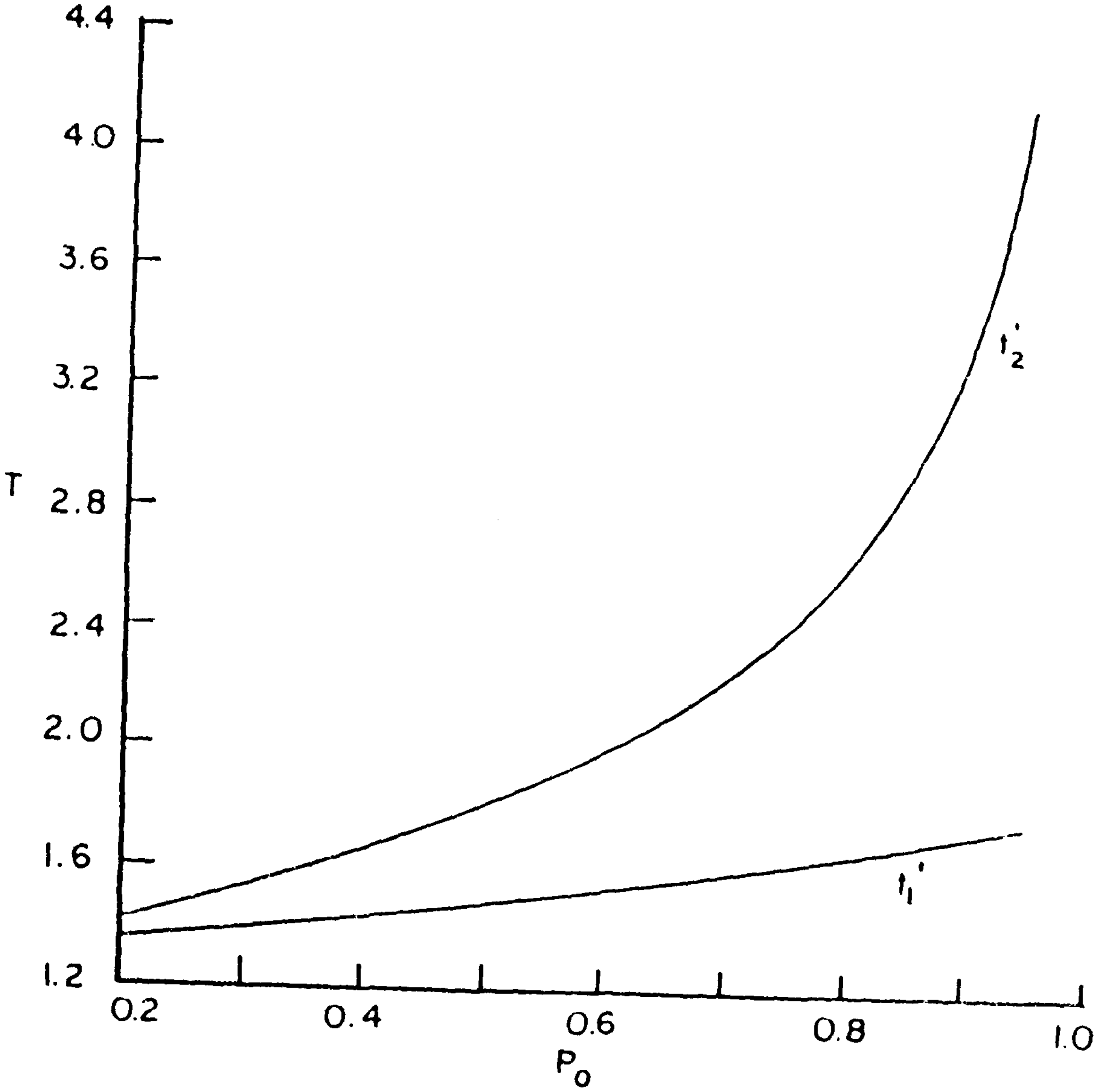


Figure 5.14 t'_1 and t'_2 Versus P_0 for Example 5.1 and $(\theta_1, \theta_2) = (2,3)$ (DD Configuration)

Figure 5.15. When $(\theta_1,\theta_2) = (2,2)$, the curve of $I(H;x)$ versus P_o is shown in Figure 5.16 and the curves of t_1' and t_2' are shown in Figure 5.17.

For the one-sensor case and $\theta = 5$, the curve of $I(H;u)$ is shown in Figure 5.18 and the curve of t' is shown in Figure 5.19.

Example 5.2

Now, let us assume that it is desired to maximize $I(\underline{x};u)$, for the system shown in Figure 5.9. We are given the following probabilities

$$P(u_1=0,u_2=0) = \frac{2}{5},\ P(u_1=0,u_2=1) = \frac{1}{15},\ P(u_1=1,u_2=0) = \frac{1}{5}$$

and,

$$P(u_1=1,u_2=1) = \frac{1}{3}.$$

We maximize the mutual information $I(\underline{x};u)$ by setting the following probabilities which set the output probabilities close to 1/2

$P_{o00} = 1$, $P_{o01} = 1$, $P_{o10} = 0$ and, $P_{o11} = 0$

This implies that

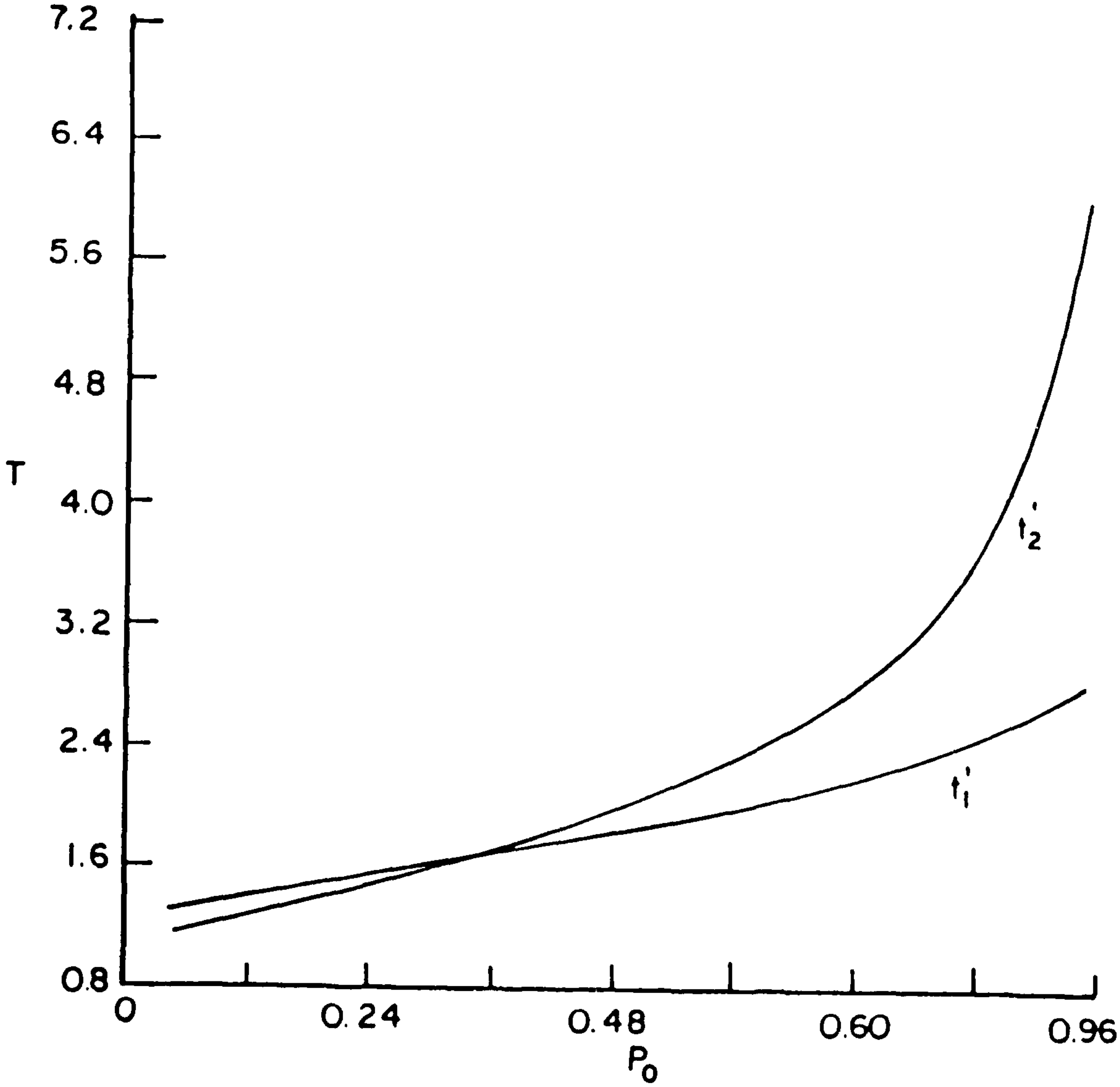


Figure 5.15 t'_1 and t'_2 Versus P_0 for Example 5.1 and $(\theta_1, \theta_2) = (3,6)$
(DD Configuration)

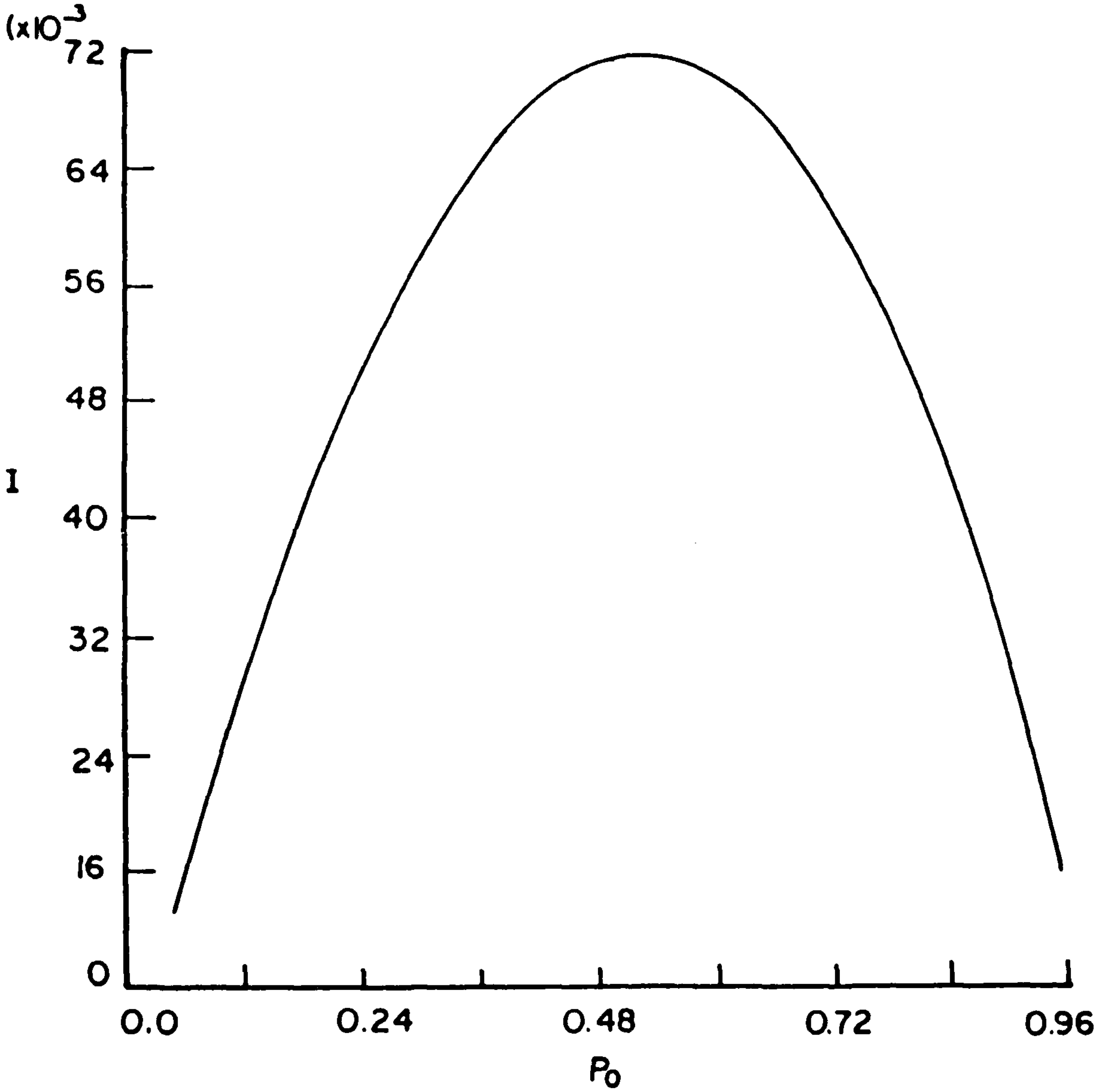


Figure 5.16 $I(H;\underline{x})$ Versus P_0 for Example 5.1 and $(\theta_1,\theta_2) = (2,2)$ (DD Configuration)

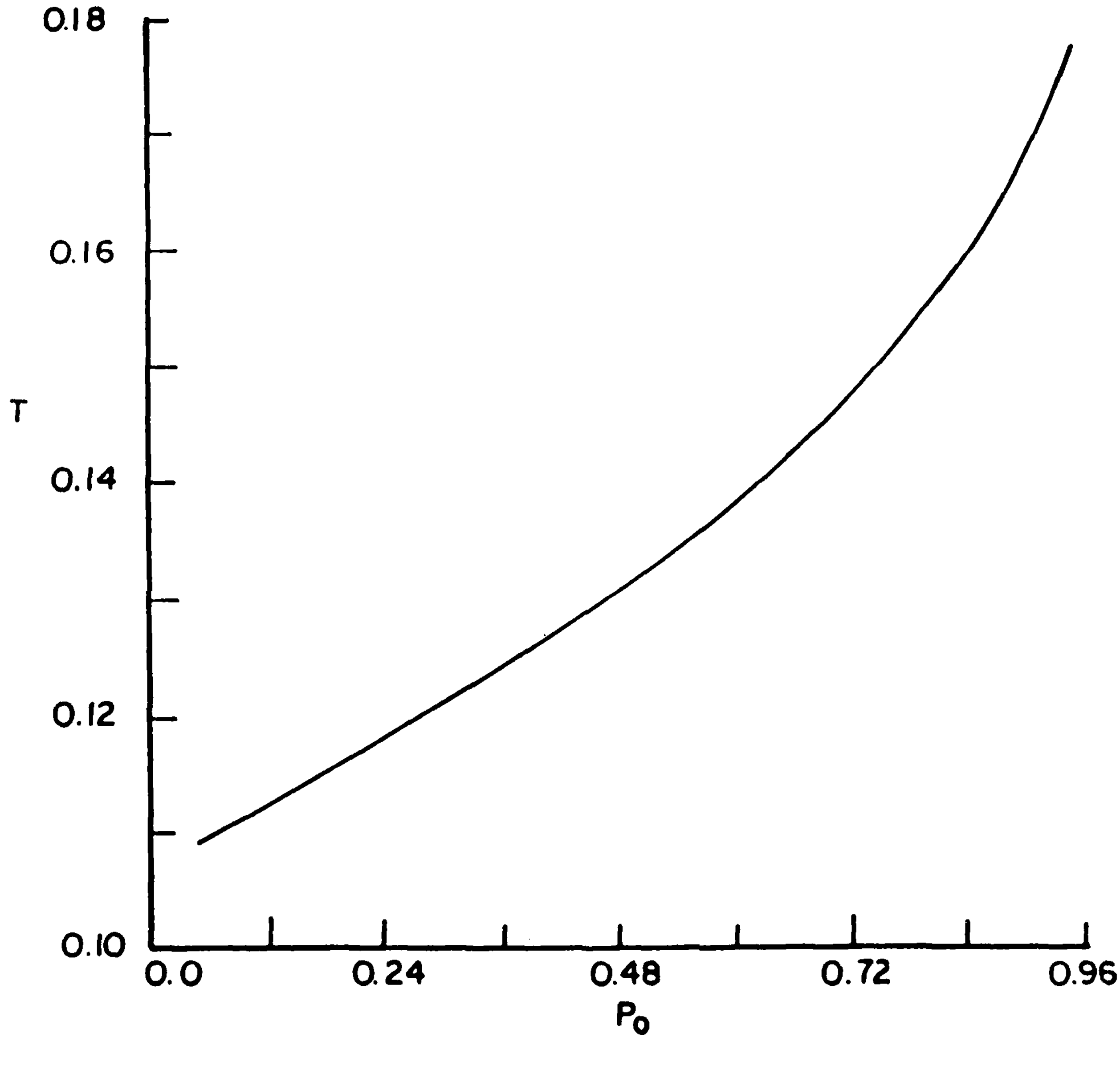


Figure 5.17 t'_1 and t'_2 Versus P_0 for Example 5.1 and $(\theta_1,\theta_2) = (2,2)$ (DD Configuration)

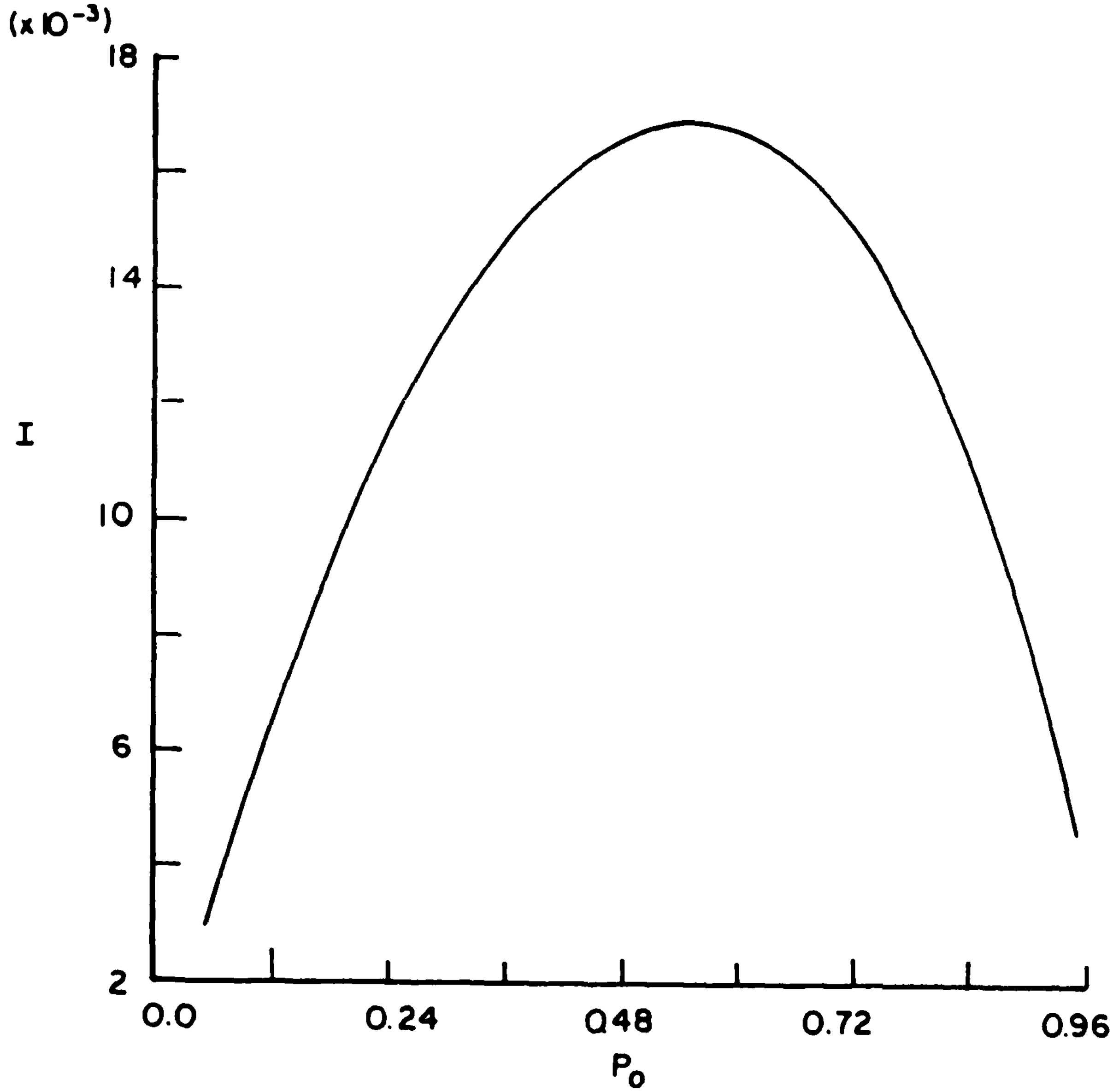


Figure 5.18 I(H;u) Versus P_0 for Example 5.1 and θ = 5. (Single Sensor)

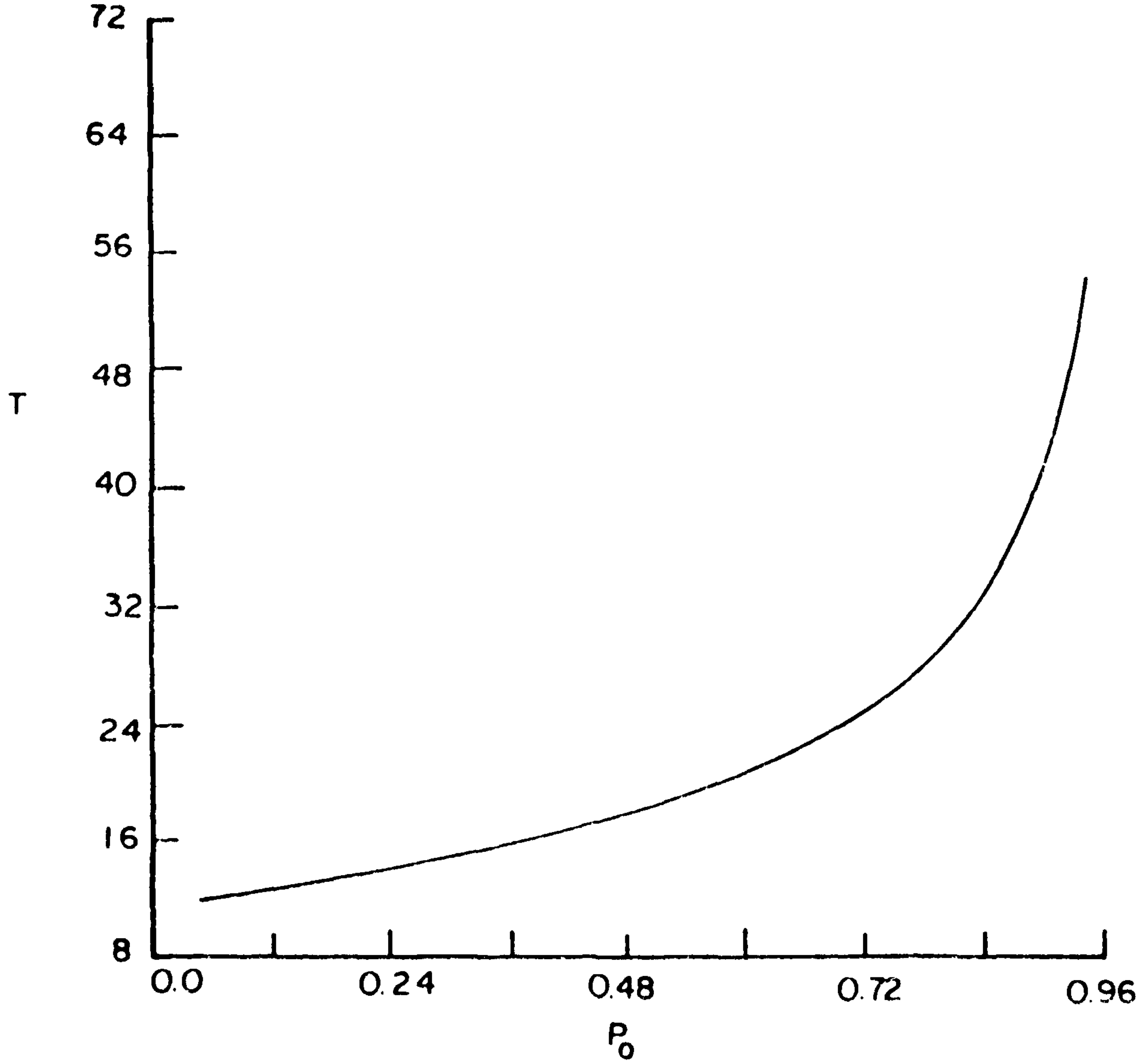


Figure 5.19 t' Versus P_0 for Example 5.1 and $\theta = 5$. (Single Sensor)

$$P(u=0) = 14/30$$

and,

$$P(u=1) = 16/30$$

Therefore, the maximum, I_{max}, of $I(\underline{x};u)$ is

$$I_{max} = (7/15) \ln (15/7) + (8/15) \ln (15/8) = 0.69.$$

In this example, we have illustrated the optimization procedure outlined in the previous section. As expected the procedure yields the maximum mutual information.

VI. Distributed Bayesian Parameter Estimation

6.1 Introduction

The Bayesian approach to the estimation of random parameters in a centralized framework is well known [1]. In this chapter, we consider the same problem in a distributed framework. Local parameter estimates are obtained at the individual sensors and are transmitted to the fusion center where they are combined to yield the global estimate. Almost all of the work reported in the literature on distributed estimation, deals with decentralized state estimation problems [19-26]. Here we develop the theory of decentralized random parameter estimation.

In Section 6.2, we formulate and solve the Bayesian estimation problem for the DPEF system. Three different criteria are considered: the minimum mean-square-error (MMSE) criterion, the absolute error criterion and the uniform cost criterion. In Section 6.3, we present the solution for the case when the combining rule is assumed to be linear. Section 6.4 presents a simple numerical example. In Section 6.5, we present a brief summary of the chapter.

6.2 Distributed Bayesian Parameter Estimation with Estimate Combining

6.2.1 Problem Statement

In this section, we consider the Bayesian DPEF problem, where the objective is to estimate a random parameter "a" with a known density function p(a). The parameter a is defined on segments of the real line. We consider the system shown in Figure 6.1. Each local estimator receives a set of observations denoted by an observation vector.

$$\underline{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})^T \qquad i = 1,2,\ldots,N,$$

where n is the number of observations at each estimator and N is the number of local estimators. Then, based on its observation vector, $\underline{y}_i$, the i^{th} local estimator estimates the parameter a, by assigning a point, $\hat{a}_i = h_i(\underline{y}_i)$ in the parameter space. The estimate combiner at the data fusion center collects the estimates from the individual (local) estimators and generates a global estimate $\hat{a}$, of the parameter a. This global estimate $\hat{a}$ depends only on the estimate vector $\underline{h} = (h_1(\underline{y}_1), h_2(\underline{y}_2), \ldots, h_N(\underline{y}_N))^T$.

The goal of this section is to develop Bayesian estimation theory for the DPEF system, i.e., design both the optimal combining rule and the estimation rules for the local estimators so as to minimize the Bayesian risk, R,

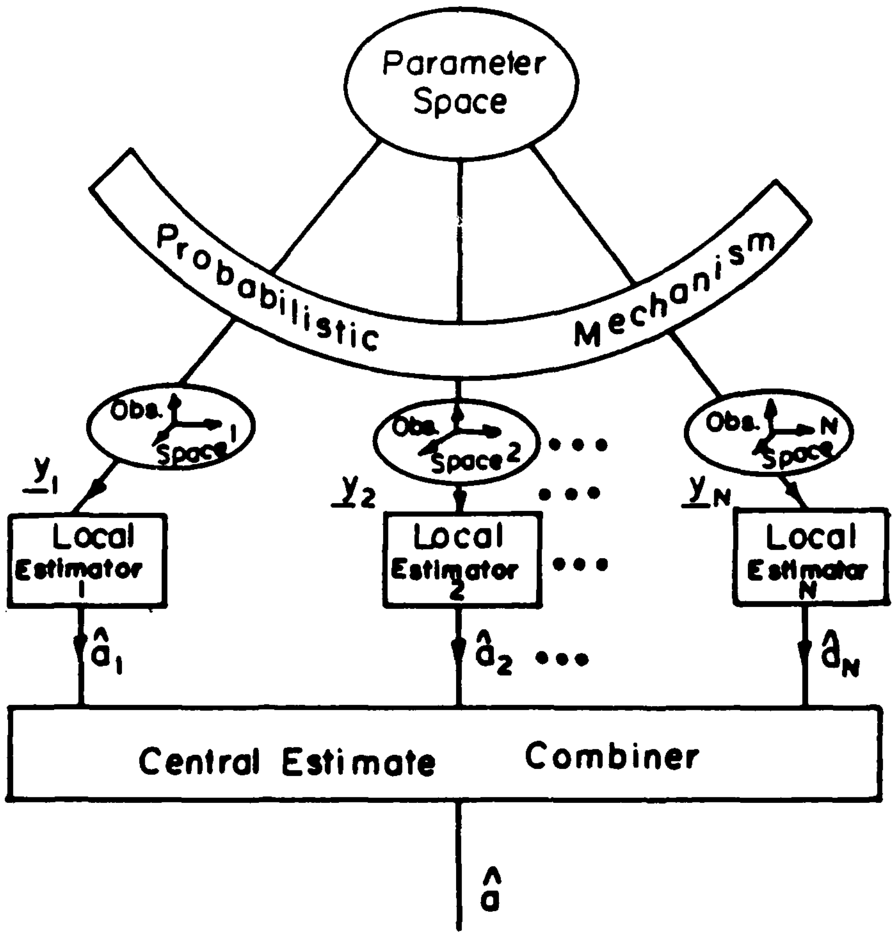


Figure 6.1 Distributed Estimation System with Estimate Combining

given by

$$R = E(C(a,\hat{a})) \tag{6-1}$$

where

$C(a,\hat{a})$ is the cost of estimating a as $\hat{a}$.

The risk R may also be written as

$$R = \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} dY\ C(a,\hat{a}(h_1(\underline{y}_1),h_2(\underline{y}_2),\ldots,h_N(\underline{y}_N))\ p(a,\underline{Y}) \tag{6-2}$$

where

$p(a,\underline{Y})$ is the joint density function of the parameter a and the observation vector $\underline{Y}$.

The cost functions that we will consider in this chapter will be functions of the error, $a_\epsilon(\underline{Y})$ defined by $a_\epsilon(\underline{Y}) = a-\hat{a}$. Three different cost criteria will be considered. The first one is minimum mean-square-error (MMSE) where

$$C(a_\epsilon)=(a-\hat{a})^2 \tag{6-3-a}$$

The second one is the minimum absolute error criterion, where

$$C(a_\epsilon) = |a-\hat{a}| \tag{6-3-b}$$

Finally, we will consider the uniform cost criterion, where

$$C(a_\epsilon) = \begin{cases} 0 & \text{if } |a_\epsilon| \leq \tau \\ 1 & |a_\epsilon| > \tau \end{cases} \tag{6-3-c}$$

and, τ is a suitably small interval.

The risk function can be expressed as

$$R = \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} d\underline{Y}\; C[a-\hat{a}(\underline{Y})] \;\; p(a,\underline{Y}) \tag{6-4}$$

or

$$R = \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} d\underline{h}\; C[a-\hat{a}(\underline{h})] \; p(a,\underline{h}) \tag{6-5}$$

The objective is to obtain the solution to the DPEF problem so as to minimize R. We obtain the combining rule $\hat{a} = f_\epsilon(\underline{h})$ and, the estimation rules $h_i(y_i)$ for each estimator i, $i = 1,2,\ldots,N$, so as to minimize R when the cost function is one of the functions discussed above. Next, we proceed with the solution for each of the cost functions.

6.2.2 Minimum Mean-Square-Error Criterion

Theorem 6.1

Given the probability density $p(a)$ of a random parameter a, the optimum DPEF system which minimizes the MMSE, i.e., the optimum combining rule and the optimum estimation rules at the local estimators, is obtained by solving the following equations simultaneously.

$$\hat{a}_{MS}(\underline{h}) = \int_{-\infty}^{+\infty} a\ p(a|\underline{h})\ da \qquad (6\text{-}6)$$

and,

$$E\left[\left(a.\frac{\delta\hat{a}}{\delta\hat{a}_i}\right)\Bigg|\, y_i\right] = E\left[\left(\hat{a}.\frac{\delta\hat{a}}{\delta\hat{a}_i}\right)\Bigg|\, y_i\right] \qquad i=1,2,\ldots,N \qquad (6\text{-}7)$$

where

$\hat{a}_{MS}$ = the minimum mean square error estimate.

Proof

The risk function for this problem is

$$R_{MS} = \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} d\underline{Y}\ [a-\hat{a}(\underline{h}(\underline{Y}))]^2\ p(a,\underline{Y}) \qquad (6\text{-}8)$$

which may be written as a function of the local estimates,

$\hat{a}_i = h_i(\underline{y}_i)$, $i = 1,\ldots,N$,

$$R_{ms} = \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} d\underline{h}\ [a-\hat{a}(\underline{h})]^2\ p(a,\underline{h}) \qquad (6\text{-}9)$$

<u>Combining Rule ($\hat{a}(\underline{h})$)</u>

While deriving $\hat{a}(\underline{h})$, we assume that the local estimation rules, $a_i = h_i(\underline{y}_i)$, $i = 1,2,\ldots,N$, are known. The joint density function of a and $\underline{h}$ may be written as

$$p(a,\underline{h}) = p(\underline{h}).p(a|\underline{h}) \qquad (6\text{-}10)$$

Using (6-10), R_{ms} becomes

$$R_{ms} = \int_{-\infty}^{+\infty} d\underline{h}\ p(\underline{h}) \int_{-\infty}^{+\infty} da\ [a-\hat{a}(\underline{h})]^2\ p(a|\underline{h}) \qquad (6\text{-}11)$$

Since the inner integral and $p(\underline{h})$ are nonnegative, minimizing the inner integral for any value of $\underline{h}$, minimizes R_{ms}. We denote the minimum mean square error estimate, by $\hat{a}_{ms}(\underline{h})$. Next, we will show that $\hat{a}_{ms}(\underline{h})$ is given by

$$\hat{a}_{ms}(\underline{h}) = \int_{-\infty}^{+\infty} da\ a\ p(a|\underline{h}) \qquad (6\text{-}12)$$

which is the conditional mean of a given $\underline{h}$.

Taking the derivative of the inner integral of (6-11), we get

$$\frac{\delta}{\delta \hat{a}(\underline{h})} \int_{-\infty}^{+\infty} da\, [a - \hat{a}(\underline{h})]^2\, p(a|\underline{h}) = -2 \int_{-\infty}^{+\infty} a\, p(a|\underline{h})\, da$$

$$+2\, \hat{a}(\underline{h}) \int_{-\infty}^{+\infty} p(a|\underline{h})\, da \qquad (6\text{-}13)$$

Setting the result in (6-13) equal to zero and using

$$\int_{-\infty}^{+\infty} p(a|\underline{h})\, da = 1$$

we obtain the result given in (6-12). Note that this corresponds to a minimum (not a maximum), because the 2nd derivative is equal to 2.

Next, we solve for the estimation rule, at the local estimators.

Local Estimation Rules

Now, we assume that the combining rule is known and, we obtain the optimum estimation rules $h_i(y_i)$ $i = 1,2,\ldots,N$ at the local estimators so as to minimize R_{MS}. While deriving the local estimate at estimator i, $i = 1,2,\ldots,N$, we assume that all other estimators have already been designed, i.e.,

$\hat{a}_j = h_j(\underline{y}_j)$, $i \neq j$; $j = 1,2,\ldots,N$, are known.

The joint probability density of a and $\underline{y}$ may be expressed as

$$p(a,\underline{Y}) = p(\underline{y}_i)\; p(a,\underline{y}_1,\underline{y}_2,\ldots,\underline{y}_{i-1},\underline{y}_{i+1},\ldots,\underline{y}_N \mid \underline{y}_i) \tag{6-15}$$

Substituting from (6-15), into (6-8), we have

$$R_{ms} = \int_{-\infty}^{+\infty} d\underline{y}_i\, p(\underline{y}_i) \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} d\underline{Y}^i\, [a-\hat{a}(h_1(\underline{y}_1),h_2(\underline{y}_2),\ldots,h_N(\underline{y}_N))]^2\, p(a,\underline{Y}^i \mid \underline{y}_i) \tag{6-16}$$

where

$$\underline{Y}^i = (\underline{y}_1^T,\ \underline{y}_2^T,\ldots,\underline{y}_{i-1}^T,\underline{y}_{i+1}^T,\ldots,\underline{y}_N^T)^T \tag{6-17}$$

and,

$$\int_{-\infty}^{+\infty} d\underline{Y}^i \quad \text{is the integral over all elements of } \underline{Y}^i \tag{6-18}$$

Let

$$I_i = \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} d\underline{Y}^i\, [a-\hat{a}(\underline{h})]^2\, p(a,\underline{Y}^i \mid \underline{y}_i) \tag{6-19}$$

In (6-16), it is obvious that $p(\underline{y}_i)$ and I_i are nonnegative. Therefore, minimizing the inner integral I_i with respect to a_i minimizes R_{ms}. Taking the derivative of I_i with respect to a_i and setting it equal to zero, we have

$$\frac{\delta I_i}{\delta \hat{a}_i} = \frac{\delta}{\delta \hat{a}_i(y_i)} \int_{-\infty}^{+\infty} da \int_{-\infty}^{+\infty} d\underline{Y}^i [a-\hat{a}(\hat{a}_1,\hat{a}_2,\ldots,\hat{a}_N)]^2 p(a,Y^i|y_i)$$

$$= 0 \qquad (6\text{-}20)$$

Interchanging the two operations, namely differentiation with respect to $\hat{a}_i$, and integration over a and Y^i, (6-20) becomes

$$0 = \int_{-\infty}^{+\infty} d\underline{Y}^i \int_{-\infty}^{+\infty} da\ p(a,\underline{Y}^i \big| y_i)\ \frac{\delta}{\delta \hat{a}_i} \{ [a-\hat{a}(\hat{a}_1,\hat{a}_2,\ldots,\hat{a}_N)]^2 \}$$

$$(6\text{-}21)$$

which is equivalent to

$$E[(a.\frac{\delta \hat{a}}{\delta \hat{a}_i}) \big| y_i] = E[(\hat{a}.\frac{\delta \hat{a}}{\delta \hat{a}_i}) \big| y_i] \qquad i = 1,2,\ldots,N \quad (6\text{-}22)$$

We repeat this procedure for all of the N estimators. A simultaneous solution of the resulting N equations yields the desired local estimation rules $\hat{a}_i = h_i(y_i)$ $i = 1,2,\ldots,N$

Overall solution

The overall solution to this problem is obtained by solving (6-12) and (6-22) simultaneously, which is the desired result.

Q.E.D.

Special Case of Two Estimators

In the case where the system has only two local estimators, the equations to be solved are

$$E[(a.\frac{\delta\hat{a}}{\delta\hat{a}_1})\Big| y_1] = E[(\hat{a}.\frac{\delta\hat{a}}{\delta\hat{a}_1})\Big| y_1] \tag{6-23}$$

$$E[a.\frac{\delta\hat{a}}{\delta\hat{a}_2})\Big| y_2] = E[(\hat{a}.\frac{\delta\hat{a}}{\delta\hat{a}_2})\Big| y_2] \tag{6-24}$$

and,

$$\hat{a} = E(a|\hat{a}_1,\hat{a}_2] \tag{6-25}$$

Solving (6-23), (6-24) and (6-25) simultaneously, yields the optimum solution. These equations are complicated to solve due to the coupling and nonlinearities. Later in Section 5.3, we solve the problem when the combining rule is linear.

Next, we solve the problem for the absolute error criterion.

6.2.3 Minimum Absolute Error Criterion

Theorem 6.2

Given the probability density $p(a)$ of a random parameter a, the optimum DPEF system which minimizes the absolute error function, i.e., the optimum combining rule and the optimum estimation rules at the local estimators is obtained by solving the following equations simultaneously.

$$\int_{-\infty}^{\hat{a}_{abs}(\underline{h})} da\; p(a|\underline{h}) = \int_{\hat{a}_{abs}(\underline{h})}^{+\infty} da\; p(a|\underline{h}) \tag{6-26}$$

and,

$$\int_{-\infty}^{+\infty} d\underline{Y}^i \int_{-\infty}^{\hat{a}_{abs}} da\left(\frac{\delta \hat{a}_{abs}}{\delta \hat{a}_i}\right) p(a,\underline{Y}^i|y_i) = \int_{-\infty}^{+\infty} d\underline{Y}^i \int_{\hat{a}_{abs}}^{+\infty} da\left(\frac{\delta \hat{a}_{abs}}{\delta \hat{a}_i}\right) p(a,\underline{Y}^i|y_i)$$

$$i = 1,2,\ldots,N \tag{6-27}$$

Proof

The risk function for the absolute error criterion is

$$R_{abs} = E(|a-\hat{a}|) \tag{6-28}$$

For the cost function defined in (6-3-b), the risk function is given by

$$R_{abs} = \int_{-\infty}^{+\infty} d\underline{h}\; p(\underline{h}) \int_{-\infty}^{+\infty} da[\,|a-\hat{a}(\underline{h})|\,]\; p(a|\underline{h}) \tag{6-29}$$

First, we obtain the optimum combining rule.

Optimum Combining Rule

In (6-29), we define $I(\underline{h})$ as being the inner integral which can be written as

$$I(\underline{h}) = \int_{-\infty}^{\hat{a}(\underline{h})} da\,[\hat{a}(\underline{h})-a]\,p(a|\underline{h}) + \int_{\hat{a}(\underline{h})}^{+\infty} da\,[a-\hat{a}(\underline{h})]\,p(a|\underline{h}) \tag{6-30}$$

Again, we assume that while deriving the optimum combining rule which minimizes R_{abs}, the estimation rules $h_1(\underline{y}_1), h_2(\underline{y}_2), \ldots,$ and $h_N(\underline{y}_N)$ are known. If a minimum exists and is interior to the parameter space, we obtain the minimum by differentiating $I(\underline{h})$ with respect to $\hat{a}(\underline{h})$ and setting it equal to zero. Therefore,

$$\frac{\delta I}{\delta \hat{a}(\underline{h})} = \int_{-\infty}^{\hat{a}_{abs}(\underline{h})} da\,p(a|\underline{h}) - \int_{\hat{a}_{abs}(\underline{h})}^{+\infty} da\,p(a|\underline{h}) = 0 \tag{6-31}$$

where $\hat{a}_{abs}(\underline{h})$ is the optimum estimate. Solving (6-31), we see that, $\hat{a}_{abs}(\underline{h})$ which minimizes R_{abs}, is the median of the a posteriori probability density of a given $\underline{h}$, i.e.,

$$\int_{-\infty}^{\hat{a}_{abs}(h)} da\ p(a|\underline{h}) = \int_{\hat{a}_{abs}(\underline{h})}^{+\infty} da\ p(a|\underline{h}) \tag{6-32}$$

Next, we obtain the estimation rules at the local estimators.

<u>Local Estimation Rules</u>

At this point, we assume that all but estimator i, have been designed and, that the combining rule is known. We obtain the estimation rule for the i^{th} estimator which minimizes R_{abs}. Using (6-28), we may rewrite (6-4) as

$$R_{abs} = \int_{-\infty}^{+\infty} dY \int_{\infty}^{+\infty} da\ [|a - \hat{a}(\hat{a}_1,\hat{a}_2,\ldots,\hat{a}_N)|]\ p(a,Y) \tag{6-33}$$

Using (6-15), (6-33) becomes

$$R_{abs} = \int_{-\infty}^{+\infty} dy_i\ p(y_i)\ \{\int_{-\infty}^{+\infty} d\underline{Y}^i [\int_{-\infty}^{\hat{a}} da\ (\hat{a} - a)\ p(a,\underline{Y}^i|y_i)$$

$$+ \int_{\hat{a}}^{+\infty} da\ (a - \hat{a})\ p(a,\underline{Y}^i|y_i)]\} \tag{6-34}$$

where, $\underline{Y}^i$ is given by (6-17). Let

$$I_{abs}(\underline{y}_i) = \int_{-\infty}^{+\infty} d\underline{Y}^i \left[\int_{-\infty}^{\hat{a}} da \; (\hat{a}-a) \; p(a,\underline{Y}^i|\underline{y}_i) \right.$$

$$\left. + \int_{\hat{a}}^{+\infty} da \; (a - \hat{a}) \; p(a,\underline{Y}^i|\underline{y}_i) \right] \qquad (6\text{-}35)$$

Assuming that a minimum exists and is interior to the parameter space, the value of $\hat{a}$ which minimizes $I_{abs}(\underline{y}_i)$, also minimizes R_{abs}. Taking the derivative of $I_{abs}(\underline{y}_i)$ with respect to $\hat{a}_i$, we have

$$\frac{\delta I_{abs}(\underline{y}_i)}{\delta \, \hat{a}_i} = \int_{-\infty}^{+\infty} d\underline{Y}^i \left\{ \int_{-\infty}^{\hat{a}} da\left(\frac{\delta\hat{a}}{\delta\hat{a}_i}\right) p(a,\underline{Y}^i|\underline{y}_i) + \frac{\delta\hat{a}}{\delta\hat{a}_i} \; (\hat{a}-\hat{a}) p(a,\underline{Y}^i|\underline{y}_i) \right.$$

$$\left. - \int_{\hat{a}}^{+\infty} da\left(\frac{\delta\hat{a}}{\delta\hat{a}_i}\right) p(a,\underline{Y}^i|\underline{y}_i) + \frac{\delta\hat{a}}{\delta\hat{a}_i} \; (\hat{a}-\hat{a}) p(a,\underline{Y}^i|\underline{y}_i) \right\}$$

$$(6\text{-}36)$$

Setting the derivative equal to zero, we have

$$\int_{-\infty}^{+\infty} d\underline{Y}^i \int_{-\infty}^{\hat{a}_{abs}} da \; \left(\frac{\delta\hat{a}_{abs}}{\delta\hat{a}_i}\right) p(a,\underline{Y}^i|\underline{y}_i) = \int_{-\infty}^{\infty} d\underline{Y}^i \int_{\hat{a}_{abs}}^{\infty} da\left(\frac{\delta\hat{a}_{abs}}{\delta\hat{a}_i}\right) \; p(a,\underline{Y}^i|\underline{y}_i)$$

$$i = 1,2,\ldots,N \qquad (6\text{-}37)$$

This procedure is repeated for all of the N estimators. The resulting N equations are solved simultaneously to yield the optimum local estimation rules which minimize R_{abs}.

Overall Solution

The overall solution to this problem is obtained by solving (6-37) and (6-32) simultaneously, which is the desired result.

Q.E.D.

Special Case of two estimators

For the two-estimator system, the equations to be solved are

$$\int_{-\infty}^{+\infty} dy_2 \int_{-\infty}^{\hat{a}_{abs}} da\left(\frac{\delta \hat{a}_{abs}}{\delta \hat{a}_1}\right) p(a, y_2 | y_1) = \int_{-\infty}^{+\infty} dy_2 \int_{\hat{a}_{abs}}^{\infty} da\left(\frac{\delta \hat{a}_{abs}}{\delta \hat{a}_1}\right) p(a, y_2 | y_1)$$

(6-38)

$$\int_{-\infty}^{+\infty} dy_1 \int_{-\infty}^{\hat{a}_{abs}} da\left(\frac{\delta \hat{a}_{abs}}{\delta \hat{a}_2}\right) p(a, y_1 | y_2) = \int_{-\infty}^{+\infty} dy_1 \int_{\hat{a}_{abs}}^{\infty} da\left(\frac{\delta \hat{a}_{abs}}{\delta \hat{a}_2}\right) p(a, y_1 | y_2)$$

(6-39)

and,

$$\int_{-\infty}^{\hat{a}_{abs}(\underline{h})} da\; p(a | \hat{a}_1 \hat{a}_2) = \int_{\hat{a}_{abs}(\underline{h})}^{\infty} da\; p(a | \hat{a}_1, \hat{a}_2] \qquad (6\text{-}40)$$

Solving (6-38), (6-39) and (6-40) yield the desired result. Next, we study the case of uniform cost function.

6.2.4 Minimum Uniform Cost Function

We consider the problem of finding the optimum estimation rules at the estimators and the optimum combining rules which minimize the average cost in this case for the DPEF system.

Theorem 6.3

For the DPEF system, the optimum combining rule and the optimum estimation rules at the local estimators which minimize the uniform cost function for a given p(a), are obtained by solving the following equations simultaneously

$$\left.\frac{\delta \ln p(a|h)}{\delta \hat{a}}\right|_{a=\hat{a}_{unif}} = 0 \qquad (6\text{-}41)$$

and,

$$\int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_i}\, p(a = \tau+\hat{a},\ \underline{Y}^i|y_i)\,dY^i = \int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_i}\, p(a= -\tau+\hat{a},\ \underline{Y}^i|y_i)\,dY^i$$

$$i = 1,2,\ldots,N \qquad (6\text{-}42)$$

Proof

The risk function in this case is given by

$$R_{unif} = \int_{-\infty}^{+\infty} d\underline{h}\; p(\underline{h})\; [1 - \int_{\hat{a}_{unif}-T}^{\hat{a}_{unif}+T} p(a|\underline{h})\, da\;] \tag{6-43}$$

which is equivalent to

$$R_{unif} = 1 - \int_{-\infty}^{+\infty} dy_i\; p(\underline{y}_i) \int_{-\infty}^{+\infty} d\underline{Y}^i \int_{-T+\hat{a}}^{T+\hat{a}} da\; p(a,\underline{Y}^i|\underline{y}_i) \tag{6-44}$$

First, we derive the optimum combining rule.

Optimum Combining rule

We assume that all estimators have known estimation rules and we obtain the optimum *combining* rule so as to minimize R_{unif}. The risk function is

$$R_{unif} = \int_{-\infty}^{+\infty} d\underline{h}\; p(\underline{h})\; [1 - \int_{\hat{a}_{unif}-T}^{\hat{a}_{unif}+T} p\,(a|\underline{h})\, da\;] \tag{6-43}$$

We also assume that a minimum exists and, is interior to the allowable range of a. Let

$$I(\underline{h}) = \int_{\hat{a}_{unif}-T}^{\hat{a}_{unif}+T} p(a|\underline{h})\, da \tag{6-45}$$

Then, maximizing $I(\underline{h})$ minimizes R_{unif} and this is achieved by taking the derivative of $I(\underline{h})$ with respect to $\hat{a}_{unif}$ and, setting it equal to zero. This results in

$$\left.\frac{\delta\, p(a|\underline{h})}{\delta a}\right|_{a=\hat{a}_{unif}(\underline{h})} = 0 \qquad (6\text{-}46\text{-}a)$$

or

$$\left.\frac{\delta \ln p(a|\underline{h})}{\delta a}\right|_{a=\hat{a}_{unif}(\underline{h})} = 0 \qquad (6\text{-}46\text{-}b)$$

Having obtained the optimum combining rule, next, we obtain the optimum estimation rules at the estimators so as to minimize R_{unif}.

Estimation Rules

We assume that the combining rule $\hat{a}(\underline{h})$ is known. We also assume that while deriving the estimation rule $\hat{a}_i=h_i(y_i)$ at the i^{th} estimator, $i = 1,2,\ldots,N$, the other estimators have known estimation rules.

In this case, R_{unif} can be expressed as

$$R_{unif} = 1 - \int_{-\infty}^{+\infty} dy_i\; p(\underline{y}_i) \int_{-\infty}^{+\infty} d\underline{Y}^i \int_{-T+\hat{a}}^{T+\hat{a}} da\; p(a,\underline{Y}^i|y_i) \qquad (6\text{-}47)$$

Let

$$I_i(\underline{y}_i) = \int_{-\infty}^{+\infty} d\underline{Y}_i \int_{-T+\hat{a}}^{T+\hat{a}} da\; p(a,\underline{Y}^i|y_i) \qquad (6\text{-}48)$$

Then, if a minimum exists and is interior to the range of a, minimizing R_{unif} with respect to $\hat{a}_i = h_i(\underline{y}_i)$ is equivalent to maximizing $I_i(\underline{y}_i)$. Taking the derivative of $I_i(\underline{y}_i)$ with respect to $\hat{a}_i$ and interchanging the derivative with respect to a_i and the integral over $\underline{Y}^i$, we get

$$\frac{\delta\ I_i(\underline{y}_i)}{\delta\ \hat{a}_i} = \int_{-\infty}^{+\infty} d\underline{Y}^i \frac{\delta}{\delta \hat{a}_i} \left(\int_{-\tau+\hat{a}}^{\tau+\hat{a}} da\ p(a, \underline{Y}^i | y_i) \right) \qquad (6\text{-}49)$$

This may be written as

$$\frac{\delta\ I_i(y_i)}{\delta\ \hat{a}_i} = \int_{-\infty}^{+\infty} dY^i \left(\int_{-\tau+\hat{a}}^{\tau+\hat{a}} da\ (0) + \frac{\delta(\tau+\hat{a})}{\delta \hat{a}_i} p(a=\tau+\hat{a},\ \underline{Y}^i | y_i) - \frac{\delta(-\tau+\hat{a})}{\delta \hat{a}_i} p(a=-\tau+\hat{a},\ Y^i | y_i) \right)$$

(6-50)

or

$$\frac{\delta I_i(\underline{y}_i)}{\delta \hat{a}_i} = \int_{-\infty}^{+\infty} d\underline{Y}^i \left(\frac{\delta \hat{a}}{\delta \hat{a}_i} [p(a=\tau+\hat{a}, \underline{Y}^i | y_i) - p(a=-\tau+\hat{a}, Y^i | y_i)] \right)$$

$$i = 1, 2, \ldots, N \qquad (6\text{-}51)$$

Setting the right hand side of (6-51) equal to zero and rearranging the result, we get

$$\int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_i} p(a=\tau+\hat{a},\ \underline{Y}^i|y_i)\ d\underline{Y}^i = \int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_i} p(a=-\tau+\hat{a},\ \underline{Y}^i|y_i)\ d\underline{Y}^i$$

$$i = 1,2,\ldots,N \qquad (6\text{-}52)$$

We may repeat this procedure for all N estimators. Thus, we obtain N equations of the form (6-52), whose solution yields the estimation rules at the local estimators which minimize R_{unif}.

Overall Solution

The overall solution to the DPEF problem with minimum R_{unif} criterion is obtained by solving (6-46) and (6-52) simultaneously, which yields the desired result.

Q.E.D.

Special Case of Two Estimators

Again, in the case of two-estimator DPEF system, the equations to be solved are

$$\left.\frac{\delta \ln (p(a|\underline{h}))}{\delta a}\right|_{a=\hat{a}(\hat{a}_1,\hat{a}_2)} = 0 \qquad (6\text{-}53)$$

$$\int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_1} p(a=\tau+\hat{a},\ y_2|y_1)\ dy_2 = \int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_1} p(a=-\tau+\hat{a},\ y_2|y_1)\ dy_2$$

(6-54)

and,

$$\int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_m}\, p(a=\tau+\hat{a},\ y_1 | y_m)\, dy_1 = \int_{-\infty}^{+\infty} \frac{\delta \hat{a}}{\delta \hat{a}_m}\, p(a=-\tau+\hat{a},\ \underline{y}_1 | \underline{y}_m)\, dy_1 \tag{6-55}$$

6.3 Suboptimum Solution- Linear Combining Rule

In the estimation problem studied in section 6.2, the optimum design requires the simultaneous solution of (N + 1) coupled nonlinear equations. These equations are hard to solve. Therefore, it is desirable to consider some suboptimum solutions. In this section, we restrict the estimation combining rule to be a linear one, i.e.,

$$\hat{a}(h) = \sum_{i=1}^{N} b_i\ h_i(\underline{y}_i) + C \tag{6-56}$$

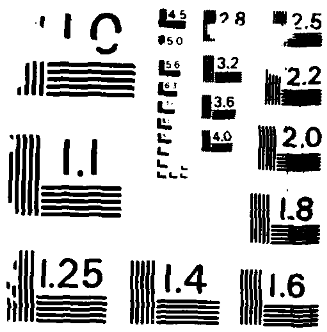
where b_i and C are constants to be found, i = 1,2,...,N. This combining rule may be simplified further by incorporating b_i in the estimate $h_i(\underline{y}_i)$ itself, i.e.,

$$\hat{a}(\underline{h}) = \sum_{i=1}^{N} h_i(\underline{y}_i) + C \tag{6-57}$$

The use of linear combining rules reduces the computational difficulty considerably. In the rest of this section, we present results for the DPEF problem using the three cost criteria for the linear combining rule.

AD-A193 934 ON THE DESIGN AND OPTIMIZATION OF DISTRIBUTED SIGNAL DETECTION AND PARAME..(U) SYRACUSE UNIV NY DEPT OF ELECTRICAL AND COMPUTER ENGINEERING.. 3/3

UNCLASSIFIED I Y HOBALLAH ET AL. SEP 87 RADC-TR-87-130 F/G 12/9 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

6.3.1 Minimum Mean-Square-Error Criterion

For this criterion, substituting (6-57) in (6-8) the mean square error becomes

$$R_{ms} = \int_{-\infty}^{+\infty} d\underline{Y} \int_{-\infty}^{+\infty} da\ [\ a - (\sum_{j=1}^{N} h_j(\underline{y}_j) + C)]^2\ p(a,\underline{Y}) \qquad (6\text{-}58)$$

Substituting (6-15) in (6-58), we get

$$R_{ms} = \int_{-\infty}^{+\infty} d\underline{y}_i p(\underline{y}_i) \int_{-\infty}^{+\infty} d\underline{Y}^i \int_{-\infty}^{+\infty} da\ [a - \sum_{j=1}^{N} h_j(\underline{y}_j) - C]^2\ p(a,\underline{Y}^i|\underline{y}_i)$$

$$(6\text{-}59)$$

We follow the same procedure as used in Section 6.2 to minimize R_{ms}, i.e., we set the derivative of the inner integral in (6-59) equal to zero and solve for the estimation rule, $h_i(\underline{y}_i)$, of estimator i, i = 1,2,...,N. The resulting $h_i(\underline{y}_i)$ is

$$h_i(\underline{y}_i) = E\ \{a|\underline{y}_i\} - \sum_{\substack{j=1 \\ j\neq i}}^{N} E\ \{(h_j(\underline{y}_j))|\underline{y}_i\} - C$$

$$i = 1,2,\ldots,N \qquad (6\text{-}60)$$

Now, we only need to find the value of C. Taking the derivative of R_{ms} with respect to C and setting it equal to zero, we obtain the desired result.

$$C = E\ \{[a - \sum_{j=1}^{N} h_j(\underline{y}_j)]\} = E\{a\} - \sum_{j=1}^{N} E\{h_j(\underline{y}_j)\} \qquad (6\text{-}61)$$

Next, we present the above result for the case of two estimators. In this case, equations (6-60) and (6-61) become

$$h_1(\underline{y}_1) = E\{a|\underline{y}_1\} - E\{(h_2(\underline{y}_2))|\underline{y}_1\} - C \tag{6-62}$$

$$h_2(\underline{y}_2) = E\{a|\underline{y}_2\} - E\{(h_1(\underline{y}_1))|\underline{y}_2\} - C \tag{6-63}$$

and,

$$C = E\{a\} - E\{h_1\} - E\{h_2\} \tag{6-64}$$

Substituting (6-63) in (6-62), we get

$$h_1(\underline{y}_1) = E\{a|\underline{y}_1\} - E\{(E\{a|\underline{y}_2\})|\underline{y}_1\} + E\{(E\{h_1(\underline{y}_1)|\underline{y}_2\})|\underline{y}_1\} \tag{6-65}$$

6.3.2 Minimum Absolute Error Criterion

Now, we solve the DPEF problem when a linear combining rule is used at the fusion center and the absolute error criterion is used for optimization. Substituting (6-57) in (6-8) and using an expression similar to (6-12), we may rewrite R_{abs} as

$$R_{abs} = \int_{-\infty}^{+\infty} d\underline{y}_i \; p(\underline{y}_i) \int_{-\infty}^{+\infty} d\underline{Y}^i$$

$$\times \left\{ \int_{-\infty}^{\sum_{j=1}^{N} h_j(\underline{y}_j)+C} da\left[\sum_{j=1}^{N} h_j(\underline{y}_j)+C-a\right] p(a,\underline{Y}^i|\underline{y}_i) \right.$$

$$\left. + \int_{\sum_{j=1}^{N} h_j(\underline{y}_j)+C}^{+\infty} da\left[a-\left(\sum_{j=1}^{N} h_j(\underline{y}_j)+C\right)\right] p(a,\underline{Y}^i|\underline{y}_i)\right\} . \qquad (6\text{-}66)$$

We use the same procedure as used in Section 6.2. We minimize the risk function with respect to $h_i(\underline{y}_i)$ by minimizing the inner integral I_i (in bracket). We take the derivative of I_i with respect to h_i and set the result equal to zero. We get

$$\frac{\delta I_i}{\delta h_i} = \int_{-\infty}^{+\infty} dY^i \left\{ \int_{-\infty}^{\sum_{j=1}^{N} h_j(\underline{y}_j)+C} da \; p(a,\underline{Y}^i|y_i) - \int_{\sum_{j=1}^{N} h_j(\underline{y}_j)+C}^{+\infty} da \; p(a,\underline{Y}^i|\underline{y}_i) \right\}$$

$$(6\text{-}67)$$

But

$$\int_{-\infty}^{\infty} d\underline{Y}^i \int_{-\infty}^{h_i(\underline{y}_i)} da\ p(a,\underline{Y}^i|\underline{y}_i) = \int_{-\infty}^{h_i(\underline{y}_i)} da\ p(a|\underline{y}_i) \qquad (6\text{-}68)$$

Rearranging (6-67) and substituting (6-68) into the resulting equation, we get

$$2\int_{-\infty}^{h_i(y_i)} da\ p(a|y_i) + 2\int_{-\infty}^{+\infty} d\underline{Y}^i \int_{h_i(\underline{y}_i)}^{\sum_{j=1}^{N} h_j(\underline{y}_j)+C} da\ p\ (a,\underline{Y}^i|\underline{y}_i) = 1 \qquad (6\text{-}69)$$

or equivalently,

$$\int_{-\infty}^{h_i(y_i)} da\ p(a|\underline{y}_i) + \int_{-\infty}^{+\infty} d\underline{Y}^i \int_{h_i(y_i)}^{\sum_{j=1}^{N} h_j(y_j)+C} da\ P(a,\underline{Y}^i|\underline{y}_i) = 1/2 \qquad i=1,2,\ldots,N.$$

(6-70)

Now, we obtain the optimum value of C. We assume that all estimators have already been designed. R_{abs} may be rewritten as

$$R_{abs} = \int_{-\infty}^{+\infty} d\underline{Y}\left(\int_{-\infty}^{\sum_{j=1}^{N} h_j(\underline{y}_j)+C} da\left[\sum_{j=1}^{N} h_j(\underline{y}_j)+C-a\right]p(a,\underline{Y}) + \int_{\sum_{j=1}^{N} h_j(\underline{y}_j)+C}^{+\infty} da\left[a-C-\sum_{j=1}^{N} h_j(\underline{y}_j)\right)p(a,\underline{Y})\right)$$

(6-71)

Taking the derivative of R_{abs} with respect to C, we get

$$\frac{\delta R_{abs}}{\delta C} = \int_{-\infty}^{+\infty} dY \left(\int_{-\infty}^{\sum_{j=1}^{N} h_j(\underline{y}_j)+C} da\ p(a,\underline{Y}) - \int_{\sum_{j=1}^{N} h_j(\underline{y}_j)+C}^{\infty} da\ p(a,\underline{Y}) \right) \qquad (6\text{-}72)$$

Rearranging (6-72), and setting the derivative equal to zero, we can write

$$\frac{\delta R_{abs}}{\delta C} = \int_{-\infty}^{+\infty} dY \left(\int_{-\infty}^{\sum_{j=1}^{N} h_j(y_j)} da\ p(a,Y) - \int_{\sum_{j=1}^{N} h_j(\underline{y}_j)}^{\infty} da\ p(a,\underline{Y}) \right) + 2\int_{-\infty}^{+\infty} d\underline{Y} \int_{\sum_{j=1}^{N} h_j(\underline{y}_j)}^{\sum_{j=1}^{N} h_j(\underline{y}_j)+C} da\ p(a,\underline{Y})$$

$$= 0 \qquad (6\text{-}73)$$

Or equivalently, C is obtained by solving the following

$$\int_{-\infty}^{+\infty} dY \int_{\sum_{j=1}^{N} h_j(\underline{y}_j)}^{\sum_{j=1}^{N} h_j(y_j)+C} da\ p(a,Y) + \int_{-\infty}^{\infty} d\underline{Y} \int_{-\infty}^{\sum_{j=1}^{N} h_j(y_j)} da\ p(a,\underline{Y}) = 1/2 \qquad (6\text{-}74)$$

The results obtained by solving N equations of the form (6-70) and one equation of the form (6-69) simultaneously minimizes R_{abs}.

In the special case of two estimators the combining rule becomes

$$\hat{a}(\underline{h}) = h_1(\underline{y}_1) + h_2(\underline{y}_2) + C \tag{6-75-a}$$

The following set of equations need to be solved to yield the desired result

$$\int_{-\infty}^{h_1(y_1)} da\; p(a|\underline{y}_1) + \int_{-\infty}^{+\infty} d\underline{y}_2 \int_{h_1(\underline{y}_1)}^{h_1(y_1)+h_2(\underline{y}_2)+C} da\; p(a,\underline{y}_2|\underline{y}_1) = 1/2 \tag{6-75-b}$$

$$\int_{-\infty}^{h_2(\underline{y}_2)} da\; p(a|\underline{y}_2) + \int_{-\infty}^{+\infty} d\underline{y}_1 \int_{h_2(\underline{y}_2)}^{h_1(\underline{y}_1)+h_2(\underline{y}_2)+C} da\; p(a,\underline{y}_1|\underline{y}_2) = 1/2 \tag{6-75-c}$$

and,

$$\int_{-\infty}^{+\infty} d\underline{y}_1 \int_{-\infty}^{+\infty} d\underline{y}_2 \int_{h_1(\underline{y}_1)+h_2(\underline{y}_2)}^{h_1(\underline{y}_1)+h_2(\underline{y}_2)+C} da\; p(a,\underline{y}_1,\underline{y}_2) + \int_{-\infty}^{+\infty} d\underline{y}_1 \int_{-\infty}^{+\infty} d\underline{y}_2 \int_{-\infty}^{h_1(\underline{y}_2)+h_2(\underline{y}_2)} da\; p(a,\underline{y}_1,\underline{y}_2) = 1/2$$

(6-75-d)

6.3.3 Minimum Uniform Cost Function

In this section, again we use the same approach as used in Subsection 6.2.4. Substituting (6-57) in (6-52), the equation describing the estimation rule $h_i(\underline{y}_i)$ of the i^{th} estimator, $i = 1, 2, \ldots, N$, becomes

$$\int_{-\infty}^{+\infty} dY^i \; p(a=\tau + \sum_{j=1}^{N} h_j(y_j) + C,\; Y^i | y_i))$$

$$= \int_{-\infty}^{+\infty} dY^i \; p(a= -\tau + \sum_{j=1}^{N} h_j(y_j) + C,\; Y^i | y_i))$$

$$i = 1,2,\ldots,N \qquad (6\text{-}76)$$

We solve for C by taking the derivative of $R_{u,m,i,\tau}$ with respect to C and setting the result equal to zero. Then, we get

$$\int_{-\infty}^{+\infty} dY \; p(a= \sum_{j=1}^{N} h_j(y_j) + C,\; Y) = 0 \qquad (6\text{-}77)$$

(6-76) and (6-77) describe the complete solution to the problem.

In the case of two estimators the equations are

$$\int_{-\infty}^{+\infty} dy_1 \; p(a=\tau+h_1+h_2+C, y_1 | y_2) = \int_{-\infty}^{+\infty} dy_1 \; p(a=-\tau+h_1+h_2+C, y_1 | y_2)$$

$$(6\text{-}78)$$

$$\int_{-\infty}^{+\infty} d\underline{y}_2 \; p(a=\tau+h_1+h_2+C, \underline{y}_2 | \underline{y}_1) = \int_{-\infty}^{+\infty} d\underline{y}_2 \; p(a=-\tau+h_1+h_2+C, \underline{y}_2 | \underline{y}_1)$$

(6-79)

and,

$$\int_{-\infty}^{+\infty} d\underline{y}_1 \int_{-\infty}^{+\infty} d\underline{y}_2 \; p(a=h_1(y_1)+h_2(y_2)+C, \; \underline{y}_1, \; \underline{y}_2) = 0 \qquad (6\text{-}80)$$

For illustration purposes, we present a simple example in the next section.

6.4 Example

In this section we consider the system shown in Figure 6.2. Two local estimators are used along with a fusion center which employs a linear combining rule. It is desired to estimate a parameter " a "whose density function is

$$p(a) = \frac{1}{(2\pi)^{1/2}\,\sigma_a} \exp\left(-\frac{a^2}{2\sigma_a^2}\right) \qquad (6\text{-}81)$$

Each estimator i receives a single observation y_i, $i = 1, 2$, corrupted by additive Gaussian noise n_i, $i = 1, 2$. Noises at the sensors are assumed to be independent of each other as well as independent of the parameter a. The two observations at the estimators and their conditional densities are

$$y_1 = a + n_1 \tag{6-82-a}$$

$$y_2 = a + n_2 \tag{6-82-b}$$

and,

$$p(y_i|a) = \frac{1}{\sigma_n \; (2\pi)^{1/2}} \; \left(\frac{(y_i - a)^2}{2 \; \sigma_n^2} \right) \qquad i=1,2 \tag{6-83}$$

Our goal is to obtain the optimum local estimation rules $h_1(y_1)$ and $h_2(y_2)$ so as to minimize R_{MS} when a linear combining rule is used at the fusion center.

When the combining rule is known to be linear , $\hat{a}$ is

$$\hat{a} = h_1(y_1) + h_2(y_2) + C \tag{6-84}$$

$h_1(y_1)$ and $h_2(y_2)$ are obtained by solving the following set of equations

$$h_1(y_1) = E[a|y_1] - E[h_2(y_2)|y_1] - C \tag{6-85-a}$$

$$h_2(y_2) = E[a|y_2] - E[h_1(y_1)|y_2] - C \tag{6-85-b}$$

$$C = E(a) - E[h_1(y_1)] - E[h_2(y_2)] \tag{6-85-c}$$

Substituting (6-85-a) into (6-85-b), we get

$$h_1(y_1) = E[a|y_1] - E[E(a|y_2)|y_1] - E[E[h_1(y_1)|y_2]|y_1] \tag{6-86}$$

For this example the conditional density function of a given y_i, $i = 1, 2$, is given by

$$p(a|y_i) = \left(\frac{\dfrac{1}{(2\pi)\sigma_n\sigma_a}}{p(y_i)} \right) \exp\left(-(1/2)\left[\frac{(y_i-a)^2}{\sigma_n^2} + \frac{a^2}{\sigma_a^2}\right]\right) \qquad i=1,2 \qquad (6\text{-}87)$$

or

$$p(a|y_i) = k_i(y_i) \exp\left(-(1/2)\,\frac{\sigma_a^2+\sigma_n^2}{\sigma_a^2\,\sigma_n^2}\left[a-\frac{\sigma_a^2}{\sigma_a^2+\sigma_n^2}y_i\right]^2\right) \qquad i=1,2 \qquad (6\text{-}88)$$

which is the probability density function of a Gaussian random variable. The mean and the variance for the conditional distribution are [30]

$$\text{mean} = \frac{\sigma_a^2}{\sigma_a^2+\sigma_n^2}\,(y_i) = m_{a|y_i} \qquad (6\text{-}89)$$

$$\text{variance} = \frac{\sigma_a^2\sigma_n^2}{\sigma_a^2+\sigma_n^2} = \sigma^2_{a|y_i} \qquad (6\text{-}90)$$

Since y_1 and y_2 are jointly Gaussian random variables, the mean of y_i, $E(y_i)$, and the covariance of y_1 and y_2 and the variance of y_i are given by

$$E(y_i) = E(a) + E(n_i) = 0 \qquad (6\text{-}91)$$

$$E(y_1y_2) = E(a^2) = \sigma_a^2 \qquad (6\text{-}92)$$

and,

$$\sigma_i^2 = \sigma_a^2 + \sigma_n^2 \tag{6-93}$$

The correlation coefficient between the two random variables y_1 and y_2, is

$$\phi = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_n^2} \tag{6-94}$$

Therefore,

$$p(y_i|y_j) = \frac{1}{\left(\dfrac{2\pi\sigma_n^2(\sigma_n^2+2\sigma_a^2)}{1 + \sigma_a^2}\right)^{1/2}} \times \exp\left(\frac{\left(y_i - y_j \dfrac{\sigma_a^2}{\sigma_a^2 + \sigma_n^2}\right)^2}{2\,\dfrac{\sigma_n^2(\sigma_n^2 + 2\sigma_a^2)}{1 + \sigma_a^2}}\right) \tag{6-95}$$

which yields

$$E(a|y_1) = \frac{\sigma_a^2}{\sigma_a^2+\sigma_n^2}\, y_1 \tag{6-96-a}$$

and,

$$E(a|y_2) = \frac{\sigma_a^2}{\sigma_a^2+\sigma_n^2}\, y_2 \tag{6-96-b}$$

Substituting (6-95) and (6-96) into (6-65), we have

$$h_1(y_1) = \frac{\sigma_a^2}{\sigma_a^2+\sigma_n^2} - E([\frac{\sigma_a^2}{\sigma_a^2+\sigma_n^2} y_2]|y_1) + E([E(h_1(y_1)|y_2]|y_1) \tag{6-97}$$

which may be rearranged as

$$h_1(y_1) = \frac{\sigma_a^2 \; \sigma_n^2}{(\sigma_a^2 + \sigma_n^2)^2} y_1 + E(E(h_1(y_1))|y_2]|y_1) \tag{6-98}$$

In order to attempt a solution to this problem, we make a simplifying assumption that $h_1(y_1)$ is a linear combination of y_1, namely $h_1(y_1) = k_1 \, y_1$. This assumption is appealing because the centralized solution is of the same form. Thus, we only need to find the coefficient of y_1 in the linear function.

$$k_1 \; y_1 = \frac{\sigma_a^2 \; \sigma_n^2}{(\sigma_a^2 + \sigma_n^2)^2}$$

$$+ \int_{-\infty}^{+\infty} dy_2 \; p(y_2|y_1) \int_{-\infty}^{+\infty} k_1 y_1 \; p(y_1|y_2) \; dy_1 \tag{6-99}$$

which becomes

$$k_1 y_1 = \frac{\sigma_a^2 \; \sigma_n^2}{(\sigma_a^2 + \sigma_n^2)} y_1 + k_1 \left(\frac{\sigma_a^2}{\sigma_a^2 + \sigma_n^2}\right)^2 y_1 \tag{6-100}$$

Eliminating y_1 and solving for k_1 , we get

$$k_1 = 1/2 \frac{\sigma_a^2}{\sigma_a^2 + (\sigma_n^2/2)} \tag{6-101}$$

Similar expression can be obtained for k_2. Then,

$$C = E[a - h_1(y_1) - h_2(y_2)] = 0 \tag{6-102}$$

and the optimum estimate is

$$\hat{a}(y_1,y_2) = (1/2)\,\frac{\sigma_a^2}{\sigma_a^2 + (\sigma_n^2/2)}\,(y_1 + y_2) \tag{6-103}$$

which is the same value as obtained for the centralized case.

Next, we obtain the estimation rules for the DPEF problem so as to minimize R_{unif}.

Minimum Uniform Error Function

In this case, the density functions $p(y_1,y_2,a)$, and $p(y_j,a|y_i)$ $j \neq i$, $i = 1, 2$ are given by

$$p(y_1,y_2,a) = \frac{1}{(2\pi)^{3/2}\sigma_a\sigma_n^2}\exp\left(-\frac{\sigma_a^2[(y_1-a)^2+(y_2-a)^2]+a^2\sigma_n^2}{2(\sigma_n\,\sigma_a)^2}\right)$$

(6-104)

$$p(y_j,a|y_i) = k(y_i)\exp\left(-\frac{\left[a - (y_1+y_2)\dfrac{\sigma_a^2}{\sigma_n^2 + 2\sigma_a^2}\right]^2}{\left[2\,\dfrac{\sigma_n^2\,\sigma_a^2}{\sigma_n^2 + 2\sigma_a^2}\right]}\right)$$

(6-105)

and for τ very small,

$$p(y_j, a=h_1+h_2+C \mid y_i) =$$

$$k(y_i)\exp\left(- \frac{\left[h_1+h_2-C-(y_1+y_2) \dfrac{\sigma_a^2}{\sigma_n^2 + 2\sigma_a^2} \right]^2}{\left[2 \dfrac{\sigma_n^2 \sigma_a^2}{\sigma_n^2 + 2\,\sigma_a^2}\right]} \right)$$

$$i \neq j \quad i,j=1,2 \qquad (6\text{-}106)$$

where $k(y_i)$ is a function of y_i.

Note that the exponential term may be easily rearranged to maximize $P(y_j, a \mid y_i)$. Its maximum is obtained by setting the following

$$h_i(y_i) = \frac{\sigma_a^2}{\sigma_a^2 + 2\sigma_n^2} y_i \qquad i=1,2. \qquad (6\text{-}107)$$

and,

$$C = 0 \qquad (6\text{-}108)$$

Then,

$$\hat{a}_{unif} = \frac{\sigma_a^2}{\sigma_a^2 + 2\sigma_n^2}(y_1 + y_2) \qquad (6\text{-}109)$$

which is the desired result.

VII. Summary and Suggestions for Future Work

7.1 Summary

In this report, we have considered the problems of hypothesis testing and parameter estimation when multiple sensors are used and a global decision (or estimate) is desired. For the hypothesis testing problem, we have considered Neyman-Pearson detection, Bayesian detection and minimum equivocation detection. In all cases, local decisions are fed to a data fusion center where a global decision is obtained. When the fusion rule is given, the decision rules at the individual detectors were derived. When the decision rules at the individual detectors are known, we derived the optimum fusion rule. We also derived the overall solution to the problem, i.e., obtain both the optimum fusion rule and the optimum decision rules at the detectors simultaneously. We obtained the results for the general problem of distributed hypothesis testing with data fusion. With an appropriate choice of the fusion rule and the cost assignment, the DD problem becomes a special case of our problem. For the Bayesian formulation, we have considered the case of identical detectors with independent observations and obtained the value of K in the "K out of N" fusion rule. Distributed postdetection integration has also been considered where two schemes were proposed. The corresponding optimum rules have been obtained.

We have also considered the distributed Bayesian parameter estimation problem, when local estimators transmit their estimates to a combiner where a global estimate is obtained. Optimum estimation rules at the local estimators and optimum combining rules have been derived for three different cost critera.

In all of the above cases, the equations that specify the solution are coupled and highly nonlinear and are therefore hard to solve. The level of computational difficulty increases rapidly with the number of decision makers (or estimators), especially for the minimum equivocation and parameter estimation problems.

7.2 Suggestions for Future Work

One area of research that should be pursued is to find efficient ways to solve the set of highly nonlinear coupled equations that are characteristic of the decentralized detection and estimation problems.

Other fertile areas will be to solve signal detection and estimation problems in an uncertain environment in a distributed framework. Knowledge based approaches to multisensor integration and data fusion can also be considered.

References

[1] H. L. Van Trees, Detection, Estimation and Modulation Theory, Volume 1, J. Wiley, 1969.

[2] J. V. Difranco and W. L. Rubin, Radar Detection, Englewood Cliffs, N. J. Prentice-Hall, 1968.

[3] R. R. Tenney and N. R. Sandell, "Detection With Distributed Sensors," IEEE Trans. on Aerospace and Electronic Systems, Vol. 17, No. 4, pp. 501-509, July 1981.

[4] F. A. Sadjadi, "Hypothesis Testing in Distributed Environment," IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-22, pp. 134-137, March 1986.

[5] G. Lauer and N. R. Sandell, Jr., "Distributed Detection for Known Signals in Correlated Noise," TP-131, ALPHATECH Inc. Burlington, MA, March 1982.

[6] L. K. Ekchian and R. R. Tenney, "Detection Networks," Proceedings of the 21st IEEE conference on Decision and Control, Orlando, Fla, pp. 686-691, December 1982.

[7] H. J. Kushner and A. Pacut, "A Simulation Study of a Decentralized Detection Problem," IEEE Trans. on Automatic Control, Vol. 27, No. 5, pp. 1116-1119, October 1982.

[8] D. Teneketzis, "The Decentralized Wald Problem," Proceedings-1983, American Control Conference, San Fransisco, CA, 1983.

[9] D. Teneketzis, "The Decentralized Quickest Detection Problem," Proceedings of the 21st IEEE Conference on Decision and Control, Orlando, Fla, December 1982.

[10] J. N. Tsitsiklis and M. Athans, "On The Complexity of Decentralized Decision Making and Detection Problems," IEEE Trans. on Automatic Control, Vol. AC-30, No. 5, pp. 440-446, May 1985.

[11] E. Conte, E. D'Addio, A. Farina and M. Longo, "Multistatic Radar Detection: Synthesis and Comparaison of Optimum and Suboptimum Receivers," IEE Proceedings, Vol. 130, Part F, No.6, October 1983.

[12] S. D. Stearns, "Optimal Detection Using Multiple Sensors," G.T.E. Products Corporation.

[13] N. R. Sandell and M. Athans, "Solution of Some Nonclassical LQG Stochastic Decision Problems," IEEE Trans. on Automatic Control, Vol. AC-19, pp. 108-116 April 1974.

[14] R. Radner, "Team Decision Problems," Am. Math. Statistics, Vol.33, No. 3, pp. 857-881, 1962.

[15] Y. C. Ho, "Team Decision Theory and Information Structures," Proceedings of the IEEE, Vol. 68, No. 6, pp. 644-654, June 1980.

[16] Y. C. Ho and K. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems Part-1," IEEE Trans. on Automatic Control, Vol. AC-17, pp. 15-22, 1972.

[17] Y. C. Ho and T. S. Chang, "Another Look at The Nonclassical Information Structure Problems," IEEE Trans. on Automatic Control, pp. 537-540, June 1980.

[18] N. R. Sandell, P. P. Varaiya, M. Athans and M. G. Safanov, "A Survey of Decentralized Control Methods for Large Scale Systems," IEEE Trans. on Automatic Control, Vol. AC-23, pp. 108-128, 1978.

[19] V. Borkar and P. P. Varaiya, "Asymptotic Agreement in Distributed Estimation," IEEE Trans. on Automatic Control, Vol. AC-27, No.3 pp. 650-655, June 1982.

[20] D. Teneketzis and P. Varaiya, "Consensus in Distributed Estimation With Inconsistent Beliefs," ALPHATECH and the Department of Electrical Engineering and Computer Sciences and Electronics Research Laboratory, University of California, Berkley, January 1984.

[21] C. Chang and J. A. Tabaczynski, "Applications of State Estimation to Target Tracking," IEEE Trans. on Automatic Control, Vol. AC-29, No. 2, pp. 98-109, February 1984.

[22] A. Willsky, M. G. Bello, D. A. Castanon, B. C. Levy and G. G. Verghese, "Combining and Updating of Local Estimates and Regional Maps Along Sets of One-Dimensional Tracks." IEEE Trans. on Automatic Control, Vol. AC-27, Mo.4, pp. 799-811, August 1982.

[23] B. C. Levy, D. C. Castanon, G. C. Verghese and A. S. Willsky, "A Scattering Framework for Decentralized Estimation Problems," Automatica, Vol. 19, No. 4, pp. 373-384, 1983.

[24] B. C. Levy, D. C. Castanon, G. C. Verghese and A. S. Willsky, "Smoothing Error Dynamics and Their Use in the Solution of Smoothing and Mapping Problems," IEEE Trans. On Information Theory, Vol IT-32, No. 4, pp. 483-495, July 1986.

[25] D. A. Castanon and D. Teneketzis, "Distributed Estimation Algorithms for Nonlinear Systems," IEEE Trans. on Automatic Control, Vol. AC-30, No. 5, pp. 418-425, May 1985.

[26] R. B. Washburn and D. Teneketzis, "Performance Analysis For Hybrid-State Estimation Problems," IEEE Trans. On Automatic Control, 1985.

[27] R. K. Varshney and P. K. Varshney, "Recursive Estimation With Uncertain Observations," IEE Proceedings part F, pp. 527-533, October 1986.

[28] M. I. Skolnik, Radar Handbook, Mc Graw-Hill, 1970.

[29] D. Middleton, An Introduction to Statistical Communication Theory, Mc Graw-Hill, 1960.

[30] A. Papoulis, Probability, Random Variables, and Stochastic Processes, Mc Graw-Hill, 1984.

[31] J. M. Miller, "An Alternative Method for Designing the Double Threshold Detector," IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-21, No. 4, July 1985.

[32] J. V. Harrington, "An Analysis of the Detection of Repeated Signals in Noise by Binary Integration," IRE Trnas. on Information Theory, IT-1, pp. 1-9, March 1955.

[33] M. Schwartz, "A Coincidence Procedure for Signal Detection," IRE Trans. on Information Theory, IT-2, pp. 135-139, December 1956.

[34] G. M. Dillard, "A Moving Window Detector for Binary Integration," IEEE Trans. on Information Theory, IT-13, pp. 2-6, January 1967.

[35] R. Worley, "Optimum Thresholds for Binary Integration," IEEE Trans. on Information Theory, IT-14, pp. 349-353, March 1968.

[36] J. F. Walker, "Performance Data for a Double-Threshold Detection Radar," IEEE Trans. on Aerospace and Electronic Systems, AES-7, pp. 142-146, January 1971.

[37] T. L. Gabrielle, "Information Criteria for Threshold Determination," IEEE Trans. on Information Theory Vol. IT, No, pp. 484-486, October 1966.

[38] R. E. Blahut, "Hypothesis Testing and Information Theory, "IEEE Trans. on Information Theory Vol. IT 20, No 4, pp. 405-417, July 1974.

[39] J. M. Van Campenhort and T.M. Cover, "Maximum Entropy and Conditional Probability," IEEE Trans. On Information Theory Vol. It-27, No. 4, pp. 483-484, July 1981.

[40] M. E. Hellman and Josef Raviv, "Probability of Error, Equivocation, and the Chernoff Bound." IEEE Trans. on Information Theory, Vol. IT-10, No. 4, pp. 368-377, July 1970.

[41] R. J. McEliece, The Theory of Information and Coding, Encyclopedia of Math., Vol. 3, Addison-Wesley, 1977.

[42] R. G. Gallager, Information Theory and Reliable Communications, Wiley, 1968.

☆U.S. GOVERNMENT PRINTING OFFICE 1988-511-117/60654

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control, Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C^3I systems. The areas of technical competence include communications, command and control, battle management, information processing, surveillance sensors, intelligence data collection and handling, solid state sciences, electromagnetics, and propagation, and electronic, maintainability, and compatibility.

END
DATE
FILMED
DTIC
JULY 88