PERFORMANCE ANALYSIS OF ACTIVE SONAR CLASSIFIERS

A Dissertation Presented to

The Faculty of the College of

Engineering and Technology

Ohio University

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

by

Nicholas K. Haddad

June, 1990

ACKNOWLEDGEMENTS

I wish to express my deep appreciation to Dr. John Tague without whom this research could not have been accomplished. I would like to thank him for his patience, invaluable guidance and encouragement throughout this study. His counselling and advice have also made a tremendous impact on my academic career.

Special gratitude goes to Dr. J.G. Kelly, Dr. R.N. Carpenter and Dr. D. Childs for their cooperation. I would also like to thank Dr. Joe Essman, Dr. Jeffrey Giesey, Dr. Dennis Irwin and Dr. Larry Snyder for serving as my graduate committee members.

Then, I would like to express my deepest gratitude to my parents, brothers and sister whose continuous encouragement and financial and emotional support made it possible for me to pursue higher education. I thank them for instilling in me the discipline and drive needed for success.

Finally, I greatly appreciate the opportunity that the United States provided me throughout my college studies. I am also grateful to Lebanon, my country, who despite its misery and hardship, is always in my heart.

TABLE OF CONTENTS

TABLES OF FIGURES

Chapter 1

INTRODUCTION

1.1 Problem Statement and Its Importance

a. General Problem Statement

Considerable attention has been given already to optimum detection for determining the presence or absence of a signal in noise. In such systems, two decisions only are made: (1) a signal (as well as noise) is present or (2) only noise occurs. For example, in a radar detection problem, we might select two hypotheses; a target is present or no target is present.

Our purpose in this dissertation is to study the optimum multiple alternative detection or classification of signals in noise, as applied to active sonars. Active sonar classification is statistical in nature due to the uncertainty associated with the scattering medium and the background noise present in the measurement data. Consequently, the problem is best posed and solved within the framework of Bayesian decision theory.

As in binary cases, the methods of decision theory are also applied to multiple alternative situations involving the possible presence of more than one signal during an observation interval T. This is done using the notion of risk, which associates certain costs with each possible decision, and uses the average risk to evaluate system performance. An optimum decision among the multiple alternatives is obtained by minimizing the average risk, and this minimum average is known as the Bayes risk.

In this dissertation, we are concerned with the implementation issues arising in active sonars. We will display receiver operating characteristics (ROC) curves in function of important parameters such as target resolution,

signal–to–noise ratio, signal design, etc. We will also study the sensitivity of the classifier when wrong a priori knowledge of the target's scattering properties is processed. Finally, we will tackle the problem of acoustic target imaging. An optimal estimate of the target's scattering coefficients will be derived and investigated.

b. Specific Problem Statement

The specific issues to be studied in this dissertation are summarized as follows:

(1) To study the classifier performance as a function of important parameters such as target resolution, signal design, etc., for meaningful target geometries.

(2) To investigate classifier sensitivity and performance when incorrect a priori knowledge of certain parameters, such as target range and target orientation, is infused into the processor.

(3) To study the imaging problem in the context of optimum classification.

c. Importance of the Problem

Most active sonar classifiers are based on accurate statistical description of the illuminated target. In other words, perfect a priori knowledge of the target's scattering properties, range, orientation, etc. is almost always assumed in classifier design. In practice, scattering models are obtained from experimental data and calculations performed on hypothesized targets. The fidelity of this approach is questionable, since the assumed nominal model is never a perfect replica of the actual target. Moreover, there will always be uncertainty about the target's range and orientation. Consequently, unexpected

and severe degradations in performance may result, and the need to investigate classifier performance when an incorrect or imprecise statistical description is used is essential.

A good classifier is one that can discriminate between two slightly different targets. An intuitive approach suggests more point scatterers modeling the target in order to exploit finer details. However, when more scatterers are considered, more noise contaminated data will be processed and performance is at stake. One way to get around this problem is to increase the signal–to–noise ratio in order to make noise less significant to the processor. Here again, there is a cost to be paid. Therefore, these different design tradeoffs constitute another important aspect of classifier design.

The last issue to be investigated is acoustic target imaging. If the "strength" or energy of each point scatterer forming the target can be estimated, one can reconstruct the target. As a result, better estimation of the target's statistics implies better classification. This leads to the importance of optimal estimation techniques in active sonars, and an investigation of the parameters affecting such techniques proves to be worthwhile.

1.2 Problem Formulation

A schematic diagram of an active sonar system is shown in figure 1.1. Short pulses of acoustic energy are radiated periodically, usually from a highly directional transmitter. If an object is located in the beam of the transmitter, the energy in the pulse will be scattered in many directions. Some of this energy will be reflected back towards an array of sensors. The return energy picked up by the sensors has random parameters that are mainly dependent on the scattering properties of the object and the transmission medium.

All signals not related to the signal backscattered from the object constitute undesired interference we call noise. It has two components: system noise and ambient noise. System noise arises within the processor and ambient noise originates from the thermal energy within the array. We will combine the system and ambient noises into a single process modeled by a zero–mean, Gaussian distributed random vector. Furthermore, we will assume that the noise process is spatially and temporally white.

The classification problem is formulated as a multiple hypothesis test on the array data. The signal model for the array data given hypotheses H_0, H_1, . . . , H_I is:

$$H_0 : \mathbf{x}(t) = \mathbf{n}(t) \quad \text{(noise only)}$$

$$H_i : \mathbf{x}(t) = \mathbf{s}_i(t) + \mathbf{n}(t) \; ; \; i = 1, 2, \ldots, I \tag{1.2.1}$$

where $\mathbf{x}(t)$ denotes the output from an array of N sensors at time t, $\mathbf{n}(t)$ represents system and ambient noise, and $\mathbf{s}_i(t)$ represents the return from the i^{th} target. The i^{th} target will be modeled with a finite number of independent point scatterers and $\mathbf{s}_i(t)$ is a function of these scatterers and the time–delayed replicas of the transmitted waveform. It will be shown in chapter 2 that $\mathbf{s}_i(t)$ can be modeled as a zero–mean stochastic process with Gaussian statistics.

In this dissertation, we will study the ternary problem where we consider three hypotheses: the null hypothesis H_0 and two alternative hypotheses H_1 and H_2 (refer to figure 1.2). Here hypothesis H_2 represents a target plus noise and H_1 represents a false target or possibly the same target of hypothesis H_2 with a different orientation or range. False targets may occur due to undesired reflected energy, called reverberation noise. In the ocean, this energy can be

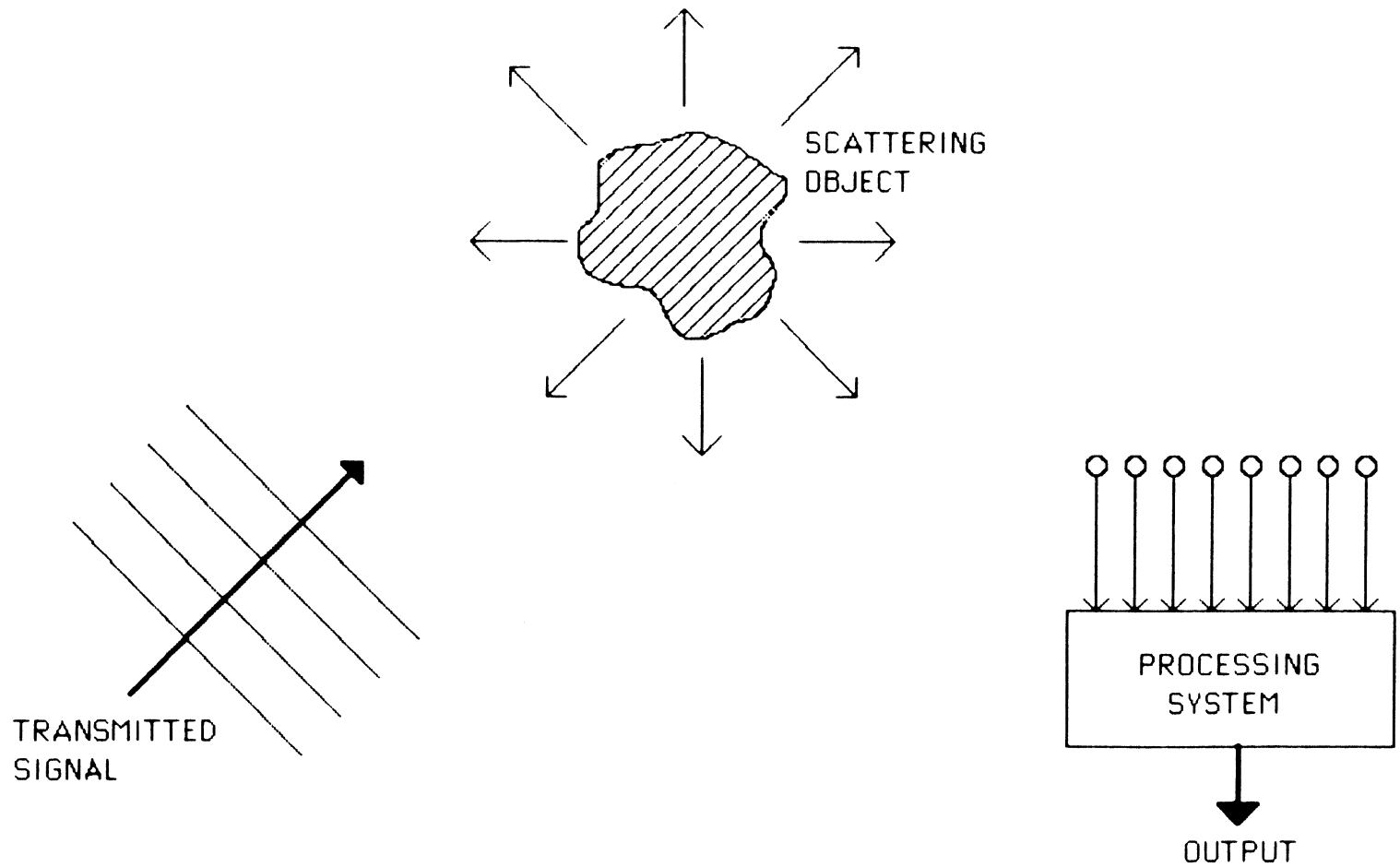


Figure 1.1: System under consideration

reflected from foreign materials such as suspended solids, sea organisms, fish, etc.

Although the ternary problem is emphasized, we present a systematic approach for designing Bayesian classifiers capable of deciding among I+1 hypotheses ($I \geq 2$). Indeed, a decision rule designed to select between H_1 and H_2 can be used to select between H_1 and H_3 or H_2 and H_3 in a four–hypotheses case, for instance. Furthermore, design parameters arising in classifier implementation are independent of the number of hypotheses considered. Studying classifier performance in function of such parameters is the basic issue addressed in this dissertation.

1.3 Basic Approach to the Problem

We begin by describing the sonar model in more details, since it is a crucial part of the overall approach to the problem. As we mentioned before, in active classification, a known signal is transmitted into a propagating medium and directed towards a region called the test volume. We will subdivide the test volume into K disjoint regions called cells. Each cell contains at most one point scatterer which models the net reflectivity properties of the object within that cell.

Measurements from an array of sensors immersed in the medium are processed in order to decide if an object is present, and if so, what kind of object is present. For the ternary problem, it can be shown that a minimum Bayesian risk decision rule for selecting between H_1 and H_2 leads to a sufficient statistic L_{12}. If L_{12} is larger than the classification threshold η, we assert that H_1 is correct; otherwise, we assert H_2 is correct.

The ability of the classifier to distinguish between H_1 and H_2 depends on

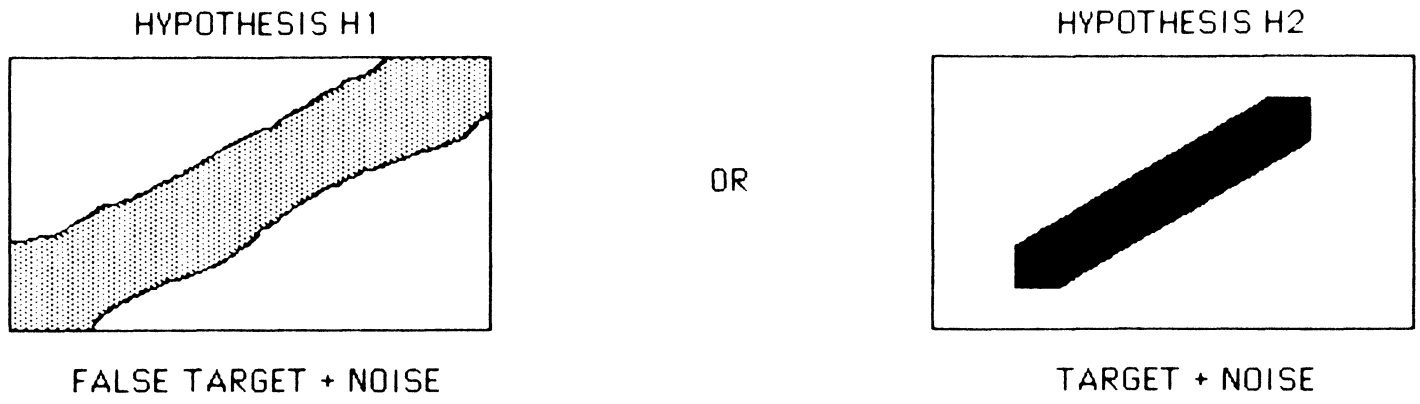


Figure 1.2 : Classification concept for the ternary problem

the probability density functions $f(L_{12}|H_1)$ and $f(L_{12}|H_2)$ of L_{12} conditioned on H_1 and H_2, respectively. Throughout this dissertation, hypothesis H_2 represents a target of interest and hypothesis H_1 represents a false target (reverberation). The performance measures we shall use are the classification and false alarm probabilities, $P(H_2|H_2)$ and $P(H_2|H_1)$, respectively. It turns out that the sufficient statistic L_{12} has a quadratic form. Distributions of quadratic forms in complex multinormal random variables have been studied by Turin [20], Kharti [21] and Johnson and Kotz [22]. They derived an expression for the characteristic function of a quadratic form. Therefore, the basic approach in evaluating the classifier performance is summarized in the following steps:

(1) Evaluate the characteristic functions $M_1(\omega)$ and $M_2(\omega)$ starting with the quadratic form L_{12}.

(2) Using the characteristic functions derived in (1), obtain the conditional probability density function $f(L_{12}|H_1)$ and $f(L_{12}|H_2)$.

(3) Evaluate the classification and false alarm probabilities, $P(H_2|H_2)$ and $P(H_2|H_1)$, based on conditional probability density functions derived in (2).

The second issue to be investigated in this dissertation is classifier sensitivity to incorrect a priori knowledge of the test volume scattering properties. The scattering properties of the test volume will be described by the scattering covariance matrices $\mathbf{K}_{a_1}$ and $\mathbf{K}_{a_2}$ corresponding to hypotheses H_1 and H_2, respectively. In order to compute the sufficient statistic L_{12}, the classifier must know $\mathbf{K}_{a_1}$ and $\mathbf{K}_{a_2}$ a priori. Perfect knowledge of these matrices is usually a difficult task. If, for instance, there is uncertainty concerning the target range or orientation under hypothesis H_2, the corresponding scattering covariance

matrix, now labeled $\mathbf{K}'_{a_2}$, will be mismatched to the actual test volume geometry. Hence, performance analysis of the "mismatched" classifier is accomplished by repeating steps (1), (2) and (3), discussed earlier in the section, with the wrong scattering covariance matrix ($\mathbf{K}'_{a_2}$ in this case), infused into the processor. In both matched and mismatched processing, residue theory will be employed extensively to evaluate $f(L_{12}|H_1)$ and $f(L_{12}|H_2)$.

In this dissertation, we will also study the issue of acoustic target imaging. This problem will be approached by evaluating the minimum variance linear unbiased estimator $\tilde{\mathbf{a}}_i$ ($i = 1, 2$) of the test volume scattering coefficients under each hypothesis. To do this, we will use basic properties of estimation theory as well as the Wiener–Hopf equation.

In conclusion, our approach to the problem is based on Bayesian decision theory. Within this framework, the classifier performance will be analyzed. Many ideas of random processes, linear algebra and complex variable theory will be employed in solving the problem.

1.4 Review of Literature

Classification and estimation, discussed in this dissertation, are a combination of the classical techniques of decision theory and random process characterization. Decision theory was studied in the middle of the eighteenth century by Thomas Bayes [1], and work in the area of estimation theory was published by Legendre [2] and Gauss [3] in the early nineteenth century. Significant contributions to the classical decision theory were developed by Fisher [4] and Neyman and Pearson [5] more than 50 years ago.

Statistical decision theory was later on applied to multiple–alternative

detection of signals in noise. Two fundamental papers in this area were published by Middleton and Van Meter [6,7] in 1955. They tackled the problem based on Bayesian decision theory. In 1962, Thomas and Wolf [8] studied the multiple detection issue using two criteria: Bayes procedure and the Neyman–Pearson approach. They found that, for both criteria, the decision is based on the likelihood ratios for the various signals.

Likelihood ratios were studied by Van Trees [13] in a classic report published more than three decades ago. More specifically, he examined the Gaussian signal in Gaussian noise detection problem and derived several equivalent forms of the optimal structure used to implement likelihood ratios. One of these forms is the estimator–correlator structure that we shall apply to our problem.

Multiple detection applied to active sonar systems is commonly known as classification. Active sonars have been used by the U.S. Navy as part of antisubmarine warfare [9]. One of the relevant applications is automatic torpedo guidance via active sonar. Due to the military nature of the application, very little work in this area has been published. Nevertheless, two reports, made available by the Naval Underwater Systems Center (NUSC), have been used extensively throughout this dissertation. The first report, written by Middleton and Pugliese [10] in 1987, provides a framework for attacking problems that prevent achieving the full operational potential of active sonar systems. Kelly and Carpenter [17] wrote the second report in which they discuss the problem from a more technical point of view.

1.5 Overview of the Dissertation

Chapter 2 introduces the concepts, definitions, notation and theorems

that will be used in subsequent chapters. In particular, we derive the sufficient statistic for the ternary problem by minimizing the Bayesian risk. We also derive an expression for the likelihood ratio based on the estimator–correlator structure.

In chapter 3, classifier performance is studied for high resolution signals, i.e., when targets are illuminated by signals capable of perfect resolution among cells. Closed form expressions for the classification and false alarm probabilities are derived. Classifier performance will be analyzed with respect to: (1) changes in target strength, (2) range or orientation mismatch, (3) signal–to–noise ratio, and (4) the number of test volume cells.

Classifier performance when targets are illuminated by linear FM signals is discussed in chapter 4. In addition to the parameters consider in Chapter 3, processor performance will be studied with respect to: (1) transmitter constant k_f and (2) target resolution.

Estimating the test volume scattering coefficients is the goal of chapter 5. First, we derive the MVLU estimator $\tilde{\mathbf{a}}_i$, then we examine the parameters needed to reduce the error variance of the estimator. These parameters turn out to be: (1) transmitter constant k_f, (2) received signal–to–noise ratio, (3) angular resolution, and (4) number of cells in the test volume. Finally, conclusions and recommendations for further research are given in chapter 6.

Chapter 2

FORMULATION OF THE PROBLEM AND MATHEMATICAL BACKGROUND

2.1 Introduction

This chapter introduces the concepts, definitions, notation, and theorems that will be used in subsequent chapters. Results will be stated, but some proofs and derivations will be omitted. The interested reader can find careful developments of detection theory in references [11,12,14], and of the estimator–correlator structure in references [12,13].

In section 2.2 we develop the mathematical model and state the assumptions made about the signal and the noise processes. Section 2.3 describes the classification problem in which the Bayes risk is introduced. Section 2.4 discusses the estimator–correlator structure. Section 2.5 ties in the previous sections together in order to derive a sufficient statistic for the ternary problem.

2.2 Signal/Noise Model

a. Signal Model

In active classification, a known signal with a duration of T seconds (finite energy) is transmitted into a propagating medium and directed towards a region called the test volume (figure 2.1). Measurements from an array of sensors immersed in the medium are processed in order to decide if an object is present, and if so, what kind of object is present.

The i^{th} signal vector (i referring to hypothesis H_i) is given by

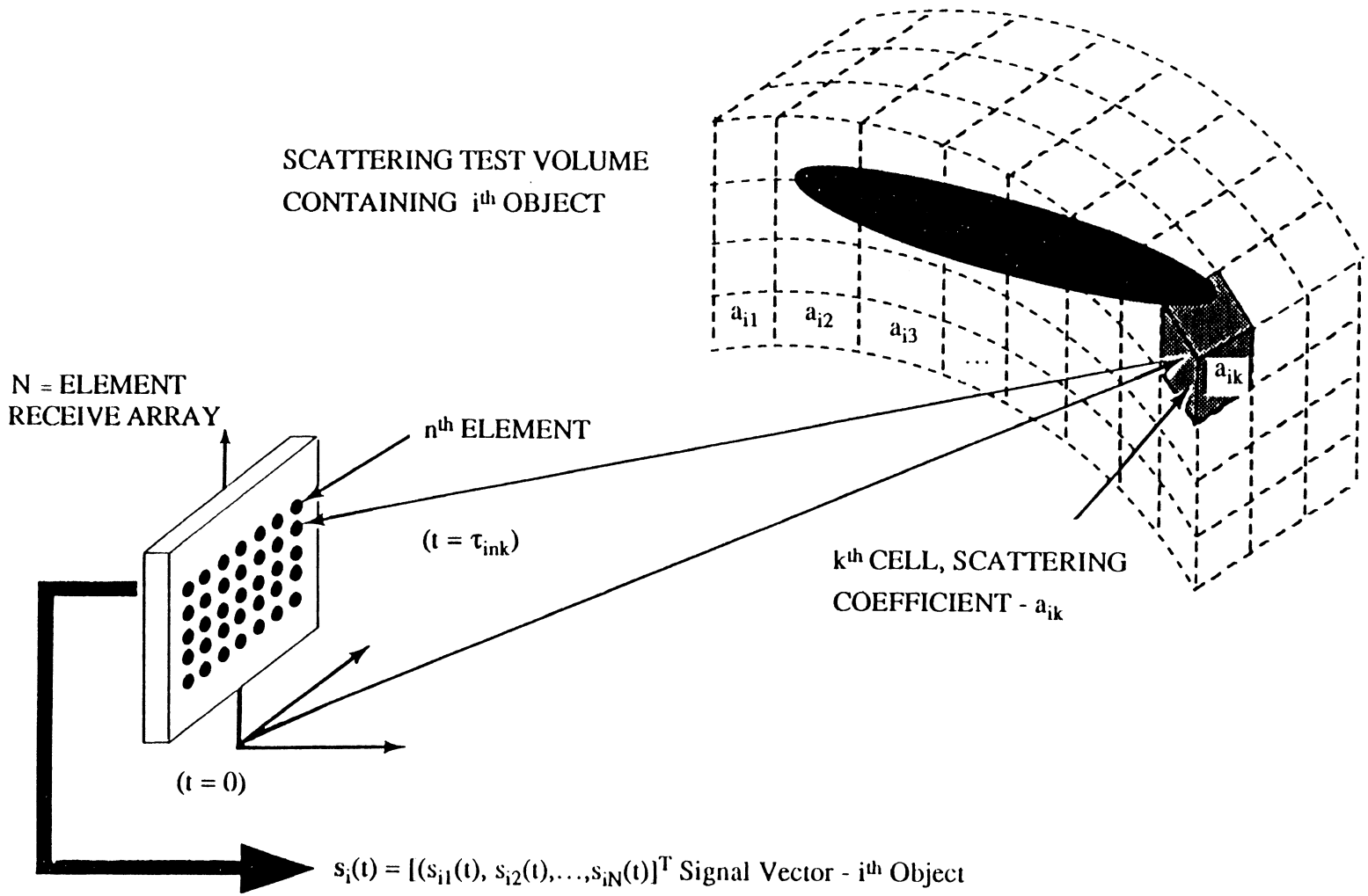


Figure 2.1 Signal Model

$$\mathbf{s}_i(t) = [s_{i1}(t), s_{i2}(t), \ldots, s_{iN}(t)]^T \quad , \quad i = 1, 2, \ldots, I \tag{2.2.1}$$

where I is the number of hypotheses considered (excluding the null hypothesis) and N is the total number of sensors in the array. The test volume is subdivided into K_i disjoint regions called cells. Each cell contains at most one point scatterer which models the net reflectivity properties of the object within that cell. The n^{th} component of the i^{th} signal vector, denoted by $s_{in}(t)$, is found by summing the returns from individual cells:

$$s_{in}(t) = \sum_{m=1}^{K_i} a_{im}\, f(t-\tau_{inm}) \, , \quad 0 < t < T \, . \tag{2.2.2}$$

In this equation, the a_{im} coefficients are random variables modeling the scattering effects within each cell, and $f(t-\tau_{inm})$ is a time–delayed replica of the transmitted waveform. τ_{inm} is the delay from the instant ($t = 0$) of transmission (from a reference transducer element in the array) to reception by the n^{th} element of the array, of the waveform scattered by the k^{th} cell. The delays are deterministic, depending only on the array and test volume geometries. Using vector–matrix notation:

$$\mathbf{s}_i(t) = \mathbf{F}_i(t)\, \mathbf{a}_i \tag{2.2.3}$$

where $\mathbf{F}_i(t)$ is the $N \times K_i$ matrix of time–delayed transmitted waveforms, and $\mathbf{a}_i$ is the $K_i \times 1$ vector of scattering coefficients. $\mathbf{F}_i(t)$ is expressed as follows:

$$\mathbf{F}(t) = \begin{bmatrix} f(t-\tau_{11}), & f(t-\tau_{12}), & \ldots, & f(t-\tau_{1K}) \\ f(t-\tau_{21}), & f(t-\tau_{22}), & \ldots, & f(t-\tau_{2K}) \\ \vdots & & & \\ f(t-\tau_{N1}), & f(t-\tau_{N2}), & \ldots, & f(t-\tau_{NK}) \end{bmatrix} \tag{2.2.4}$$

For simplicity of notation, we omit the index i in both $\mathbf{F}_i(t)$ and the time delays τ_{inm} since these are function of the test volume geometry regardless of the hypothesis considered. The index i will be omitted where appropriate throughout this dissertation. The randomness of the test volume is characterized solely by the random vector a_i, which will be modeled as a complex, zero–mean, normally distributed vector with covariance matrix K_{a_i}. The covariance matrices are the a priori knowledge required for optimal classification. Mathematically, we can write the following equations:

$$\mathbf{a}_i = [a_{i1}, a_{i2}, \ldots, a_{iK}]^T \tag{2.2.5}$$

$$E[\mathbf{a}_i] = 0 \tag{2.2.6}$$

$$\mathbf{K}_{a_i} = E[\mathbf{a}_i \mathbf{a}_i^H] . \tag{2.2.7}$$

Note that the test volume in figure 2.1 is fixed to K cells with spherical coordinates with respect to some reference point. The reference point is conveniently chosen as a transducer element in the receiving array for example. The figure also illustrates an example with a uniform planar array in a non–refractive medium. In practice, the dimensions and decomposition of test regions and estimates of $\mathbf{K}_{a_i}$ will be determined directly from both experimental data and calculations from scattering models of the hypothesized targets. Of course, the dimensions of the test regions would match roughly at least the characteristic

dimensions of the targets. We are also assuming that neither the scatterers nor the sonar platform are in motion.

b. Noise Model

The thermal noise $\mathbf{n}(t)$ at the receive array will be modeled as a zero–mean stochastic process with Gaussian statistics. Practically speaking, it is usually sufficient to assume that the noise is spatially and temporally white, meaning its covariance matrix is given by:

$$\mathbf{K}_n(t_1,t_2) = E[\mathbf{n}(t_1)\,\mathbf{n}^H(t_2)] = \frac{N_0}{2}\,\delta(t_1-t_2)\,\mathbf{I} \qquad (2.2.8)$$

where $\delta(\cdot)$ is the Dirac delta function and $\mathbf{I}$ is an N×N identity matrix. Note that $\mathbf{n}(t)$ is an N×1 random vector such that

$$E[\mathbf{n}(t)] = \mathbf{0}\,. \qquad (2.2.9)$$

c. Overall model

Let us consider I hypotheses H_i (i = 1, . . . , I) where each hypothesis refers to a specific target in the test volume. Different hypotheses could refer to the same object with possibly different orientations. We will call H_0 the null hypothesis, referring to the situation where no target is present in the test volume. Calling $\mathbf{x}(t)$ the N×1 random vector of array measurements for $0 < t < T$ (T is the signal duration) and adding the signal vector $\mathbf{s}_i(t)$ (equation 2.2.3) to the thermal noise process $\mathbf{n}(t)$ at the array, we can describe the overall model by the following equations:

$$H_0 : \mathbf{x}(t) = \mathbf{n}(t) \tag{2.2.10a}$$

$$H_i : \mathbf{x}(t) = \mathbf{F}(t)\,\mathbf{a}_i + \mathbf{n}(t) \quad i = 1, \ldots, I \tag{2.2.10b}$$

where $\mathbf{F}(t)$ and $\mathbf{a}_i$ are given in equations (2.2.4) and (2.2.5) and the measurement $\mathbf{x}(t)$ (array output) is such that $0 < t < T$. We shall further assume that the received signal vector $\mathbf{s}_i(t)$ and the noise vector $\mathbf{n}(t)$ are uncorrelated. In other words:

$$E[\mathbf{s}_i(t_1)\,\mathbf{n}^H(t_2)] = E[\mathbf{F}(t_1)\mathbf{a}_i\,\mathbf{n}^H(t_2)] = \mathbf{0} \tag{2.2.11}$$

From equations (2.2.6), (2.2.9) and (2.2.10), we can easily show that

$$H_i : E[\mathbf{x}(t)] = 0 \quad i = 0, \ldots, I\,. \tag{2.2.12}$$

Therefore, the covariance matrix of $\mathbf{x}(t)$ is given by

$$\mathbf{K}_{x_i}(t_1,t_2) = E[\mathbf{x}(t_1)\,\mathbf{x}^H(t_2)\,|\,H_i] \quad i = 1, \ldots, I\,. \tag{2.2.13}$$

Using equations (2.2.10b), (2.2.11) and Hermitian properties:

$$E[\mathbf{x}(t_1)\,\mathbf{x}^H(t_2)\,|\,H_i] = E[(\mathbf{F}(t_1)\mathbf{a}_i + \mathbf{n}(t_1)(\mathbf{n}^H(t_2) + \mathbf{a}_i^H\mathbf{F}^H(t_2))]$$

$$= E[\mathbf{F}(t_1)\mathbf{a}_i\,\mathbf{a}_i^H\mathbf{F}^H(t_2) + E[\mathbf{F}(t_1)\mathbf{a}_i\,\mathbf{n}^H(t_2)]$$

$$+ E[\mathbf{n}(t_1)\,\mathbf{a}_i^H\mathbf{F}^H(t_2)] + E[\mathbf{n}(t_1)\,\mathbf{n}^H(t_2)]\,.$$

Therefore,

$$\mathbf{K}_{x_i}(t_1,t_2) = \mathbf{F}(t_1)\,\mathbf{K}_{a_i}\mathbf{F}^H(t_2) + \mathbf{K}_n(t_1,t_2) \quad i = 1, \ldots, I\,. \tag{2.2.14}$$

The covariance matrix of $\mathbf{s}_i(t)$ has implicitly been derived in the step above and is expressed as

$$\mathbf{K}_{s_i}(t_1,t_2) = \mathbf{F}(t_1)\,\mathbf{K}_{a_i}\mathbf{F}^H(t_2) \quad i = 1, 2, \ldots, I\,. \tag{2.2.15}$$

2.3 Optimal Bayes Classifier

Beginning with well–established results from statistical decision theory, the classification problem is formulated as a multiple hypothesis test on the array data. An optimal Bayes test is one that minimizes the risk in applying a given decision rule. More details and proofs can be found in references [11,15,16]. The Bayesian risk R is defined as [15]

$$R = \sum_{i=0}^{I} \sum_{j=0}^{I} c_{ij}\, p_j\, \Pr[H_i \mid H_j] \quad i,j = 0, 1, \ldots I \tag{2.3.1}$$

where c_{ij} is the cost of deciding between hypothesis H_i when H_j is true; p_j is the a priori probability of H_j; $\Pr[H_i \mid H_j]$ is the probability that a decision rule chooses H_i when H_j is true. Note that $\Pr[H_i \mid H_j]$ can be expressed as follows:

$$\Pr[H_i \mid H_j] = \int_{z_i} p(\mathbf{x}(t) \mid H_j)\, d\mathbf{x}(t) \tag{2.3.2}$$

where $p(\mathbf{x}(t) \mid H_j)$ is the conditional probability density function of a measurement vector $\mathbf{x}(t)$ given H_j; z_i is a region in the sample space of $\mathbf{x}(t)$ such that H_i

is chosen if a realization of H_i is contained in z_i; the regions z_i ($i = 0, \ldots, I$) are disjoint and cover the entire sample space.

It can be shown [15] that a minimum Bayesian risk decision rule for selection the k^{th} alternative is to choose H_k is

$$E_k(\mathbf{x}(t)) \max \{E_0(x(t)), E_1(\mathbf{x}(t)), \ldots, E_I(\mathbf{x}(t))\} \tag{2.3.3}$$

where

$$E_i(\mathbf{x}(t)) = \sum_{j=0}^{I} \lambda_{ij} \Lambda_i(\mathbf{x}(t)) \quad \lambda_{ij} = -c_{ij}p_j, \quad i = 0, 1, \ldots, I \tag{2.3.4}$$

and $\Lambda_j(\mathbf{x}(t))$ is the likelihood ratio of H_j with respect to the null hypothesis. $\Lambda_j(\mathbf{x}(t))$ can be expressed as follows:

$$\Lambda_j(\mathbf{x}(t)) = \frac{p(\mathbf{x}(t)|H_j)}{p(\mathbf{x}(t)|H_0)}. \tag{2.3.5}$$

In this dissertation we call the constants λ_{ij} classifier weights. Strict application of the Bayesian rule requires values for each cost c_{ij} and each a priori probability p_j in order to compute the classifier weights λ_{ij}. Although specification of these values is not usually possible, assuming equal a priori probabilities may be reasonable in some applications. Furthermore, each cost c_{ij} may be set to zero if it is associated with a correct decision. It would not be reasonable to assume that all costs associated with incorrect decisions are equal. These difficulties become even greater as the number of possible hypotheses becomes larger. Nevertheless, the Bayesian decision rule given by equations (2.3.3) and (2.3.4)

gives the structure to design the classifier in function of likelihood ratios $\Lambda_i(\mathbf{x}(t))$. When applied to experimental data, this structure offers a method of determining the coefficients $\{\lambda_{ij}\}$ empirically (given subsequent constraints such as equal a priori probabilities for hypotheses and zero costs for correct decisions, for example). One possible approach is to [17]: (1) specify values of the probabilities of each type of error (except one); (2) determine from experimental data empirical operating characteristics curves by iterative variation of the coefficients $\{\lambda_{ij}\}$, selecting values corresponding to the assigned error probabilities.

A theoretical approach to determining the coefficients $\{\lambda_{ij}\}$ is to apply a generalization of the Neyman–Pearson Lemma [8,15,16] to the multiple hypotheses test directly instead of using the Bayes criterion. However, this approach leads to a decision rule equivalent to that of equations (2.3.3) and (2.3.4) except that the coefficients $\{\lambda_{ij}\}$ could, in principle, be expressed directly as functions of the specified error probabilities.

2.4 Derivation of the Likelihood Ratio

a. The "estimator–correlator" structure

One can show that the optimal array processor structure (references [12,13,18])can be implemented in an "estimator–correlator" structure (figure 2.2). In other words, the array measurements are directed into two branches, each of which is described by a matrix filter acting on the data. The matrix filters $\mathbf{Q}(\cdot,\cdot)$ and $\mathbf{G}(\cdot,\cdot)$ are found by solving the following integral equations:

$$\int_T \mathbf{K}_n(t_1,t_2)\,\mathbf{Q}(t_2,t_3)\,dt_2 = \delta(t_1-t_3)\,\mathbf{I} \qquad (2.4.1)$$

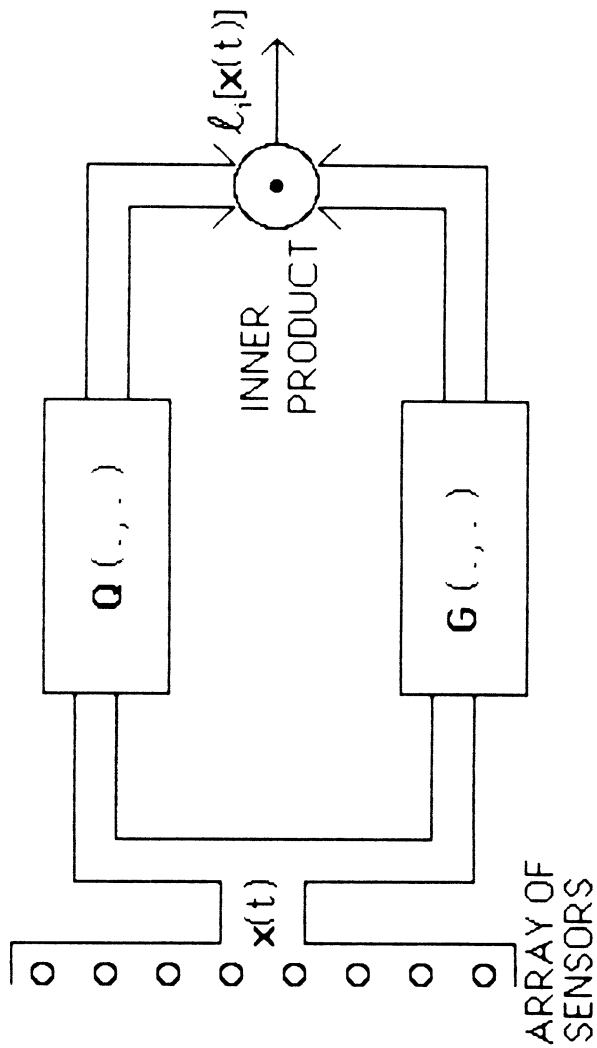


Figure 2.2: The Estimator-Correlator canonical structure.

$$\int_T \mathbf{K}_{x_i}(t_1,t_2)\, \mathbf{G}(t_2,t_3)\, dt_2 = \mathbf{K}_{s_i}(t_1,t_3) \qquad (2.4.2)$$

for i = 1,2, . . . , I, where $\mathbf{K}_n(\cdot,\cdot)$ is the covariance kernel of the measurement noise, $\mathbf{K}_{s_i}(\cdot,\cdot)$ is the covariance kernel of the returned signal alone, $\mathbf{K}_{x_i}(\cdot,\cdot)$ is the covariance kernel of the combined returned signal plus noise, T is the observation interval, and I is the identity matrix. The lower branch containing $\mathbf{G}(\cdot,\cdot)$ will be called the estimator branch: it calculates an optimal estimate of the backscattered signal based on the array data $\mathbf{x}(t)$. The upper branch will be referred to as the inverse filter branch, because it represents the inverse of $\mathbf{K}_n(\cdot,\cdot)$. The outputs of each branch are correlated to form the scalar–valued test statistic $\ell_i(\mathbf{x}(t))$ such that the likelihood ratio $\Lambda_i(\mathbf{x}(t))$ is expressed as [17]

$$\Lambda_i[\mathbf{x}(t)] = \gamma_i \exp(\ell_i[\mathbf{x}(t)]\,. \qquad (2.4.3)$$

Note that $\mathbf{x}(t)$ is a random process, γ_i is a positive constant and i is the hypothesis in question.

b. Application of the "estimator–correlator" to our model

In this section, we will implement the optimal detector by solving equations (2.4.1) and (2.4.2) for the model discussed in section (2.2). Using the sifting property, equation (2.4.1) has the obvious solution given by

$$\mathbf{Q}(t_2,t_3) = \frac{2}{N_0}\,\delta(t_2-t_3)\,\mathbf{I}\,. \qquad (2.4.4)$$

In order to find $\mathbf{G}(t_2,t_3)$, equations (2.2.8), (2.2.14) and (2.2.15) will be

used in equation (2.4.2). Let us assume that $G(t_2,t_3)$ can be expressed as follows:

$$G(t_2,t_3) = \mathbf{F}(t_2) \left[\frac{2}{N_0} \mathbf{H}_i\right] \mathbf{F}^H(t_3) \tag{2.4.5}$$

where $\mathbf{H}_i$ is a constant matrix to be found. Therefore, equation (2.4.2) becomes

$$\frac{2}{N_0} \int_T (\mathbf{F}(t_1)\, \mathbf{K}_{a_i} \mathbf{F}^H(t_2) + \frac{N_0}{2} \delta(t_1-t_2)\, \mathbf{I})\, \mathbf{F}(t_2)\, \mathbf{H}_i \mathbf{F}^H(t_3)\, dt_2$$

$$= \mathbf{F}(t_1)\, \mathbf{K}_{a_i} \mathbf{F}^H(t_3) \tag{2.4.6}$$

which can be rewritten as

$$\frac{2}{N_0} \mathbf{F}(t_1)\, \mathbf{K}_{a_i} \left[\int_T \mathbf{F}^H(t_2)\, \mathbf{F}(t_2)\, dt_2\right] \mathbf{H}_i\, \mathbf{F}^H(t_3)$$

$$+ \left[\int_T \delta(t_1-t_2)\, \mathbf{F}(t_2)\, dt_2\right] \mathbf{H}_i \mathbf{F}^H(t_3) = \mathbf{F}(t_1)\, \mathbf{K}_{a_i} \mathbf{F}^H(t_3)\,. \tag{2.4.7}$$

Let us define the signal correlation matrix Φ as

$$\Phi = \frac{1}{NE_f} \int_T \mathbf{F}^H(t_2)\, \mathbf{F}(t_2)\, dt_2 \tag{2.4.8}$$

where N is the total number of sensors in the array (figure 2.1) and E_f is the signal energy of duration T sec. E_f is expressed as

$$E_f = \int_T |f(t)|^2 \, dt \, . \tag{2.4.9}$$

Using the sifting property

$$\int_T \delta(t_1 - t_2) \, \mathbf{F}(t_2) \, dt_2 = \mathbf{F}(t_1) \, . \tag{2.4.10}$$

Equation (2.4.7) can be simplified using equations (2.4.8) and (2.4.10) such that

$$\mathbf{F}(t_1) \, \underline{\mathbf{K}}_{a_i} \left[\frac{2NE_f}{N_0}\right] \Phi \, \mathbf{H}_i \mathbf{F}^H(t_3) + \mathbf{F}(t_1) \, \mathbf{H}_i \mathbf{F}^H(t_3)$$

$$= \mathbf{F}(t_1) \, \mathbf{K}_{a_i} \mathbf{F}^H(t_3) \, . \tag{2.4.11}$$

Let us define the signal to noise ratio ρ_N as

$$\rho_N = \frac{2NE_f}{N_0} \tag{2.4.12}$$

and factor out (2.4.11), therefore,

$$\mathbf{F}(t_1) \, [\rho_N \, \mathbf{K}_{a_i} \, \Phi \, \mathbf{H}_i + \mathbf{H}_i] \, \mathbf{F}^H(t_3) = \mathbf{F}(t_1) \, \mathbf{K}_{a_i} \mathbf{F}^H(t_3) \tag{2.4.13}$$

which implies by analogy that

$$(\rho_N \, \mathbf{K}_{a_i} \Phi + \mathbf{I}) \, \mathbf{H}_i = \mathbf{K}_{a_i} \tag{2.4.14}$$

or

$$\mathbf{H}_i = (\mathbf{I} + \rho_N \mathbf{K}_{a_i} \Phi)^{-1} \mathbf{K}_{a_i} . \tag{2.4.15}$$

The estimator branch output in figure 2.2 is given by

$$\int_T \mathbf{G}(t_1,t_2)\, \mathbf{x}(t_2)\, dt_2 \tag{2.4.16}$$

which can be rewritten using (2.4.5) as

$$\frac{2}{N_0} \int_T \mathbf{F}(t_1)\, \mathbf{H}_i\, \mathbf{F}^H(t_2)\, \mathbf{x}(t_2)\, dt_2 \tag{2.4.17}$$

or

$$\frac{2}{N_0} \mathbf{F}(t_1)\, \mathbf{H}_i \int_T \mathbf{F}^H(t_2)\, \mathbf{x}(t_2)\, dt_2 . \tag{2.4.18}$$

Similarly the noise branch output is given by

$$\int_T \mathbf{Q}(t_1,t_2)\, \mathbf{x}(t_2)\, dt_2 = \frac{2}{N_0} \int_T \delta(t_1-t_2)\, \mathbf{x}(t_2)\, dt_2 = \frac{2}{N_0} \mathbf{x}(t_1) . \tag{2.4.19}$$

The outputs of each branch are correlated to form the scalar valued test statistic $\ell_i(\mathbf{x}(t))$. Using equations (2.4.18) and (2.4.19), $\ell_i(\mathbf{x}(t))$ is expressed as follows:

$$\ell_i(\mathbf{x}(t)) = \left[\frac{2}{N_0}\right]^2 \int_T \mathbf{x}^H(t_1)\, \mathbf{F}(t_1)\, dt_1\, \mathbf{H}_i \int_T \mathbf{F}^H(t_2)\, \mathbf{x}(t_2)\, dt_2 . \tag{2.4.20}$$

Equation (2.4.3) can be written using equation (2.4.20) in an alternative form as

follows:

$$\Lambda_i[\mathbf{x}(t)] = \gamma_i \exp(\mathbf{y}^H \mathbf{H}_i \mathbf{y}) \tag{2.4.21}$$

where $\mathbf{y}$ is expressed as

$$\mathbf{y} = \frac{2}{N_0} \int_T \mathbf{F}^H(t)\, \mathbf{x}(t)\, dt \tag{2.4.22}$$

is the output of a matrix matched filter operation.

2.5 Derivation of the Sufficient Statistic for the Ternary Problem

Throughout this dissertation, we address the ternary problem. There are three hypotheses; the null hypothesis H_0 and two alternative hypotheses H_1 and H_2. We shall assume that the null hypothesis has been eliminated [19] since our interest lies in choosing between H_1 and H_2 rather than detecting whether a target exists or not. The decision rule is based on minimizing the Bayesian risk described in section 2.3 and summarized by equations (2.3.3) and (2.3.4). First let $c_{ii} = 0$ and $p_j = p$ where $0 \leq p \leq 1$ and $i = 0, 1, 2$. In other words, it is assumed that each hypothesis is equally probable a priori and that the costs of correct decisions are zero. Moreover, it is assumed that the cost of deciding H_i given H_0 is independent of i; therefore, $c_{i0} = c_0$ where c_0 is a positive constant for $i = 1, 2$. With these assumptions, equation (2.3.4) becomes for the ternary case ($I = 2$):

$$E_0[\cdot] = \lambda_{01} \Lambda_1[\cdot] + \lambda_{02} \Lambda_2[\cdot] \tag{2.5.1a}$$

$$E_1[\cdot] = \lambda_{10} \Lambda_0[\cdot] + \lambda_{12} \Lambda_2[\cdot] \tag{2.5.1b}$$

$$E_2[\cdot] = \lambda_{20}\,\Lambda_0[\cdot] + \lambda_{21}\,\Lambda_1[\cdot] \qquad (2.5.1c)$$

where $\lambda_{10} = \lambda_{20} = -\,c_0 p$ and $\Lambda_0[\cdot] = 1$ (equation 2.3.5). In the above equations, $E_i[\cdot] > E_0[\cdot]$, i = 1,2 are tests for detection of the signals $\mathbf{s}_i(t)$. Equation (2.3.3) gives the following classifier rule:

$$\text{Choose } H_1 \text{ if } E_1[\cdot] > E_2[\cdot], \text{ otherwise, choose } H_2 .$$

By using equations (2.5.1b) and (2.5.1c) and remembering that λ_{12}, $\lambda_{21} < 0$, the classifier rule becomes:

$$\text{Choose } H_1 \text{ if } \frac{\Lambda_1[\cdot]}{\Lambda_2[\cdot]} > \frac{\lambda_{12}}{\lambda_{21}}, \text{ otherwise choose } H_2 .$$

Now using the expression of the likelihood ratio given in equation (2.4.21) in the above decision rule, we can write:

$$\text{Choose } H_1 \text{ if } \gamma_1 \exp(\mathbf{y}^H \mathbf{H}_1 \mathbf{y})/(\gamma_2 \exp(\mathbf{y}^H \mathbf{H}_2 \mathbf{y})) > \lambda_{12}/\lambda_{21} ,$$
$$\text{otherwise choose } H_2.$$

Taking the logarithm of the above equation, we can derive the following classifier rule:

$$\text{Choose } H_1 \text{ if } L_{12} > \eta, \text{ otherwise choose } H_2 ,$$

where

$$L_{12} = \mathbf{y}^H(\mathbf{H}_1 - \mathbf{H}_2)\mathbf{y} \qquad (2.5.2)$$

and

$$\eta = \mathrm{Ln}(\lambda_{12}\gamma_2/\lambda_{21}\gamma_1) \; . \qquad (2.5.3)$$

L_{12} is called the sufficient statistic for choosing between H_1 and H_2 while η is a threshold value to be optimized. $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{y}$ are given by equations (2.4.15) and (2.4.22).

Our discussion of the ternary problem is easily generalized to more hypotheses. Conceptually, choosing between H_1 and H_2 is no different than choosing between H_1 and H_3 or H_2 and H_3, in a four–hypotheses problem for example. In other words, the design parameters involved in a ternary classifier remain the same when more hypotheses are considered. From here stems the importance of the ternary problem and the subsequent chapters further justify our confinement.

Chapter 3

CLASSIFIER PERFORMANCE FOR HIGH RESOLUTION SIGNALS

3.1 Introduction

This chapter investigates classifier performance when targets are illuminated by high resolution signals capable of perfect resolution among individual test volume cells. Mathematically, this translates into $\Phi = \mathbf{I}$. Although in practice such signals are not realizable, they provide us with insight into the processor performance. Also, this assumption enables us to write closed form solutions for the classification and false alarm probabilities.

Section 3.2 summarizes fundamental results obtained from chapter 2. They will be needed in this chapter as well as in subsequent chapters. In section 3.3, we investigate classifier performance when a perfect statistical description of each object to be classified is available. This is known as matched processing and is optimal in that it defines the upper bound on performance. Section 3.4 examines the relationship between the performance of the classifier and the fidelity of a priori knowledge made available to it. In section 3.5 we display receiver operating characteristic (ROC) curves illustrating the classifier performance with respect to: (1) increased target strength, (2) mismatched processing, (3) signal–to–noise ratio, and (4) the number of cells in the test volume.

3.2 Summary of Fundamental Equations

This section presents a brief review of the fundamental equations of the problem. We also derive more relationships that will be used throughout this

dissertation. In chapter 2, we showed that the sufficient statistic for deciding between H_1 and H_2 (assuming H_0 has been eliminated) is given by

$$L_{12} = \mathbf{y}^H (\mathbf{H}_1 - \mathbf{H}_2)\, \mathbf{y} \tag{3.2.1}$$

where

$$\mathbf{y} = \frac{2}{N_0} \int_T \mathbf{F}^H(t)\, \mathbf{x}(t)\, dt \tag{3.2.2}$$

and

$$\mathbf{H}_i = (\mathbf{I} + \rho_N\, \mathbf{K}_{a_i} \Phi)^{-1}\, \mathbf{K}_{a_i} \tag{3.2.3}$$

for i = 1, 2. Classification is performed by comparing L_{12} to a threshold η. If L_{12} is larger than η, we assert that H_1 is correct; otherwise, we assert H_2 is correct. Note that ρ_N is the defined signal–to–noise ratio and is expressed in function of the noise level N_0, the number of sensors N and the signal energy E_f

$$\rho_N = \frac{2NE_f}{N_0} \tag{3.2.4}$$

where

$$E_f = \int_T |f(t)|^2\, dt\ . \tag{3.2.5}$$

Φ is the signal correlation matrix and is given by

$$\Phi = \frac{1}{NE_f} \int_T \mathbf{F}^H(t)\, \mathbf{F}(t)\, dt\ . \tag{3.2.6}$$

The ability of the classifier to distinguish between H_1 and H_2 depends on the probability density function of L_{12} conditioned on H_1 and H_2. The performance measures we shall use are the classification and false alarm probabilities. If H_2 is a target of interest and H_1 is a false target, then the classification probability is the probability that H_2 is chosen given H_2 is correct:

$$P(H_2|H_2) = \int_{-\infty}^{\eta} f(L_{12}|H_2)\, dL_{12} \tag{3.2.7}$$

where η is the classification threshold. The false alarm probability is:

$$P(H_2|H_1) = \int_{-\infty}^{\eta} f(L_{12}|H_1)\, dL_{12}\,. \tag{3.2.8}$$

Note that $f(L_{12}|H_i)$ is the probability density function of L_{12} conditioned on H_i. The density functions are found as follows. Since $\mathbf{y}$ (equation 3.2.2) is a linear function of the Gaussian process $\mathbf{x}(t)$, it too is normally distributed under either hypothesis. $\mathbf{y}$ is zero mean with covariance matrix

$$K_{y_i} = E[\mathbf{y}\mathbf{y}^H|H_i] \tag{3.2.9}$$

which implies

$$K_{y_i} = \left[\frac{2}{N_0}\right]^2 E\left[\int_T \int_T \mathbf{F}^H(t_1)\, \mathbf{x}(t_1)\, \mathbf{x}^H(t_2)\, \mathbf{F}(t_2)\, dt_1 dt_2\right] \tag{3.2.10}$$

or

$$\mathbf{K}_{y_i} = \left[\frac{2}{N_0}\right]^2 \int_T \int_T \mathbf{F}^H(t_1)\, \mathbf{K}_{x_i}(t_1,t_2)\, \mathbf{F}(t_2)\, dt_1 dt_2 \,. \tag{3.2.11}$$

Using equations (2.2.8) and (2.2.14), equation (3.2.11) becomes

$$\mathbf{K}_{y_i} = \left[\frac{2}{N_0}\right]^2 \int_T \int_T \mathbf{F}^H(t_1)\, \mathbf{F}(t_1)\, \mathbf{K}_{a_i}\, \mathbf{F}^H(t_2)\, \mathbf{F}(t_2)\, dt_1 dt_2$$

$$+ \left[\frac{2}{N_0}\right]^2 \int_T \int_T \mathbf{F}^H(t_1)\, \frac{N_0}{2}\, \delta(t_1-t_2)\, \mathbf{I}\, \mathbf{F}(t_2)\, dt_1 dt_2 \,. \tag{3.2.12}$$

Using the fact that $2/N_0 = \rho_N/(NE_f)$ (equation 3.2.4) and applying the sifting property, equation (3.2.12) becomes

$$\mathbf{K}_{y_i} = \rho_N^2\, \mathbf{\Phi}\, \mathbf{K}_{a_i}\, \mathbf{\Phi}^H + \rho_N \mathbf{\Phi} \quad (i = 1,2) \,. \tag{3.2.13}$$

One can easily verify that $\mathbf{\Phi}$ is a Hermitian matrix (equation 3.2.6). This property will be used in the remaining chapters.

Distributions of quadratic forms in complex multinormal random variables have been studied by Turin [20], Kharti [21], and Johnson and Kotz [22]. They showed that the characteristic function $M_i(\omega)$ of the quadratic form (3.2.1) is

$$M_i(\omega) = 1/\det[\mathbf{V}_i(\omega)] \tag{3.2.14}$$

where

$$\mathbf{V}_i(\omega) = \mathbf{I} - j\omega\, \mathbf{K}_{y_i}(\mathbf{H}_1 - \mathbf{H}_2) \,. \tag{3.2.15}$$

Then

$$f(L_{12}|H_i) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} M_i(\omega)\, e^{-j\omega L_{12}}\, d\omega \qquad (3.2.16)$$

for i = 1,2. This summarizes the fundamental equations that will be used extensively in the remaining discussion.

3.3 Evaluating Classifier Performance (Matched Processing)

a. Matched Processing

In section 3.3 we will study the performance of the "matched" classifier. Matched processing means that the classifier has perfect a priori knowledge of each object to be classified; in other words, the processing is exactly matched to the physical characteristics of the target and environment enclosed by the test volume.

b. Derivation of Closed Form Expressions

In this section we will derive closed form expressions for the classification and false alarm probabilities in function of the threshold η. The scenario of interest is depicted in figure 3.1, where a horizontal cut of the test volume is shown for both H_1 and H_2. The test volume is assumed to be identical for each hypothesis, and it comprises K cells with an independent scatterer in each cell. If under hypothesis H_1 no target is present and the scattered return is due to reverberation alone, all cells in the test volume will possess uniform scattering strength. Reverberation is returned energy from suspended solids in the medium. If, in addition to reverberation, a target is present under hypothesis H_2, the uniform scattering strength of a number of cells J ($J \leq K$) in the test volume may be replaced with new scattering strength. Therefore, it is assumed that each

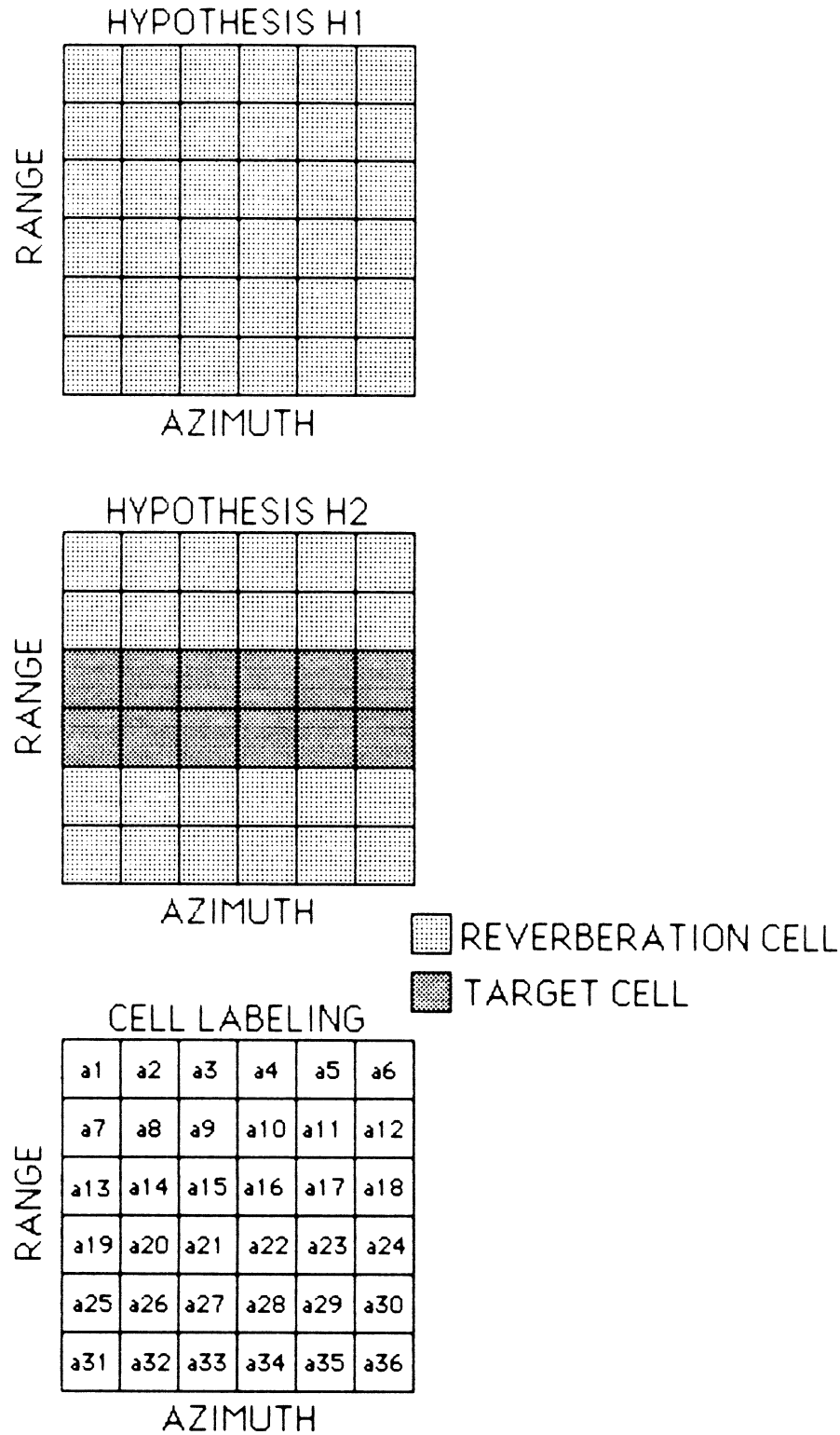


Figure 3.1: Example of a two-dimensional test volume.
K = 36 test cells and J = 12 target-like cells.

scatterer falls into one of two categories: target–like or reverberation. It is also assumed that reverberation scatterers are of equal strength under both H_1 and H_2. Hence, the a priori scattering matrices K_{a_i} (i = 1,2) are obtained as follows:

$$E[a_{im}\, a_{in}] = 0 \quad (m \neq n, \text{ uncorrelated scatterers}) \tag{3.3.1}$$

$$E[a_{ik}^2] = \begin{cases} \sigma_R^2 \text{ (reverberation cell)} \\ \sigma_0^2 \text{ (target–like cell)} \end{cases} \quad k = 1, \ldots, K \tag{3.3.2}$$

where a_{im} is the m^{th} component of the scattering vector $\mathbf{a}_i$ (i = 1,2). Equations (3.3.1) and (3.3.2) imply that $\mathbf{K}_{a_1}$ and $\mathbf{K}_{a_2}$ are diagonal and of full rank. We shall further assume that the waveform f(t) interrogating the test volume is capable of perfect resolution among cells. This reduces the signal correlation matrix to the identity matrix

$$\Phi = \mathbf{I}\,. \tag{3.3.3}$$

These assumptions diagonalize $\mathbf{V}_1(\omega)$ and $\mathbf{V}_2(\omega)$, making their determinant calculations easy.

Let us proceed with the calculations leading to solutions for $f(L_{12} \mid H_i)$, i = 1,2. Using the assumption, that $\Phi = \mathbf{I}$, equations (3.2.3) and (3.2.13) reduce to

$$\mathbf{H}_i = (\mathbf{I} + \rho_N \mathbf{K}_{a_i})^{-1}\, \mathbf{K}_{a_i} \tag{3.3.4}$$

and

$$\mathbf{K}_{y_i} = \rho_N^2 \mathbf{K}_{a_i} + \rho_N \mathbf{I} \tag{3.3.5}$$

for i = 1,2. Therefore, matrices $\mathbf{K}_{y_1}$ and $\mathbf{H}_1$ are diagonal and have identical elements each

$$[\mathbf{K}_{y_1}]_{k,k} = \rho_N (1 + \rho_N \sigma_R^2);\ k = 1 , \ldots , K \tag{3.3.6}$$

$$[\mathbf{H}_1]_{k,k} = \sigma_R^2/(1 + \rho_N \sigma_R^2);\ k = 1, \ldots , K . \tag{3.3.7}$$

A similar result for $\mathbf{H}_2$ is obtained. If there are J target cells, $\mathbf{H}_2$ is expressed as follows:

$$[\mathbf{H}_2]_{k,k} = \begin{cases} \sigma_R^2/(1 + \rho_N \sigma_0^2) & \text{(J target-like cells)} \\ \sigma_R^2/(1 + \rho_N \sigma_R^2) & \text{(K–J reverberation cells)} \end{cases} \quad k = 1, \ldots , K \tag{3.3.8}$$

Clearly, $\mathbf{H}_1$–$\mathbf{H}_2$ is diagonal and only J of its main diagonal elements, corresponding to cells with different scattering energies under H_1 and H_2, are non–zero. In other words

$$[\mathbf{H}_1–\mathbf{H}_2]_{k,k} = \begin{cases} \dfrac{\sigma_R^2-\sigma_0^2}{(1+\rho_N\sigma_0^2)(1+\rho_N\sigma_R^2)} & \text{(J target-like cells)} \\ 0 & \text{(K–J reverberation cells)} \end{cases} \quad k = 1, \ldots , K \tag{3.3.9}$$

Therefore, $\mathbf{K}_{y_1}(\mathbf{H}_1–\mathbf{H}_2)$ has J non–zero terms:

$$[\mathbf{K}_{y_1}(\mathbf{H}_1–\mathbf{H}_2)]_{k,k} = \begin{cases} \beta & \text{(J target-like cells)} \\ 0 & \text{(K–J reverberation cells)} \end{cases} \quad k = 1, \ldots , K \tag{3.3.10}$$

where

$$\beta = \frac{\rho_N(\sigma_R^2 - \sigma_0^2)}{1+\rho_N\sigma_0^2}. \tag{3.3.11}$$

This yields the result for $M_1(\omega)$ in function of β and J:

$$M_1(\omega) = \frac{1}{(1-j\omega\beta)^J}. \tag{3.3.12}$$

Then

$$f(L_{12}|H_1) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-j\omega L_{12}}}{(1-j\omega\beta)^J}\, d\omega. \tag{3.3.13}$$

The integral can be evaluated by means of residues (Appendix A). If $\beta < 0$ ($\sigma_R^2 < \sigma_0^2$), then

$$f(L_{12}|H_1) = -\frac{1}{\Gamma(J)}\frac{L_{12}^{J-1}}{\beta^J} e^{-\frac{L_{12}}{\beta}}, \quad L_{12} \le 0 \tag{3.3.14}$$

and it equals 0 for $L_{12} > 0$. Note that $\Gamma(J) = (J-1)!$

The probability density function $f(L_{12}|H_2)$ is obtained by similar means. Matrix $\mathbf{K}_{y_2}$ is diagonal and its diagonal elements are given by

$$[\mathbf{K}_{y_2}]_{k,k} = \begin{cases} \rho_N(1+\rho_N\sigma_0^2) & (J \text{ target-like cells}) \\ \rho_N(1+\rho_N\sigma_R^2) & (K-J \text{ reverberation cells}) \end{cases} \quad k = 1, \ldots, K \tag{3.3.15}$$

Next $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$ is evaluated and the result is a diagonal matrix whose elements are

$$[\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)]_{k,k} = \begin{cases} \alpha & \text{(J target-like cells)} \\ 0 & \text{(K–J reverberation cells)} \end{cases} \qquad k = 1, \ldots, K \tag{3.3.16}$$

where

$$\alpha = \frac{\rho_N(\sigma_R^2-\sigma_0^2)}{1+\rho_N\sigma_R^2}. \tag{3.3.17}$$

This yields the result for $M_2(\omega)$ in function of α and J:

$$M_2(\omega) = \frac{1}{(1-j\omega\alpha)^J}. \tag{3.3.18}$$

If $\alpha < 0$ ($\sigma_R^2 < \sigma_0^2$), then $f(L_{12}|H_2)$ is given by analogy with equation (3.3.14) by

$$f(L_{12}|H_2) = -\frac{1}{\Gamma(J)}\frac{L_{12}^{J-1}}{\alpha^J}e^{-\frac{L_{12}}{\alpha}}, \quad L_{12} \le 0 \tag{3.3.19}$$

and it equals 0 elsewhere. These results are used to compute the classification and false alarm probabilities as a function of the classification threshold η. Using equations (3.2.7) and (3.3.19):

$$P(H_2|H_2) = 1 - \int_\eta^0 f(L_{12}|H_2)\,dL_{12} = 1 + \frac{1}{\Gamma(J)}\int_\eta^0 \frac{L_{12}^{J-1}}{\alpha^J}e^{-\frac{L_{12}}{\alpha}}\,dL_{12}, \quad \eta \le 0. \tag{3.3.20}$$

Let $L_{12}/\alpha = u$, therefore

$$P(H_2|H_2) = 1 + \frac{1}{\Gamma(J)} \int_{\eta/\alpha}^{0} \frac{\alpha^{J-1} u^{J-1} e^{-u}}{\alpha^{J}} (\alpha\, du), \quad \eta \leq 0 \tag{3.3.21}$$

or

$$P(H_2|H_2) = 1 - \frac{1}{\Gamma(J)} \int_{0}^{\eta/\alpha} e^{-u} u^{J-1}\, du, \quad \eta \leq 0 \tag{3.3.22}$$

which can be written

$$P(H_2|H_2) = 1 - \gamma(J,\eta/\alpha)/\Gamma(J), \quad \eta \leq 0 \tag{3.3.23}$$

where [23]

$$\gamma(a,y) = \int_{0}^{y} e^{-t} t^{a-1}\, dt \tag{3.3.24a}$$

and

$$\Gamma(N) = (N-1)! \tag{3.3.24b}$$

are the incomplete and complete Gamma functions, respectively. By the same token, we can show that the false alarm probability is

$$P(H_2|H_1) = 1 - \gamma(J,\eta/\beta)/\Gamma(J), \quad \eta \leq 0\,. \tag{3.3.25}$$

This concludes the analysis of the matched classifier.

3.4 Evaluating Classifier Performance (Mismatched Processing)

a. Mismatched Processing

Our purpose in this section is to investigate the robustness of the

classifier when wrong a priori knowledge is processed. We will consider two types of mismatch: (1) object misorientation in which the only uncertainty is in the object orientation and (2) range misidentification in which the uncertainty lies in the object range.

b. Derivation of Closed Form Expressions

(1) Object Misorientation

The scenario of interest is shown in figure 3.2. Here, perfect a priori knowledge of H_1 is assumed: no object is present and energy returned is caused by reverberation only. The only uncertainty in H_2 is the object orientation. In other words, the processor assumes the object has a bow or stern aspect rather than the correct broadside aspect. We will also assume that the correct number of cells is used; in other words, the assumed length of the object is correct.

In order to study the performance of the mismatched classifier, we first construct an assumed scattering matrix $\mathbf{K}'_{a_2}$ which is used to construct $\mathbf{H}_2$. Note that $\mathbf{H}_1$ remains the same since hypothesis H_1 has not changed. Next we compute $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$. Here $\mathbf{K}_{y_2}$ is computed from the actual scattering matrix $\mathbf{K}_{a_2}$, because $\mathbf{y}$ comes directly from array measurements (equation 3.2.2).

Given these assumptions, the matrix $\mathbf{H}_1-\mathbf{H}_2$ is diagonal with J non–zero entries such that

$$[\mathbf{H}_1-\mathbf{H}_2]_{k,k} = \begin{cases} \dfrac{\sigma_R^2-\sigma_0^2}{(1+\rho_N\sigma_0^2)(1+\rho_N\sigma_R^2)} & \text{(J target-like cells)} \\ 0 & \text{(K–J reverberation cells)} \end{cases} \qquad k = 1, \ldots, K \tag{3.4.1}$$

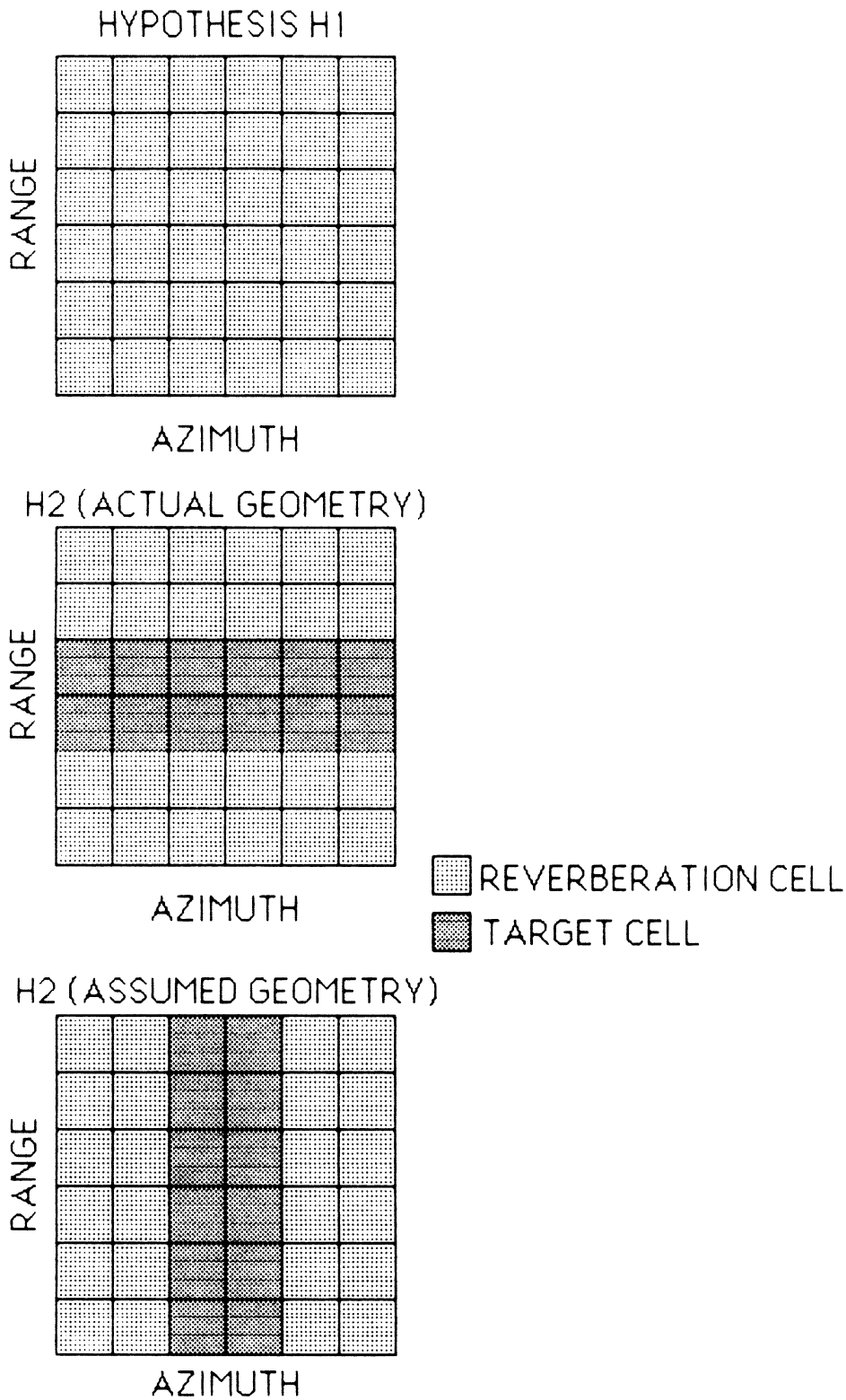


Figure 3.2: Example illustating object misorientation for K = 36 cells, J = 12 target cells and L = 4 intersection cells.

$\mathbf{K}_{y_1}$ is diagonal, and since all of its main diagonal elements are $\rho_N(1+\rho_N\sigma_R^2)$ (equation 3.3.6):

$$M_1(\omega) = \frac{1}{(1-j\omega\beta)^J} \tag{3.4.2}$$

where β is given in equation (3.3.11). Therefore, the false alarm probability is given by

$$P(H_2|H_1) = 1 - \gamma(J,\eta/\beta)/\Gamma(J) \, . \tag{3.4.3}$$

The expression for $M_2(\omega)$ is derived by examining how different amounts of overlap between the actual and assumed target geometries affect $\mathbf{V}_2(\omega)$. The first step is to find $\mathbf{K}_{y_2}$, which is obtained from the actual target geometry. It is diagonal and can be written as

$$[\mathbf{K}_{y_2}]_{k,k} = \begin{cases} \rho_N(1+\rho_N\sigma_R^2) & (J \text{ target-like cells}) \\ \rho_N(1+\rho_N\sigma_R^2) & (K-J \text{ reverberation cells}) \end{cases} \quad k = 1, \ldots, K \tag{3.4.4}$$

We should note here that the J target–like cells of equation (3.4.1) come from the assumed geometry while those of equation (3.4.4) come from the actual geometry of H_2. Bearing this in mind, the product $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$ is diagonal with J non–zero diagonal terms. If L is the number of intersection cells between the actual and the assumed geometries of H_2, then

$$[K_{y_2}(H_1-H_2)]_{k,k} = \begin{cases} \dfrac{\rho_N(\sigma_R^2-\sigma_0^2)}{1+\rho_N\sigma_R^2} & (L \text{ intersection cells}) \\ \dfrac{\rho_N(\sigma_R^2-\sigma_0^2)}{1+\rho_N\sigma_0^2} & (J-L \text{ non-intersection target-like cells}) \\ 0 & (K-J \text{ remaining cells}) \end{cases} \tag{3.4.5}$$

Using α and β defined in section 3.3 in the above equation, the characteristic function $M_2(\omega)$ can be written using (3.2.14) and (3.2.15) as

$$M_2(\omega) = \frac{1}{(1-j\omega\alpha)^L(1-j\omega\beta)^{J-L}}. \tag{3.4.6}$$

In order to obtain $f(L_{12}|H_2)$, it is necessary to evaluate

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{-j\omega L_{12}}}{(1-j\omega\alpha)^L(1-j\omega\beta)^{J-L}}\,d\omega \tag{3.4.7}$$

for α and β negative and $J \geq L$. The form above has been solved in appendix B. Substituting N by L, M by J–L and x by L_{12} in (B.1) and (B.19), the result for $\sigma_R^2 < \sigma_0^2$ ($\alpha, \beta < 0$) is

$$f(L_{12}|H_2) = f_1(L_{12}|H_2) + f_2(L_{12}|H_2) \tag{3.4.8}$$

where

$$f_1(L_{12}|H_2) = -\frac{\alpha^{J-2L}L_{12}^{L-1}}{\Gamma(L)(\alpha-\beta)^{J-L}} \sum_{n=0}^{L-1} \begin{bmatrix} L-1 \\ n \end{bmatrix} \frac{\Gamma(J-L+n)}{\Gamma(J-L)} \frac{\alpha^n\beta^n}{(\beta-\alpha)^n} \frac{e^{-L_{12}/\alpha}}{L_{12}^n} \tag{3.4.9}$$

and

$$f_2(L_{12}|H_2) = -\frac{\beta^{2L-J}\, L_{12}^{J-L-1}}{\Gamma(J-L)(\beta-\alpha)^L} \sum_{m=0}^{J-L-1} \left[\begin{matrix} J-L-1 \\ m \end{matrix} \right] \frac{\Gamma(L+m)}{\Gamma(L)} \frac{\alpha^m \beta^m}{(\alpha-\beta)^m} \frac{e^{-L_{12}/\beta}}{L_{12}^m} \tag{3.4.10}$$

for $L_{12} \leq 0$. Note that $f(L_{12}|H_2) = 0$ for $L_{12} > 0$. From equation 3.2.7, we can write

$$P(H_2|H_2) = 1 - \int_{\eta}^{0} f(L_{12}|H_2)\, dL_{12}\,, \quad \eta \leq 0 \tag{3.4.11}$$

or

$$P(H_2|H_2) = 1 + (I_1+I_2)\,, \quad \eta \leq 0 \tag{3.4.12}$$

where

$$I_1 = -\int_{\eta}^{0} f_1(L_{12}|H_2)\, dL_{12} \tag{3.4.13}$$

and

$$I_2 = -\int_{\eta}^{0} f_2(L_{12}|H_2)\, dL_{12}\,. \tag{3.4.14}$$

In order to solve for I_1, we need to solve for the following integral (refer to equation 3.4.9)

$$\int_{\eta}^{0} \frac{L_{12}^{L-1}\, e^{-L_{12}/\alpha}}{L_{12}^n}\, dL_{12} = \int_{\eta}^{0} L_{12}^{L-n-1} e^{-L_{12}/\alpha}\, dL_{12}\,. \tag{3.4.15}$$

Letting $L_{12}/\alpha = u$,

$$L_{12}^{L-n-1}\, e^{-L_{12}/\alpha}\, dL_{12} = \alpha^{L-n-1} u^{L-n-1} e^{-u}\, (\alpha du)\,. \tag{3.4.16}$$

Therefore,

$$\int_{\eta}^{0} \frac{L_{12}^{L-1}\, e^{-L_{12}/\alpha}}{L_{12}^{n}}\, dL_{12} = \alpha^{L-n} \int_{\eta/\alpha}^{0} u^{L-n-1} e^{-u}\, du$$

$$= -\alpha^{L-n}\, \gamma(L-n,\eta/\alpha)\ . \tag{3.4.17}$$

We can now solve for I_1:

$$I_1 = -\frac{\alpha^{J-2L}}{\Gamma(L)(\alpha-\beta)^{J-L}} \sum_{n=0}^{L-1} \binom{L-1}{n} \frac{\Gamma(J-L+n)}{\Gamma(J-L)} \frac{\alpha^n \beta^n}{(\beta-\alpha)^n}\, \alpha^{L-n}\, \gamma(L-n,\eta/\alpha) \tag{3.4.18}$$

$$I_1 = -\frac{\alpha^{J-L}}{(\alpha-\beta)^{J-L}} \sum_{n=0}^{L-1} \left[\binom{L-1}{n} \frac{\Gamma(J-L+n)}{\Gamma(J-L)\,\Gamma(L)} \right] \frac{\beta^n}{(\beta-\alpha)^n}\, \gamma(L-n,\eta/\alpha)\ . \tag{3.4.19}$$

Note that the term between brackets in (3.4.19) becomes

$$\frac{(L-1)!}{(L-1-n)!\,n!} \frac{\Gamma(J-L+n)}{\Gamma(J-L)\,\Gamma(L)} = \frac{1}{(L-1-n)!} \frac{\Gamma(J-L+n)}{\Gamma(J-L)\,n!} \tag{3.4.20}$$

Therefore,

$$I_1 = \frac{-\alpha^{J-L}}{(\alpha-\beta)^{J-L}} \sum_{n=0}^{L-1} \frac{F(1-J+L,-n,1,1)}{(L-1-n)!} \frac{\beta^n}{(\beta-\alpha)^n}\, \eta(L-n,\eta/\alpha) \tag{3.4.21}$$

where $F(\cdot,\cdot,\cdot,\cdot)$ is the hypergeometric function [23] and is defined as follows

$$F(a,b,c,1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}. \tag{3.4.22}$$

In order to evaluate I_2, we will follow a similar approach. From (3.4.10) and (3.4.14), we can write

$$\int_{\eta}^{0} \frac{L_{12}^{J-L-1}\, e^{-L_{12}/\beta}}{L_{12}^{m}}\, dL_{12} = -\beta^{J-L-m}\, \gamma(J-L-m,\eta/\beta) \tag{3.4.23}$$

which can be obtained by analogy with equation (3.4.17). We can now solve for I_2:

$$I_2 = \frac{-\beta^{2L-J}}{\Gamma(J-L)(\beta-\alpha)^{L}} \sum_{m=0}^{J-L-1} \begin{bmatrix} J-L-1 \\ m \end{bmatrix} \frac{\Gamma(L+m)}{\Gamma(L)} \frac{\alpha^m \beta^m}{(\alpha-\beta)^m}\, \beta^{J-L-m}\, \gamma(J-L-m,\eta/\beta) \tag{3.4.24}$$

$$I_2 = \frac{-\beta^{L}}{(\beta-\alpha)^{L}} \sum_{m=0}^{J-L-1} \left[\begin{bmatrix} J-L-1 \\ m \end{bmatrix} \frac{\Gamma(L+m)}{\Gamma(J-L)\,\Gamma(L)} \right] \frac{\alpha^m}{(\alpha-\beta)^m}\, \gamma(J-L-m,\eta/\beta). \tag{3.4.25}$$

The term between brackets is written as

$$\frac{(J-L-1)!}{(J-L-m-1)!m!} \frac{\Gamma(L+m)}{\Gamma(J-L)\,\Gamma(L)} = \frac{1}{(J-L-m-1)!} \frac{\Gamma(L+m)}{\Gamma(L)\,m!}. \tag{3.4.26}$$

Therefore,

$$I_2 = \frac{-\beta^L}{(\beta-\alpha)^L} \sum_{m=0}^{J-L-1} \frac{F(1-L,-m,1,1)}{(J-L-m-1)!} \frac{\alpha^m}{(\alpha-\beta)^m} \gamma(J-L-m,\eta/\beta) \tag{3.4.27}$$

where $F(\cdot,\cdot,\cdot,\cdot)$ is defined in (3.4.22). As a result the classification probability is given by equation (3.4.12), (3.4.21) and (3.4.27) while the false alarm probability is given by equation (3.4.3) for α, β and η negative.

(2) Range Misidentification

The scenario of interest is shown in figure 3.3. We assume that perfect a priori knowledge of H_1 is known and the only uncertainty is in the target range under hypothesis H_2. The analysis of the processor performance can be repeated as in part (1). However, a careful look at the analysis in part (1) allows us to directly evaluate the classification and false alarm probabilities. For instance, since hypothesis H_1 remains the same, then $M_1(\omega)$ is given as in (3.4.2) by

$$M_1(\omega) = \frac{1}{(1-j\omega\beta)^J}. \tag{3.4.28}$$

In order to evaluate $M_2(\omega)$, we can consider the range mismatch as an "object misorientation" with no intersection cells (L=0). Setting L = 0 in equation (3.4.6), we get

$$M_2(\omega) = \frac{1}{(1-j\omega\beta)^J}. \tag{3.4.29}$$

Therefore, $M_1(\omega) = M_2(\omega)$ and we can deduce that the classification probability $P(H_2|H_2)$ equals the false alarm probability $P(H_2|H_1)$. It turns out that when the range is misestimated, and there is no overlap between the actual and

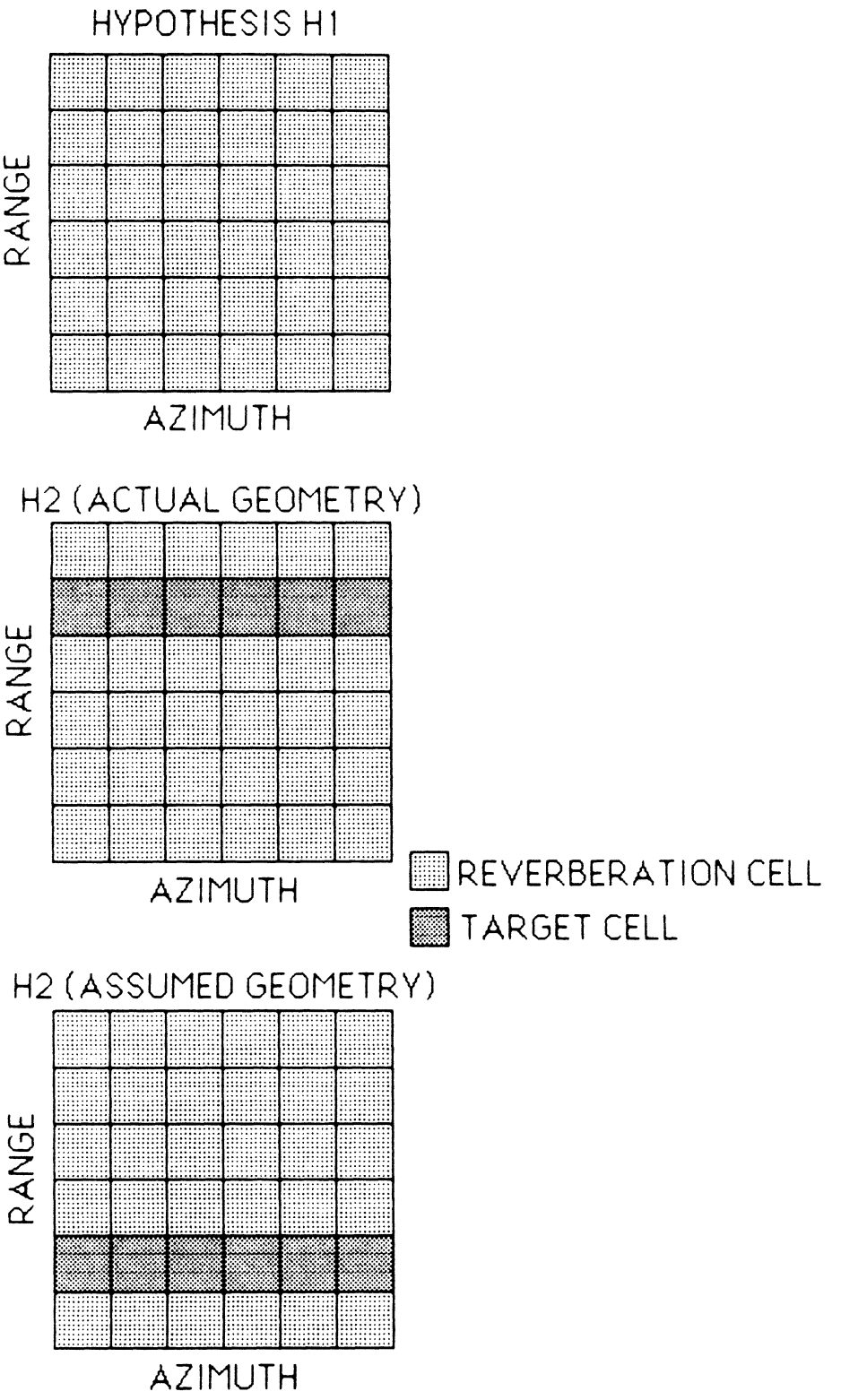


Figure 3.3: Example illustating range misorientation for K = 36 cells, and J = 6 target cells.

assumed test volume geometries (under H_2), the processor is unable to distinguish between H_1 and H_2 at all. This conclusion can be reached with an intuitive argument. We have presupposed perfect cell resolution; therefore, when an incorrect range estimate is infused into the classifier, only return from reverberation cells in processed. But the only statistical difference between H_1 and H_2 is the difference between total scattering strengths. When reverberation energy along is processed, it is not possible to distinguish between target–like and reverberation–like energy; consequently, classification is not possible.

3.5 Simulation Results and Conclusions

In order to evaluate the classification performance, we shall display the receiver operating characteristic (ROC) curves. A ROC curve is a plot of the classification probability $P(H_2|H_2)$ versus the false alarm probability $P(H_2|H_1)$. $P(H_2|H_2)$ and $P(H_2|H_1)$ are computed in function of the classification threshold η. If, for instance, $P(H_2|H_2)$ and $P(H_2|H_1)$ are equal, the processor cannot distinguish between H_1 and H_2 at all and one might as well "toss a coin" to classify the target. A good processor performance translates into a high classification probability and a low false alarm probability for a certain value of the classification threshold η. ROC curves provide the design engineer in this field with a criterion to select the classification threshold η with respect to some parameters.

The equations derived in this chapter were evaluated numerically on a personal computer (IBM PS/2) using Turbo Pascal programming language. Figure 3.4 displays ROC curves for the matched processor using equations (3.3.23) and (3.3.25). These equations were derived for α, β, η negative. In this example, we set the signal–to–noise ratio ρ_N to unity. Next, we set the

reverberation strength per cell σ_R^2 to 10 while we vary the target strength per cell from 20 to 50 by steps of 10. We remind the reader that the geometry in question is shown in figure 3.1. It is clear from figure 3.4 that as the target to reverberation contrast in the test volume increases, the processor performance is improved. In other words, the performance of the classifier degrades tremendously as the reverberation increases. We draw the first conclusion: the classifier performance is a direct function of the reverberation noise. Good performance occurs when reverberation noise is small.

Next, we examine the classifier performance when the wrong scattering covariance matrix is infused into the processor. The closed form expressions for the ROC curves of figure 3.5 are shown in equations (3.4.3) and (3.4.12), (3.4.21), (3.4.27). Note that the geometry of interest is shown in figure 3.2. In this example, the number of target–like cells J and the signal–to–noise ratio ρ_N are set to 10 and 1, respectively. Also, we set σ_R^2 to 10 and σ_0^2 to 20. The only parameter which was varied is L, the number of overlapping target–like cells between the assumed and the actual test volume geometries (under H_2). Clearly, classification performance falls off as the degree of "mismatch" gets worse. Therefore, we can draw the second conclusion: the ability of the classifier to distinguish between objects decreases as the difference between actual and assumed test volume geometries increases. The performance of the mismatched classifier is directly proportional to the number and energy of the target like cells which overlap between the assumed and actual geometries. If an incorrect range is assumed and there is no overlap between target–like cells, the classifier cannot distinguish between H_1 and H_2 at all.

Third, we investigate the classifier performance with respect to the signal–to–noise ratio ρ_N. In this example, we consider the matched processor

case (figure 3.1) for J = 4 cells, $\sigma_R^2 = 10$ and $\sigma_0^2 = 30$. The curves in figure 3.6 display ROC curves for $\rho_N = 1.0, 0.25, 0.1$ and 0.01. Experimentation has shown that $\rho_N = 1.0$ (0 dB) is a high signal–to–ratio in our model. A higher signal–to–noise ratio would not improve the performance by much. Also, the graph shows that a signal–to–noise ratio of 0.01 (–20 dB) degrades the performance tremendously. A signal–to–noise ratio of 0.25 or 0.1 (– 6 or –10 dB) provides an acceptable performance. It is up to the application engineer to specify the signal–to–noise ratio taking into consideration the performance required, other parameters (i.e., J, σ_R^2, σ_0^2) and the cost to be paid with high SNR.

Finally, we demonstrate the relationship between performance and number of cells. In this example, the test volume geometry does not change (figure 3.1); however, the cell resolution does. In other words, each cell is subdivided into two cells such that the sum of their scattering strength has the same value as the scattering strength of the "root" cell. Figure 3.7 displays the classifier performance for a small target to reverberation contrast ($\sigma_R^2 = 10$, $\sigma_0^2 = 15$) when J = 4 (lower curve), J = 8 (upper curve) and J = 16 (middle curve). This is an example in which increasing the cell resolution degrades the classifier performance. This effect is not as pronounced for a higher target to reverberation contrast (figure 3.8). Here, performance increases as the number of cells increases. This is expected up to a point. It turns out that beyond an optimum number, performance begins to fall off as J is increased. The reason for this is because more white noise is processed along with the test volume return, and eventually the signal–to–noise ratio begins to fall off.

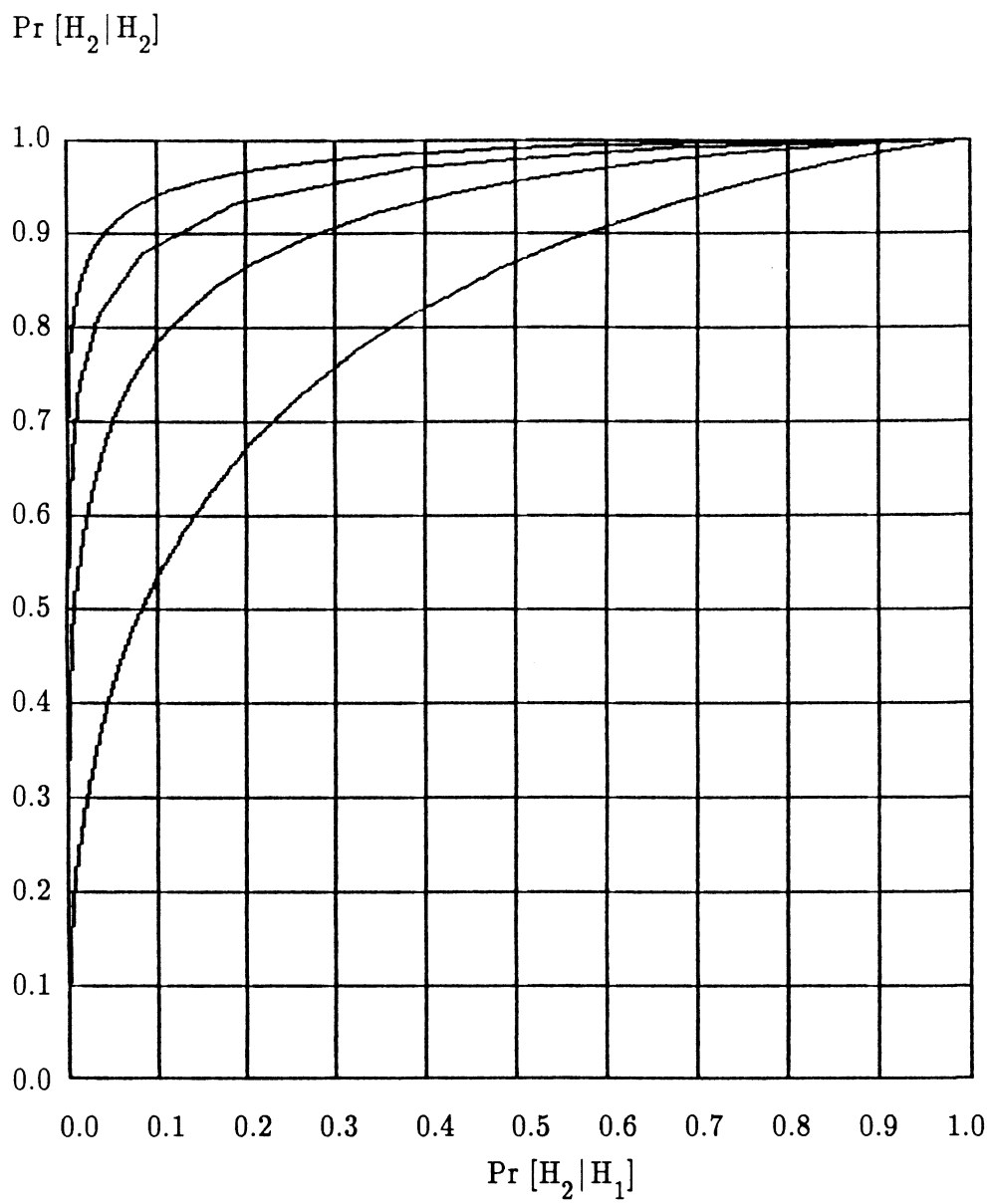


Figure 3.4: ROC curves for the matched classifier. Here J = 4 target cells, $\rho_N = 1.0$ and $\sigma_R^2 = 10$. σ_0^2 varies from 20 (lowest curve) to 50 (highest curve) by steps of 10.

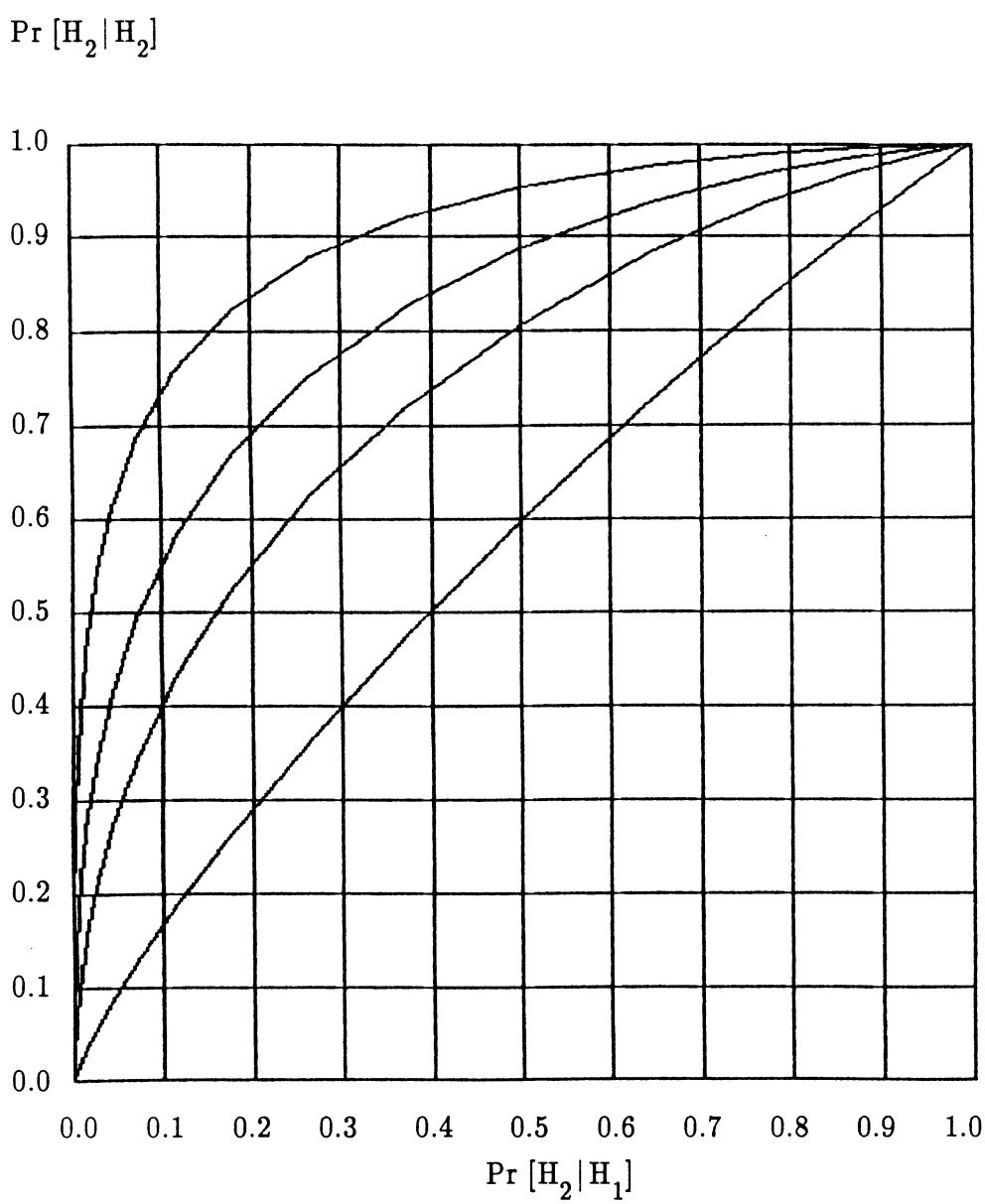


Figure 3.5 ROC curves for the mismatched classifier. Here J = 10 target cells, $\rho_N = 1.0$, $\sigma_R^2 = 10$ and $\sigma_0^2 = 20$. L takes the values 1, 4, 6 and 9 (lowest to highest curve, respectively).

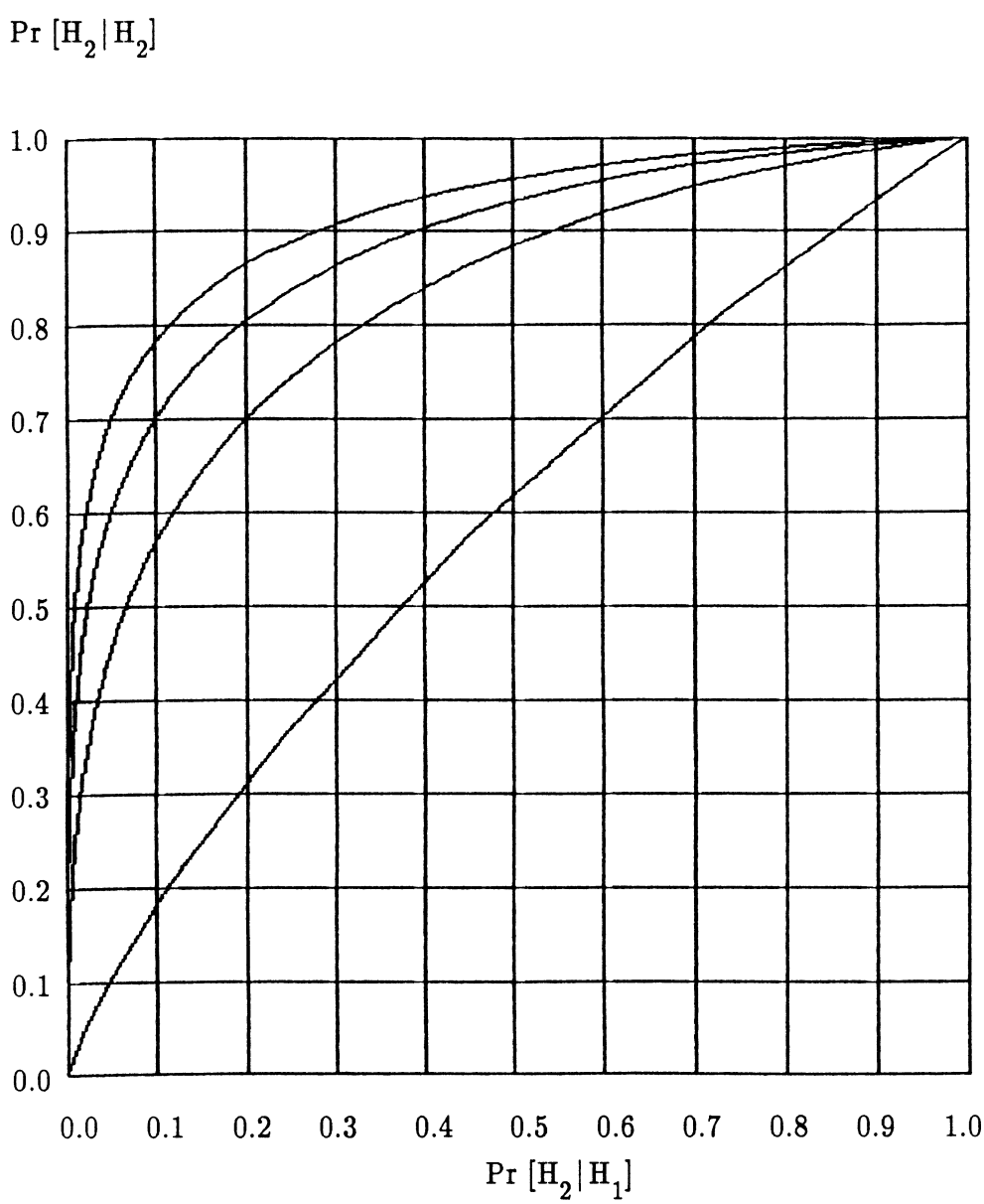


Figure 3.6 ROC curves displaying performance versus signal–to–noise ratio for the matched processor case. Here J = 4 cells, $\sigma_R^2 = 10$ and $\sigma_0^2 = 30$. ρ_N takes the values 1.0, 0.25, 0.1 and 0.01 (highest to lowest curve, respectively).

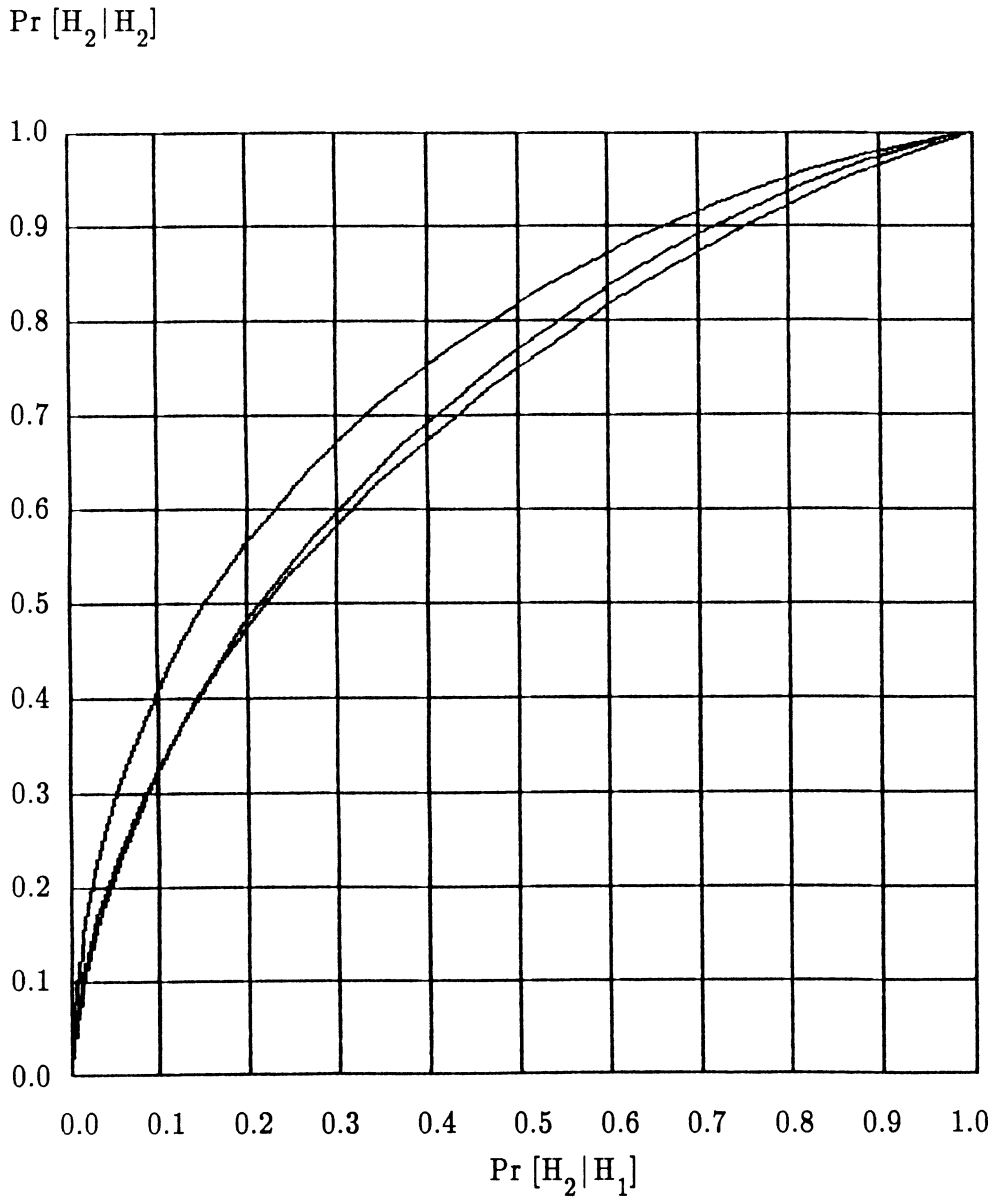


Figure 3.7 ROC curves displaying performance versus the number of cells in the test volume for $\rho_N = 1.0$. The triplet (J, σ_R^2, σ_0^2) takes the following values: a) lowest curve (4, 10, 15), b) highest curve (8, 5, 7.5), and c) middle curve (16, 2.5, 3.25).

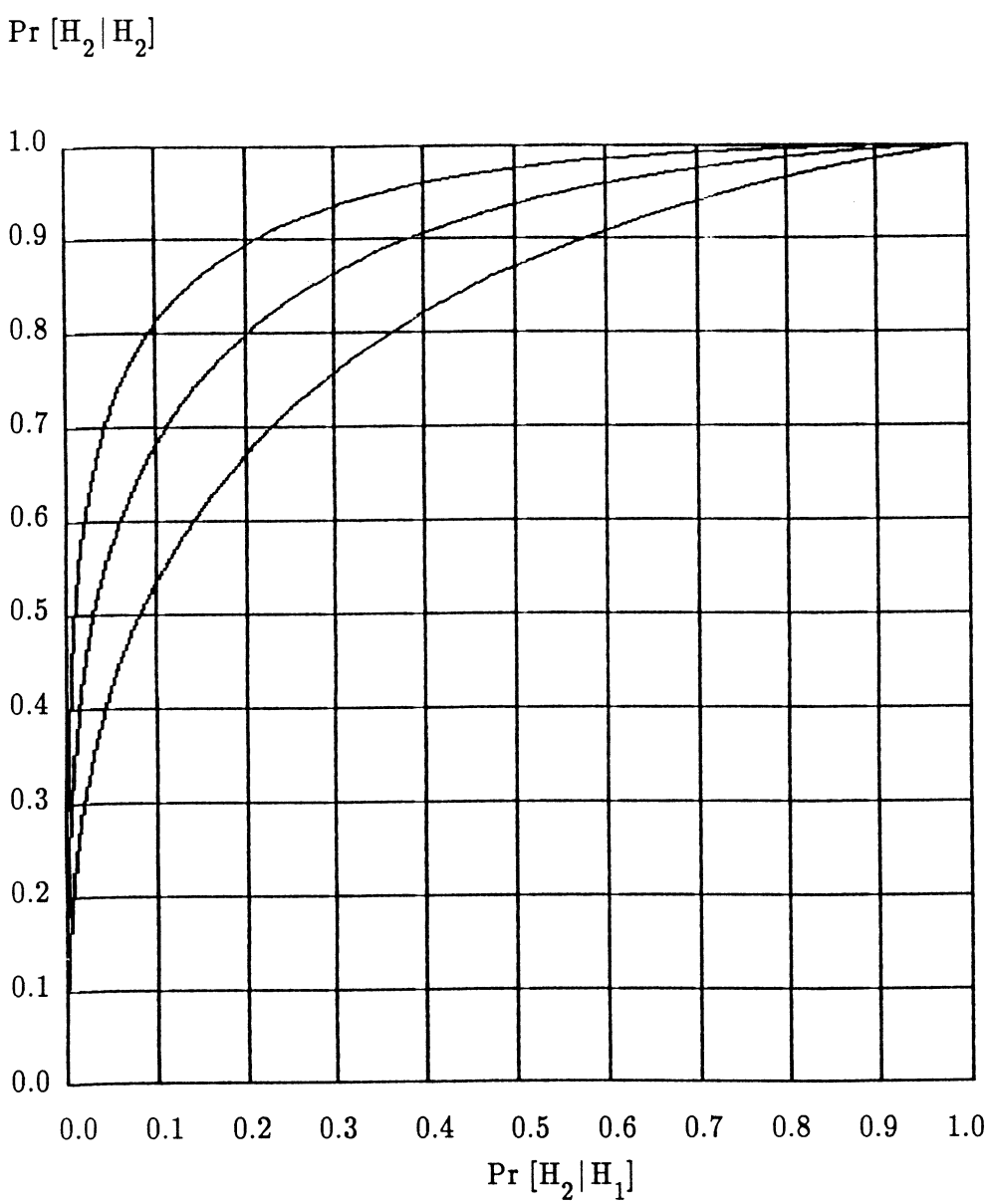


Figure 3.8 ROC curves displaying performance versus the number of cells in the test volume for $\rho_N = 1.0$. The triplet (J, σ_R^2, σ_0^2) takes the following values: a) lowest curve (4, 10, 20), b) middle curve (8, 5, 10), and c) highest curve (16, 2.5, 5).

Chapter 4

CLASSIFIER PERFORMANCE FOR LINEAR FM SIGNALS

4.1 Introduction

In this chapter, we investigate the classifier performance when targets are illuminated by linear FM pulses of the form

$$f(t) = A \cos (\omega_0 t + \pi k_f t^2) , \ 0 < t < T \qquad (4.1.1)$$

where k_f is the transmitter constant. If k_f is set to zero, f(t) describes a single frequency pulse of duration T seconds. In this case the signal correlation matrix Φ will no longer be the identity matrix and closed form expressions for the false alarm and classification probabilities will be expressed in terms of the eigenvalues of $\mathbf{K}_{y_i}(\mathbf{H}_1-\mathbf{H}_2)$, for i = 1, 2.

In section 4.2, we develop these expressions, starting with characteristic functions. The need to diagonalize $\mathbf{K}_{y_i}(\mathbf{H}_1-\mathbf{H}_2)$ for i = 1, 2 turns out to be an important concern in evaluating characteristic functions. Section 4.3 discusses how the signal correlation matrix Φ is computed for a real situation in which we assume a two–dimensional test region and a linear array of eight elements. In the process, we evaluate the time delays needed to compute Φ. In section 4.4, we display (ROC) curves illustrating the classifier performance with respect to: (1) transmitter constant k_f, (2) target strength, (3) signal–to–noise ratio, (4) angular width, (5) range misidentification, and (6) target misorientation.

4.2 Evaluating Classifier Performance ($\Phi \neq I$)

a. Diagonalization of $\mathbf{K}_{y_i}(\mathbf{H}_1-\mathbf{H}_2)$

In section 3.2 we summarized the fundamental equations needed to evaluate the classifier performance. The probability density function $f(L_{12}|H_i)$ of the test statistic L_{12} conditioned on H_i is obtained from the characteristic function $M_i(\omega)$ expressed as

$$M_i(\omega) = 1/\det[\mathbf{V}_i(\omega)] \tag{4.2.1}$$

where

$$\mathbf{V}_i(\omega) = \mathbf{I} - j\omega\, \mathbf{K}_{y_i}(\mathbf{H}_1-\mathbf{H}_2) \tag{4.2.2}$$

for i = 1, 2. $\mathbf{K}_{y_i}$ and $\mathbf{H}_i$ are given by

$$\mathbf{H}_i = (\mathbf{I} + \rho_N \mathbf{K}'_{a_i}\, \Phi)^{-1}\, \mathbf{K}'_{a_i} \tag{4.2.3}$$

$$\mathbf{K}_{y_i} = \rho_N^2\, \Phi\, \mathbf{K}_{a_i} \Phi^H + \rho_N\, \Phi \tag{4.2.4}$$

for i = 1, 2. Note that the scattering covariance matrix in equation (4.2.3) has been labeled $\mathbf{K}'_{a_i}$. This is due to the fact that $\mathbf{H}_1$ and $\mathbf{H}_2$ are computed by the processor and the a priori covariance matrix $\mathbf{K}'_{a_i}$ is usually mismatched to the actual test volume geometry. $\mathbf{K}_{y_i}$ is obtained from the output $\mathbf{y}$ of a matrix matched filter and is a direct function of the array measurement. Therefore, $\mathbf{K}_{a_i}$ in equation (4.2.4) perfectly describes the statistical scattering properties of

the test volume under hypothesis H_i. In the computer simulation, we shall treat this problem in the general case where the processor assumes wrong a priori information. If the a priori information matches exactly the test volume geometry, we will set $\mathbf{K}'_{a_i} = \mathbf{K}_{a_i}$.

It is now clear that the characteristic function $M_i(\omega)$ cannot be expressed analytically, since Φ is no longer a diagonal matrix. One approach to get around this problem is to diagonalize $\mathbf{K}_{y_i}(\mathbf{H}_1-\mathbf{H}_2)$ for i = 1, 2. Let, for instance, T be the matrix of eigenvectors of $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$ assuming for the moment that the eigenvalues are real and distinct. Therefore,

$$\mathbf{T}^{-1}\,\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)\mathbf{T} = \Lambda \tag{4.2.5}$$

where Λ is the K×K diagonal matrix of eigenvalues of $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$. Pre–multiplying and postmultiplying the above equation by T and $\mathbf{T}^{-1}$, respectively, we get

$$\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2) = \mathbf{T}\,\Lambda\,\mathbf{T}^{-1}\,. \tag{4.2.6}$$

Using this result and substituting I in (4.2.2) by $\mathbf{T}\,\mathbf{I}\,\mathbf{T}^{-1}$, $\mathbf{V}_2(j\omega)$ is expressed as

$$\mathbf{V}_2(j\omega) = \mathbf{T}\,\mathbf{I}\,\mathbf{T}^{-1} - j\omega\,\mathbf{T}\,\Lambda\,\mathbf{T}^{-1} \tag{4.2.7}$$

or

$$\mathbf{V}_2(j\omega) = \mathbf{T}(\mathbf{I} - j\omega\,\Lambda)\mathbf{T}^{-1}\,. \tag{4.2.8}$$

Therefore,

$$M_2(j\omega) = 1/\det[\mathbf{T}(\mathbf{I}-j\omega\,\Lambda)\mathbf{T}^{-1}]\,. \tag{4.2.9}$$

Recalling the following determinant properties

$$\det(\mathbf{A}\,\mathbf{B}\,\mathbf{C}) = \det(\mathbf{A})\det(\mathbf{B})\det(\mathbf{C}) \tag{4.2.10}$$

and

$$\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A}) \tag{4.2.11}$$

$M_2(j\omega)$ is expressed as

$$M_2(j\omega) = 1/\det(\mathbf{I}-j\omega\,\Lambda) \tag{4.2.12}$$

which is equivalent to

$$M_2(\omega) = \prod_{k=1}^{K}\left[\frac{1}{1-j\omega\lambda_k}\right] \tag{4.2.13}$$

where $\{\lambda_k\}$, $k = 1, \ldots, K$ is the set of eigenvalues of $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$ and K is again the number of cells in the test volume. A similar approach is used to diagonalize $\mathbf{K}_{y_1}(\mathbf{H}_1-\mathbf{H}_2)$. Hence, the characteristic function $M_1(\omega)$ is expressed as

$$M_1(\omega) = \prod_{k=1}^{K}\left[\frac{1}{1-j\omega\lambda'_k}\right] \tag{4.2.14}$$

where $\{\lambda'_k\}$, $k = 1, \ldots, K$ is the set of eigenvalues of $\mathbf{K}_{y_1}(\mathbf{H}_1-\mathbf{H}_2)$.

Next, we discuss the sign and type of these eigenvalues. In appendix C, we showed that $\mathbf{H}_1$ and $\mathbf{H}_2$ are Hermitian and that $\mathbf{K}_{y_1}$ and $\mathbf{K}_{y_2}$ are positive definite matrices. This implies that $\mathbf{H}_1-\mathbf{H}_2$ is Hermitian and has real eigenvalues. Now let $\mathbf{A} = \mathbf{H}_1-\mathbf{H}_2$ and $\mathbf{B} = \mathbf{K}_{y_1}^{-1}$ (i = 1, 2). It is stated in reference [25] that if $\mathbf{B}$ is positive definite, which is the case here since the inverse of a positive definite matrix is positive definite as well, then the problem $\mathbf{Ax} = \lambda\ \mathbf{Bx}$ has eigenvalues of the same sign as the conventional $\mathbf{Ax} = \lambda\mathbf{x}$. In other words, $\mathbf{B}^{-1}\mathbf{A} = \mathbf{K}_{y_i}(\mathbf{H}_1-\mathbf{H}_2)$ has eigenvalues of the same sign as $\mathbf{H}_1-\mathbf{H}_2$. Since $\mathbf{H}_1-\mathbf{H}_2$ is Hermitian but not necessarily positive or negative definite, the eigenvalues $\{\lambda_k'\}$ and $\{\lambda_k\}$ (k = 1, . . . , K) of $\mathbf{K}_{y_1}(\mathbf{H}_1-\mathbf{H}_2)$ and $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$ are real.

b. Classification and False Alarm Probabilities

In this section, we will derive expressions for the classification and false alarm probabilities in terms of the classification threshold η and the eigenvalues $\{\lambda_k\}$ and $\{\lambda_k'\}$ discussed in part a. We will not confine our discussion to a specific geometry for hypotheses H_1 and H_2. Nevertheless, we still assume that the test volume is identical (same number of cells) under both hypotheses. Also, each cell contains at most one scatterer. These scatterers are uncorrelated and follow equations (3.3.1) and (3.3.2).

Let us now evaluate the classification probability from the characteristic function $M_2(\omega)$. First, we find the probability density function $f(L_{12}|H_2)$:

$$f(L_{12}|H_2) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \prod_{k=1}^{K} \left[\frac{1}{1-j\omega\lambda_k}\right] e^{-j\omega L_{12}}\, d\omega \tag{4.2.15}$$

or

$$f(L_{12}|H_2) = \frac{1}{2\pi}\left[\prod_{k=1}^{K}\left[\frac{1}{-j\lambda_k}\right]\right]\int_{-\infty}^{\infty}\prod_{k=1}^{K}\left[\frac{1}{\omega+j/\lambda_k}\right] e^{-j\omega L_{12}}\, d\omega\,. \tag{4.2.16}$$

The integral in the above equation can be written as:

$$\int_C f(z)\, dz = \int_{-\infty}^{\infty} f(z)\, dz + \int_{C_R} f(z)\, dz \tag{4.2.17}$$

where C is the counterclockwise contour (figure 4.1) enclosing K′ poles of f(z), z is a complex variable and

$$f(z) = \frac{e^{-jzL_{12}}}{(z+j/\lambda_1)(z+j/\lambda_2)\ \ldots\ (z+j/\lambda_k)}\,. \tag{4.2.18}$$

It can be shown reference [24] that $\int_{C_R} f(z)\, dz = 0$ as R goes to infinity. Note that the K′ poles $-j/\lambda_1, -j/\lambda_2, \ldots, j/\lambda_{K'}$ enclosed by the contour correspond to K′ negative eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{K'}$. Residues at these poles are computed as follows:

$$\text{Res}\Big|_{z=-j/\lambda_1} = \frac{e^{-L_{12}/\lambda_1}}{(-j/\lambda_1+j/\lambda_2)(-j/\lambda_1+j/\lambda_3)\ \ldots\ (-j/\lambda_1+j/\lambda_K)} \tag{4.2.19}$$

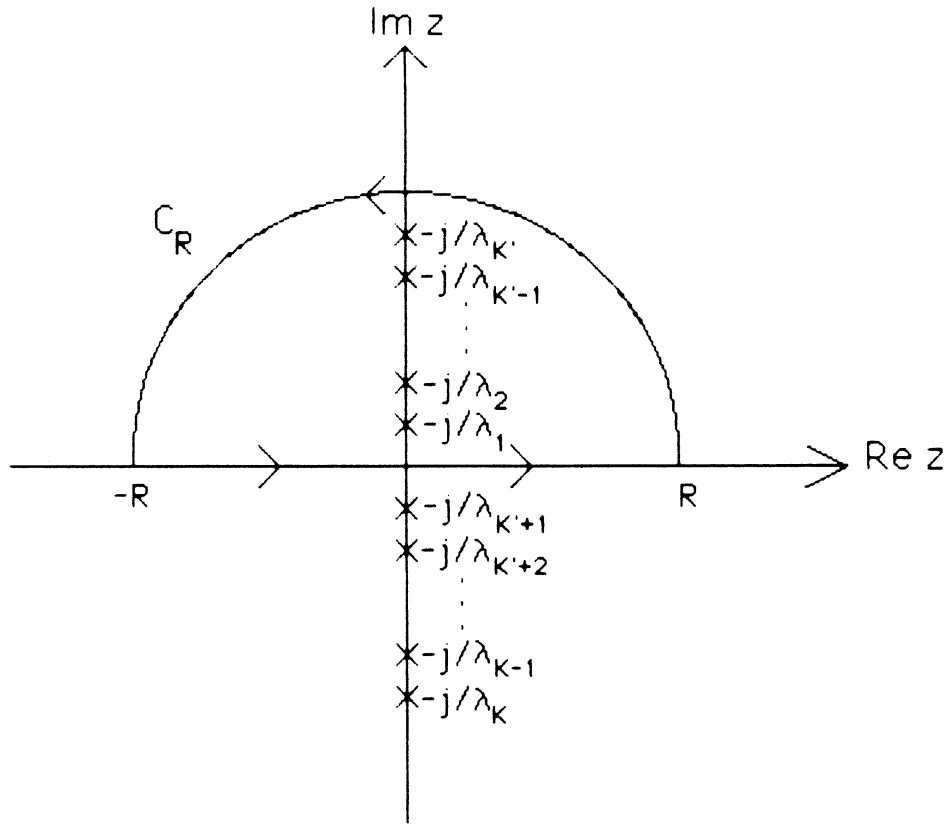


Figure 4.1: Integration contour. The K' poles inside the contour correspond to K' negative eigenvalues, and the remaining K - K' poles outside the contour correspond to positive eigenvalues.

or

$$\text{Res}\Big|_{z=-j/\lambda_1} = \frac{e^{-L_{12}/\lambda_1}}{\prod_{\substack{k=1 \\ k\neq 1}}^{K} (-j/\lambda_1 + j/\lambda_k)} . \tag{4.2.20}$$

Similarly

$$\text{Res}\Big|_{z=-j/\lambda_2} = \frac{e^{-L_{12}/\lambda_2}}{\prod_{\substack{k=1 \\ k\neq 2}}^{K} (-j/\lambda_2 + j/\lambda_k)} . \tag{4.2.21}$$

The sum of the K′ residues is given by

$$\sum_{i=1}^{K'} \frac{e^{-L_{12}/\lambda_i}}{\prod_{\substack{k=1 \\ k\neq i}}^{K} (-j/\lambda_i + j/\lambda_k)} = \sum_{i=1}^{K'} \frac{e^{-L_{12}/\lambda_i}}{j^{K-1} \prod_{\substack{k=1 \\ k\neq i}}^{K} (-j/\lambda_k + j/\lambda_i)} . \tag{4.2.22}$$

Equation (4.2.16) can be written in function of the residues as:

$$f(L_{12}|H_2) = \frac{1}{2\pi} \left[\prod_{k=1}^{K} \left[-\frac{1}{\lambda_k} \right] \right] \frac{1}{j^K} \frac{2\pi j}{j^{K-1}} \sum_{i=1}^{K'} \frac{e^{-L_{12}/\lambda_i}}{\prod_{\substack{k=1 \\ k\neq i}}^{K} (1/\lambda_k - 1/\lambda_i)} . \tag{4.2.23}$$

The term outside the summation can be rearranged as

$$\left[\prod_{k=1}^{K}\left[-\frac{1}{\lambda_k}\right]\right]\frac{j}{j^{K}j^{K-1}}=\frac{(-1)^K}{\prod_{k=1}^{K}(\lambda_k)}\frac{-1}{j^{2K}}=\frac{-(-1)^K}{\left[\prod_{k=1}^{K}(\lambda_k)\right](-1)^K}=\frac{-1}{\prod_{k=1}^{K}(\lambda_k)}.$$

(4.2.24)

Therefore,

$$f(L_{12}|H_2)=\frac{-1}{\prod_{k=1}^{K}(\lambda_k)}\sum_{i=1}^{K'}\frac{e^{-L_{12}/\lambda_i}}{\prod_{\substack{k=1\\k\neq i}}^{K}(1/\lambda_k-1/\lambda_i)} \tag{4.2.25}$$

for L_{12}, $\lambda_i < 0$. The classification probability is given by

$$P(H_2|H_2)=\frac{-1}{\prod_{k=1}^{K}(\lambda_k)}\int_{-\infty}^{\eta}\left[\sum_{i=1}^{K'}\frac{e^{-L_{12}/\lambda_i}}{\prod_{\substack{k=1\\k\neq i}}^{K}(1/\lambda_k-1/\lambda_i)}\right]dL_{12}.$$

(4.2.26)

Since

$$\int_{-\infty}^{\eta}e^{-L_{12}/\lambda_i}\,dL_{12}=-\lambda_i\left[e^{-L_{12}/\lambda_i}\right]_{-\infty}^{\eta}=-\lambda_i e^{-\eta/\lambda_i},$$

(4.2.27)

then

$$P(H_2|H_2)=\frac{1}{\prod_{k=1}^{K}(\lambda_k)}\sum_{i=1}^{K'}\frac{\lambda_i e^{-\eta/\lambda_i}}{\prod_{\substack{k=1\\k\neq i}}^{K}(1/\lambda_k-1/\lambda_i)} \tag{4.2.28}$$

for $\eta < 0$. the summation above can be rewritten as

$$\sum_{i=1}^{K'} \frac{\left[\prod_{\substack{k=1 \\ k\neq i}}^{K} (\lambda_i\lambda_k)\right]\lambda_i\, e^{-\eta/\lambda_i}}{\prod_{\substack{k=1 \\ k\neq i}}^{K} (\lambda_i-\lambda_k)} = \sum_{i=1}^{K'} \frac{\left[\prod_{\substack{k=1 \\ k\neq i}}^{K} (\lambda_k)\right]\lambda_i^{K-1}\, e^{-\eta/\lambda_i}}{\prod_{\substack{k=1 \\ k\neq i}}^{K} (\lambda_i-\lambda_k)}. \tag{4.2.29}$$

Therefore,

$$P(H_2|H_2) = \sum_{i=1}^{K'} \frac{\lambda_i^{K-1}\, e^{-\eta/\lambda_i}}{\prod_{\substack{k=1 \\ k\neq i}}^{K} (\lambda_i-\lambda_k)} \tag{4.2.30}$$

for $\eta < 0$, $\lambda_i < 0$ and λ_k' real. Following the same steps, the false alarm probability is given as

$$P(H_2|H_1) = \sum_{i=1}^{K'} \frac{(\lambda_i')^{K-1}\, e^{-\eta/\lambda_i'}}{\prod_{\substack{k=1 \\ k\neq i}}^{K} (\lambda_i'-\lambda_k')} \tag{4.2.31}$$

for $\eta < 0$, $\lambda_i' < 0$ and λ_k' real. Note that the number of negative eigenvalues of the set $\{\lambda_k'\}$ is also K′ because, as mentioned before, $\mathbf{K}_{y_1}(\mathbf{H}_1-\mathbf{H}_2)$ has eigen-values of the same sign as $\mathbf{H}_1-\mathbf{H}_2$. In the numerical results, it was observed

that the positive eigenvalues have much lower magnitude (of the order 10^{-8}) than the negative eigenvalues (of the order 10^{-1}). Therefore, the positive eigenvalues can be neglected with respect to the negative eigenvalues. $P(H_2|H_2)$ in equation (4.2.30) becomes

$$P(H_2|H_2) = \sum_{i=1}^{K'} \frac{\lambda_i^{K-1} \, e^{-\eta/\lambda_i}}{\lambda_i^{K-K'} \prod_{\substack{k=1 \\ k\neq i}}^{K'} (\lambda_i - \lambda_k)} \tag{4.2.32}$$

or

$$P(H_2|H_2) = \sum_{i=1}^{K'} \frac{\lambda_i^{K'-1} \, e^{-\eta/\lambda_i}}{\prod_{\substack{k=1 \\ k\neq i}}^{K'} (\lambda_i - \lambda_k)} \tag{4.2.33}$$

for λ_i, λ_k and η negative. Similarly the false alarm probability becomes

$$P(H_2|H_1) = \sum_{i=1}^{K'} \frac{(\lambda_i')^{K'-1} \, e^{-\eta/\lambda_i'}}{\prod_{\substack{k=1 \\ k\neq i}}^{K'} (\lambda_i' - \lambda_k')} \tag{4.2.34}$$

for λ_i', λ_k' and η negative.

4.3 Computation of the Signal Correlation Matrix

a. Elements of Φ in Function of Time Delays

In this section we propose to evaluate the elements of the signal correlation matrix Φ in function of time delays for a linear FM pulse. A linear FM pulse of duration T seconds and amplitude A is expressed as (references [26] and [27]):

$$f(t) = A \cos(\omega_0 t + \pi k_f t^2), \quad 0 < t < T \tag{4.3.1}$$

where ω_0 is the center frequency in radians per seconds and k_f is the transmitter constant. The instantaneous frequency of f(t) is given by

$$\frac{1}{2\pi}\frac{d(\pi k_f t^2)}{dt} = k_f t . \tag{4.3.2}$$

If $k_f = 0$, f(t) reduces to

$$f(t) = A \cos(\omega_0 t), \quad 0 < t < T . \tag{4.3.3}$$

This pulse is commonly known as the single–frequency pulse or simply a pure tone. In this dissertation, we shall treat a pure tone as a particular case of the linear FM pulse ($k_f = 0$). Using equation (2.24) and (2.4.8), each element of the K×K matrix Φ is expressed as

$$[\Phi]_{ij} = \frac{1}{NE_f}\int_0^T \left[\sum_{k=1}^{N} f(t-\tau_{ki})\, f(t-\tau_{kj})\right] dt \tag{4.3.4}$$

which evaluates to the summation of the integral:

$$\int_0^T f(t-\tau_{ki})\, f(t-\tau_{kj})\, dt \,. \tag{4.3.5}$$

Changing the variable $t-\tau_{kj}$ to t', (4.3.5) becomes

$$\int_0^T f(t'+\tau_{kj}-\tau_{ki})\, f(t')\, dt' \,. \tag{4.3.6}$$

Therefore,

$$[\Phi]_{ij} = \frac{1}{NE_f} \sum_{k=1}^{N} \int_0^T f(t)\, f(t-\tau)\, dt \tag{4.3.7}$$

where

$$\tau = \tau_{ki} - \tau_{kj} \,. \tag{4.3.8}$$

If f(t) is a linear FM pulse, the analytical solution of the integral in (4.3.7) cannot be easily obtained. Hence, we resorted to numerical integration in the computer simulation. If, however, f(t) is a pure tone, then

$$[\Phi]_{ij} = \frac{A^2}{NE_f} \sum_{k=1}^{N} \int_0^T \cos(\omega_0 t)\, \cos[\omega_0(t-\tau)]\, dt \,. \tag{4.3.9}$$

Using trigonometric properties, we can write

$$\int_0^T \cos(\omega_0 t)\cos[\omega_0(t-\tau)]\,dt = I_1 + I_2 \tag{4.3.10}$$

where

$$I_1 = \frac{1}{2}\int_0^T \cos[\omega_0(2t-\tau)]\,dt = \frac{1}{4\omega_0}[\sin(\omega_0(2T-\tau)) + \sin\,\omega_0\tau] \tag{4.3.11}$$

and

$$I_2 = \frac{1}{2}\int_0^T \cos(\omega_0\tau)\,dt = \frac{T}{2}\cos(\omega_0\tau)\,. \tag{4.3.12}$$

Therefore,

$$[\Phi]_{ij} = \frac{A^2}{NE_f}\sum_{k=1}^{N}\left\{\frac{1}{4\omega_0}[\sin(\omega_0(2T-\tau)) + \sin(\omega_0\tau)] + \frac{T}{2}\cos(\omega_0\tau)\right\}. \tag{4.3.13}$$

The signal energy E_f for a pure tone is

$$E_f = \int_0^T (A\cos\omega_0 t)^2\,dt = \frac{A^2}{2}\int_0^T (1+\cos 2\omega_0 t)\,dt\,. \tag{4.3.14}$$

Remembering that T is not the period of the signal but rather its pulse duration, then

$$E_f = \frac{A^2}{2}\left[T + \frac{\sin\ 2\omega_0 T}{2\omega_0}\right]. \tag{4.3.15}$$

The signal energy E_f for a linear FM signal is computed using numerical integration. This concludes our discussion of how Φ can be computed in terms of time delays.

b. Time Delay Computations

In this section, we propose to evaluate time delays for specific test volume geometries. Each time delay τ_{nk} represent the time it takes the waveform f(t) to travel from the transmitter to the k^{th} cell and back to the n^{th} sensor in the array. Figure (4.2.a) illustrates this scenario. One can see that the travel time is proportional to the distance traveled

$$c\tau_{nk} = \|\mathbf{r}_k\| + \|\mathbf{r}_k - \mathbf{d}_n\| \; ; \; k = 1, 2, \ldots, K; \; n = 1, 2, \ldots, N \tag{4.3.16}$$

where $\mathbf{r}_k$ is the position vector fixing the origin of the k^{th} cell of the i^{th} object's test region; and $\mathbf{d}_n$ is the position vector for the n^{th} element in the receiving array. A nonrefractive medium with sound speed c is assumed.

A Taylor series expansion in terms of $\mathbf{d}_n$ about 0 (equation D.9) gives, showing only first order terms,

$$c\tau_{nk} = 2\|\mathbf{r}_k\| - \boldsymbol{\alpha}_k^T \mathbf{d}_n + \ldots \tag{4.3.17}$$

where $\boldsymbol{\alpha}_k = \mathbf{r}_k/\|\mathbf{r}_k\|$ is the unit vector along $\mathbf{r}_k$. To neglect higher order terms is to assume that each of the K cells is in the Fraunhofer zone of the array. here it is assumed that for the array dimensions and distances of interest, the Fraunhofer zone applies and equation (4.3.17) without the higher order terms is taken to be exact. Therefore,

$$\tau_{nk} = \tau_k - \frac{1}{c} \boldsymbol{\alpha}_k^T \cdot \mathbf{d}_n \; ; \; k = 1, 2, \ldots, K; \; n = 1, 2, \ldots, N \tag{4.3.18}$$

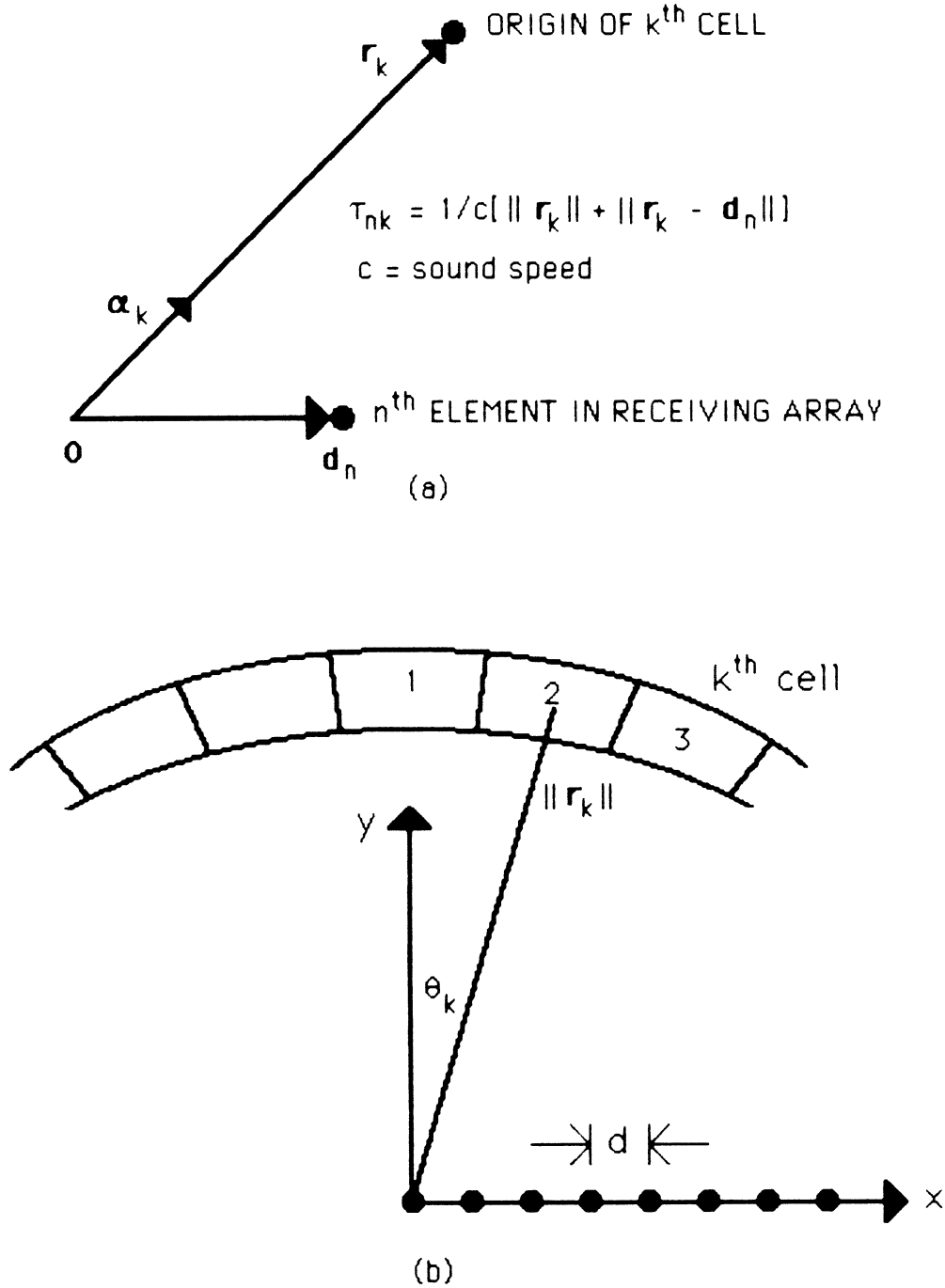


Figure 4.2 . a. Time delay configuration for nonrefractive medium.
b. Single annulus and linear array.

where

$$\tau_k = \frac{2}{c} \|\mathbf{r}_k\| \ . \tag{4.3.19}$$

Let us now consider the case depicted in figure (4.2.b) where we assume a uniform linear array and a single range annulus. In this example, the transmitter location at the origin of the x–y axis coincides with the first array element and the center of the annulus. Therefore, τ_k is a constant for all k and the distance $\|\mathbf{d}_n\|$ from the n^{th} array element to the origin is given by

$$\|\mathbf{d}_n\| = (n-1)d \ ; \ \ n = 1, 2, \ldots, N \tag{4.3.20}$$

where d is the spacing between sensors. Now let the cells be uniform in angular extent. In other words,

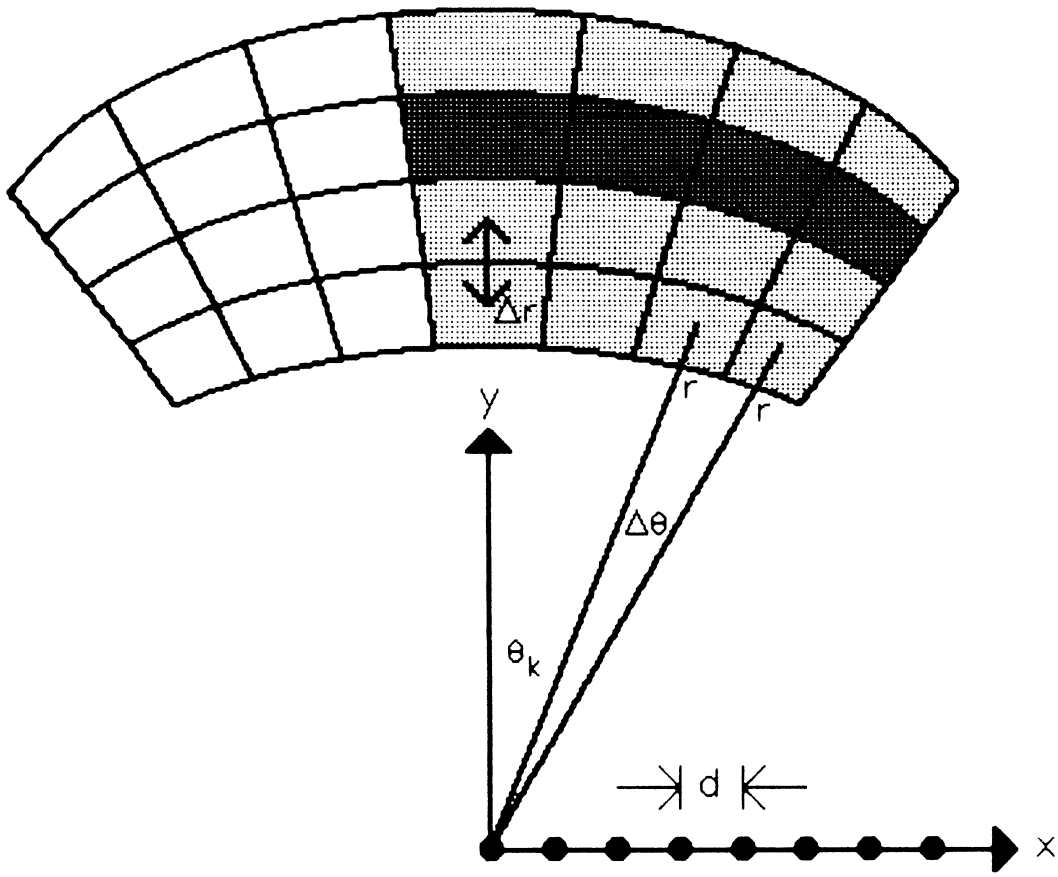
$$\theta_k = (k-1)\ \Delta\theta \ ; \ \ k = 1, 2, \ldots, K \ . \tag{4.3.21}$$

Using these assumptions, the time delay τ_{nk} is expressed as

$$\tau_{nk} = \tau_k - \frac{1}{c}(n-1)d \sin\left[(k-1)\ \Delta\theta\right] \tag{4.3.22}$$

for n = 1, 2, . . . , N and k = 1, 2, . . . , K.

The real situation of interest is when the target is contained in a two–dimensional test region of K cells as shown in figure 4.3. The target occupies one annulus of J cells. It is easily seen that if the length of the target is small compared to the range of interest, then the distance between cell k and cell k+1



CELL LABELING k

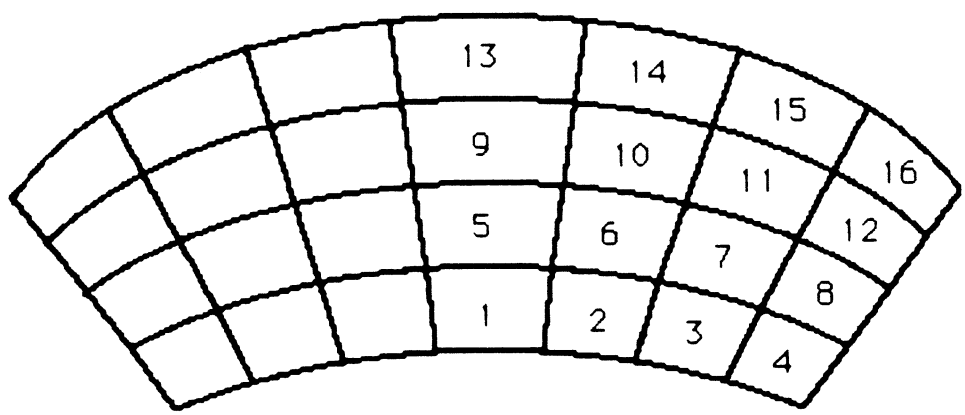


Figure 4.3: Two dimensional test region and linear array.

within one annulus is equal to the distance between cell k and cell k+J in the next annulus. Let r be the range of the first annulus and Δr the distance between two neighboring cells (in azimuth or in range). The range of the second annulus is given by

$$r + \Delta r = r + r\Delta\theta = r(1+\Delta\theta) . \tag{4.3.23}$$

Similarly the range of the third annulus is

$$r + 2\Delta r = r + 2r\Delta\theta = r(1 + 2\Delta\theta) . \tag{4.3.24}$$

Since all cells within one annulus are at the same range, the range of each cell in function of its index k is given by

$$r[1 + ((k-1) \text{ div } J)\Delta\theta] , k = 1, \ldots , K \tag{4.3.25}$$

where a div b symbolizes the integer division of a by b. The time delay τ_k defined in (4.3.19) is now expressed as

$$\tau_k = \frac{2}{c} r[1 + ((k-1) \text{ div } J) \Delta\theta] , \; k = 1, \ldots , K . \tag{4.3.26}$$

We can also generalize the expression of θ_k given in (4.3.21) to accommodate the two–dimensional test region of figure 4.3. Therefore,

$$\theta_k = ((k-1) \bmod J) \Delta\theta ; \; k = 1, \ldots , K \tag{4.3.27}$$

where a mod b equals the remainder of the integer division of a by b. The time delay τ_{nk} for the two–dimensional test region can be summarized using (4.3.18), (4.3.26) and (4.3.27) as

$$\tau_{nk} = \frac{2r}{c}\left[1 + ((k-1) \text{ div } J)\,\Delta\theta\right] - \frac{(n-1)d}{c}\sin\left[((k-1) \bmod J)\,\Delta\theta\right] \tag{4.3.28}$$

for $k = 1, \ldots, K$ and $n = 1, \ldots, N$.

4.4 Simulation Results and Conclusions

In this section we will display ROC curves for two–dimensional test region illuminated by linear FM pulses. The pure tone pulse is treated as a particular FM signal with a transmitter constant k_f equal to zero. The array of sensors is linear with eight elements (figure 4.3) separated by a distance d equal to half of the signal wavelength ($d = \lambda/2$). The center frequency f_0 of the transmitter is fixed to 10 Khz and the test volume range r in equation (4.3.28) equals 1000 meters. The speed of sound c in water is 1500 m/sec. The angular width $\Delta\theta$ will be expressed in function of the null–to–null beamwidth (figure 4.4) of the array. The null–to–null beamwidth is given by

$$BW = 2\sin^{-1}\left[\frac{\lambda}{Nd}\right] = 29^\circ\ . \tag{4.4.1}$$

In the simulation, we first evaluate the time delays (equation 4.3.28). Then we construct the signal correlation matrix Φ using equations (4.3.1) and (4.3.7). If $k_f = 0$ (pure tone), Φ is evaluated analytically using equation (4.3.13). Knowing σ_R^2, σ_0^2 and Φ, we can compute $\mathbf{K}_{y_i}(\mathbf{H}_1 - \mathbf{H}_2)$ for $i = 1, 2$. The

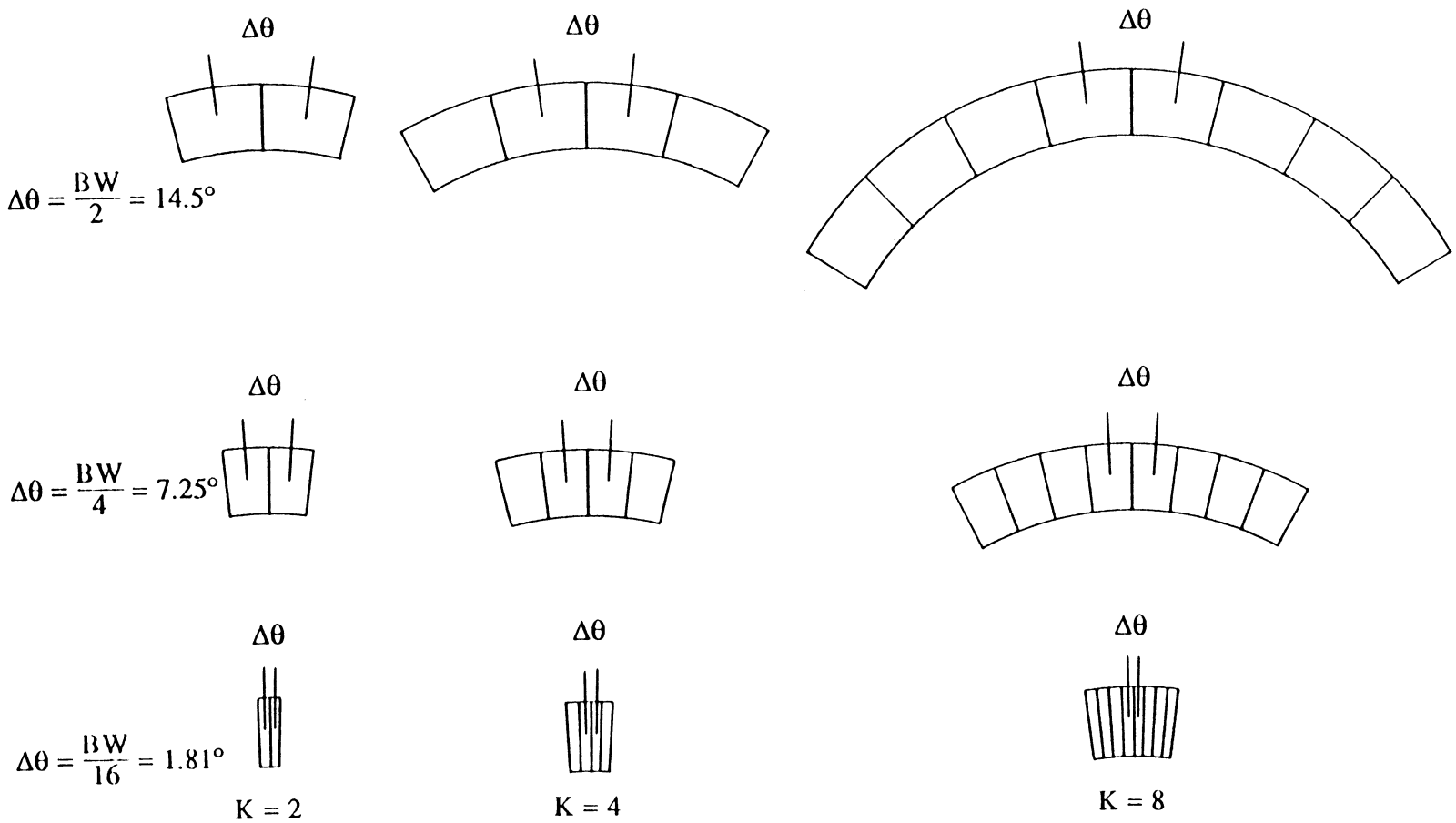
eigenvalues $\{\lambda_k'\}$ of $\mathbf{K}_{y_1}(\mathbf{H}_1-\mathbf{H}_2)$ are used to compute the false alarm probability given in equation (4.2.34). Also, the eigenvalues $\{\lambda_k\}$ of $\mathbf{K}_{y_2}(\mathbf{H}_1-\mathbf{H}_2)$ are used to compute the classification probability given in equation (4.2.33). The results will demonstrate the performance of the classifier with respect to (1) the transmitter constant k_f, (2) the target–to–reverberation contrast, (3) the signal–to–noise ratio ρ_N, and (4) the angular width $\Delta\theta$ for the matched processor case (figure 4.5). For the mismatched processor (figure 4.11), we display the classifier performance for both (5) range misidentification and (6) object misorientation.

(1) Varying the transmitter constant k_f. In this example, we vary the transmitter constant k_f. Here the number of cells J occupied by the target is 4, the signal– to–noise ratio ρ_N is 1.0, the angular width $\Delta\theta$ is equal to one eighth of the beamwidth ($\Delta\theta$ = BW/8 = 3.625°), and the pulse duration T is set to 1 second. The reverberation and target strength per cell σ_R^2 and σ_0^2 are set to 10 and 40 (low contrast) in figure 4.6 and to 10 and 80 (high contrast) in figure 4.7. In both cases k_f takes the following values: 0, 1, 10, 100, 1000 and 10000. It is clear that as k_f increases the performance improves noticeably. This improvement is more pronounced for high contrast (figure 4.7). For instance, for k_f greater than or equal to 100, we can choose η such that $P(H_2|H_1) = 0.1$ and $P(H_2|H_2) > 0.9$. The cost one has to pay for large values of k_f is expensive transducers because of the bandwidth required to transmit these signals.

(2) Varying the target to reverberation contrast. Here J = 4 cells, $\rho_N = 1.0$, $\Delta\theta = 3.625°$, $k_f = 0$ and T = 30 ms. The reverberation strength per cell σ_R^2 is set to 10 while the target strength per cell σ_0^2 is varied from 20 to 70 by steps of 10. It is clear that performance improves as the target–to–reverberation contrast increases (figure 4.8).

(3) Varying the signal–to–noise ratio ρ_N. In this example, we examine the classifier performance as we increase the signal–to–noise ratio. Here J = 4, $\Delta\theta = 3.625°$, T = 30 ms, $k_f = 0$, $\sigma_R^2 = 10$ and $\sigma_0^2 = 80$. In figure 4.9, the signal–to–noise ratio ρ_N takes the following values: 0.001, 0.01, 0.1, 0.25, 1, 10 and 100 (lowest to highest curve, respectively). The performance is improved for signal–to–noise ratios greater than 0.1 (–10 dB). For $\rho_N = 0.001$ (–30 dB), the performance degrades drastically and one can no longer distinguish between H_1 and H_2.

(4) Varying the angular width $\Delta\theta$. Our variable is now the angular width $\Delta\theta$. Here $\sigma_R^2 = 10$, $\sigma_0^2 = 70$, J = 4, $k_f = 0$, T = 30 ms and $\rho_N = 1$. $\Delta\theta$ takes the following values: BW/2, BW/4, BW/16, BW/32 and BW/64 (highest to lowest curve, respectively). We see that the performance degrades as the classifier attempts finer angular resolution (figure 4.10). Practically speaking, trying to distinguish finer details of a target is not an easy task. One cost we have to pay at high angular resolution can, for instance, be a large signal–to–noise ratio.



K = NUMBER OF CELLS

BW = NULL TO NULL BEAMWIDTH = $2\ \mathrm{SIN}^{-1}\frac{\lambda}{Nd}$ FOR N = 8, d = $\frac{\lambda}{2}$: BW = 29°

Figure 4.4 Test Regions for Example: Two-Dimensional Linear Array.

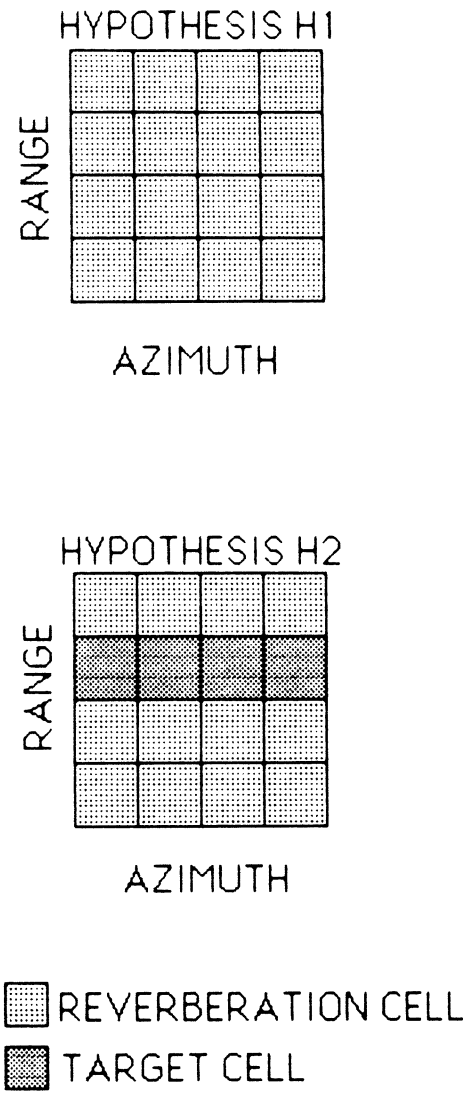


Figure 4.5: Example of a two-dimensional test volume. K = 16 test cells and J = 4 target-like cells.

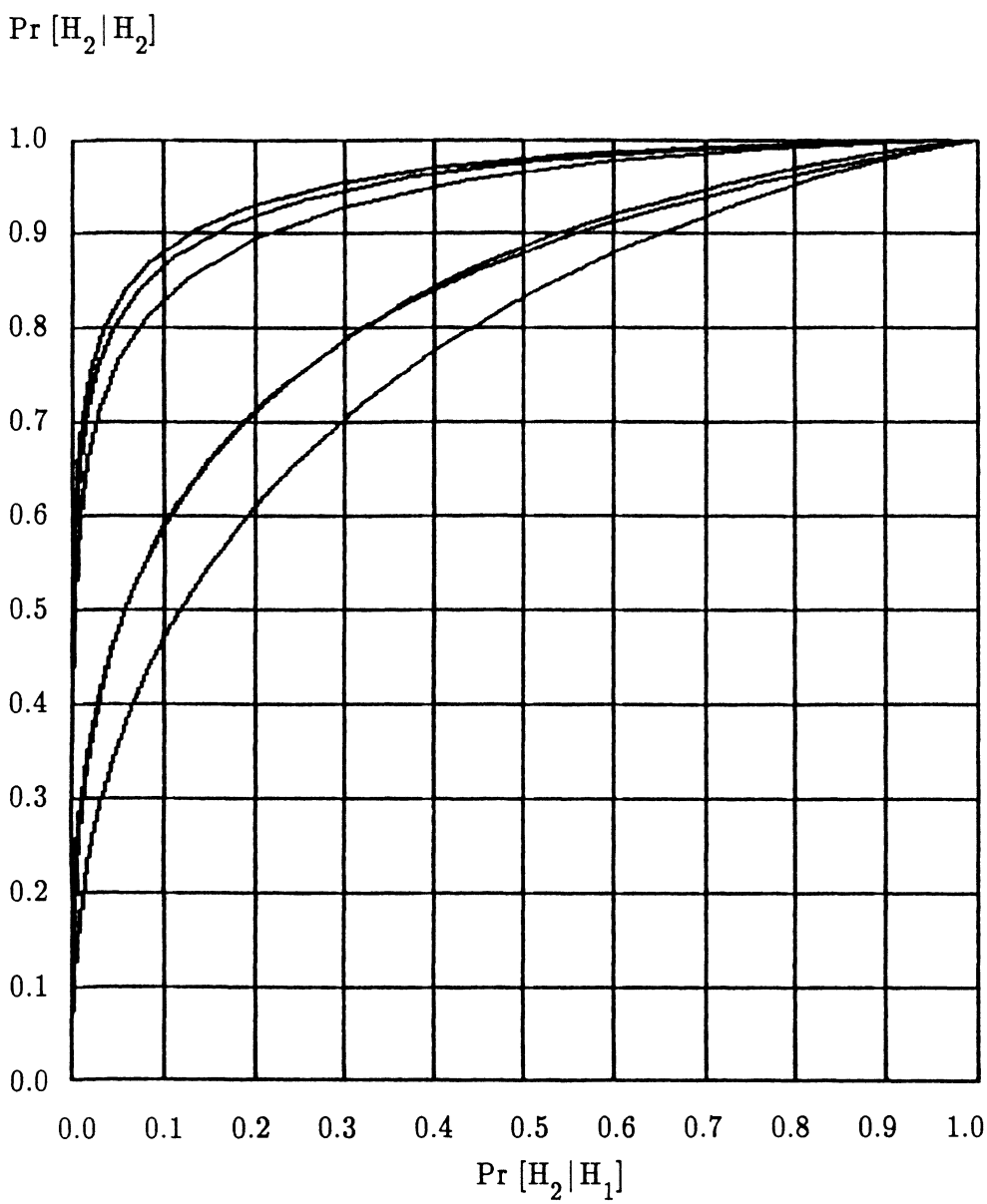


Figure 4.6: ROC curves displaying classifier performance in function of the transmitter constant k_f. Here J = 4 target cells, $\Delta\theta = 3.625°$, T = 1 second, $\rho_N = 1.0$, $\sigma_R^2 = 10$ and $\sigma_0^2 = 40$. k_f takes the values 0, 1, 10, 100, 1000 and 10000 (lowest to highest curve, respectively).

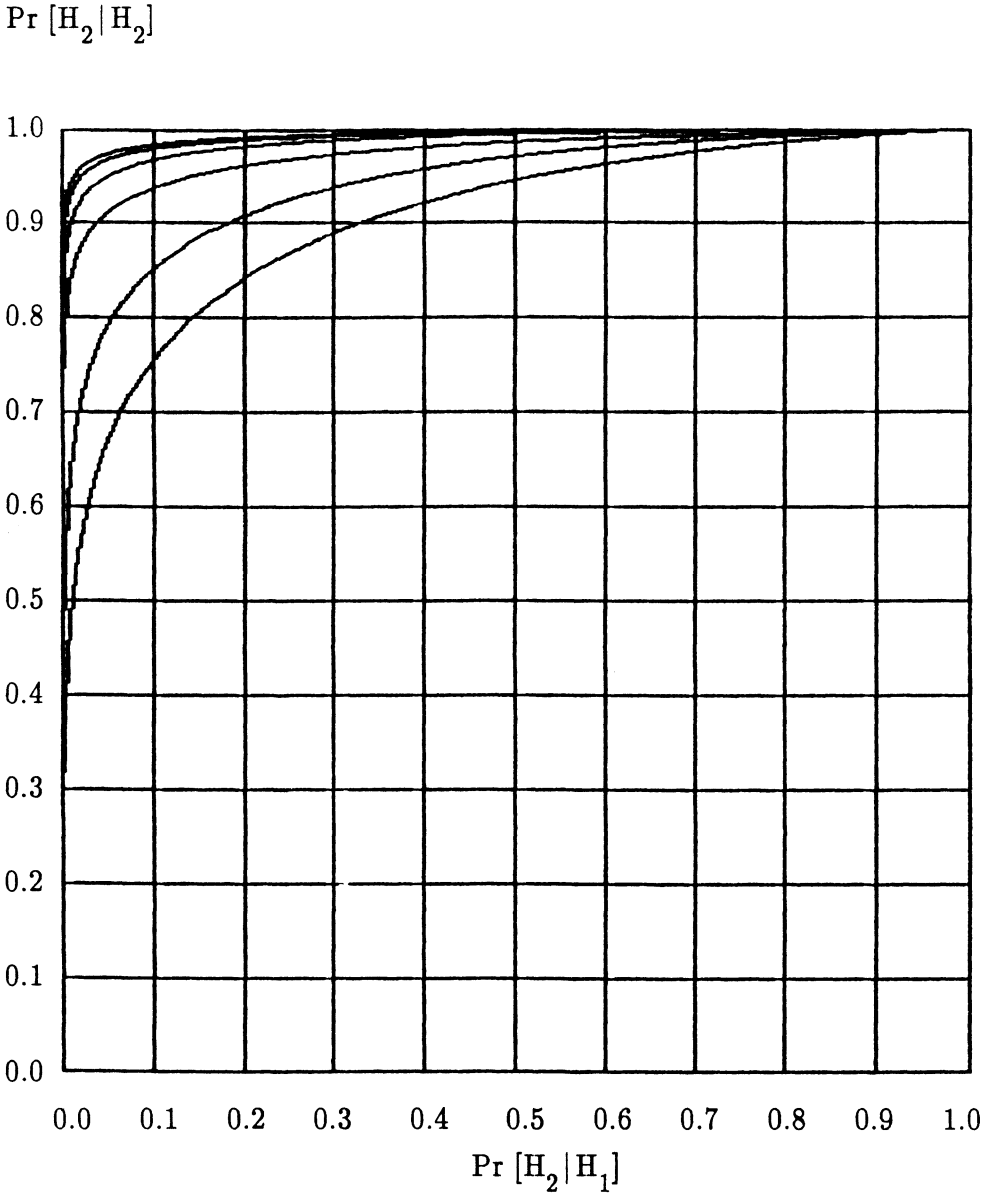


Figure 4.7: ROC curves displaying classifier performance in function of the transmitter constant k_f. Here J = 4 target cells, $\Delta\theta = 3.625°$, T = 1 second, $\rho_N = 1.0$, $\sigma_R^2 = 10$ and $\sigma_0^2 = 80$. k_f takes the values 0, 1, 10, 100, 1000 and 10000 (lowest to highest curve, respectively.)

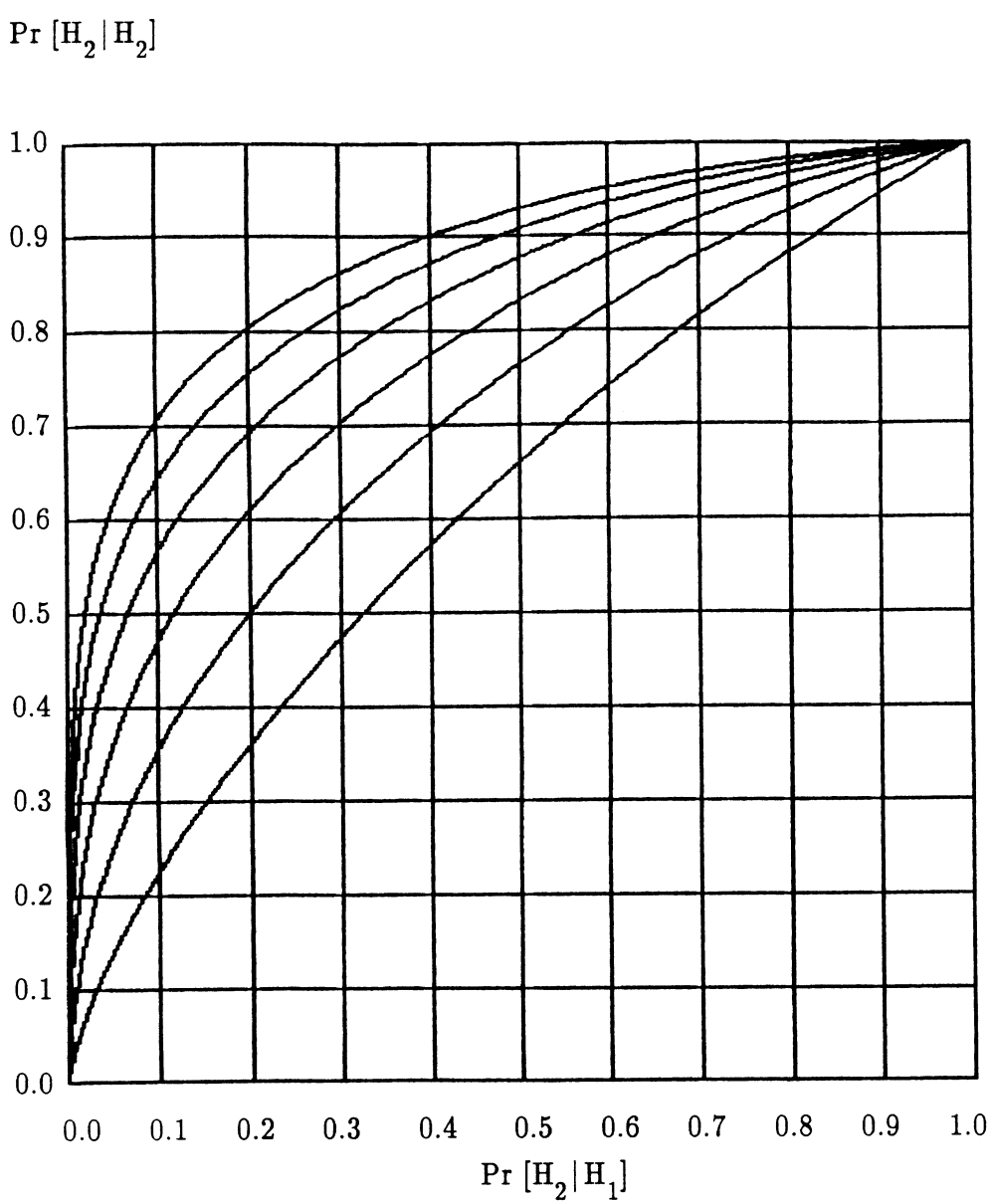


Figure 4.8: ROC curves displaying classifier performance in function of the target–to–reverberation contrast. Here J = 4 target cells, $\Delta\theta$ = 3.625°, T = 30 ms, ρ_N = 1.0, k_f = 0 and σ_R^2 = 10. σ_0^2 is varied from 20 (lowest curve) to 70 (highest curve) by steps of 10.

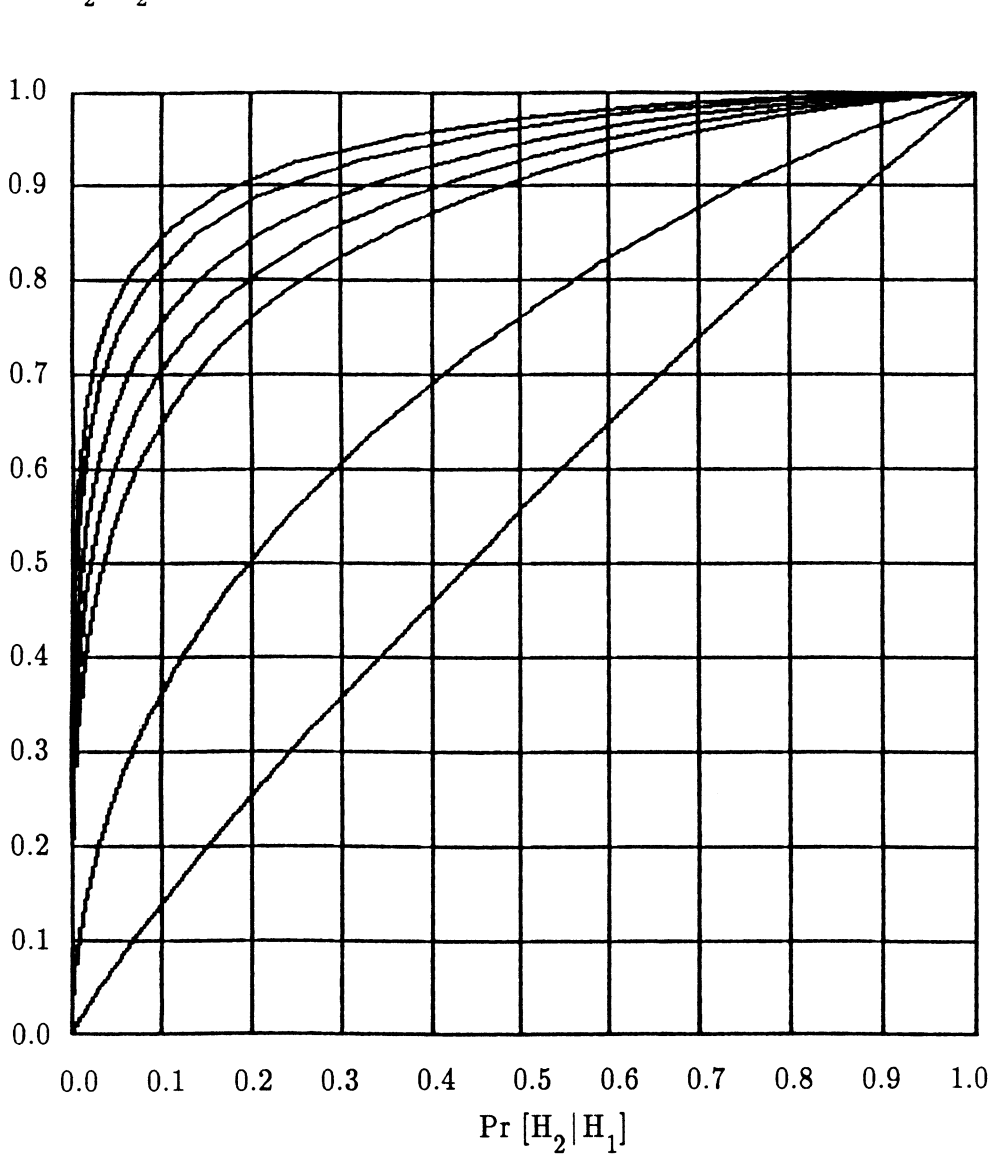


Figure 4.9: ROC curves displaying classifier performance in function of the signal–to–noise ratio ρ_N. Here J = 4 target cells, $\Delta\theta = 3.625°$, T = 30 ms, $k_f = 0$, $\sigma_R^2 = 10$ and $\sigma_0^2 = 80$. ρ_N takes the values 0.001, 0.01, 0.1, 0.25, 1, 10, 100 (lowest to highest curve, respectively).

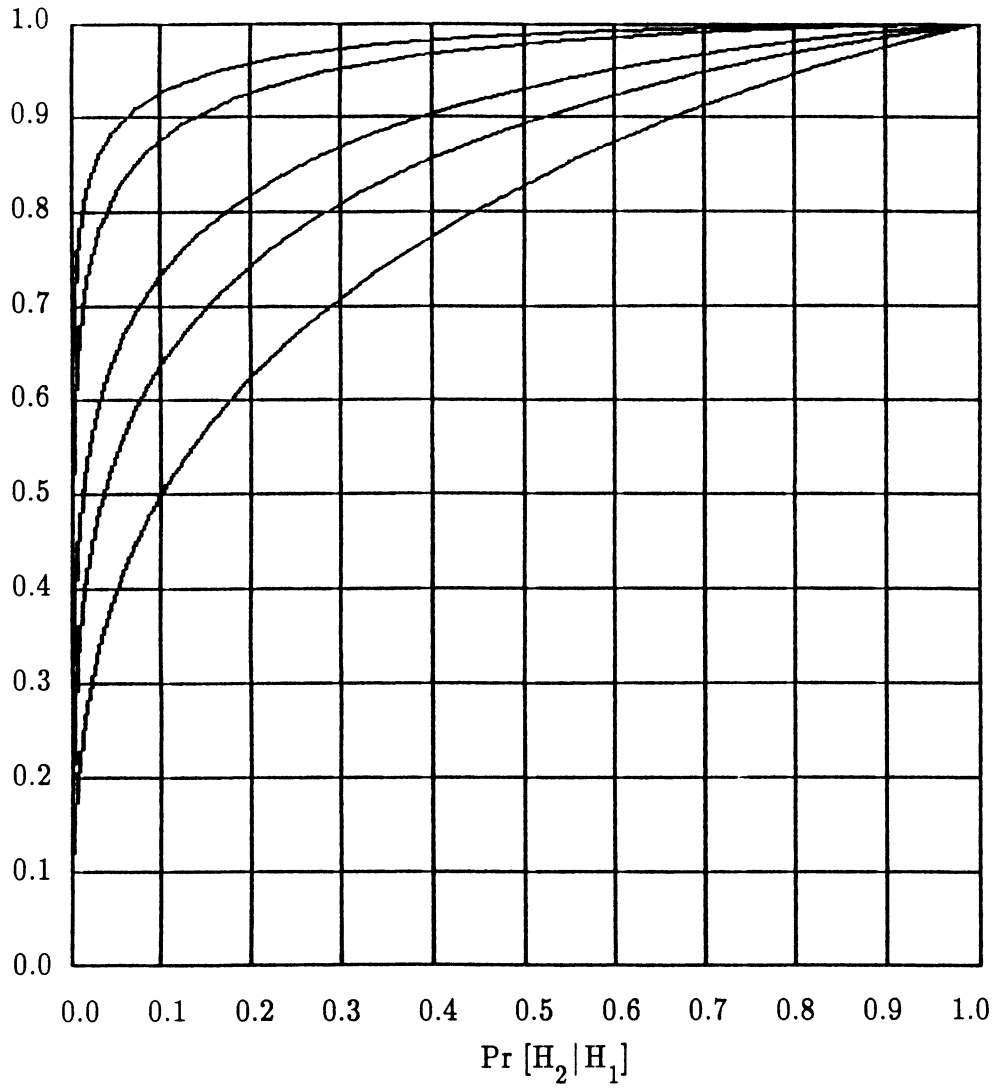


Figure 4.10: ROC curves displaying classifier performance in function of the angular width $\Delta\theta$. Here J = 4 target cells, T = 30 ms, $k_f = 0$, $\sigma_R^2 = 1\ 0$ and $\sigma_0^2 = 70$. $\Delta\theta$ takes the values BW/2, BW/4, BW/16, BW/32 and BW/64 (highest to lowest curve, respectively).

(5) Range misidentification. Now we examine the classifier performance when the wrong scattering covariance matrix is infused into the processor. In this example, the processor assumes the wrong range (refer to figure 4.11). The number of target–like cells J and the signal–to–noise ratio ρ_N are set to 4 and 1, respectively. Also, $\Delta\theta = 3.625°$, $k_f = 0$, $\sigma_R^2 = 10$, $\sigma_0^2 = 70$ and T = 30 ms. Here the farthest range is named row 1 and the closest range is named row 4. The target is located in row 2. In figure 4.12, the highest curve is obtained when the processor assumes perfect a priori knowledge of the test region (matched case). The remaining curves display the classifier performance when the processor assumes the target is in row 1, row 3 or row 4 (second highest to lowest curve, respectively). It is obvious that the better the target range is estimated, the better the classifier performance is. The assumption that the target is in row 1 and row 3 provides about the same classifier performance while the assumption that the target is in row 4 gives the worst performance.

In the case of high resolution signals ($\Phi = I$), we showed that the classifier cannot distinguish between H_1 and H_2 if the range is misestimated. This is not true for non–high resolution signals (particularly FM signals) since the trans– mitted signal picks up energy from neighboring cells.

(6) Object Misorientation. In this example, the processor assumes the wrong orientation (figure 4.11). J, ρ_N, $\Delta\theta$, k_f, T, σ_R^2 and σ_0^2 are assigned the same values as in (5). Here, we vary the azimuth of the target. For instance, if the target is assumed to be in column 1, the angle it makes with the transmitter is the smallest. If, however, the target is assumed to be in column

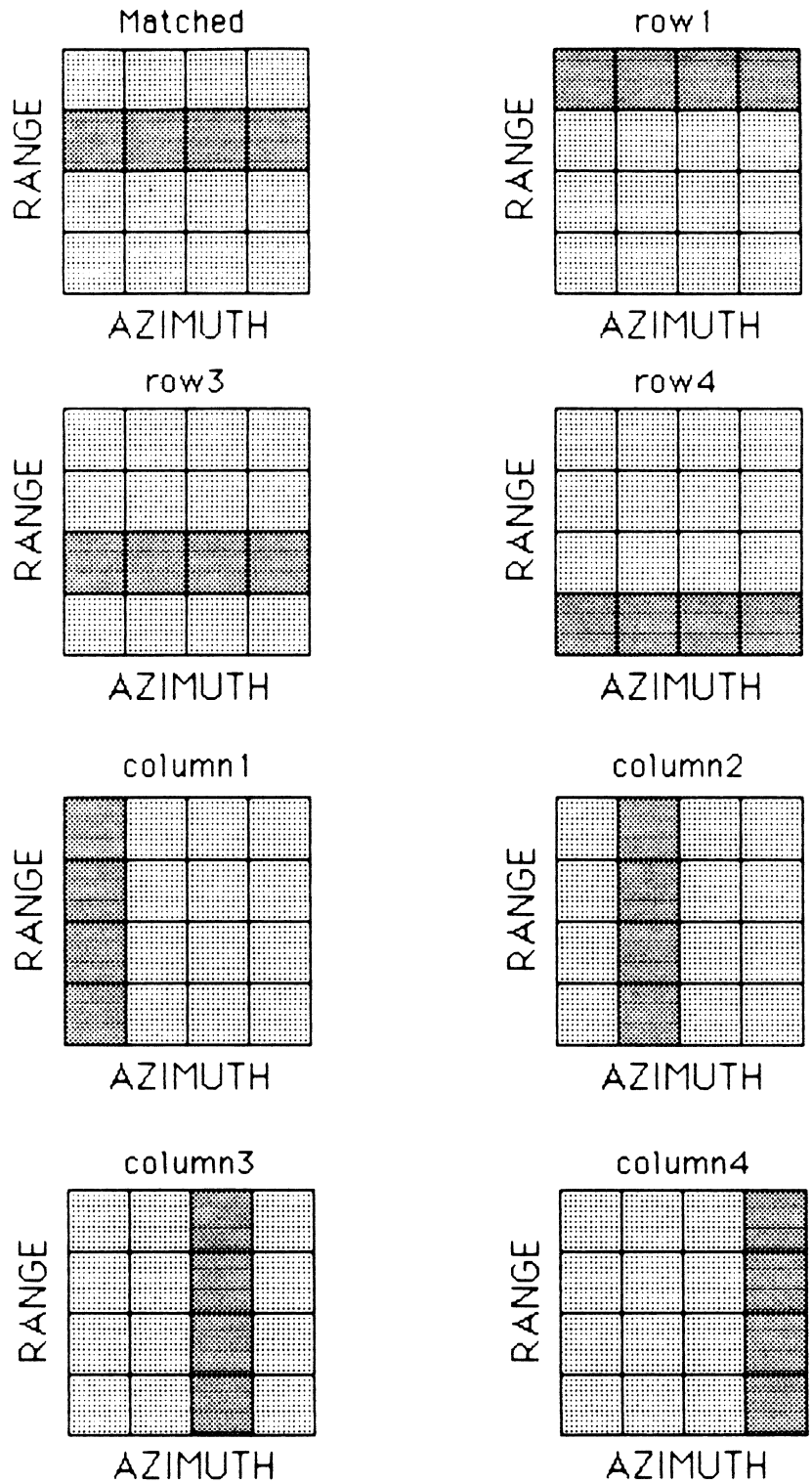


Figure 4.11: Scenario illustrating range and orientation mismatch for K = 16 cells and J = 4 target-like cells.

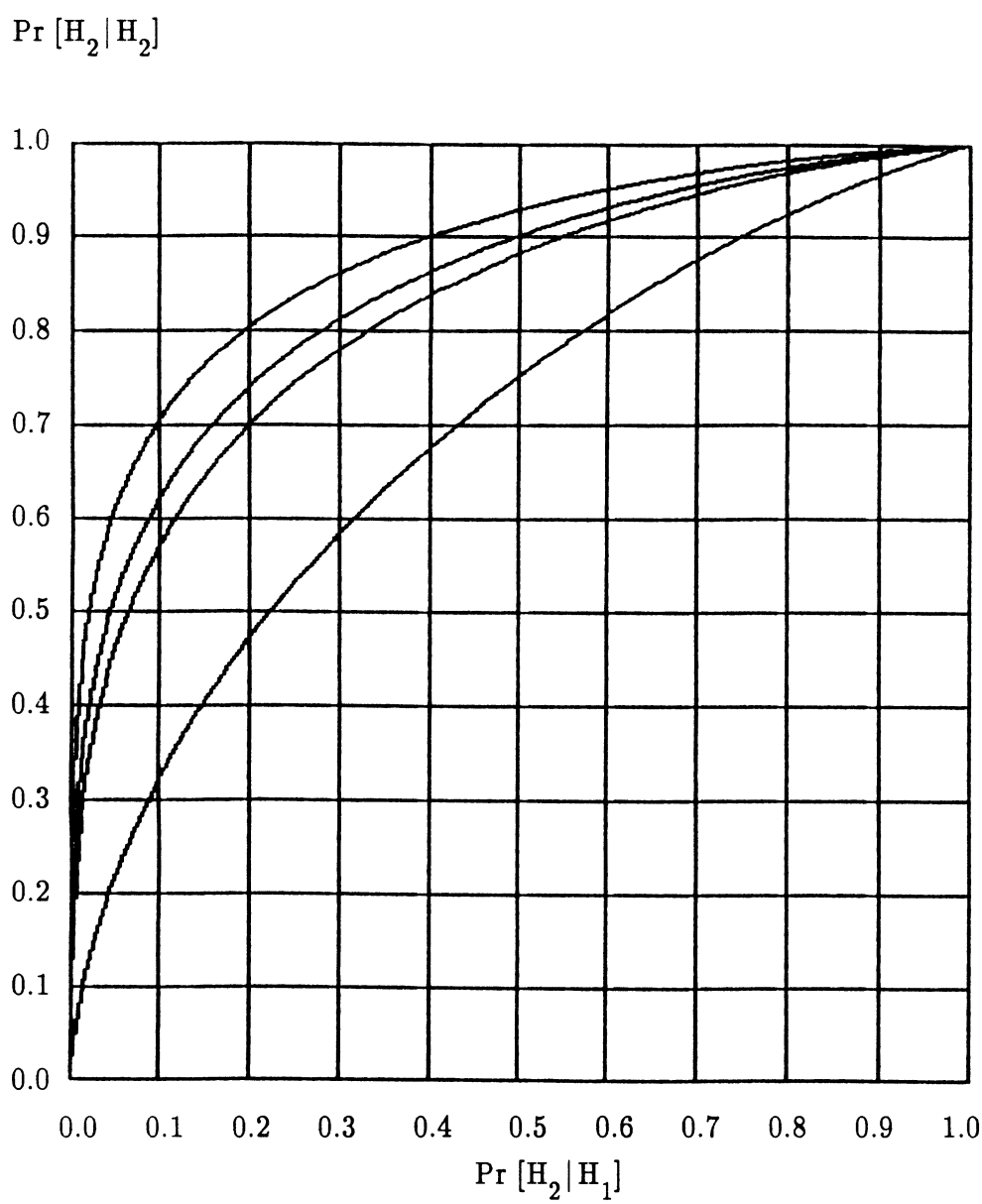


Figure 4.12: ROC curves for the range mismatch. Here J = 4 target cells, $\rho_N = 1.0$, T = 30 ms, $\Delta\theta = 3.625°$, $k_f = 0$, $\sigma_R^2 = 10$ and $\sigma_0^2 = 70$. Highest curve is for the matched processor. In remaining curves, the target is assumed to be row 1, 3 and 4 (second highest to lowest curve, respectively).

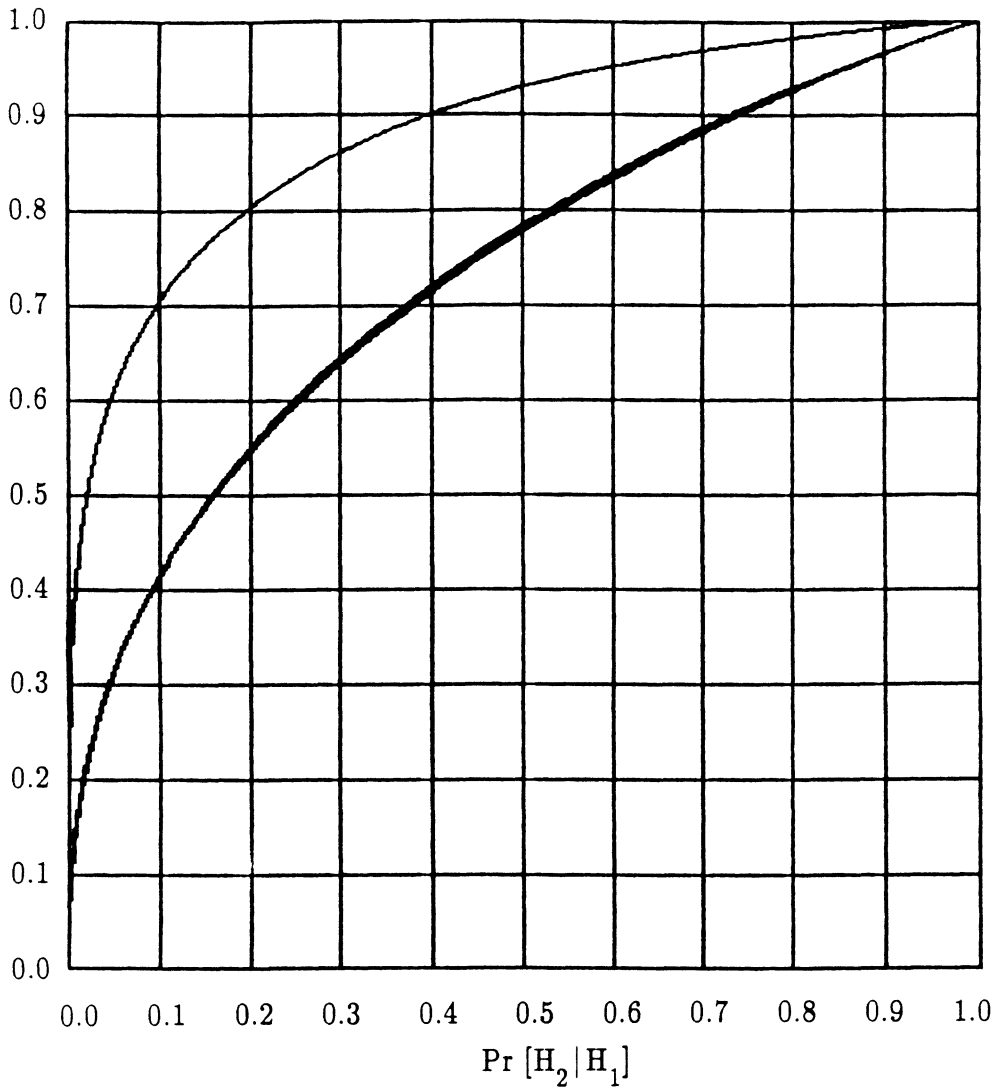


Figure 4.13: ROC curves for the orientation mismatch. Here J = 4 target cells, $\rho_N = 1.0$, T = 30 ms, $\Delta\theta = 3.625°$, $k_f = 0$, $\sigma_R^2 = 10$ and $\sigma_0^2 = 70$. Highest curve is for the matched processor. In the remaining curves, the target is assumed to be in columns 1, 2, 3 and 4 (over–laping curves).

4, the angle it makes with the transmitter is the largest. The target is actually located in row 2. In figure 4.13, the highest curve is obtained when the processor assumes perfect a priori knowledge of the test volume (matched case). The remaining curves display the classifier performance when the processor assumes that the target is located in column 1, 2, 3 or 4. It is clear that the classifier performance degrades when the object is misoriented. Interestingly, this degradation is independent of the target azimuth. Note that the curves coincide together in this case. Hence we can draw the following conclusion: when targets are illuminated by linear FM signals, a correct knowledge of the target orientation is very important for the classifier. On the contrary, the classifier is not as sensitive (within limits) to a wrong estimated range. This concludes our discussion of simulation results.

Chapter 5

ACOUSTIC TARGET IMAGING

5.1 Introduction

This chapter is concerned with an estimation problem. In order to image the test volume encompassing the target, we would like to estimate the scattering coefficients. The estimator $\tilde{\mathbf{a}}_i$ is sought to be a minimum variance linear unbiased (MVLU) estimator of the random vector $\mathbf{a}_i$.

In section 5.2, we derive the MVLU estimator $\tilde{\mathbf{a}}_i$ in function of the data $\mathbf{x}(t)$ and the a priori scattering covariance matrix $\mathbf{K}_{a_i}$ for i = 1, 2. This will be done using the Wiener–Hopf equation (reference 17). We shall also prove that classification is a function of the properties of the estimator $\tilde{\mathbf{a}}_i$.

This issue will be discussed further in section 5.3. There, we will derive an expression for the variance of the estimator error. In part a) of section 5.3, the variance of the estimation error is derived for the general case of mismatched processing and in part b), it is simplified for the specific case of matched processing.

In section 5.4, we display graphic results illustrating variance reduction. We will see that the variance is reduced by increasing the so–called received signal–to–noise ratio and decreasing the angular resolution of the test region. These results will be discussed for both matched and mismatched processing and the test region of concern is either a single range annulus or a two–dimensional grid. We will also illustrate variance reduction with increased transmitter constant k_f for linear FM signals.

5.2 Estimation of $\mathbf{a}_i$

In this section, we derive the minimum variance linear unbiased (MVLU) estimator of the random vector $\mathbf{a}_i$ from the data $\mathbf{x}(t)$, which under H_i ($i = 1, 2$), are given by the linear model

$$\mathbf{x}(t) = \mathbf{F}(t)\mathbf{a}_i + \mathbf{n}(t) , \quad 0 < t < T . \tag{5.2.1}$$

We are seeking an estimator $\tilde{\mathbf{a}}_i$, under H_i, to be a minimum variance estimator of $\mathbf{a}_i$. $\tilde{\mathbf{a}}_i$ is also unbiased and linear. Therefore

$$E[\tilde{\mathbf{a}}_i | H_i] = E[\mathbf{a}_i] \tag{5.2.2}$$

and

$$\tilde{\mathbf{a}}_i = \int_T \mathbf{H}_i(t)\, \mathbf{x}(t)\, dt , \qquad i = 1, 2 \tag{5.2.3}$$

where $\mathbf{H}_i(t)$ is a K×N unknown matrix function. Since the measurement process $\mathbf{x}(t)$ has a mean of zero, then

$$E[\tilde{\mathbf{a}}_i | H_i] = 0 , \qquad (i = 1, 2) \tag{5.2.4}$$

let $\epsilon_i = \tilde{\mathbf{a}}_i - \mathbf{a}_i$ represent the estimation error, and $q(\tilde{\mathbf{a}})$ the quadratic form given by

$$q(\tilde{\mathbf{a}}_i) = E[\epsilon_i^H \mathbf{V}\, \epsilon_i | H_i] \tag{5.2.5}$$

where $\mathbf{V}$ is any K×K positive definite matrix and $\tilde{\mathbf{a}}_i$ is any estimator that satisfies equations (5.2.2) and (5.2.3). $q(\tilde{\mathbf{a}}_i)$ is minimized when (references 17,

28 and 29) $H_i(t)$ satisfies the Wiener–Hopf equation given by

$$E[\mathbf{a}_i\, \mathbf{x}^H(t)\,|\,H_i] = \int_T \mathbf{H}_i(\tau)\, E[\mathbf{x}(\tau)\, \mathbf{x}^H(t)\,|\,H_i]\, d\tau \tag{5.2.6}$$

for i = 1, 2 and $0 < \tau, t < T$. Using equation (5.2.1),

$$E[\mathbf{a}_i\, \mathbf{x}^H(t)\,|\,H_i] = E[\mathbf{a}_i\mathbf{a}_i^H\, \mathbf{F}^H(t)] + E[\mathbf{a}_i\mathbf{n}^H(t)] \qquad (i = 1, 2) \tag{5.2.7}$$

since $E[\mathbf{a}_i\mathbf{n}^H(t)] = 0$ for $0 < t < T$, therefore,

$$E[\mathbf{a}_i\, \mathbf{x}^H(t)\,|\,H_i] = \mathbf{K}_{a_i}\mathbf{F}^H(t)\,, \qquad i = 1, 2\,. \tag{5.2.8}$$

From equations (2.2.8) and (2.2.14)

$$\mathbf{K}_{x_i}(t_1,t_2) = \mathbf{F}(t_1)\mathbf{K}_{a_i}\mathbf{F}^H(t_2) + \frac{N_0}{2}\,\delta(t_1-t_2)\mathbf{I}\,, \quad i = 1, 2. \tag{5.2.9}$$

Equation (5.2.6) becomes

$$\mathbf{K}_{a_i}\mathbf{F}^H(t) = \int_T \mathbf{H}_i(\tau)\, \mathbf{F}(\tau)\, \mathbf{K}_{a_i}\mathbf{F}^H(t)\, d\tau + \frac{N_0}{2}\int_T \mathbf{H}_i(\tau)\, \delta(\tau-t)\mathbf{I}\, d\tau\,. \tag{5.2.10}$$

Using the sifting property, equation (5.2.10) becomes

$$\mathbf{K}_{a_i}\mathbf{F}^H(t) = \int_T \mathbf{H}_i(\tau)\, \mathbf{F}(\tau)\, \mathbf{K}_{a_i}\mathbf{F}^H(t)\, d\tau + \frac{N_0}{2}\,\mathbf{H}_i(t)\,, \tag{5.2.11}$$

for i = 1, 2 and 0 < t < T. In order to solve the matrix integral equation above, let us assume $\mathbf{H}_i(t)$ can be written in the form

$$\mathbf{H}_i(t) = \frac{2}{N_0} \mathbf{G}_i \mathbf{F}^H(t) , \qquad i = 1, 2 \tag{5.2.12}$$

where $\mathbf{G}_i$ is a K×K constant matrix to be determined. Substituting equation (5.2.12) in equation (5.2.11), we get

$$\mathbf{K}_{a_i} \mathbf{F}^H(t) = \frac{2}{N_0} \int_T \mathbf{G}_i \mathbf{F}^H(\tau)\, \mathbf{F}(\tau)\, \mathbf{K}_{a_i} \mathbf{F}^H(t)\, d\tau + \mathbf{G}_i \mathbf{F}^H(t) \tag{5.2.13}$$

or

$$\mathbf{K}_{a_i} \mathbf{F}^H(t) = \frac{2}{N_0} \mathbf{G}_i \left[\int_T \mathbf{F}^H(\tau)\, \mathbf{F}(\tau)\, d\tau \right] \mathbf{K}_{a_i} \mathbf{F}^H(t) + \mathbf{G}_i \mathbf{F}^H(t) . \tag{5.2.14}$$

Recalling the expressions of the signal correlation matrix Φ and the signal–to–noise ratio:

$$\Phi = \frac{1}{NE_f} \int_T \mathbf{F}^H(t)\, \mathbf{F}(t)\, dt \tag{5.2.15}$$

and

$$\rho_N = \frac{2NE_f}{N_0} , \tag{5.2.16}$$

equation (5.2.14) becomes

$$\mathbf{K}_{a_i} \mathbf{F}^H(t) = \rho_N\, \mathbf{G}_i\, \Phi\, \mathbf{K}_{a_i} \mathbf{F}^H(t) + \mathbf{G}_i \mathbf{F}^H(t) \tag{5.2.17}$$

or

$$\mathbf{K}_{a_i}\mathbf{F}^H(t) = \mathbf{G}_i(\mathbf{I} + \rho_N \, \mathbf{\Phi} \, \mathbf{K}_{a_i})\mathbf{F}^H(t) \tag{5.2.18}$$

which implies

$$\mathbf{K}_{a_i} = \mathbf{G}_i(\mathbf{I} + \rho_N \, \mathbf{\Phi} \, \mathbf{K}_{a_i}) \, . \tag{5.2.19}$$

Therefore,

$$\mathbf{G}_i = \mathbf{K}_{a_i}(\mathbf{I} + \rho_N \, \mathbf{\Phi} \, \mathbf{K}_{a_i})^{-1} \, . \tag{5.2.20}$$

Remembering that $\mathbf{H}_i$ defined in chapter 2 and given by

$$\mathbf{H}_i = (\mathbf{I} + \rho_N \, \mathbf{K}_{a_i} \, \mathbf{\Phi})^{-1} \, \mathbf{K}_{a_i} \tag{5.2.21}$$

is a Hermitian matrix (appendix C), then

$$\mathbf{H}_i = \mathbf{H}_i^H = \mathbf{K}_{a_i}(\mathbf{I} + \rho_N \, \mathbf{\Phi} \, \mathbf{K}_{a_i})^{-1} = \mathbf{G}_i \, . \tag{5.2.22}$$

Equation (5.2.12) becomes

$$\mathbf{H}_i(t) = \frac{2}{N_0} \mathbf{H}_i\mathbf{F}^H(t) = \frac{2}{N_0} \mathbf{K}_{a_i}(\mathbf{I} + \rho_N \, \mathbf{\Phi} \, \mathbf{K}_{a_i})^{-1} \, \mathbf{F}^H(t) \, , \tag{5.2.23}$$

and the optimal MVLU estimator $\tilde{\mathbf{a}}_i$ (under $\mathbf{H}_i$) is given by (from equations 5.2.3 and 5.2.23):

$$\tilde{\mathbf{a}}_i = \frac{2}{N_0} \mathbf{K}_{a_i} (\mathbf{I} + \rho_N \, \Phi \, \mathbf{K}_{a_i})^{-1} \int_T \mathbf{F}^H(t) \, \mathbf{x}(t) \, dt \, , \quad i = 1, 2 \tag{5.2.24}$$

or more simply

$$\tilde{\mathbf{a}}_i = \mathbf{H}_i \, \mathbf{y} \, , \qquad i = 1, 2 \tag{5.2.25}$$

where $\mathbf{H}_i$ is given in equation 5.2.22 and $\mathbf{y}$ is the output of a matrix matched filter operation and is expressed as

$$\mathbf{y} = \frac{2}{N_0} \int_T \mathbf{F}^H(t) \, \mathbf{x}(t) \, dt \, . \tag{5.2.26}$$

Let us now express the likelihood ratio derived in chapter 2 in function of the estimator $\tilde{\mathbf{a}}_i$. Recalling equation 2.4.21, the i^{th} likelihood ratio is expressed in the following form:

$$\Lambda_i[\mathbf{x}(t)] = \gamma_i \exp(\mathbf{y}^H \mathbf{H}_i \mathbf{y}) \, . \tag{5.2.27}$$

This equation can be written using (5.2.25) and (5.2.26) as follows:

$$\Lambda_i[\mathbf{x}(t)] = \gamma_i \exp \left[\frac{2}{N_0} \int_T \mathbf{x}^H(t) \, \mathbf{F}(t) \, \tilde{\mathbf{a}}_i \, dt \right] \tag{5.2.28}$$

or

$$\Lambda_i[\mathbf{x}(t)] = \gamma_i \exp \left[\frac{2}{N_0} \int_T \mathbf{x}^H(t) \, \tilde{\mathbf{s}}_i(t) \, dt \, \right] \tag{5.2.29}$$

where

$$\tilde{\mathbf{s}}_i(t) = \mathbf{F}(t) \, \tilde{\mathbf{a}}_i \, . \tag{5.2.30}$$

Under H_i, $\tilde{\mathbf{a}}_i$ is the MVLU estimator of $\mathbf{a}_i$ and is given by equation (5.2.24). Thus, $\tilde{\mathbf{s}}_i(t)$ is the optimal estimator of $\mathbf{s}_i(t)$ and equation (5.2.29) is the estimator correlator structure used in chapter 2 to derive the likelihood ratio.

5.3 Variance of $\tilde{\mathbf{a}}_i$

a. Mismatched Classifier

Although classifier performance is a function of the distribution function of the likelihood ratio as given by equation (5.2.27), the result in equation (5.2.28) indicates that performance is ultimately a function of the properties of the optimal estimator of $\mathbf{a}_i$ in equation (5.2.24). Our purpose in this section is to find an expression for the variance of the estimation error given by

$$\text{var}(\tilde{\epsilon}_i) = E[\tilde{\mathbf{a}}_i - \mathbf{a}_i)(\tilde{\mathbf{a}}_i - \mathbf{a}_i)^H | H_i] , \qquad i = 1, 2 \tag{5.3.1}$$

and try to minimize it. We shall treat the problem in the general case where the processor assumes the wrong a priori scattering covariance matrix in order to find $\tilde{\mathbf{a}}_i$ (mismatched processing). For this reason, we will label the scattering covariance matrix in 5.2.24 by $\mathbf{K}'_{\mathbf{a}_i}$. The matrix $\mathbf{K}'_{\mathbf{a}_i}$ is assumed by the processor while the matrix $\mathbf{K}_{\mathbf{a}_i}$ perfectly describes the scattering properties of the test region. For matched processing discussed in part b), we will set $\mathbf{K}'_{\mathbf{a}_i} = \mathbf{K}_{\mathbf{a}_i}$. Hence, equation (5.2.24) is now expressed as follows:

$$\tilde{\mathbf{a}}_i = \frac{2}{N_0} \mathbf{K}'_{\mathbf{a}_i} (\mathbf{I} + \rho_N \, \mathbf{\Phi} \, \mathbf{K}'_{\mathbf{a}_i})^{-1} \int_T \mathbf{F}^H(t) \, \mathbf{x}(t) \, dt , \qquad i = 1, 2 \tag{5.3.2}$$

or

$$\tilde{\mathbf{a}}_i = \mathbf{H}_i' \, \mathbf{y} \,, \qquad i = 1, 2 \tag{5.3.3}$$

where $\mathbf{H}_i'$ is a Hermitian matrix given by

$$\mathbf{H}_i' = \mathbf{K}_{\mathbf{a}_i}' (\mathbf{I} + \rho_N \, \mathbf{\Phi} \, \mathbf{K}_{\mathbf{a}_i}')^{-1}, \qquad i = 1, 2 \,. \tag{5.3.4}$$

Using Hermitian properties and knowing that $\mathbf{K}_{\mathbf{a}_i} = E[\mathbf{a}_i \mathbf{a}_i^H | H_i]$, equation (5.3.1) can be expanded in the following form:

$$\text{var}(\tilde{\epsilon}_i) = E[(\tilde{\mathbf{a}}_i - \mathbf{a}_i)(\tilde{\mathbf{a}}_i - \mathbf{a}_i)^H | H_i] = E[\tilde{\mathbf{a}}_i \tilde{\mathbf{a}}_i^H | H_i] - 2E[\tilde{\mathbf{a}}_i \mathbf{a}_i^H | H_i] + \mathbf{K}_{\mathbf{a}_i} \,, \; i = 1, 2 \,. \tag{5.3.5}$$

Let us now evaluate each unknown term separately

$$E[\tilde{\mathbf{a}}_i \mathbf{a}_i^H | H_i] = E[\mathbf{H}_i' \, \mathbf{y} \, \mathbf{y}^H \, \mathbf{H}_i'^H | H_i] = \mathbf{H}_i' \, \mathbf{K}_{\mathbf{y}_i} \, \mathbf{H}_i'^H \,. \tag{5.3.6}$$

Note that $\mathbf{K}_{\mathbf{y}_i}$ is given in equation (3.2.13) by

$$\mathbf{K}_{\mathbf{y}_i} = \rho_N^2 \, \mathbf{\Phi} \, \mathbf{K}_{\mathbf{a}_i} \mathbf{\Phi}^H + \rho_N \, \mathbf{\Phi} \,, \qquad i = 1, 2 \tag{5.3.7}$$

or

$$\mathbf{K}_{\mathbf{y}_i} = \rho_N \, \mathbf{\Phi} \, (\mathbf{I} + \rho_N \, \mathbf{K}_{\mathbf{a}_i} \, \mathbf{\Phi}) \,, \qquad i = 1, 2 \,. \tag{5.3.8}$$

In this equation $\mathbf{K}_{a_i}$ is the correct a priori scattering covariance matrix since $\mathbf{y}$ is a direct function of the array measurement $\mathbf{x}(t)$. From equations (5.3.4), (5.3.6) and (5.3.8), we can write

$$E[\tilde{\mathbf{a}}_i \tilde{\mathbf{a}}_i^H | H_i] = \rho_N \mathbf{K}'_{a_i} (\mathbf{I} + \rho_N \Phi \mathbf{K}'_{a_i})^{-1} \Phi (\mathbf{I} + \rho_N \mathbf{K}_{a_i} \Phi)(\mathbf{I} + \rho_N \mathbf{K}'_{a_i} \Phi)^{-1} \mathbf{K}'_{a_i} . \tag{5.3.9}$$

Next we evaluate $E[\tilde{\mathbf{a}}_i \mathbf{a}_i^H | H_i]$. From equation (5.3.3)

$$E[\tilde{\mathbf{a}}_i \mathbf{a}_i^H | H_i] = E[\mathbf{H}'_i \mathbf{y}\, \mathbf{a}_i^H | H_i] = \mathbf{H}'_i E[\mathbf{y}\, \mathbf{a}_i^H | H_i] \tag{5.3.10}$$

where

$$E[\mathbf{y}\, \mathbf{a}_i^H | H_i] = \frac{2}{N_0} \int_T \mathbf{F}^H(t) E[\mathbf{x}(t)\, \mathbf{a}_i^H | H_i]\, dt . \tag{5.3.11}$$

From equation (2.2.10b)

$$E[\mathbf{x}(t)\, \mathbf{a}_i^H | H_i] = E[\mathbf{F}(t)\, \mathbf{a}_i \mathbf{a}_i^H + \mathbf{n}(t)\, \mathbf{a}_i^H] \tag{5.3.12}$$

or

$$E[\mathbf{x}(t)\, \mathbf{a}_i^H | H_i] = \mathbf{F}(t)\, \mathbf{K}_{a_i} . \tag{5.3.13}$$

Equation (5.3.11) becomes

$$E[\mathbf{y}\, \mathbf{a}_i^H | H_i] = \frac{2}{N_0} \left[\int_T \mathbf{F}^H(t)\, \mathbf{F}(t)\, dt \right] \mathbf{K}_{a_i} \tag{5.3.14}$$

which can be written using equations (5.2.15) and (5.2.16) as

$$E[\mathbf{y}\,\mathbf{a}_i^H \mid H_i] = \rho_N\,\mathbf{\Phi}\,\mathbf{K}_{a_i}\,. \tag{5.3.15}$$

Therefore,

$$E[\tilde{\mathbf{a}}_i\mathbf{a}_i^H \mid H_i] = \rho_N\,\mathbf{K}'_{a_i}(\mathbf{I} + \rho_N\,\mathbf{\Phi}\,\mathbf{K}'_{a_i})^{-1}\,\mathbf{\Phi}\,\mathbf{K}_{a_i}\,. \tag{5.3.16}$$

As a result, the variance of the estimation error for the mismatched classifier is expressed in equations (5.3.5), (5.3.9) and (5.3.16). In section 5.4, we simulate these results for specific test volume geometries.

b. Matched Classifier

In this part, we assume the processor has perfect a priori knowledge of the statistical properties of the test volume under hypothesis H_i. In other words, the assumed a priori covariance matrix $\mathbf{K}'_{a_i}$ is now correct. Mathematically speaking, $\mathbf{K}'_{a_i}$ should be substituted by $\mathbf{K}_{a_i}$. Equations (5.3.9) and (5.3.16) become

$$E[\tilde{\mathbf{a}}_i\tilde{\mathbf{a}}_i^H \mid H_i] = \rho_N\,\mathbf{K}_{a_i}(\mathbf{I} + \rho_N\,\mathbf{\Phi}\,\mathbf{K}_{a_i})^{-1}\,\mathbf{\Phi}\,\mathbf{K}_{a_i} \tag{5.3.17}$$

and

$$E[\tilde{\mathbf{a}}_i\mathbf{a}_i^H \mid H_i] = \rho_N\,\mathbf{K}_{a_i}(\mathbf{I} + \rho_N\,\mathbf{\Phi}\,\mathbf{K}_{a_i})^{-1}\,\mathbf{\Phi}\,\mathbf{K}_{a_i}\,. \tag{5.3.18}$$

Since for the matched classifier $E[\tilde{\mathbf{a}}_i\tilde{\mathbf{a}}_i^H \mid H_i] = E[\tilde{\mathbf{a}}_i\mathbf{a}_i^H \mid H_i]$, the right–hand side of equation (5.3.5) becomes

$$\text{var}(\tilde{\epsilon}_i) = \mathbf{K}_{a_i} - \rho_N \mathbf{K}_{a_i}(\mathbf{I} + \rho_N \mathbf{\Phi}\, \mathbf{K}_{a_i})^{-1} \mathbf{\Phi}\, \mathbf{K}_{a_i} . \tag{5.3.19}$$

Factoring out $\mathbf{K}_{a_i}(\mathbf{I} + \rho_N \mathbf{\Phi}\, \mathbf{K}_{a_i})^{-1}$, we get

$$\text{var}(\tilde{\epsilon}_i) = \mathbf{K}_{a_i}(\mathbf{I} + \rho_N \mathbf{\Phi}\, \mathbf{K}_{a_i})^{-1} [(\mathbf{I} + \rho_N \mathbf{\Phi}\, \mathbf{K}_{a_i}) - \rho_N \mathbf{\Phi}\, \mathbf{K}_{a_i}] \tag{5.3.20}$$

which implies

$$\text{var}(\tilde{\epsilon}_i) = \mathbf{H}_i = \mathbf{K}_{a_i}(\mathbf{I} + \rho_N \mathbf{\Phi}\, \mathbf{K}_{a_i})^{-1} \tag{5.3.21}$$

for i = 1, 2. Equation 5.3.21 shows the direct dependence of the estimation error covariance matrix on a priori information regarding the i^{th} scattering target's spatial properties. Moreover, the estimation error is a function of parameters characteristics of the array, signal design, etc. Investigation of $\mathbf{H}_i$ will help us understand the classifier performance. These results will be shown in the next section.

5.4 Simulation Results and Conclusions

a. Single Range Annulus

In this example, we assume that all cells are within a single range annulus and are uniform in angular extent as shown in figure 4.2.b. In chapter 4, we showed that the time delay τ_{nk} it takes the waveform f(t) to travel from the transmitter (located at the origin) to the k^{th} cell and back to the n^{th} sensor in an N–element linear array is given by

$$\tau_{nk} = \frac{2r}{c} - \frac{(n-1)d}{c} \sin[(k-1)\,\Delta\theta]\ ;\ n = 1, \ldots, N;\ k = 1, \ldots, K \tag{5.4.1}$$

where r is the annulus range, $\Delta\theta$ is the angular extent between cells and c is the speed of sound in the medium. The values of $\Delta\theta$ specify the angular resolution of the processor. To proceed with calculations, let us further assume that the scattering covariance matrix $\mathbf{K}_a$ is given by

$$\mathbf{K}_a = \sigma_0^2\,\mathbf{I}\,. \tag{5.4.2}$$

In other words, the scattering strength of the target is uniformly distributed among each of the K cells of the annulus. For the matched classifier, the variance of $\tilde{\mathbf{a}}$ is given by equation (5.3.21) as

$$\mathrm{var}(\tilde{\epsilon}) = \mathbf{H} = \mathbf{K}_a(\mathbf{I} + \rho_N\,\Phi\,\mathbf{K}_a)^{-1} \tag{5.4.3}$$

which, using equation (5.4.2), becomes

$$\mathrm{var}(\tilde{\epsilon}) = \sigma_0^2(\mathbf{I} + \rho_N\,\sigma_0^2\,\Phi)^{-1}\,. \tag{5.4.4}$$

We define the variance ratio as the trace of the error variance matrix normalized by the trace of $\mathbf{K}_a$. For this example, it is given by

$$\frac{\mathrm{tr}[\mathrm{var}(\tilde{\epsilon})]}{\mathrm{tr}[\mathbf{K}_a]} = \frac{1}{K}\,\mathrm{tr}[\mathbf{I} + \rho_N\,\sigma_0^2\,\Phi]^{-1}\,. \tag{5.4.5}$$

The elements of Φ in function of time delays are discussed thoroughly in chapter 4. Equation (5.4.5) shows that the variance ratio is a function of $\rho_N \sigma_0^2$ (received signal–to–noise ratio) and the resolution parameter $\Delta\theta$. For the particular case when $\Phi = \mathbf{I}$, equation (5.3.5) becomes

$$\frac{\mathrm{tr}[\mathrm{var}(\tilde{\epsilon})]}{\mathrm{tr}[\mathbf{K}_a]} = \frac{1}{1+\rho_N \sigma_0^2} . \tag{5.4.6}$$

This equation shows that minimizing the variance ratio can be achieved with increased received signal–to–noise ratio $\rho_N \sigma_0^2$. In the graphs presented in this chapter, the variance ratio is calculated for N = 8 array elements (with d = $\lambda/2$ distance between them), a pulse duration T of 30 ms, a range of 1000 m and center frequency f_0 of 10 Khz. Also, the speed of sound c in water equals 1500 m/sec and the beamwidth BW measured between the first nulls is given by

$$\mathrm{BW} = 2 \sin^{-1}\left[\frac{\lambda}{Nd}\right] = 29^\circ . \tag{5.4.7}$$

In figures 5.1 and 5.2, we display the variance ratio (in dB) versus the angular width $\Delta\theta$ and the received signal–to–noise ratio $\rho_N \sigma_0^2$ (in dB) for K = 2 and K = 8, respectively. In this case, the transmitted waveform is a pure tone pulse. It is clear that as the classifier processes high angular resolution, in other words as $\Delta\theta$ decreases, the error variance increases. The cost one has to pay for good resolution is a large signal–to–noise ratio. The figures also illustrate the performance dependence on the number of cells in the test region. For instance, if a variance ratio of –10 dB at an angular extent $\Delta\theta$ = BW/8 is desired, subdividing the test region in 2 cells requires a received signal–to–noise

ratio of about 12 dB. However, if the test region is subdivided in 8 cells, the processor requires a received signal–to–noise ratio of 45 dB. We can also deduce from the graphs that the high resolution signal (Φ = I) has the lowest error variance as compared to a pure tone pulse. This conclusion is true regardless of the angular extent, the signal–to–noise ratio, or the number of cells in the test region. Therefore, signal design is an important aspect in target imaging.

Next we discuss variance reduction as we increase the transmitter constant k_f for linear FM pulses. Figures 5.3, 5.4 and 5.5 illustrate variance reduction for a test region subdivided in 2 cells when $k_f = 10^3$, 10^5 and 10^7, respectively. We see that the estimator variance decreases with a large k_f. For $k_f = 10^3$, an angular extent of BW/64 increases the variance ratio (with respect to an angular extent of BW/32) by about 38 dB for received signal–to–noise ratios greater than 40 dB. This variance ratio is increased no more than 8 dB for received signal–to–noise ratios as low as 10 dB or higher when $k_f = 10^5$. We also notice in this case that angular widths of BW/8, BW/16 and BW/32 have a variance as low as the high resolution signal for $\Delta\theta$ = BW/32 or $\Delta\theta \leq$ BW/16. With k_f equal to 10^7, the variance ratio is more or less independent of the angular resolution (overlapping curves). Here we reduce the variance by about 7 dB with respect to the high resolution performance for received signal–to–noise ratios as low as 0 dB or higher. This performance tends to be a limit in variance reduction. In fact, further increase in k_f does not reduce the error variance noticeably compared to $k_f = 10^7$. Higher values of k_f would be impractical and the cost one has to pay for such large values is expensive transducers because of the bandwidth required to transmit these signals.

Figures 5.6, 5.7 and 5.8 illustrate a similar situation for K = 8 cells. Figure 5.6 ($k_f = 1000$) is similar to figure 5.2 ($k_f = 0$). A transmitter constant $k_f = 10^5$ (figure 5.7) reduces the variance ratio when $\Delta\theta = BW/8$. In this case, there is not much variance reduction for $\Delta\theta < BW/16$. Increasing k_f to 10^7 gives the desired variance reduction. Here also performance of linear FM signals with large k_f is better than that of high resolution signals. We conclude that variance reduction is better achieved when linear FM pulses (as compared to pure tone pulses), with a relatively large k_f (typically 10^7), illuminate a test region. Depending on the value of k_f, the processor can achieve a variance reduction as good as that of high resolution signals (if not better). The cost to be paid with more cells in the test region is a larger transmitter constant and/or a larger received signal–to–noise ratio. Hence, linear FM pulses have a better performance than pure tone pulses for target classification and imaging.

b. Two–Dimensional Test Region

We now consider the two–dimensional test region shown in figure 4.3. The receiver is also a uniform linear array of 8 elements. In chapter 4, we discussed in detail how the elements of Φ can be computed in function of time delays τ_{nk} (equation 4.3.28) for both a pure tone and a linear FM pulse. We will apply the following examples to the test volume geometries depicted in figure 4.5 (matched classifier) and figure 4.11 (mismatched classifier).

First, we examine variance reduction for a pure tone pulse in function of the angular extent $\Delta\theta$. Figure 5.9 shows that the classifier achieves a very poor variance reduction even at the cost of a small angular resolution ($\Delta\theta = BW/2$) and a large received signal–to–noise ratio (80 dB). In this case, the variance reduction achieved by high resolution signals is very desirable (–20 dB) at a

received signal–to–noise ratio of about (25 dB). Here also, we used J = 4 target–like cells and the reverberation noise energy per cell is $\sigma_R^2 = 10$. All other parameters have been assigned the same values used in part a). We can make the first conclusion: in order to obtain a desirable variance reduction of the estimator $\tilde{\mathbf{a}}_2$ for the matched classifier, the transmitted signal f(t) should approach high resolution characteristics. This can be achieved by illuminating the target with a linear FM signal having a large transmitter constant k_f.

Next, we examine variance reduction when wrong a priori knowledge of the test region scattering properties is fed to the classifier. Here also, f(t) is a pure tone pulse, J = 4 target–like cells, $\Delta\theta$ = BW/8 and $\sigma_R^2 = 10$. All other parameters are assigned the values used in part a). In figures 5.10 and 5.11, we illustrate variance reduction when the target is mismatched in range and orientation, respectively. It is clear that target imaging, when the test region is illuminated by a pure tone pulse, is very sensitive to wrong a priori information.

From the above discussion, we conclude that when a two–dimensional test region is interrogated by a pure tone pulse, the estimator $\tilde{\mathbf{a}}_i$ (estimated image of the test region) cannot be achieved with a low variance error. This conclusion holds true for both the matched and mismatched classifier. Moreover, variance reduction for misoriented targets (refer to figure 4.11) is independent of the target azimuth. Therefore, linear FM signals provide a better performance for target imaging. This concludes our discussion of chapter 5.

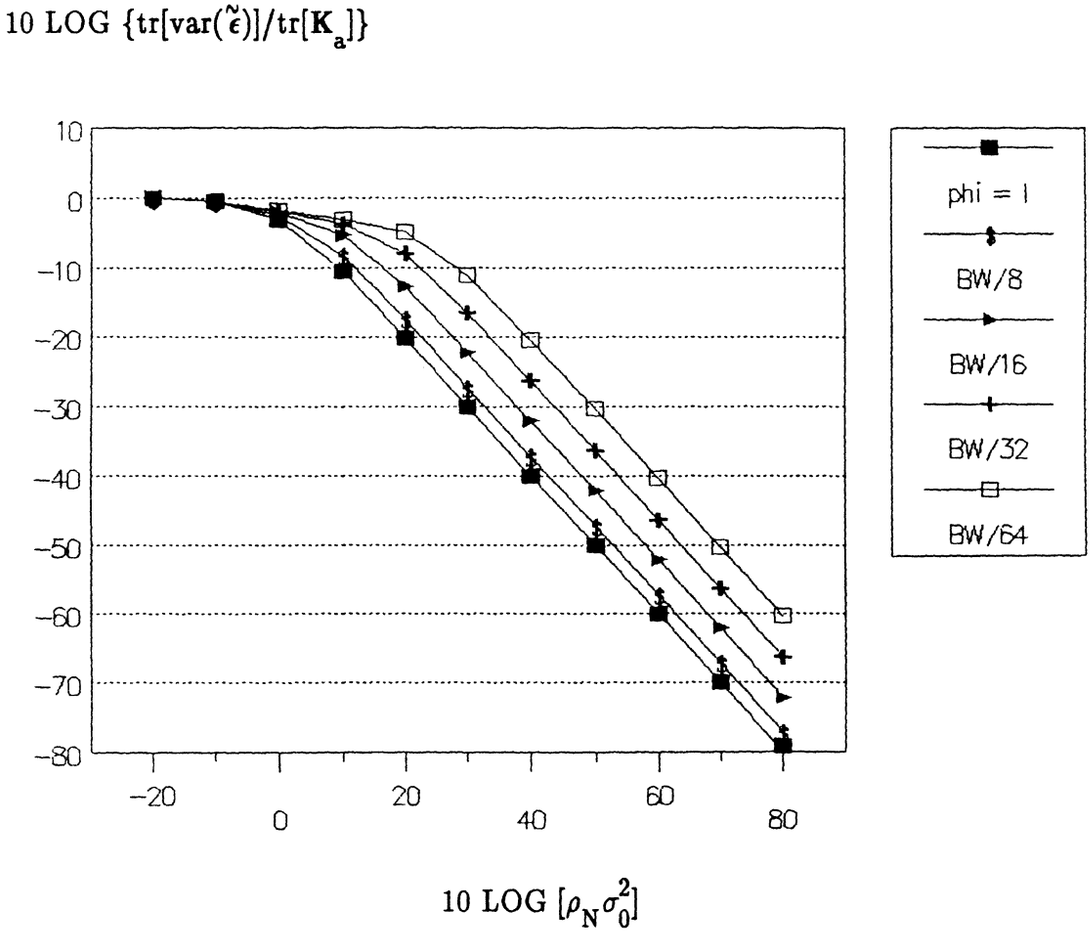


Figure 5.1 Variance ratio tr[var($\tilde{\epsilon}$)]/tr[$\mathbf{K}_a$] versus angular width $\Delta\theta$ and received signal–to–noise ratio for 2 angle cells. f(t) is a pure tone.

10 LOG $\{tr[var(\tilde{\epsilon})]/tr[\mathbf{K}_a]\}$

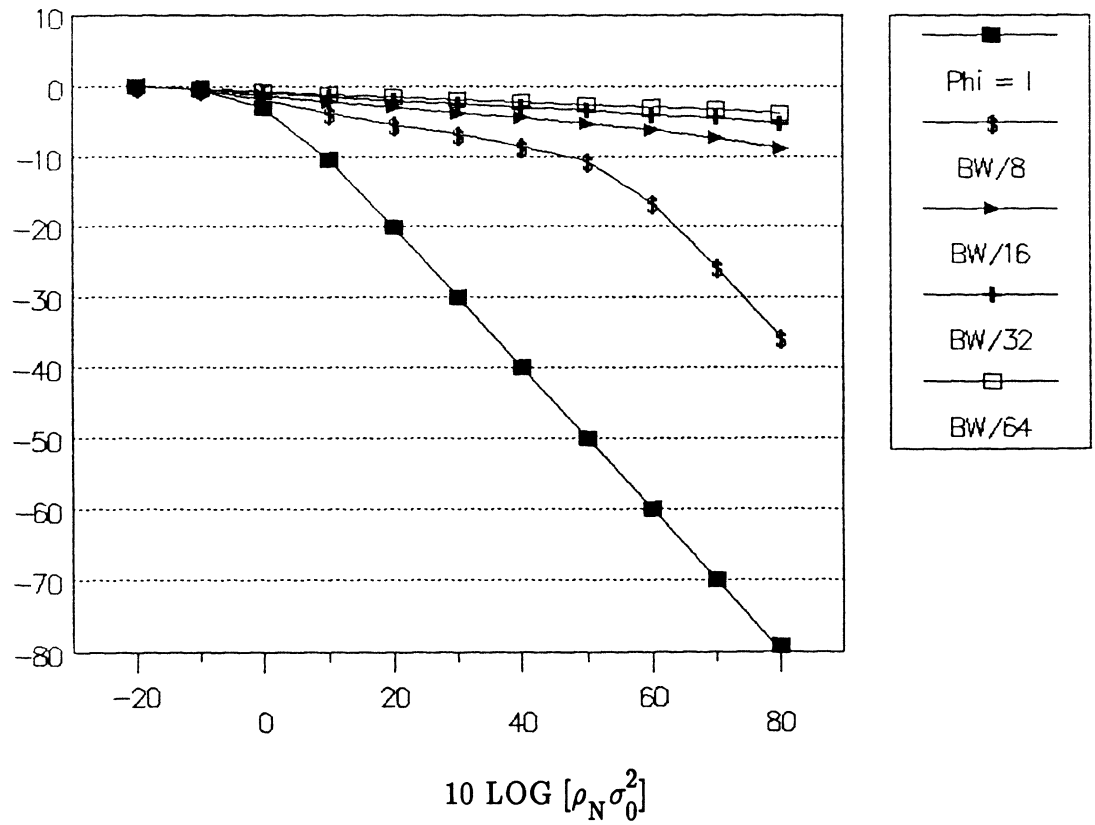


Figure 5.2 Variance ratio $tr[var(\tilde{\epsilon})]/tr[\mathbf{K}_a]$ versus angular width $\Delta\theta$ and received signal–to–noise ration for 8 angle cells. f(t) is a pure tone pulse.

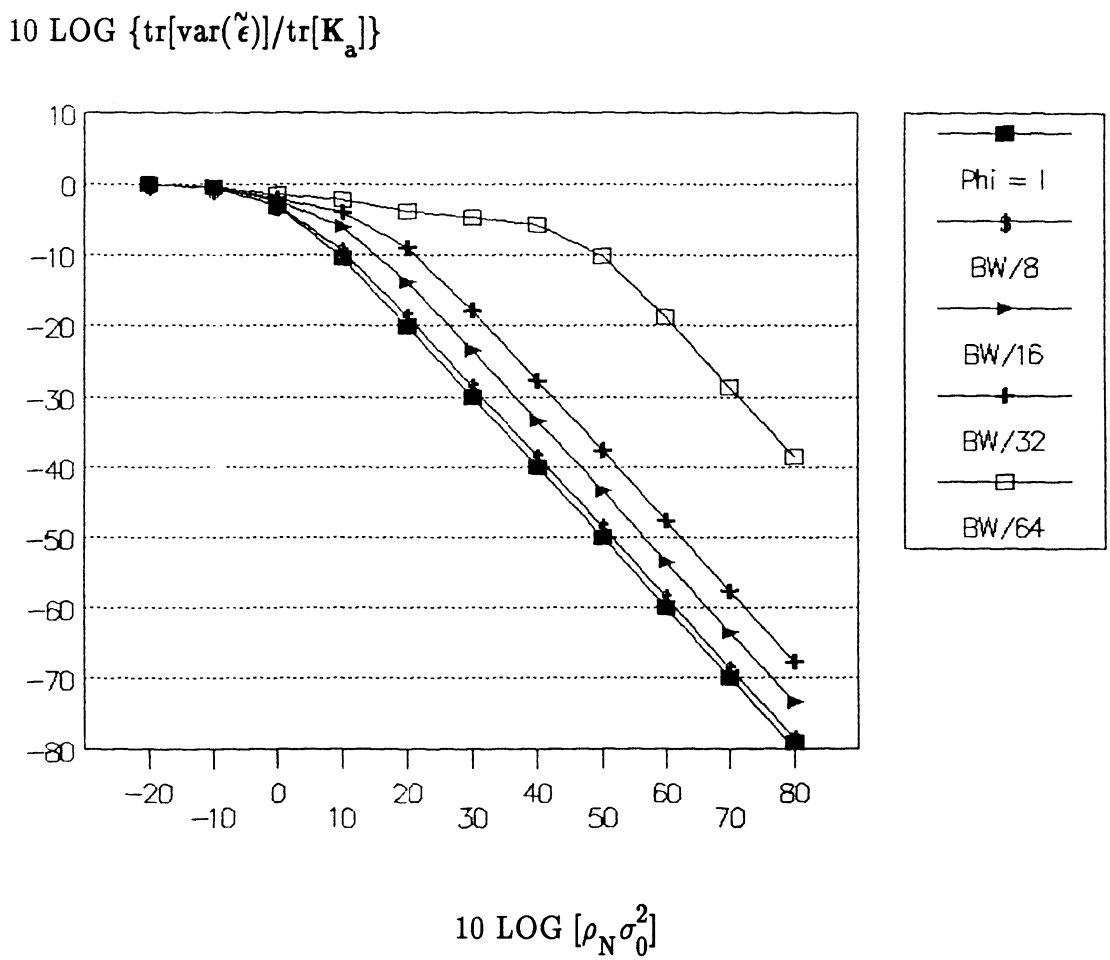


Figure 5.3 Variance ratio tr[var($\tilde{\epsilon}$)]/tr[$\mathbf{K}_a$] versus angular width $\Delta\theta$ and received signal–to–noise ratio for 2 angle cells. f(t) is a linear FM pulse with $k_f = 10^3$.

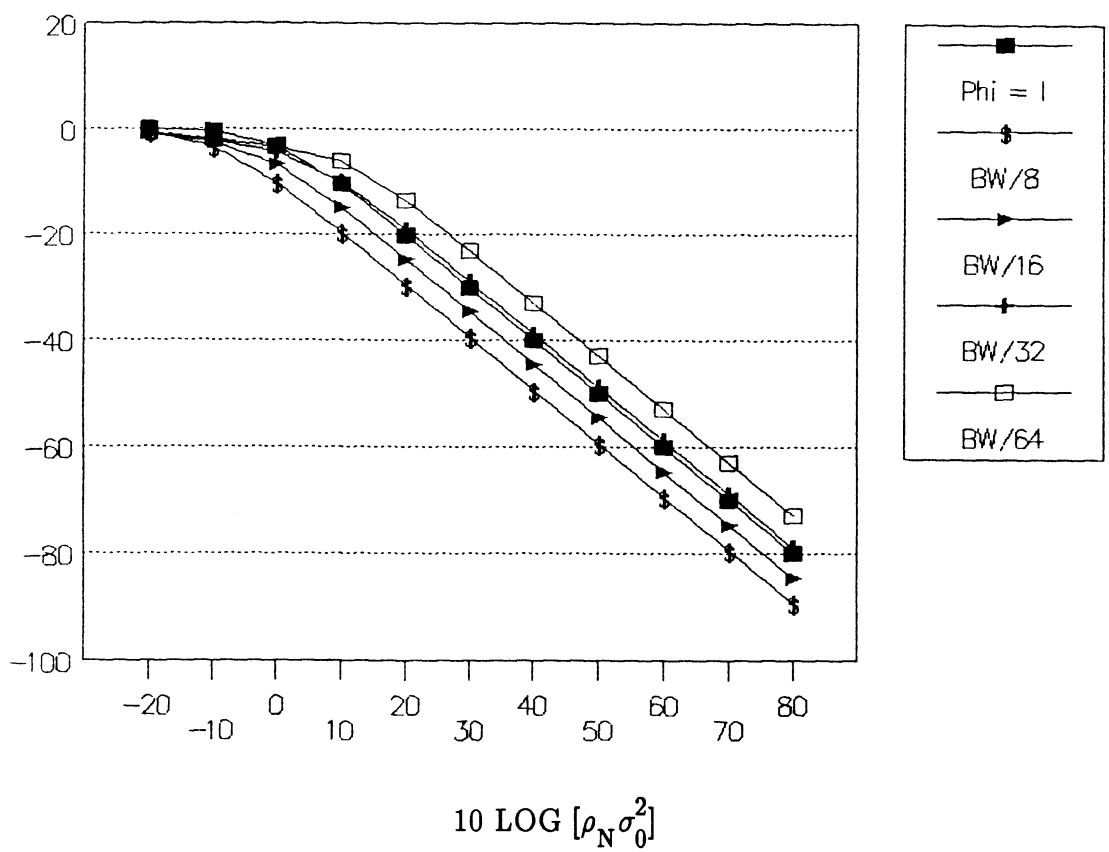


Figure 5.4 Variance ratio tr[var($\tilde{\epsilon}$)]/tr[$\mathbf{K}_a$] versus angular width $\Delta\theta$ and received signal–to–noise ratio for 2 angle cells. f(t) is a linear FM pulse with $k_f = 10^5$.

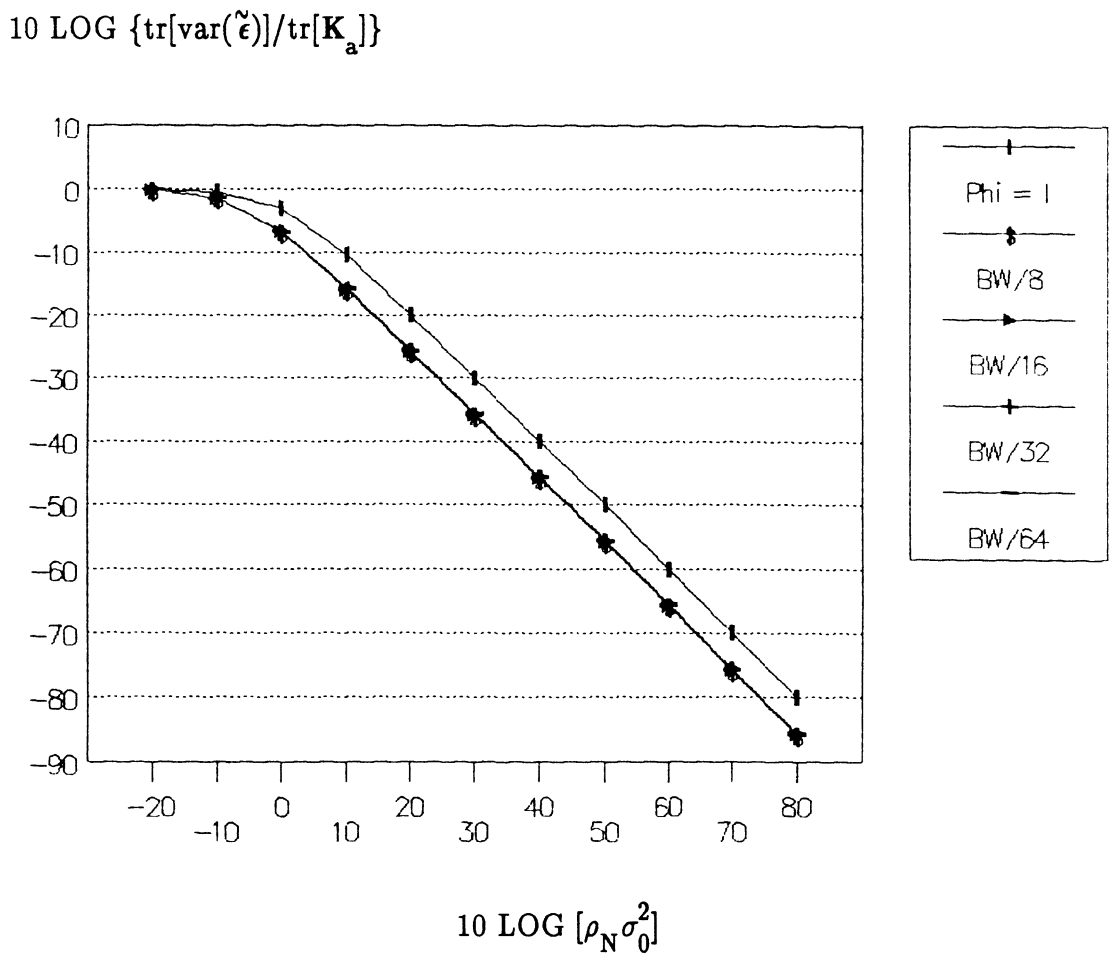


Figure 5.5 Variance ratio tr[var($\tilde{\epsilon}$)]/tr[$\mathbf{K}_a$] versus angular width $\Delta\theta$ and received signal–to–noise ratio for 2 angle cells. f(t) is a linear FM pulse with $k_f = 10^7$.

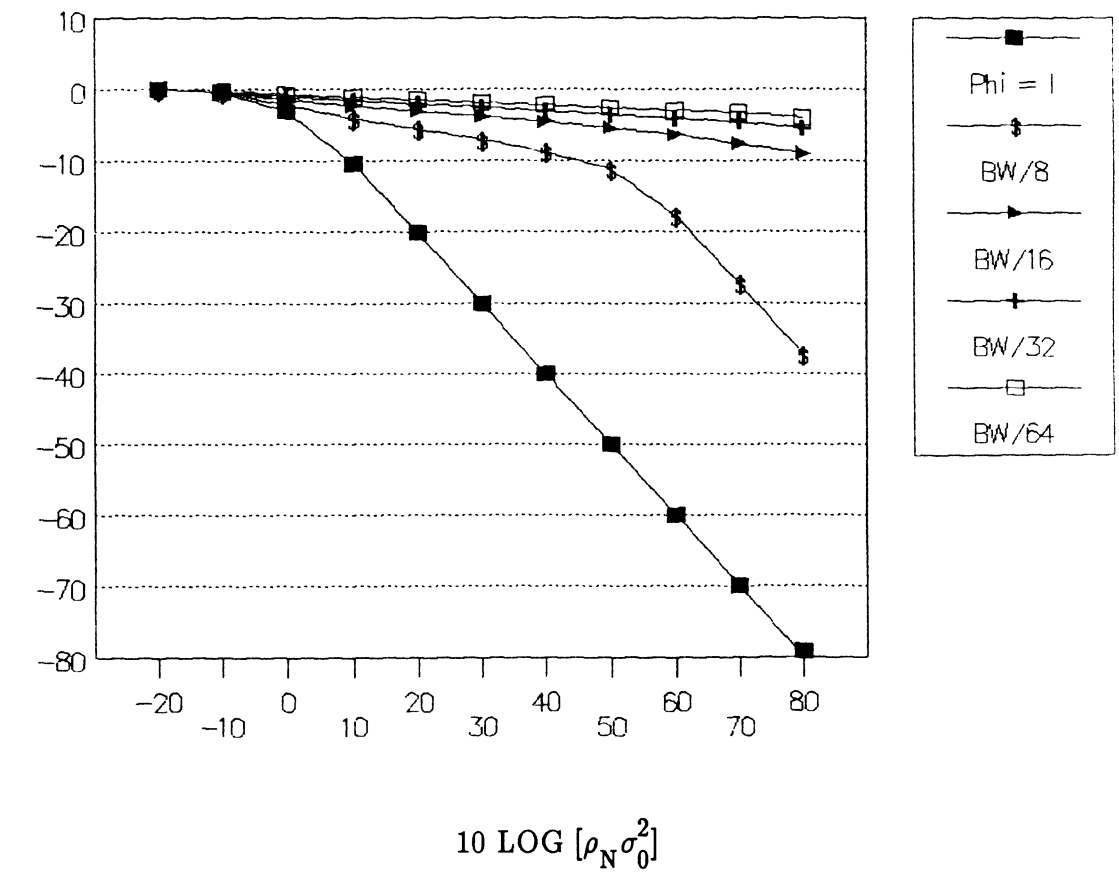


Figure 5.6 Variance ratio tr[var($\tilde{\epsilon}$)]/tr[K_a] versus angular width $\Delta\theta$ and received signal–to–noise ratio for 8 angle cells. f(t) is a linear FM pulse with $k_f = 10^3$.

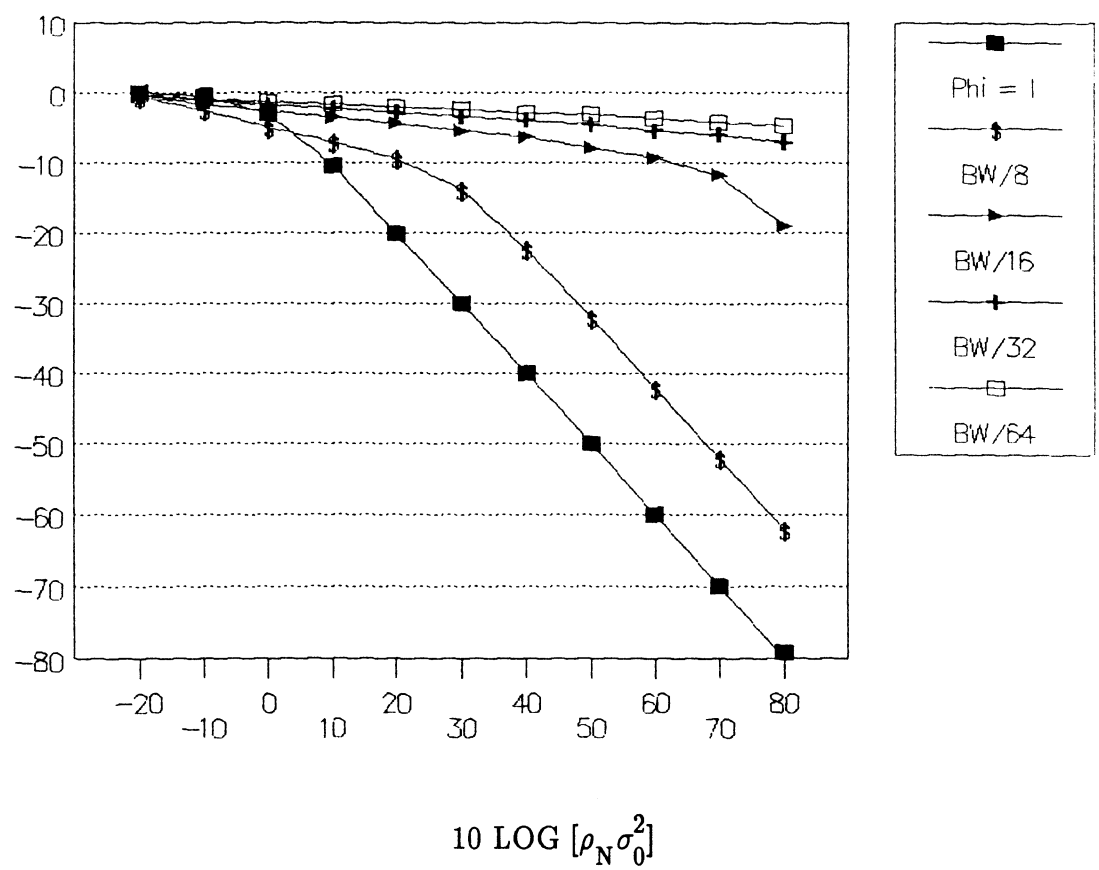


Figure 5.7 Variance ratio tr[var($\tilde{\epsilon}$)]/tr[$\mathbf{K}_a$] versus angular width $\Delta\theta$ and received signal–to–noise ratio for 8 angle cells. f(t) is a linear FM pulse with $k_f = 10^5$.

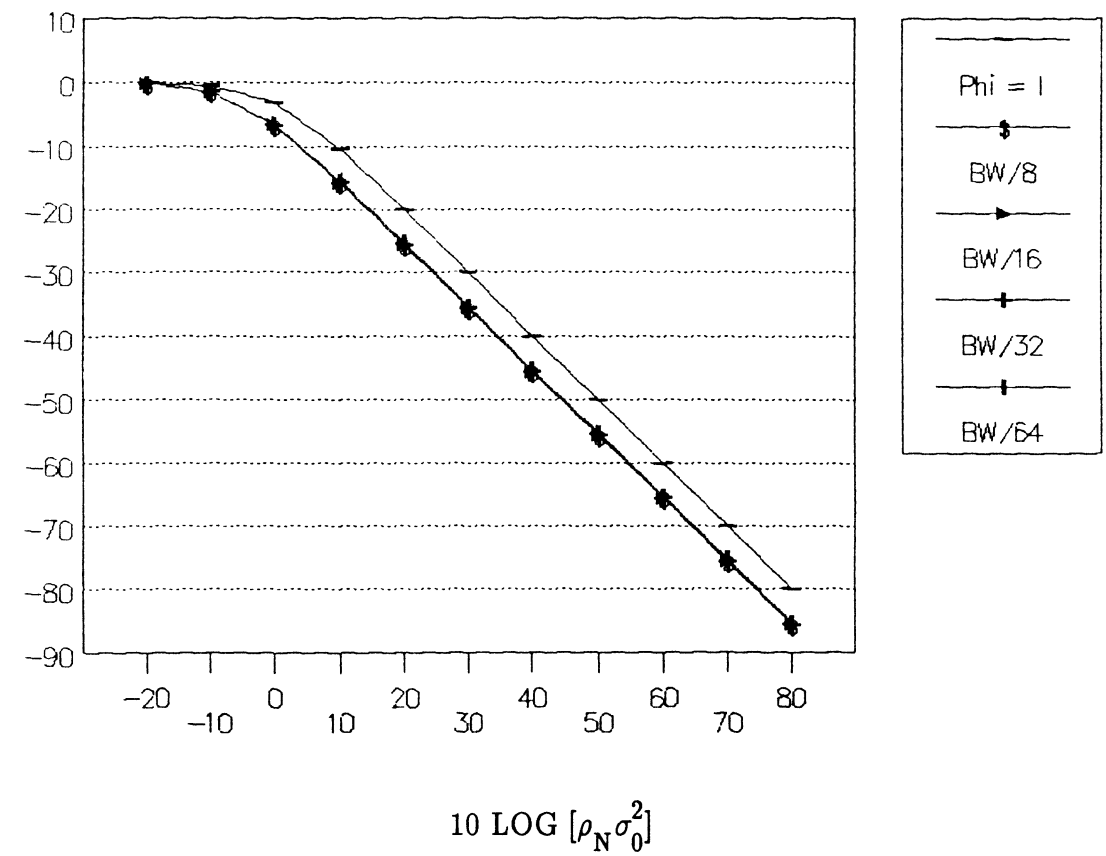


Figure 5.8 Variance ratio tr[var($\tilde{\epsilon}$)]/tr[$\mathbf{K}_a$] versus angular width $\Delta\theta$ and received signal–to–noise ratio for 8 angle cells. f(t) is a linear FM pulse with $k_f = 10^7$.

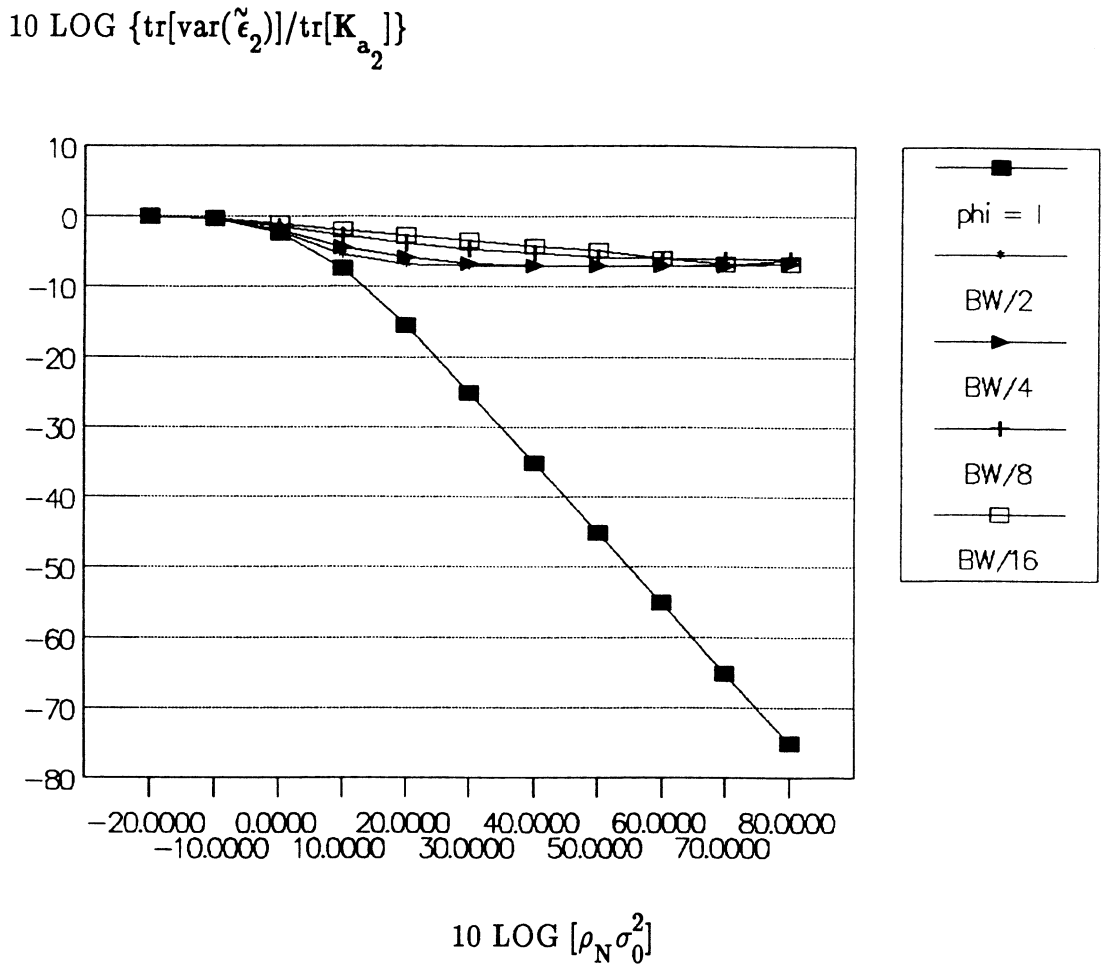


Figure 5.9 Variance ratio tr[var($\tilde{\epsilon}_2$)]/tr[$\mathbf{K}_{a_2}$] versus angular width $\Delta\theta$ and received signal–to–noise ratio for the two–dimensional test region of figure 4.5. Here J = 4 target–like cells and f(t) is a pure tone pulse.

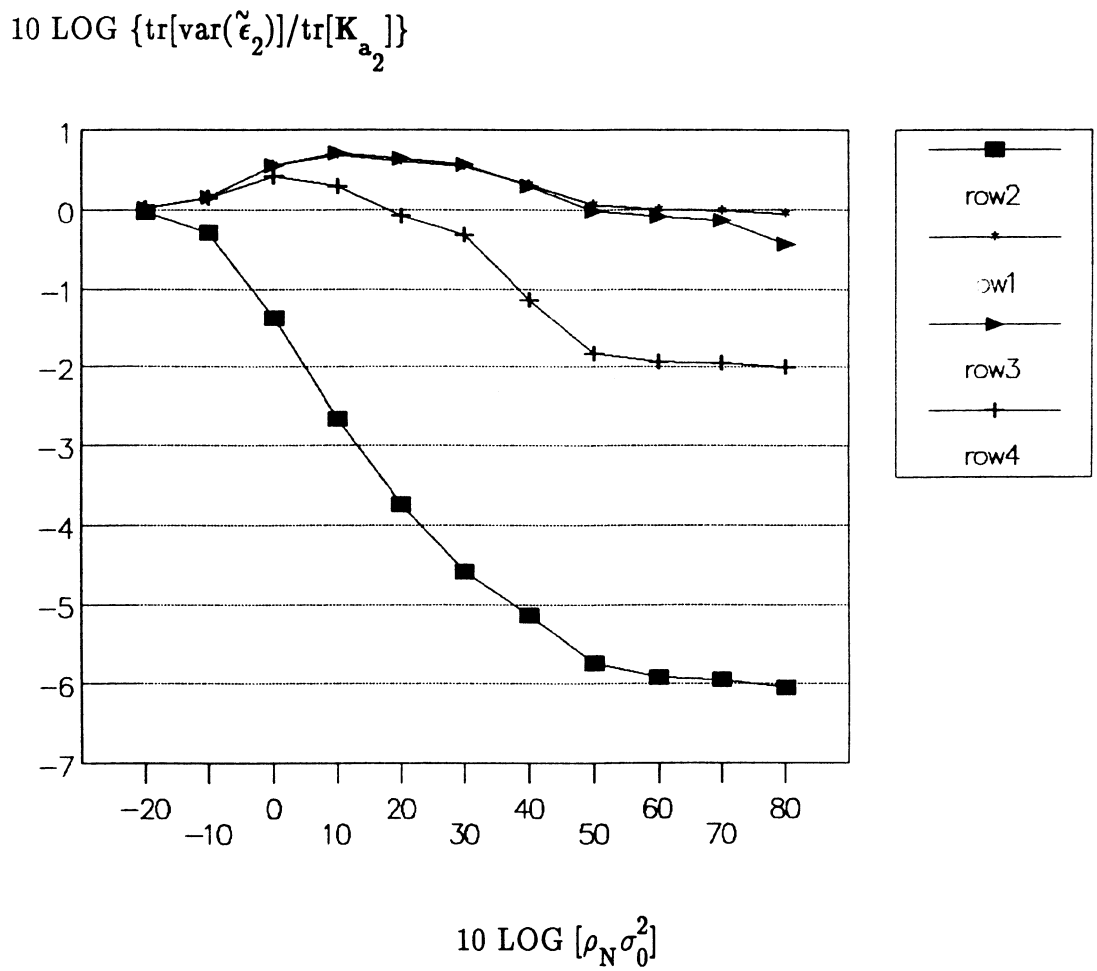


Figure 5.10 Variance ratio tr[var($\tilde{\epsilon}_2$)]/tr[$\mathbf{K}_{a_2}$] versus received signal–to–noise ratio for the two–dimensional test region of figure 4.11. Here the mismatch is in the target range, J = 4 target–like cells and f(t) is a pure tone pulse.

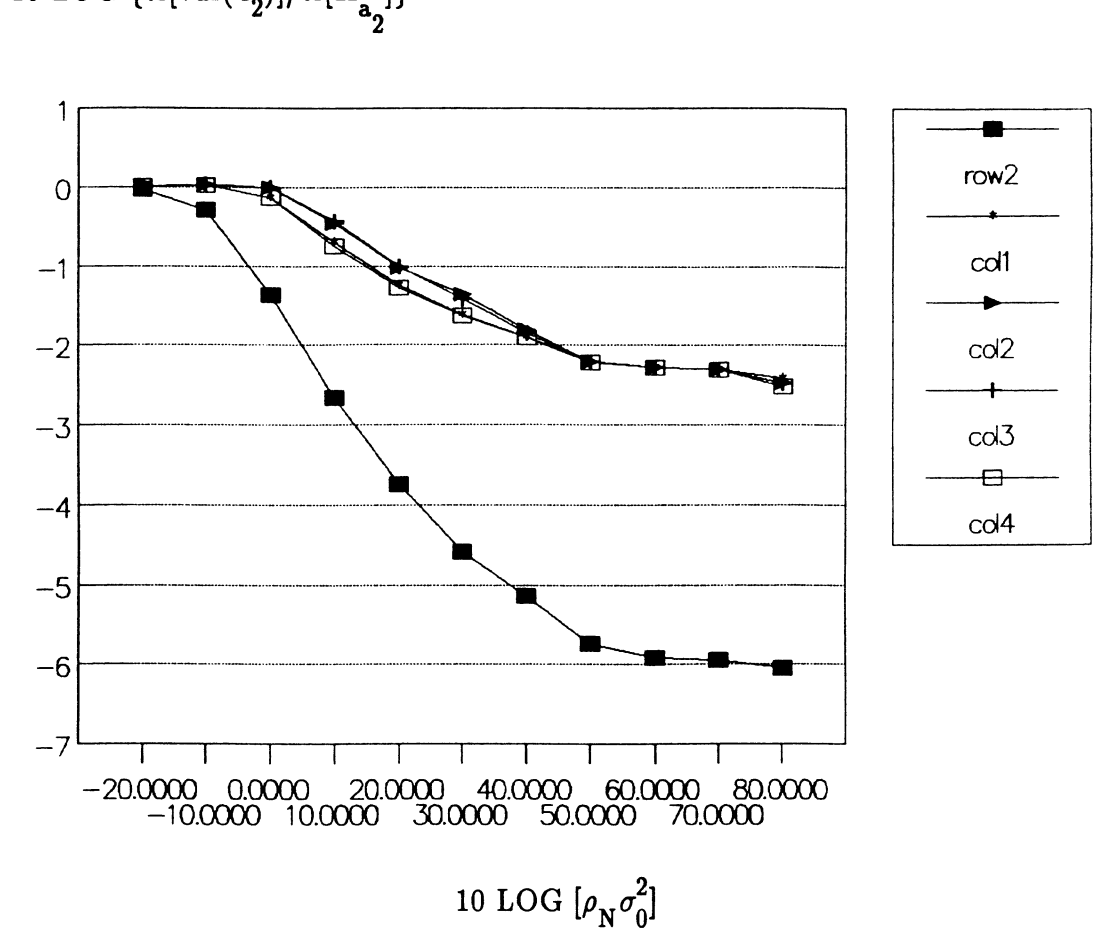


Figure 5.11 Variance ratio tr[var($\tilde{\epsilon}_2$)]/tr[$\mathbf{K}_{a_2}$] versus received signal–to–noise ratio for the two–dimensional test region of figure 4.11. Here the mismatch is in the target orientation, J = 4 target–like cells and f(t) is a pure tone pulse.

Chapter 6

CONCLUDING REMARKS

6.1 Conclusions

A systematic approach for analyzing the performance of Bayesian optimal classifiers has been presented. Closed–form solutions for classification and false alarm probabilities were obtained and several receiver operating characteristics illustrating meaningful physical scenarios were presented. Two classes of transmitted signals in the sonar system were considered: a) high resolution signals having a signal correlation matrix equal to the identity matrix and b) linear FM and pure tone pulses. Also, a minimum value linear unbiased estimator of the random vector $\mathbf{a}_1$ was developed and different methods of minimizing the estimator error were investigated.

When high resolution signals interrogate the test volume, we can make two general conclusions:

1. The classifier is very sensitive to incorrect a priori knowledge of the scattering properties of the test volume being interrogated. Its ability to distinguish between targets decreases as the difference between actual and assumed test volume geometries increases. If the classifier assumes the wrong target orientation, its performance is directly proportional to the number and energy of the target–like cells which overlap between the assumed and nominal geometries.
2. If an incorrect range is assumed and there is no overlap between target–like cells, the classifier cannot distinguish between H_1 and H_2 at all. This conclusion can be reached with an intuitive argument. Since perfect cell resolution was presupposed; therefore, when an incorrect

range estimate is infused into the processor, only return from reverberation cells is processed. However, the only statistical difference between H_1 and H_2 is the difference between total scattering strengths. When reverberation energy alone is processed, it is not possible to distinguish between target–like and reverberation–like energy; consequently, classification is not possible.

Performance of linear FM signals in active sonar systems can be summarized as follows:

1. Classifiers performance is improved with the following parameters: a) the transmitter constant k_f, b) target strength, and c) the signal–to–noise ratio. Performance degrades with finer angular resolution. An extensive quantitative discussion has been presented in chapter 4. We can make the following suggestions for implementation purposes. For a test volume subdivided into 16 cells (target occupies 4 cells), we generally suggest a linear FM pulse having a transmitter constant k_f = 100 and a pulse duration of 30 ms. The signal–to–noise ratio should be no less than 10 dB and an angular extent of 3.625° generally provides acceptable performance. These values lead to a classification probability of 0.8 or greater while keeping the false alarm probability constrained to 0.1. Here we assumed that the reverberation strength per cell (σ_R^2) equals 10 and the target strength per cell (σ_0^2) equals 40.
2. The classifier is sensitive to mismatched a priori information. If the processor assumes the wrong target orientation, the degradation in performance is independent of the assumed azimuth. In other words, performance depends on the number of intersection cells between the

assumed and actual target geometries rather than the azimuth of the mismatched target. This result was predicted when high resolution signals were discussed.

3. When the target range is misestimated, performance is directly proportional to the range's estimation error. If this error is small, the classifier performs relatively well. We conclude that the processor is not as sensitive (due to blurring) to range misidentification when targets are illuminated by linear FM signals as opposed to high resolution signals. For instance, if the target length is 100 m and its range is 1000 m, a range estimation error less than or equal to 25 m leads to a classification probability of 0.6. If the range estimation error is 50 m or more the classification probability becomes 0.3 or lower. If the exact range is known, the classification probability is 0.7. In this discussion, the false alarm probability is also constrained to 0.1.

We can also draw the following conclusions regarding acoustic target imaging:

1. Imaging the target is performed by estimating the test volume scattering coefficients (vector $\tilde{\mathbf{a}}_1$). The error variance of the estimation error depends on system characteristics such as array/test volume geometry and transmitted waveform.
2. The best images are generally obtained with high resolution signals. Good images can also be obtained with linear FM pulses having a large transmitter constant. For the single annulus geometry, physically representing a reverberation–free environment, we suggest a transmitter constant of 10^7 when the target occupies 8 cells. With such a large value

of k_f, we even achieve better variance reduction than with high resolution signals. These images can have a very fine angular resolution ($\Delta\theta$ = BW/64). If the target occupies 2 cells, we suggest a transmitter constant of 10^5. Such large values of k_f are limited by transducer bandwidth.

3. In a reverberating environment, variance reduction is not easily achieved. For instance, if $k_f = 0$ (pure tone), the variance ratio is about –6 dB at a received signal–to–noise ratio as high as 80 dB. Using the same parameters in a reverberation–free environment, we were able to obtain a variance ratio of at least –70 dB.

4. A bad image does not necessarily imply a bad classifier. A plot of the classification probability versus the variance ratio, while keeping the false alarm probability constrained to 0.1, will provide more quantitative results and needs to be done.

6.2 Recommendations for Future Research

Several interesting open problems need to be solved. First, we can relax some of the assumptions made about the noise process. The assumption of a Gaussian density for the noise stems from the central limit theorem and continues to be a popular assumption in classical detection since it leads to simpler detection structures. However, recent studies of data from a number of noise environments show that noise is often non–stationary, non–white and non–Gaussian [30, 31, 32]. Hence, detectors designed to operate under the assumption of white Gaussian noise may result in severe degradations in performance when operated in a non–white non–Gaussian noise environment. For this reason, the design of a non–parametric classifier structure, designed

with minimal knowledge of the noise characteristics, may be appropriate and ought to be researched. Although non–parametric detectors are simple in structure, their conservative performance compared with optimum detectors is a major drawback [33]. Another alternative would be an adaptive classifier which can adapt or learn the characteristics of the noise process by changing its structure with time. Such structures are usually very complex but result in nearly optimum performance.

Second, we assumed that the scattering coefficients are uncorrelated and frequency independent at least over the band of f(t). A relaxation of these assumptions will undoubtedly lead to a more complete and hence better model.

Third, the problem needs to be studied when the target and the array are in relative motion. At this point, we can predict that the only parameter that will vary is the signal correlation matrix. It will now be a function of the target's velocity. A helpful discussion of targets in motion is given in references [12, 17 and 27].

Finally, one can further investigate the system characteristics, mainly array and signal design. For instance, a non–uniform linear array or a planar array can be studied, and another class of transmitted signals may have an attractive performance. More work on the relation between number of cells and classifier performance needs to be done as well.

REFERENCES

1. Thomas Bayes, "An Essay Towards Solving a Problem in Doctrine of Chances", Phil. Trans. 53, 370–418 (1764).
2. A.M. Legendre, Nouvelles Methodes pour La Determination ces Orbites des Cometes, Paris, 1806.
3. K.F. Gauss, Theory of Motion of the Heavenly Bodies Moving About the Sun in Conic Sections, reprinted by Dover, New York, 1963.
4. R.A. Fisher, "Theory of Statistical Estimation", Proc. Cambridge Philos. Soc. 22, 700 (1925).
5. J. Neyman and E.S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses", Phil. Trans. Roy. Soc. London A231, 289 (1933).
6. D. Middleton and D. Van Meter, "Detection and Extraction of Signals in Noise from the Point of View of Statistical Decision Theory", SIAM, Vol. 3, No. 4, December 1955.
7. D. Middleton and D. Van Meter, "On Multiple–Alternative Detection of Signals in Noise", IRE Transactions on Information Theory, Vol. IT–7, September 1955, pp. 1–9.
8. J.B. Thomas and J.K. Wolf, "On the Statistical Detection Problem for Multiple Signals", IRE Transactions on Information Theory, July 1962.
9. J.B. Tucker, "Cold War in the Ocean Depths", High Technology, July 1985.
10. D. Middleton and W.N. Pugliese, "Active Underwater Acoustic Classification: Some Old Problems and New Approaches in Modeling and Processing", NUSC Technical Report 8067, August 12, 1987.

11. H.L. Van Trees, Detection, Estimation and Modulation Theory, Vol. 1 (New York: John Wiley and Sons, Inc., 1968).

12. H.L. Van Trees, Detection, Estimation and Modulation Theory, Vol. 3 (New York: John Wiley and Sons, Inc., 1968).

13. H.L. Van Trees, "A Unified Theory for Optimum Array Processing", Report 4160866, A.D. Little, Inc., August 1966.

14. A.D. Whalen, Detection of Signals in Noise (New York: Academic Press, 1971).

15. C.R. Rao, Linear Statistical Inference and Its Applications (New York: John Wiley and Sons, Inc., 1965).

16. T.S. Ferguson, Mathematical Statistics: A Decision Theoretic Approach (New York: Academic Press, 1967).

17. J.G. Kelly and R.N. Carpenter, "A Bayesian Approach to Acoustic Imaging and Object Classification by High Frequency Sonar", NUSC Technical Report 6836, May 15, 1989.

18. J.A. Tague, "Estimation–Correlation, Modeling and Identification in Adaptive Array Processors", Ph.D. Dissertation, The Pennsylvania State University, 1987.

19. J.G. Kelly and R.N. Carpenter, "Classification with Active Sonar: Theoretical Results", SACLANT Symposium on Future Uses of Towed and Static Arrays, La Spezia, Italy, October–November, 1988.

20. G. Turin, "The Characteristic Function of Hermitian Quadratic Forms", Biometrika, Vol. 47, pp. 199–201.

21. C. Kharti, "On Certain Distribution Problems Based on Positive Definite Quadratic Functions in Normal Vectors", Annals of Mathematical Statistics, Vol. 35, pp. 1807–1810.

22. N.L. Johnson and S. Kotz, Distributions in Statistics: Continuous Multivariate Distribution (New York: John Wiley and Sons, 1972).

23. A. Erdelyi, editor, Higher Transcendental Functions, Vol. I (New York: McGraw–Hill, 1953).

24. R.V. Churchill and J.W. Brown, Complex Variables and Applications (New York: McGraw–Hill, 1984).

25. G. Strang, Linear Algebra and Its Applications, Second Edition. (Academic Press, 1976), p. 256.

26. F.G. Stremler, Introduction to Communication Systems, Second Edition. (Addison–Wesley Publishing Company, 1982).

27. N. Levanon, Radar Principles (John Wiley and Sons, 1988).

28. A.P. Sage and J.L. Melsa, Estimation Theory with Applications to Communications and Control (New York: McGraw–Hill Book Co., 1971).

29. A.V. Balakrishnan, Communication Theory (New York: McGraw Hill Book Co., 1968).

30. S.L. Bernstein, M.L. Burrows, J.E. Evans, et al., "Long–Range Communications at Extremely Low Frequencies", Proc. IEEE 62(3): 292–312, March 1974.

31. A.R. Milne and J.H. Ganton, "Ambient Noise Under Arctic–Sea Ice", Journal of Acoustic Society of America 36(5): 855–863, May 1964.

32. J.G. Vietche and A.R. Wilks, "A Characterization of Arctic Undersea Noise", Journal of Acoustic Society of America 77(3): 989–999, March 1985.

33. J.B. Thomas, "Nonparametric Detection", Proc. IEEE 58: 623–631, May 1970.

APPENDIX A

In this appendix, we evaluate the following integral

$$I = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-j\omega x}}{(1-j\omega\alpha)^J} d\omega \tag{A.1}$$

using contour integration. We assume that α is a negative constant. J is a positive integer and ω and x are real variables. (A.1) can be written as

$$I = \frac{1}{2\pi} \frac{1}{(-j\alpha)^J} \int_C \frac{e^{-jzx}}{(z+j/\alpha)^J} dz \tag{A.2}$$

where C is the counterclockwise contour (figure A.1) enclosing the pole $z_1 = -j/\alpha$ and z is a complex variable. The contour integral above can be written as

$$\int_C f(z)\, dz = \int_{-\infty}^{\infty} f(z)\, dz + \int_{C_R} f(z)\, dz \tag{A.3}$$

where

$$f(z) = \frac{e^{-jzx}}{(z+j/\alpha)^J}. \tag{A.4}$$

It can be shown [24] that $\int_{C_R} f(z)\, dz = 0$ as R goes to infinity. (A.2) becomes

$$I = \frac{1}{2\pi} \frac{1}{(-j\alpha)^J} \left[\frac{2\pi j}{\Gamma(J)} \frac{d^{J-1}}{dz^{J-1}} e^{-jzx} \right]_{z=-j/\alpha} \tag{A.5}$$

or

$$I = \frac{j}{(-j\alpha)^J} \frac{1}{\Gamma(J)} (-jx)^{J-1} e - \frac{x}{\alpha}. \qquad (A.6)$$

Therefore the solution of (A.1) is

$$I = -\frac{1}{\Gamma(J)} \frac{x^{J-1}}{\alpha^J} e - \frac{x}{\alpha}, \quad x \leq 0 \qquad (A.7)$$

and I equals 0 for $x > 0$. $\Gamma(\cdot)$ is the gamma function and is defined as $\Gamma(J) = (J-1)!$ for J integer.

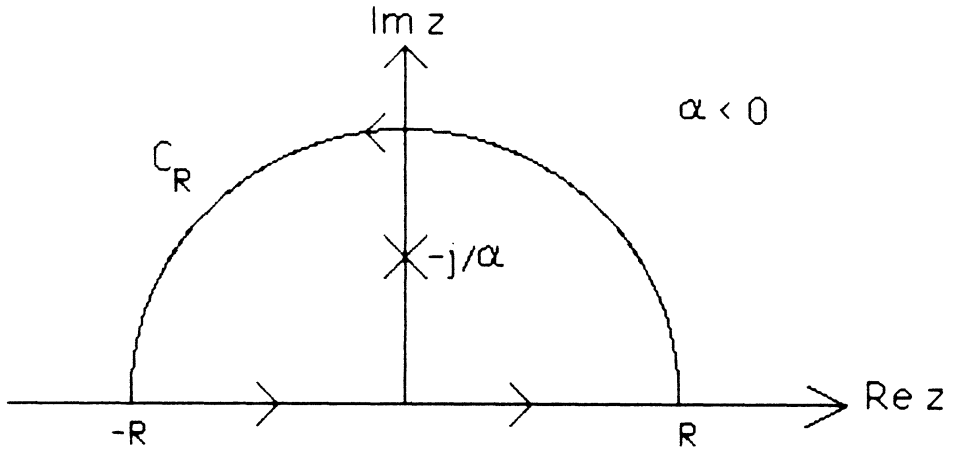


Figure A.1: Integration Contour (1 pole)

APPENDIX B

In this appendix, we propose to evaluate the following integral

$$I = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-j\omega x}}{(1-j\omega\alpha)^N (1-j\omega\beta)^M} \, d\omega \tag{B.1}$$

using contour integration. We assume that α and β are negative real constants, M and N are positive integers, and ω and x are real variables. (B.1) can be written as

$$I = \frac{1}{2\pi} \frac{1}{(-j\alpha)^N (-j\beta)^M} \int_C \frac{e^{-jz\,x}}{(z+j/\alpha)^N (z+j/\beta)^M} \, dz \tag{B.2}$$

where C is the counterclockwise contour (figure B.1) enclosing the poles $z_1 = -j/\alpha$ and $z_2 = -j/\beta$ and z is a complex variable. The contour integral above can be written as

$$\int_C f(z)\, dz = \int_{-\infty}^{\infty} f(z)\, dz + \int_{C_R} f(z)\, dz \tag{B.3}$$

where

$$f(z) = \frac{e^{-jz\,x}}{(z+j/\alpha)^N (z+j/\beta)^M} . \tag{B.4}$$

It can be shown [24] that $\int_{C_R} f(z)\, dz = 0$ as R goes to infinity. (B.2) becomes

$$I = \frac{1}{2\pi} \frac{1}{(-j\alpha)^N (-j\beta)^M} [2\pi j(B_1 + B_2)] \tag{B.5}$$

where

$$B_1 = \frac{1}{(N-1)!}\frac{d^{N-1}}{dz^{N-1}}\left[\frac{e^{-jzx}}{(z+j/\beta)^M}\right]_{z=-j/\alpha} \tag{B.6}$$

and

$$B_2 = \frac{1}{(M-1)!}\frac{d^{M-1}}{dz^{M-1}}\left[\frac{e^{-jzx}}{(z+J/\alpha)^N}\right]_{z=-j/\beta} \tag{B.7}$$

are the residues at poles z_1 and z_2, respectively. (B.5) can be written as

$$I = I_1 + I_2 \tag{B.8}$$

where

$$I_1 = \frac{j}{(-j\alpha)^N(-j\beta)^M}\frac{1}{(N-1)!}\frac{d^{N-1}}{dz^{N-1}}\left[\frac{e^{-jzx}}{(z+j/\beta)^M}\right]_{z=-j/\alpha} \tag{B.9}$$

and

$$I_2 = \frac{j}{(-j\alpha)^N(-j\beta)^M}\frac{1}{(M-1)!}\left[\frac{e^{-jzx}}{(z+j/\alpha)^N}\right]_{z=-j/\beta}. \tag{B.10}$$

In order to find I_1, we will try a recursive approach. We noticed that making the substitution $b = -\beta$ in (B.9) helps solve a sign ambiguity in later steps. Therefore,

$$I_1 = \frac{j}{(-j\alpha)^N(jb)^M}\frac{1}{(N-1)!}\frac{d^{N-1}}{dz^{N-1}}\left[\frac{e^{-jzx}}{(z-j/b)^M}\right]_{z=-j/\alpha}. \tag{B.11}$$

For N=1:

$$I_1 = \frac{j}{(-j\alpha)(jb)^M} \frac{e - \frac{x}{\alpha}}{(-j/\alpha - j/b)^M} = \frac{-1}{\alpha} \frac{\alpha^M b^M}{b^M(\alpha+b)^M} e^{-x/\alpha}$$

$$I_1 = -\frac{\alpha^{M-1}}{(\alpha+b)^M} e^{-x/\alpha}. \qquad \text{(B.12)}$$

For N=2:

$$I_1 = \frac{j}{(-j\alpha)^2(jb)^M} \left[\frac{-jx}{(z-j/b)^M} - \frac{M}{(z-j/b)^{M+1}} \right] e^{-jzx} \Bigg|_{z=-j/\alpha}$$

$$I_1 = -\frac{-j}{\alpha^2} \left[\frac{-jx}{(jb)^M(-j/\alpha - j/b)^M} - \frac{M}{(jb)^M(-j/\alpha - j/b)^{M+1}} \right] e^{-x/\alpha}$$

$$I_1 = \frac{1}{\alpha^2} \left[\frac{\alpha^M b^M x}{b^M(\alpha+b)^M} + \frac{M\alpha^{M+1} b^{M+1}}{b^M(\alpha+b)^{M+1}} \right] e^{-x/\alpha}$$

$$I_1 = -\frac{\alpha^{M-2} x}{(\alpha+b)^M} \left[1 + \frac{M\alpha b}{(\alpha+b)x} \right] e^{-x/\alpha}. \qquad \text{(B.13)}$$

For N=3:

$$I_1 = \frac{j/2}{(-j\alpha)^3(jb)^M} \left[\frac{-x^2}{(z-j/b)^M} + \frac{2jMx}{(z-j/b)^{M+1}} + \frac{M(M+1)}{(z-j/b)^{M+2}} \right] e^{-jzx} \Bigg|_{z=-j/\alpha}$$

$$I_1 = \frac{1}{2\alpha^3} \left[\frac{-x^2}{(jb)^M(-j/\alpha - j/b)^M} + \frac{2jMx}{(jb)^M(-j/\alpha - j/b)^{M+1}} \right.$$

$$\left. + \frac{M(M+1)}{(jb)^M(-j/\alpha - j/b)^{M+2}} \right] e^{-x/\alpha}$$

$$I_1 = \frac{1}{2} \left[-\frac{\alpha^{M-3} x^2}{(\alpha+b)^M} - \frac{2Mx\alpha^{M-2} b}{(\alpha+b)^{M+1}} - \frac{M(M+1)\alpha^{M-1} b^2}{(\alpha+b)^{M+2}} \right] e^{-x/\alpha}$$

$$I_1 = -\frac{1}{2}\frac{\alpha^{M-3}x^2}{(\alpha+b)^M}\left[1 + \frac{2M\alpha b}{(\alpha+b)x} + \frac{M(M+1)\alpha^2 b^2}{(\alpha+b)x^2}\right] e^{-x/\alpha}. \tag{B.14}$$

In general, we can write

$$I_1 = -\frac{\alpha^{M-N}x^{N-1}}{\Gamma(N)(\alpha+b)^M}\left[\sum_{n=0}^{N-1}\begin{bmatrix}N-1\\ n\end{bmatrix}\frac{\Gamma(M+n)}{\Gamma(M)}\frac{\alpha^n b^n}{(\alpha+b)^n x^n}\right] e^{-x/\alpha} \tag{B.15}$$

for $x \leq 0$. Note that $\Gamma(\cdot)$ is the Gamma function and $(\cdot)$ is a combination defined, respectively, as $\Gamma(N) = (N-1)!$ and $\begin{bmatrix}n\\ p\end{bmatrix} = \frac{n!}{p!(n-p)!}$. Substituting b by $-\beta$ in (B.15), we get

$$I_1 = -\frac{\alpha^{M-N}x^{N-1}}{\Gamma(N)(\alpha-\beta)^M}\left[\sum_{n=0}^{N-1}(-1)^n\begin{bmatrix}N-1\\ n\end{bmatrix}\frac{\Gamma(M+n)\alpha^n B^n}{\Gamma(M)(\alpha-\beta)^n x^n}\right] e^{-x/\alpha} \tag{B.16}$$

for $x \leq 0$. I_2 is obtained from the residue at $z_2 = -j/\beta$. This step can be done by analogy with equation (B.16). Therefore,

$$I_2 = -\frac{\beta^{N-M}x^{M-1}}{\Gamma(M)(\beta-\alpha)^N}\left[\sum_{m=0}^{M-1}(-1)^m\begin{bmatrix}M-1\\ M\end{bmatrix}\frac{\Gamma(N+m)\alpha^m\beta^m}{\Gamma(N)(\beta-\alpha)^m x^m}\right] e^{-x/\beta} \tag{B.17}$$

for $x \leq 0$. If α and β were positive the signs of I_1 and I_2 would be inverted. Using the fact that

$$(-1)^k/(a-b)^k = 1/(b-a)^k \tag{B.18}$$

then we can write I as

$$I = -\frac{\alpha^{M-N} x^{N-1}}{\Gamma(N)(\alpha-\beta)^{M}} \left[\sum_{n=0}^{N-1} \binom{N-1}{n} \frac{\Gamma(M+n)\alpha^{n}\beta^{n}}{\Gamma(M)(\beta-\alpha)^{n}x^{n}} \right] e^{-x/\alpha}$$

$$-\frac{\beta^{N-M} x^{M-1}}{\Gamma(M)(\beta-\alpha)^{N}} \left[\sum_{m=0}^{M-1} \binom{M-1}{m} \frac{\Gamma(N+m)\alpha^{m}\beta^{m}}{\Gamma(N)(\alpha-\beta)^{m}x^{m}} \right] e^{-x/\alpha} \qquad \text{(B.19)}$$

for $x \leq 0$.

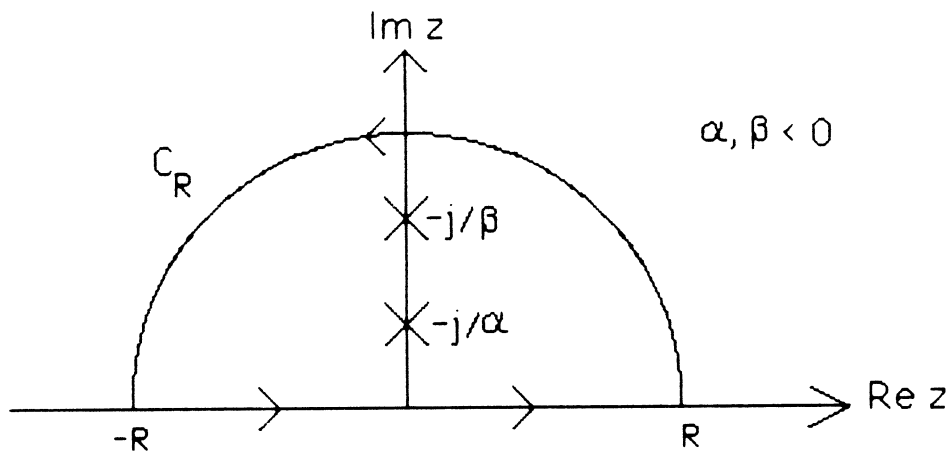


Figure B.1: Integration Contour (2 poles)

APPENDIX C

In this appendix, we demonstrate some matrix properties that are used in chapter 4. First, we prove that $\mathbf{H}_i$ (i = 1, 2) is Hermitian. It is easy to verify that $\mathbf{K}_{a_i}$ (i = 1, 2) and Φ (equation 2.4.8) are Hermitian matrices. Starting with equation (2.4.15)

$$\mathbf{H}_i = (\mathbf{I} + \rho_N \mathbf{K}_{a_i} \Phi)^{-1} \mathbf{K}_{a_i}, \tag{C.1}$$

and using Hermitian properties, we can write

$$\mathbf{H}_i^H = \mathbf{K}_{a_i}(\mathbf{I} + \rho_N \Phi \mathbf{K}_{a_i})^{-1} = \mathbf{K}_{a_i}(\mathbf{K}_{a_i}^{-1}\mathbf{K}_{a_i} + \rho_N \Phi \mathbf{K}_{a_i})^{-1} \tag{C.2}$$

or

$$\mathbf{H}_i^H = \mathbf{K}_{a_i}\mathbf{K}_{a_i}^{-1}(\mathbf{K}_{a_i}^{-1} + \rho_N \Phi)^{-1} = (\mathbf{K}_{a_i}^{-1}\mathbf{K}_{a_i} + \rho_N \Phi)^{-1} \tag{C.3}$$

which can be written as

$$\mathbf{H}_i^H = [\mathbf{K}_{a_i}^{-1}(\mathbf{I} + \rho_N \mathbf{K}_{a_i} \Phi)]^{-1} = (\mathbf{I} + \rho_N \mathbf{K}_{a_i} \Phi)^{-1} \mathbf{K}_{a_i}. \tag{C.4}$$

Therefore, $\mathbf{H}_i^H = \mathbf{H}_i$ and $\mathbf{H}_i$ (i = 1, 2) is a Hermitian matrix.

Next, we show that

$$\mathbf{K}_{y_i} = \rho_N^2 \Phi \mathbf{K}_{a_i} \Phi^H + \rho_N \Phi, \quad i = 1, 2 \tag{C.5}$$

is Hermitian and positive definite. Using Hermitian properties, it is easy to show that $\mathbf{K}_{y_i}^H = \mathbf{K}_{y_i}$. To show positive definiteness, let $\mathbf{x}$ be a nonzero vector. Writing $\mathbf{K}_{y_i}$ as defined in (3.2.9), we have

$$\mathbf{x}^H \mathbf{K}_{y_i} \mathbf{x} = \mathbf{x}^H \, E[\mathbf{y}\mathbf{y}^H | H_i]\mathbf{x} = E[\mathbf{x}^H \mathbf{y}\mathbf{y}^H \mathbf{x} | H_i] \, . \tag{C.6}$$

Letting $z = \mathbf{y}^H \mathbf{x}$, (C.6) becomes

$$\mathbf{x}^H \mathbf{K}_{y_i} \mathbf{x} = E[\mathbf{z}^H \mathbf{z}] = E[|\mathbf{z}|^2] > 0 \, . \tag{C.7}$$

(C.7) is a real and sufficient condition for the Hermitian matrix $\mathbf{K}_{y_i}$ (i = 1, 2) to be positive definite. This property holds true for all covariance matrices.

APPENDIX D

In this appendix, we would like to obtain an approximate expression of

$$f(\mathbf{d}_n) = \|\mathbf{r}_k - \mathbf{d}_n\| \tag{D.1}$$

in terms of a Taylor series about $\mathbf{d}_n = 0$. Showing only first order terms,

$$f(\mathbf{d}_n) = f(0) + (\mathbf{d}_n - 0)^T \nabla f(0) + \ldots . \tag{D.2}$$

Writing $f(\mathbf{d}_n)$ in function of the coordinates of $\mathbf{r}_k$ and $\mathbf{d}_n$ assuming three-dimensional vectors, we get

$$f(\mathbf{d}_n) = [(r_1 - d_1)^2 + (r_2 - d_2)^2 + (r_3 - d_3)^2]^{1/2} . \tag{D.3}$$

Using this equation, we can write

$$f(0) = \|\mathbf{r}_k\| \tag{D.4}$$

and

$$\frac{\partial f(\mathbf{d}_n)}{\partial d_1} = -\frac{2(\mathbf{r}_1 - d_1)}{2\|\mathbf{r}_k - \mathbf{d}_n\|} = -\frac{(r_1 - d_1)}{\|\mathbf{r}_k - \mathbf{d}_n\|} . \tag{D.5}$$

Therefore

$$\nabla f(\mathbf{d}_n) = \left[\frac{\partial f(\mathbf{d}_n)}{\partial d_1}, \frac{\partial f(\mathbf{d}_n)}{\partial d_2}, \frac{\partial f(\mathbf{d}_n)}{\partial d_3} \right] \tag{D.6}$$

or

$$\nabla f(\mathbf{d}_n) = -\frac{(\mathbf{r}_k - \mathbf{d}_n)}{\|\mathbf{r}_k - \mathbf{d}_n\|} \tag{D.7}$$

which implies

$$\nabla f(0) = -\mathbf{r}_k / \|\mathbf{r}_k\| = -\boldsymbol{\alpha}_k \tag{D.8}$$

where $\boldsymbol{\alpha}_k$ is a unit vector along $\mathbf{r}_k$. Substituting these results in (D.2), $\|\mathbf{r}_k - \mathbf{d}_n\|$ is expressed approximately as

$$\|\mathbf{r}_k - \mathbf{d}_n\| = \|\mathbf{r}_k\| - \mathbf{d}_n^T \boldsymbol{\alpha}_k = \|\mathbf{r}_k\| - \boldsymbol{\alpha}_k^T \mathbf{d}_n \,. \tag{D.9}$$

ABSTRACT

HADDAD, NICHOLAS KARIM. Ph.D. June, 1990. Electrical and Computer Engineering

Performance Analysis of Active Sonar Classifiers

Director of Dissertation: Dr. John Tague

This dissertation studies the theoretical underpinnings of active sonar classifiers. We present a systematic approach for designing optimal Bayesian classifiers and analyzing their performance. We emphasize the ternary case where three hypotheses are considered: H_0 (noise only), H_1 (reverberation plus noise) and H_2 (target plus noise).

We start by deriving a sufficient statistic to decide between H_1 and H_2, assuming H_0 has already been eliminated. Then, closed–form solutions for classification and false alarm probabilities are obtained and several receiver operating characteristics curves illustrating meaningful physical scenarios are presented. Two classes of illuminating signals are considered: high resolution and linear FM signals.

Many design parameters affecting classifier performance are studied. Perhaps the most important issue is classifier performance when incorrect a priori knowledge of the target's spatial properties is processed. Other parameters such as target resolution, signal–to–noise ratio, transmitter constant in linear FM signals, etc. are investigated as well.

The final issue presented is acoustic target imaging. A minimum variance linear unbiased estimator of the scattering coefficients of the test volume encompassing the target is derived. Furthermore, we investigate error