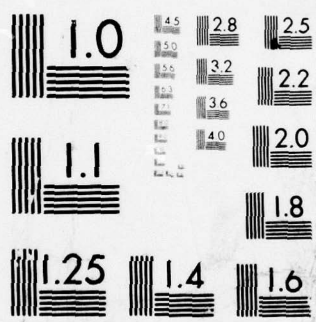
AD-A073 478 JOHNS HOPKINS UNIV BALTIMORE MD DEPT OF ELECTRICAL --ETC F/G 9/4
ROBUST NONCOHERENT DETECTION.(U)
FEB 79 H L WEINERT, A H EL-SAWY N00173-77-C-0189
UNCLASSIFIED NL

1 OF 2
AD
A073478


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

II
LEVEL

DATA SHEET
PHOTOGRAPH
THIS SHEET

1
INVENTORY

"Robust Noncoherent Detection."

N00173-77-C-0189
DOCUMENT IDENTIFICATION

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DISTRIBUTION STATEMENT

Accession For	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification per Doc.	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

DISTRIBUTION STAMP

DDC
RECEIVED
SEP 5 1979
E

DATE ACCESSIONED

79 08 20 082

DATE RECEIVED IN DDC

PHOTOGRAPH THIS COPY

THE JOHNS HOPKINS UNIVERSITY

ELECTRICAL ENGINEERING DEPARTMENT

ROBUST NONCOHERENT DETECTION

By

HOWARD L. WEINERT
ABDEL-RAHMAN H. EL-SAWY

FINAL REPORT
CONTRACT NO. N00173-77-C-0189
NAVAL RESEARCH LABORATORY

ROBUST NONCOHERENT DETECTION

by

Howard L. Weinert
Abdel-Rahman H. El-Sawy
The Johns Hopkins University
Electrical Engineering Department
Baltimore, Maryland 21218

Final Report
Contract No. N00173-77-C-0189
Naval Research Laboratory
Washington, D.C. 20375

February 28, 1979

Abstract

In this report the problem of robust detection of non-coherent signals in noise is investigated under the assumption that the noise distribution is unknown, but a member of a known class of distribution functions. This problem is divided into three main categories: partially-coherent signals, unknown frequency and phase, and completely random phase.

In the partially-coherent problem, two general methods are introduced for the design of robust detectors which guarantee a non-trivial lower bound on the probability of detection and a non-trivial upper bound on the probability of false alarm. Two recievers are designed for the special case in which the distributions of the noise inphase and quadrature phase components are members of a class of p-point distributions. Simulation results for finite sample size are also given for different distributions. The M-detector method is extended to give a general solution for the second problem.

In the third problem, three different solutions are presented. All consider that the disttribution of the envelope of the noise is a member of some class of distribution functions which is defined by quantiles. Simulation results for these three detectors are given at finite sample sizes for different distributions.

Throughout the report, the log normal and contaminated normal distributions are utilized in conducting the simulations.

Table of Contents

		Page
I.	Introduction	1
II.	Detection of Partially Coherent Signals In Noise I	7
	II.1 Introduction	7
	II.2 Problem Statement	9
	II.3 Problem Solution	13
	II.4 The Family of P-Point Distributions	24
	II.5 Numerical Results	27
	II.6 Summary	43
III.	Detection of Partially Coherent Signals In Noise II	44
	III.1 Introduction	44
	III.2 Stochastic Approximation Estimation	44
	III.3 Stochastic Approximation Detectors	49
	III.4 Numerical Results	50
	III.5 General Discussion	59
	III.6 Summary	63
IV.	Detection of Signals with Unknown Phase and Frequency	65
	IV.1 Introduction	65
	IV.2 The Gaussian Noise Case	67
	IV.3 M-Doppler Detector	71
	IV.4 Numerical Results	74

V.	Detection of Signals with Random Phase in Sea Clutter	81
	V.1 Introduction	81
	V.2 M-Detector on Envelope	83
	V.3 Robust Quantizer Detectors	95
	V.4 The Extreme-Value Theory Detectors	104
	V.5 The Trimmed-Mean Detector	110
	V.6 Detector Comparison	112
	V.7 Summary	117
VI.	Performance of the M-Detector with Dependent Data	118
VII.	Summary and Conclusion	130
	References	132

I. INTRODUCTION

The different approaches for the design of a receiver for detection of signals in additive noise fall between two extremes. On one hand is the classical parametric approach which requires an exact and complete statistical description of both the signal and the channel and utilizes the likelihood ratio as a test statistic. On the other extreme are nonparametric or distribution free approaches which require minimal information about the utilized channel and lead to constant false alarm detectors, such as sign, rank and median detectors [1].

Often in practice, the information available is only partial, either because the channel description is incomplete or because some aspects of the channel vary either temporally or spacially. This partial information is usually more than required for design of a nonparametric detector, but is insufficient for the design of a parametric detector. The treatment of such problems depends on the nature of the missing information, which can be divided into two main categories.

1) Parameter Ambiguity: In this case there is a mathematical expression for the channel model with some ambiguity about the exact value of some parameters; i.e., the actual noise is a member of a known finite-dimensional class of distribution functions. For example, the channel is normal with unknown mean and/or variance.

2) Distribution Ambiguity: This is the intrinsically more difficult case where there is no mathematical expression for the noise distribution, but some information about the distribution is available such as the first m moments. This means that the noise distribution is a member of an infinite-dimensional class of noise distributions.

The problems in the first case are easier to handle than those in the second and in some particular situations it is possible to achieve an optimal performance. At other times only a minimax or adaptive detector is possible. Actually, the problems in this category have been studied extensively as the composite hypothesis problem and many good adaptive and minimax techniques have been developed [2-4].

In treating the problems in the second category, the receiver designers divided into three groups. The first and most conservative group accepted the nonparametric techniques to at least guarantee the probability of error of the first kind or the probability of false alarm. The second group utilized the available partial information to obtain the best fit from a selected class of noise distributions, then designed a parametric detector based cn this approximate distribution. Utilization of nonparametric techniques in such cases is useful, but their use results in ignoring most of the available information. On the other hand, generation of an approximate noise distribution, although desirable since it leads to a parametric receiver, may be very dangerous since the actual observations may not come from the assumed class of distributions. This implies that the detector performance may deteriorate markedly from the expected performance. To demonstrate this point, let the approximate density function be $f_j(x)$ under H_j, then the likelihood ratio will be $\Lambda(x) = \prod_{i=1}^{N} \frac{f_1(x_i)}{f_0(x_i)}$. Now, a single factor $\frac{f_1(x_k)}{f_1(x_k)}$ equal (or almost equal) to zero or infinity will upset the test statistic and change the final decision [5].

Out of these two groups and as a result of observations similar to those above, came the third group which rejected both approaches and went further to reject the utilization of strictly parametric techniques since these systems are sensitive to bad data and there is no known procedure to protect any device completely against the presence of such outliers or bad data [6] and [7]. This group believe the best solution to the problem is in the design of detectors which make use of the maximum possible amount of the available information and perform well over the class of all distributions which possesses the properties described by this information. These detectors, which are called robust detectors, are not expected to perform as well as the optimal detectors for all distributions in this class; however, they should perform better than the nonparametric method for a large subset of this class. The search for such robust procedures has been very active during the last few years [5,8-10]. All of the above work assumes that the signal is coherent and all except [10] assumes some class of contaminated distributions. In [10], a general procedure for robust detection of known signals in additive noise was presented. These receivers are called M-detectors.

In this report, the problem of "Robust Detection of Non-Coherent Signals in Additive Noise" is considered. In general, the problem is divided into three basic cases.

1) Partially-coherent signals: In this case, it is assumed that the signal has an unknown phase which is constant over each sequence of observations, but changes randomly from one sequence to another.
2) Unknown frequency and phase: It is assumed that both the signal frequency and phase are unknown. This is usually the case in detection in the presence of Doppler effects.
3) Random Phase: In this case it is assumed that the unknown phase changes randomly from one sample to another.

Chapters II and III investigate the first problem. In Chapter II we extend the concept of M-detectors presented in [10] to the noncoherent case. We show that under some sufficient conditions on the family there exists a transformation which maps the input data into some intermediate space. Utilizing these points as new input data we construct a detector and show that under some assumptions on the signal amplitude it will be robust in a max-min sense over the family of distributions under consideration. In Chapter III we present another detector, called the stochastic-approximation detector, which has similar properties as the M-detector. We also give simulation results to show the finite sample size performance of the above detectors and to compare it to the conventional square law detector. Chapter IV extends the results of the previous two chapters to the second case. In all three chapters it is assumed that some information about the inphase and quadrature phase components is known.

In Chapter V a set of detectors is presented to treat the third problem. All the detectors considered in this chapter depend on the envelope of the received observations as input data.

In Chapter VI the performance of the M-detector with dependent data is investigated.

Throughout this report, the stability of the threshold with respect to the actual noise distribution and the power of the test were used as criteria for judging receiver performance. Also, for all simulations, the noise distribution was assumed either to be lognormal or contaminated normal. These two distributions are the most common in the above detection problems [11,12]. The normal distribution was utilized principally for comparisons.

II. Detection of Partially Coherent Signals In Noise I

II.1 Introduction

This chapter investigates the detection of signals with unknown phase in additive noise under the assumption that available information about the noise distribution is not complete. Specifically, we shall consider these situations where there is no available mathematical expression for the noise distribution function, either because it changes in shape with uncontrolled factors or just because of the absence of enough data. On the other hand, it will be assumed that some information is available about the distribution of the inphase and quadrature phase noise components. The signal component will be considered as samples from a demodulated version of the received signal which is given by

$$S_r(t) = B \operatorname{Sin}(w_c t+\theta) \qquad 0 \leq t \leq T \qquad (I.1)$$

where θ is the unknown pahse which is a random variable with uniform distribution over the interval $[-\Pi, \Pi]$, T is the total observation period and B is the signal amplitude which is equal to zero under H_0 and is greater than zero under H_1. We shall call this problem the partially coherent case.

This problem appears in many communication systems when there is uncertainity about the phase of the wave form generated by the transmitter's ossilators. In radar applications, it appears in the case of detection of fixed targets [3 pp 335]. A more important aspect of this study is that it furnishes a basis for the more important problems in radar systems such as the design of doppler processors. In such situations, both the frequency and phase will be assumed unknown.

Usually in practice, the engineer uses some version of the square law detector, not because of its optimality but because of its simplicity. In this detection procedure the detector sums the inphase component samples and the quadrature samples, and then utilizes the square root of the sume of the squares of the above two sums as a test statistic. This method has two main defects. First, since the sum is a linear operation, if we have a large noise sample, either because the actual distribution has a long tail or because of some measurement error, then this observation is more likely to upset the decision. Second, if the distribution is not known exactly then setting the threshold will require an adaptive technique. Usually, these adaptive techniques cause some loss in the power of the test [13].

In the following, we give a procedure for the design of robust detectors that guarantee an upper bound on both probability of false alarm and probability of missing under the above mentioned assumptions. We design one of these detectors for the class of p-point distributions, where only one quantile is known. We compare this detector to the square law detector for the lognormal, contaminated normal and Guassian distributions. The lognormal was chosen because of its long tail and because it appears in many practical situations such as detection in sea clutter and atmospheric noise. The contaminated normal is shorter in tail than lognormal and represents sea clutter interference in some sea states [14]. It also can be considered as a moderate representation of measurement errors [7]. The Gaussian was chosen because the square law detector is optimal against it.

II.2 Problem Statement

Consider the set of observations $\{z_i\}$ given by its inphase and quadrature phase components denoted by $\{x_i\}$ and $\{y_i\}$ respectively where

$$\begin{aligned} x_i &= A_1 + w_i \\ & \qquad\qquad i = 1, \ldots, N \qquad\qquad (II.2) \\ y_i &= A_2 + v_i \end{aligned}$$

where $\{w_i\}$ and $\{v_i\}$ are two sequences of independent identically distributed noise which are also mutually independent, $f(w_i)$ and $f(v_i)$ are the density functions of w_i and v_i respectively, which are unknown members of a class of symmetric density functions F, A_1 and A_2 are the inphase and quadrature phase signal components such that $\sqrt{A_1^2+A_2^2} = A$. We define the two hypotheses H_0 and H_1 as

$$
\begin{aligned}
H_0 &: \; A = 0 \\
H_1 &: \; A > 0
\end{aligned}
\tag{II.3}
$$

We are interested in finding a detection procedure to distinguish between H_0 and H_1 such that both the probabilities of error of the first and second types have the best achievable nontrivial upper bounds.

In the following analysis we shall consider $f(v)$ and $f(w)$ to have the same mathematical expression and we shall deal only with the case where A_1 and A_2 are constants over each group of observations. The generalization to the case where A_1 and A_2 are changing in a known manner as functions of i is straight forward [10].

Before proceeding to the solution, we shall need a few additional definitions. Define a class of functions C such that $L \varepsilon C$ if

1) L is convex, symmetric about the origin and strictly increasing for positive arguments.

2) $l(t) = dL(t)/dt$ is continuous for all t.

3) for all $f \in F$, $E_f[l^2(x)] < \infty$

4) for all $f \in F$, $\partial E_f[l(x-\theta)]/\partial\theta$ exists and is non zero in some neighborhood of the origin.

Also, define θ_{NH} as the value of θ which minimizes

$$\sum_{i=1}^{N} L(x_i-\theta) \tag{II.4}$$

and θ_{NV} as the value of θ which minimizes

$$\sum_{i=1}^{N} L(y_i-\theta) \tag{II.5}$$

or, equivalently,

$$\sum_{i=1}^{N} l(x_i-\theta_{NH}) = 0 \tag{II.6}$$

and

$$\sum_{i=1}^{N} l(y_i-\theta_{NV}) = 0 \tag{II.7}$$

and define θ_N as

$$\theta_N = \sqrt{\theta_{NH}^2+\theta_{NV}^2} \tag{II.8}$$

Define also $B_d(0|f)$ and $B_d(A|f)$ as the probability of false alarm and the probability of detection, respectively, using detection strategy d when the actual noise distribution is $f(\cdot)$.

The main results of this section will be obtained by verifying the following two inequalities:

$$B_{L_0}(0|f) \leq B_{L_0}(0|f_0) \quad \text{for all} \quad f \in F \tag{II.9}$$

and

$$B_{L_0}(v|f_0) \leq B_{L_0}(v|f) \quad \text{for all} \quad f \in F \tag{II.10}$$

where $B_d(v|f) = \lim_{N\to\infty} B_d(A|f)$ with $A = v/\sqrt{N}$, and then by showing that $B_{L_0}(v|f_0)$ is related to the power of the optimal coherent detector, when the true noise density function is f_0, by the same relation as that between the power of a coherent and noncoherent detector when the noise is Gaussian. The subscript L_0 refers to a detector based on a threshold test using θ_N drived from $L_0 \in C$ as a test statistic; i.e., a test of the form

$$\sqrt{N}\ \theta_N \underset{H_0}{\overset{H_1}{\lessgtr}} \gamma \tag{II.11}$$

we shall call this detector the M-detector on components. In fact it might seem more appropriate to show that $B_{L_0}(v|f_0)$ is greater than or equal to the power of any randomized test of hypotheses under the above assumptions and assuming that the true noise is $f_0(\cdot)$, but with the absence of an optimal procedure to design such a detector, it is the authors opinion that the proposed criterion is the best possible one.

II.3 Problem Solution

Assume for the moment that the class of density functions F contains a density $f^*(\cdot)$ of minimum Fisher information for location, that is, a density for which

$$I(F) = \int_{-\infty}^{\infty} [f'(x)/f(x)]^2 f(x)dx$$

is a minimum over all $f \varepsilon F$. Also, assume that the class C of functions contains a function L^* such that

$$L^*(x) = -\log f^*(x) .$$

For this case we will show that $(L^*, f^*) = (L_0, f_0)$; i.e., it satisfies equations (9) and (10) described above. First, we require a preliminary lemma whose proof can be found in [15] . It is also a special case of Lemma 1 [10] so we shall only outline it here.

Lemma (II.1):

Whenevery $L \varepsilon C$ and $f \varepsilon F$, $\sqrt{N}\ (\theta_{NH}-A_1)[\sqrt{N}(\theta_{NV}-A_2)]$ is asymptotically distributed as a zero-mean, normal random variable with variance

$$v^2(f,L) = \frac{\int_{-\infty}^{\infty} l^2(x)\ f(x)\ dx}{\{\frac{\partial}{\partial\theta} E_f[l(x-\theta)]|_{\theta=0}\}^2} \tag{II.12}$$

Proof:

1) As $N \to \infty$ $\theta_{NH} \to A_1$ and $\theta_{NV} \to A_2$ a.s. and in probability.

2) Assume that $A=0$ then

$$P(\theta_{NH} \leq \frac{K}{\sqrt{N}}) = P\{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} l(x_i - \frac{k}{\sqrt{N}}) \leq 0\} \qquad \text{(II.13)}$$

3) $P\{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} l(x_i - \frac{k}{\sqrt{N}})\} \underset{N\to\infty}{\Longrightarrow}$ Normal (μ, V^2)

as $\mu = \frac{\partial}{\partial\theta} E_f[l(x-\theta)]|_{\theta=0}$

and

$$V^2 = \int_{-\infty}^{\infty} l^2(x)\, f(x)dx$$

From the above lemma and under the assumption that $\{w_i\}$ and $\{v_i\}$ are independent of each other we obtain with $T_N = \sqrt{N}\,(\theta_N)$,

$$f(T_N|H_0) \underset{N\to\infty}{\to} \frac{T_N}{v^2(f,L)} e^{-\frac{T_N^2}{2v^2(f,L)}} \qquad \text{(II.14)}$$

and

$$f(T_N|H_1) \underset{N\to\infty}{\to} \frac{T_N}{v^2(f,L)} e^{-\frac{T_N^2+v^2}{2v^2(f,L)}} I_0(\frac{T_N\, v}{v^2(f,L)}) \qquad \text{(II.15)}$$

where $I_0(\cdot)$ is the modefined Bessel function of the first kind.

Then the asymptotic power of a test based on T_N satisfies

$$B_L(v|f) = \int_{\gamma}^{\infty} \frac{T_N}{v^2(f,L)} e^{-\frac{T_N{}^2 + v^2}{2v^2(f,L)}} I_0 \left(\frac{T_N v}{v^2(f,L)}\right) d\,T_N$$

$$= Q(\alpha,\gamma) \tag{II.16}$$

where $Q(.,.)$ is Marcam's Q function, $\alpha = \frac{v}{v^2(f,L)}$ and γ is the threshold of the test. Notice that for any finite falue of γ , if A is a constant then $v \to \infty$ as $N \to \infty$ and $Q(\alpha,\gamma) \to 1$; i.e., the test is always consistent.

In the above discussion it was assumed that v_i and w_i are independent for all i. Nevertheless, the results are still correct if v_i and w_i are given by

$$\begin{aligned} v_i &= u_i \cos \phi_i \\ & \qquad\qquad\qquad i = 1,\ldots, N \\ w_i &= u_i \sin \phi_i \end{aligned} \tag{II.17}$$

where $\{u_i\}$ and $\{\phi_i\}$ are sequences of independence identically distributed random variables with ϕ_i uniformly distributed between $[0,2\pi]$. To show this, we have

$$P[\theta_{NH} \le \frac{k_1}{\sqrt{N}} , \theta_{NV} \le \frac{k_2}{\sqrt{N}}]$$

$$= P\{ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} 1(x_i - \frac{k_1}{\sqrt{N}}) \le 0, \frac{1}{\sqrt{N}} \sum_{i=1}^{N} 1(y_i - \frac{k_2}{\sqrt{N}}) \le 0\}$$

$$= P\{ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i \le 0, \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Y_i \le 0\} \tag{II.18}$$

but for any two constants a and b

$[\frac{a}{\sqrt{N}} \sum_{i=1}^{N} X_i + \frac{b}{\sqrt{N}} \sum_{i=1}^{N} Y_i]$ is asymptotically distributed

as a normal random variable, then $[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i]$ and

$[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} y_i]$ are asymptotically jointly normal random

variables [16] and it is enough to show that they are uncorrelated to prove their independent. Since

$$E\{(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i)(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} Y_i)\}$$

$$= \frac{1}{N} \sum_{i=1}^{N} E(X_i Y_i) + (N-1)\, E(X_i)E(Y_i) \qquad \text{(II.19)}$$

then

$$\operatorname{cov}\,[(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i),\ (\frac{1}{\sqrt{N}} \sum_{i=1}^{N} Y_i)]$$

$$= \frac{1}{N} \sum_{i=1}^{N} E(X_i Y_i) - E(X_i)E(Y_i)$$

$$= 0(\frac{1}{N}) \qquad \text{(II.20)}$$

and since the variance of each is of order 1 then the correlation coefficient is of order $(\frac{1}{N})$ and goes to zero as $N \to \infty$; i.e., they are asymptotically independent.

Then

$$P\{\sqrt{N}\ \theta_{NH}\le k_1,\ \sqrt{N}\ \theta_{NV}\le k_2\} \Rightarrow P\{\frac{1}{\sqrt{N}}\sum_{i=1}^{N} 1(x_i-\frac{k_1}{\sqrt{N}})\le 0\}.$$

$$P\{\frac{1}{\sqrt{N}}\sum_{i=1}^{N} 1(y_i-\frac{k_1}{\sqrt{N}})\le 0$$

$$= P\{\sqrt{N}\ \theta_{NH}\le k_1\}\ P\{\sqrt{N}\ \theta_{NV}\le k_2\} \tag{II.21}$$

and the results of the previous discussion continue to hold. This completes the proof of the following lemma

Lemma II.2:

Whenever $L \varepsilon C$ and $f \varepsilon F$, v_i independent from w_i for all i=1,..., N or given by eqn. (II.17), $\sqrt{N}\theta_N$ will be asymptotically distributed under H_0 as

$$f(T_N) = \frac{T_N}{v^2(f,L)}\ e^{-\frac{T_N{}^2}{2v^2(f,L)}} \tag{II.22}$$

and under H_1

$$f(T_N) = \frac{T_N}{v^2(f,L)}\ e^{-\frac{(T_N{}^2+v^2)}{2v^2(f,L)}}\ I_0(\frac{T_N v}{v^2(f,L)}) \tag{II.23}$$

To study the power relations for different $f \varepsilon F$ we need the following lemma.

Lemma (II.3):

If there exists an $f_0 \varepsilon F$ such that $L_0 = -\log f_0 \varepsilon C$ then

$$\sup_{L\varepsilon C} B_L(v|f_0) = B_{L_0}(v|f_0) \tag{II.24}$$

and if

$$v^2(f,L_0) \leq v^2(f_0,L_0) \quad \text{for all } f \in F \tag{II.25}$$

then the following relation holds asymptotically

$$B_{L_0}(0|f) \leq B_{L_0}(0|f_0) \quad \text{for all } f \in F \tag{II.26}$$

Proof:

From equation (II.22)

$$B_{L_0}(0|f) = \exp\left(\frac{-\gamma^2}{2v^2(f,L_0)}\right)$$

$$\leq \exp\left(\frac{-\gamma^2}{2v^2(f_0,L_0)}\right)$$

by virtue of equation (II.25).

Also, for any level of false alarm α and for any $L \in C$, if f_0 is the true density

$$\gamma^2 = -2v^2(f_0,L) \ln \alpha \tag{II.27}$$

$$B_L(v|f_0) = \int_\gamma^\infty \frac{T_N}{v^2(f_0,L)} e^{-\frac{T_N{}^2 + v^2}{2v^2(f_0,L)}} I_0\left(\frac{T_N v}{v^2(f_0,L)}\right) d\, T_N$$

$$= \int_{\sqrt{-2 \ln\alpha}}^\infty R\, e^{-\frac{1}{2}\left(R^2 + \frac{v^2}{v^2(f_0,L)}\right)} I_0\left(R \frac{v}{v(f_0,L)}\right) d\, R$$

$$= Q\left[\frac{v}{v(f_0,L)}, \sqrt{-2 \ln\alpha}\right] \tag{II.28}$$

but as $v^2(f_0,L)$ decreases, $v_1 = \frac{v}{v(f_0,L)}$ increases and so does $B_L(v|f_0)$. On the other hand $v^2(f_0,L_0) \leq v^2(f_0,L)$ since this is the variance of the maximum likelihood estimates of A_1 and A_2 and is equivalent to the Cramer-Rao lower bound. This completes the proof of the lemma.

To get the relation between $B_{L_0}(v|f)$ and $B_{L_0}(v|f_0)$ we assume that $B_{L_0}(0|f_0) = \alpha$, then

$$B_{L_0}(v|f) = \int_{v(f_0,L_0)\sqrt{-2\ln\alpha}}^{\infty} \frac{z}{v^2(f,L_0)} \exp\left[-\frac{z^2+v^2}{2v^2(f,L_0)}\right] I_0\left(\frac{zv}{v^2(f,L_0)}\right) dz$$

using $R = \dfrac{z}{v(f,L_0)}$

$$B_{L_0}(v|f) = \int_{\gamma'=\frac{v(f_0,L_0)}{v(f,L_0)}\sqrt{-2\ln\alpha}}^{\infty} R\ e^{-\frac{R^2+v'^2}{2}}\ I_0(Rv')\ dR$$

$$= Q(v',\gamma') \qquad \text{(II.29)}$$

where $v' = \dfrac{v}{v(f,L_0)}$ and $\gamma' = \dfrac{\gamma}{v(f,L_0)}$.

Notice that as v(.,.) decreases both v' and γ' increases with the same ratio, which makes this relation untractable analytically, especially with the absence of an analytic expression for the Q function itself. On the other hand, we can write eq (II.29) in the form

$$Q(v',\gamma') = Q(r\,\bar{v}, r\,\gamma) \tag{II.30}$$

where $r = \dfrac{v(f_0,L_0)}{v(f,L_0)} \geq 1$, $\bar{v} = \dfrac{v}{v(f_0,L_0)}$, and if $\bar{v} > \gamma$ then [3]

$$Q(r\,\bar{v},r\gamma) = 1 - e^{-(v'^2+\gamma'^2)/2} \sum_{n=1}^{\infty} \left(\frac{\gamma'}{v'}\right)^n I_n(v'\,\gamma')$$

$$= 1 - e^{\frac{-r^2\bar{v}^2}{2}(1+k^2)} \sum_{n=1}^{\infty} (k)^n\, I_n(r^2\,\bar{v}^2 k) \tag{II.31}$$

where $k = \dfrac{\gamma}{\bar{v}} < 1$. Taking the derivative with respect to r,

$$\frac{\partial}{\partial r} Q(r\,\bar{v},r\,\gamma) = G[I_1(r^2\,\bar{v}^2\,k) - k^2\, I_0(r^2\,\bar{v}^2\,k)] \tag{II.32}$$

where G is a positive constant. Thus

$\dfrac{\partial}{\partial r} Q(r\,\bar{v},r\,\gamma) \geq 0$ iff $\dfrac{I_1(r^2\bar{v}^2\,k)}{I_0(r^2\bar{v}^2\,k)} \geq k^2$. This implies that if $\bar{v}$ and γ satisfy the above inequality for all values of $r \geq 1$ then the power of the test will be monotonic in r

$$B_{L_0}(v|f) \geq B_{L_0}(v|f_0)\ . \tag{II.33}$$

This completes the proof of the following lemma.

Lemma (II.4):

If there exists $f_0 \varepsilon F$ such that $L_0 = -\log f_0 \varepsilon C$ and

$$v^2(f_0, L_0) \geq v^2(f, L_0) \quad f \varepsilon F \qquad \text{(II.34)}$$

and if $\dfrac{I_1(r^2\,\bar{v}^2 k)}{I_0(r^2\,\bar{v}^2 k)} \geq k^2$ for all $r \geq 1$ and $\bar{v} > \gamma$ then

$$B_{L_0}(v|f) \geq B_{L_0}(v|f_0) \quad f \varepsilon F. \qquad \text{(II.35)}$$

The above inequality was solved for values of $\gamma \geq 1.9226$ which is equivalent to $P_F \leq .15$ and it was found that k should be less than or equal to (0.91166). Values of $\bar{v}$ are such that the probability of detection is greater than or equal to about 60%. Notice that this is the region of interest in most communication applications. Notice also that in the case of robust coherent detectors [10] this range was for all $P_D \geq .5$.

The above lemma gives the lower bound on the probability of detection over the class F. To investigate how good this lower bound is, we compare the proposed M-detector to the asymptotically most powerful coherent detector when the true noise density is $f_0(.)$. If we assume that the probability of false alarm is α , for the asymptotically most powerful coherent detector

$$B_{opt}(v_1|f_0) = 1 - \Phi\,\{\Phi^{-1}(1-\alpha) - v_1\, I^{\frac{1}{2}}(f_0)\} \qquad \text{(II.36)}$$

where $\Phi(.)$ is the standard normal distribution function and $I(.)$ is the Fisher information number for location. For the proposed M-detector, since $v^2(f_0,L_0)=\frac{1}{I(f_0)}$ then

$$B_{L_0}(v_2|f_0) = \int_{\sqrt{-2\ln\alpha}}^{\infty} z\, e^{-\frac{z^2+v_2^2 I(f_0)}{2}} I_0[z\, v_2\, I^{\frac{1}{2}}(f_0)]\, dz$$

$$= Q(v_2\, I^{\frac{1}{2}}(f_0), \sqrt{-2\ln\alpha}) \qquad \text{(II.37)}$$

Comparing the above two equations with the case when $f_0(.)$ is Gaussian [17] we find that if both probability of false alarm and detection are constant then $\frac{v_2}{v_1}$ is the same in both cases, i.e. in replacing a coherent scheme with an M-detector, the loss in signal to noise ratio will be the same in the M-detector case as in the Gaussian case.

All the above can be concluded in the following theorem.

<u>Theorem II.1</u>:

If there exists $_0(.)\ \varepsilon\ F \ni L_0=-\log f_0\ \varepsilon\ C$ and

$$v^2(f_0,L_0) \geq v^2(f,L_0) \quad \forall f\ \varepsilon\ F$$

then the M-detector will satisfy the following relations asymptotically

$$B_{L_0}(0|f) \leq B_{L_0}(0|f_0) \quad \forall f\ \varepsilon\ F \qquad \text{(II.38)}$$

$$B_L\ (v|f) \leq B_{L_0}(v|f_0) \quad \forall L\ \varepsilon\ C \qquad \text{(II.39)}$$

and if $\bar{v} = \frac{v}{v(f_0,L_0)} \geq \gamma$ threshold and $\frac{I_1(r^2\bar{v}\gamma)}{I_0(r^2\bar{v}\gamma)} \geq (\frac{\gamma}{\bar{v}})^2$

for all $r \geq 1$

then

$$B_{L_0}(v|f) \geq B_{L_0}(v|f_0) \quad \forall f \in F \tag{II.40}$$

The above results show that the design of the receiver will depend on finding the least favorable density function f_0 as defined in Lemma II.3. To find this density, we notice that it must have a finite Fisher information number for location according to assumption 3 on C. Thus, it will also be least favorable in the family G of density functions defined as

$$G = \{f: \ f \in F \ \text{ and } I(f) < \infty\}. \tag{II.41}$$

However, the least favorable density from this class is the one with minimum Fisher information (cf. Theorem 2 [15]). Thus, to find f_0, we follow the steps described below.

a) Find f^* such that $I(f^*) \leq I(f)$, all $f \in G$. This can usually be done using Lagrange multiplier techniques.

b) Check that $-\log f^* \in C$.

c) Check that f^* maximizes $v^2(L^*,f)$ over F. In general, if f^* does not maximize $v^2(L^*,f)$ for all $f \in F$, then f_0 will not exist since in this case we have only two alternatives. Either there is no f which maximizes the variance over F, or there exists an f which maximizes the variance but is not in G, so $L=-\log f$ is not an element of C.

In summary, the results given above show that the most robust detector may be designed by finding f_0 maximizing

$$v^2(f,-\log f) = (I(f))^{-1}, \quad \text{all } f \in F, \tag{II.42}$$

and then using $L_0=-\log(f_0)$ to define θ_N (if $L_0 \in C$) and basing the threshod detector on θ_N.

II.4 The Family of P-Point Distributions

This section deals with the above detection problem when the noise has a density f which is a member of the family F of p-point distributions defined as

$$F = \{f: \int_{-a}^{a} f(x)d(x) = p, \quad f \text{ symmetric and continuous at } \pm a\} . \tag{II.43}$$

Reasons for choosing this specific family of distributions are as follows.

1) This family covers a very wide class of distributions; for fixed a and p, F contains a scaled version of almost every symmetric distribution.

2) The required information for this family, $\int_{-a}^{a} f(x)dx$, is one of the most easily measured parameters of a distribution.

To apply our procedure to this family, we must find f_0 and L_0 which are given by the following lemma.

Lemma (II.5):

Over the family F of density functions given by (II.43), the density function

$$f_0(x) = \begin{cases} b_1 \cos^2(c_1 x), & |x| \le a \\ b_2 e^{-c_2|x|} & |x| > a, \end{cases} \qquad \text{(II.44)}$$

which, for appropriate choices of b_1, b_2, c_1, and c_2, is continuous, and has a continuous first derivative, has the following properties: a) $L_0 = -\log f_0 \in C$, b) $\inf_{L\in C} v^2(f_0,L) = v^2(f_0,L_0) = 1/I(f_0) = \sup_{f\in F} v^2(f,L_0)$. The proof of this lemma can be found in either [18],[19] or in [20].

Lemma II.5 shows that the detection strategy will be to find θ_{NH} and θ_{NV} such that

$$\sum_{i=1}^{N} l(x_i - \theta_{NH}) = 0 \qquad \text{(II.45)}$$

and

$$\sum_{i=1}^{N} l(y_i - \theta_{NV}) = 0 \qquad \text{(II.46)}$$

where

$$l(t) = \begin{cases} c_1 \tan(c_1 t) & |t| \le a \\ c_2 \operatorname{sgn}(t) & |t| > a \end{cases} \tag{II.47}$$

It can be shown that this reduces to

$$\sum_{i=1}^{N} l_1(x_i - \theta_{NH}) = 0 \tag{II.48}$$

and

$$\sum_{i=1}^{N} l_1(y_i - \theta_{NV}) = 0 \tag{II.49}$$

with

$$l_1(t) = \begin{cases} \tan(c_1 t) & |t| \le a \\ \tan(c_1 a) \operatorname{sgn}(t) & |t| > a \end{cases} \tag{II.50}$$

θ_N must then be compared to the threshold γ,

$$\sqrt{N}\,\theta_N \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{II.51}$$

where

$$B_{L_0}(0|f_0) = \alpha = e^{-\frac{\gamma^2}{2\sigma_0^2}} \tag{II.52}$$

with

$$\sigma_0^2 = \frac{1}{I(f_0)} = \frac{1}{4a\, b_1\, c_1^2} \tag{II.53}$$

Thus, we have

$$B_{L_0}(v|f) \geq B_{L_0}(v|f_0) = Q(\frac{v}{\sigma_0}, \sqrt{-2\ln\alpha}) \qquad \text{(II.54)}$$

II.5 Numerical Results

To examine the finite sample size performance of these M-detectors, we apply the proposed procedure to 3 different distributions

1) Lognormal distribution

$$f(u) = \frac{1}{(2\pi\sigma^2 u^2)^{1/2}} \exp\left[\frac{-2[\ln^2(u)]}{\sigma^2}\right]$$

where $\sigma^2 = 6$ db.

2) Contaminated normal distribution

$$f(u) = (1-r^2)\frac{u}{\sigma^2}\exp\left(\frac{-u^2}{2\sigma^2}\right) + \frac{r^2}{g^2\sigma^2}\exp\left(\frac{-u^2}{2g^2\sigma^2}\right)$$
$$+ \frac{2r(1-r)}{g\sigma^2}\exp\left(\frac{-u^2(g^2+1)}{4g^2\sigma^2}\right) I_0\left(\frac{u^2(g^2-1)}{4g^2\sigma^2}\right)$$

Notice that this is the probability density of the envelope of two orthogonal components, each having a density function of the form

$$f(w) = \frac{(1-r)}{(2\pi\sigma^2)^{1/2}}\exp\left(\frac{-w^2}{2\sigma^2}\right) + \frac{r}{(2\pi g^2\sigma^2)^{1/2}}\exp\left(\frac{-w^2}{2g^2\sigma^2}\right)$$

r and g was taken to be 0.25 and 2.25 respectively.

3) Rayleigh distribution

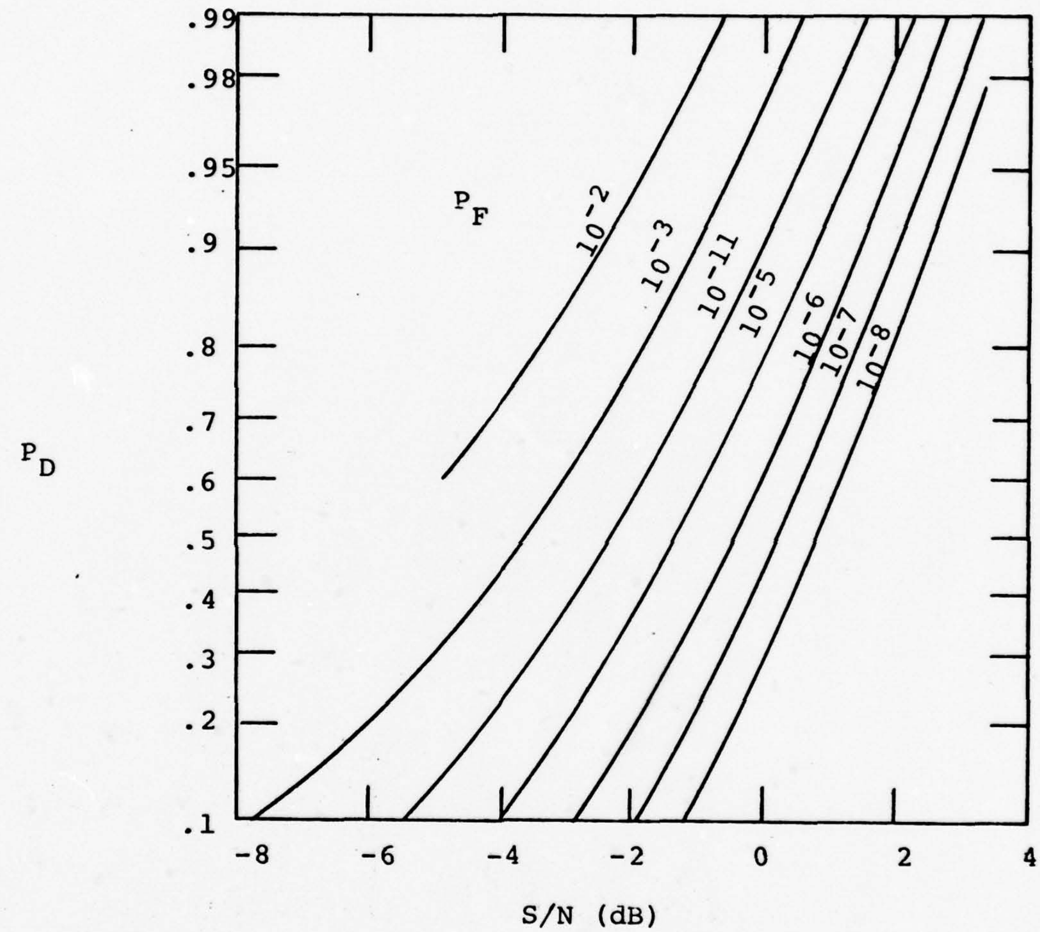
$$f(u) = u \; e^{\frac{-u^2}{2}}$$

Each of the above distributions was considered to be a member of a p-point class of distributions. We set p=.5, then found a. Having a and p the parameters of the nonlinearity $l_1(.)$ in equation (II.50) were calculated, then equations (II.48) and (II.49) were utilized to find θ_N . Using θ_N a test of the form given by (II.51) was conducted. To find the value of the threshold γ a run was made first under H_0 using 10^6 groups of samples then a minimum mean square error technique was applied to fit the output data to a curve of the form

$$- \ln P_F = a_0 + a_1 \gamma + a_2 \gamma^2.$$

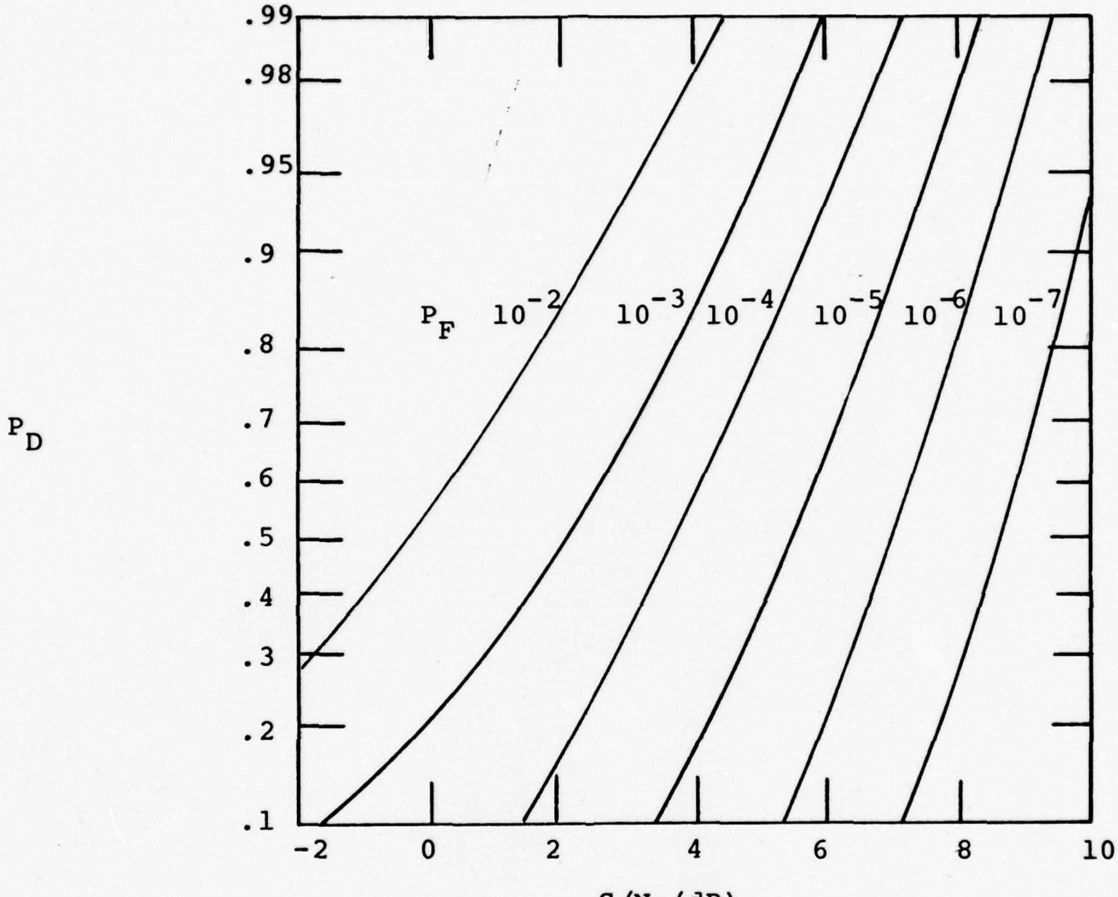
The above equation was then used to set the threshold for low probabilities of false alarm.

Figures 1 and 2 show the probability of detection P_D verses the signal to noise ratio (S/N) for different values of probability of false alarm For log-normal, (S/N) here is defined as 20 $\log_{10} \frac{A}{M}$, where M is the median of the distribution of the envelope. Figures 3 and 4 show the same for the contaminated normal and the Rayleigh distributions. Notice that the signal to noise ratio required to achieve a certain probability



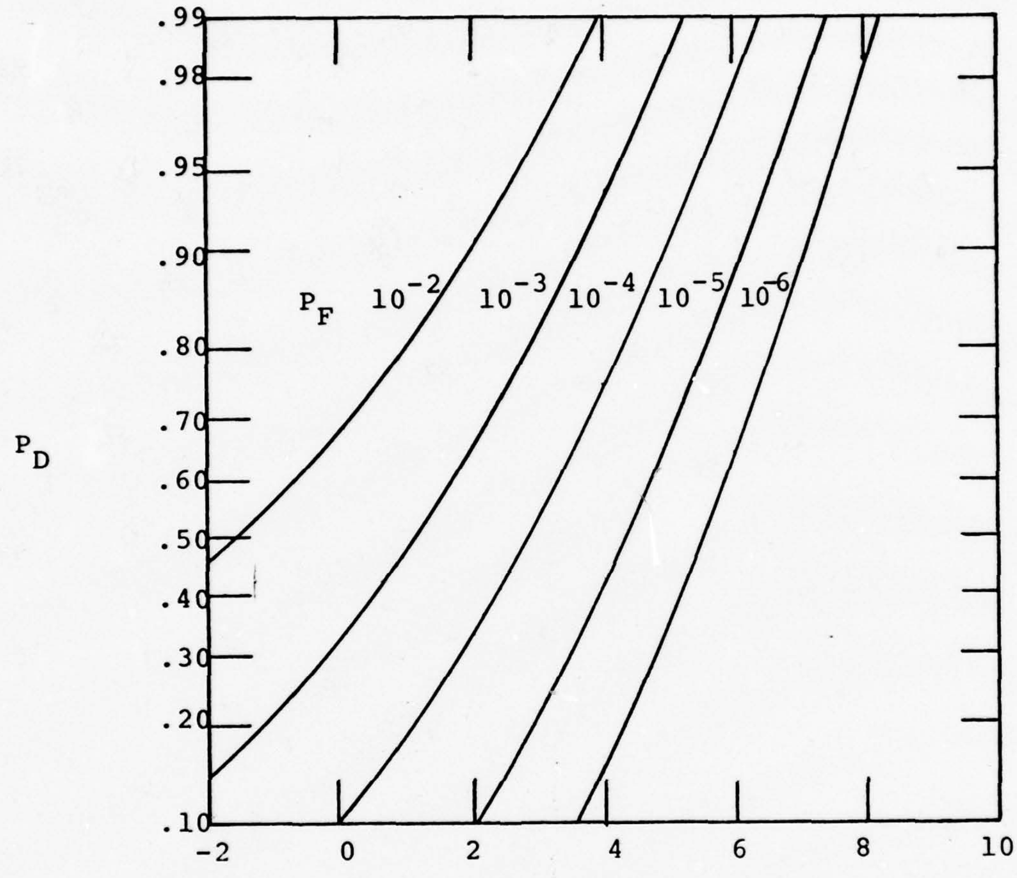
probability of detection vs S/N for the log-normal distribution (σ=6 dB) and M-detector on components: N=30.

Figure 1



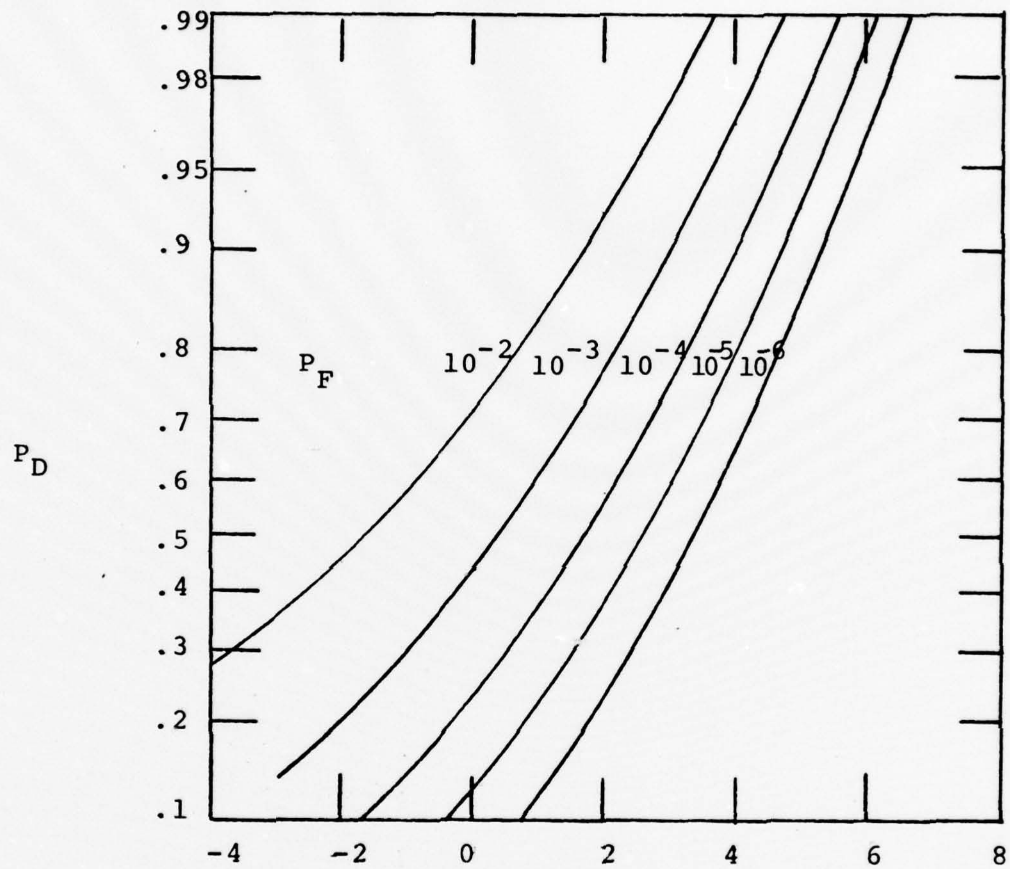
probability of detection vs S/N for the log-normal distribution (σ=6 dB) and M-detector on components: N=10.

Figure 2



probability of detection vs S/N for the contaminated normal distribution (r=0.25,g=2.25) and the M-detector on components; N=10

Figure 3

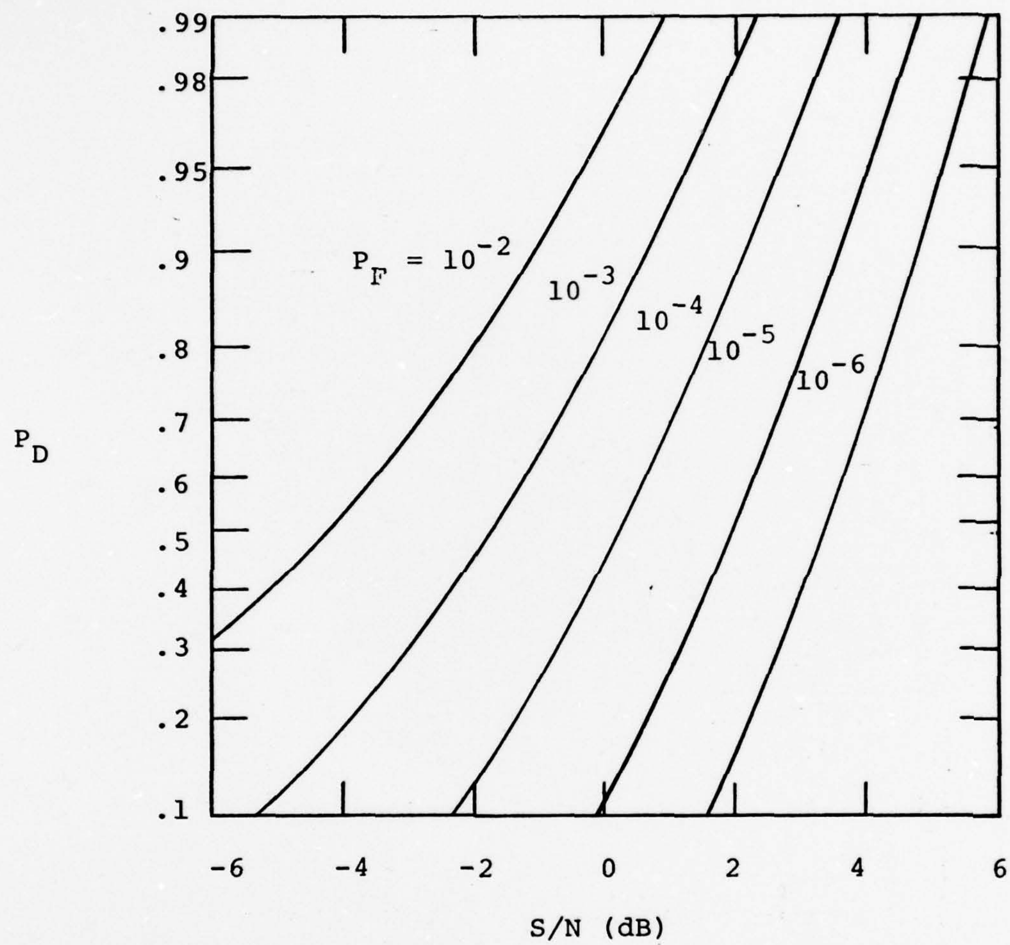


probability of detection vs S/N for the Rayleigh distribution , the M-detector on components; N=10.

Figure 4

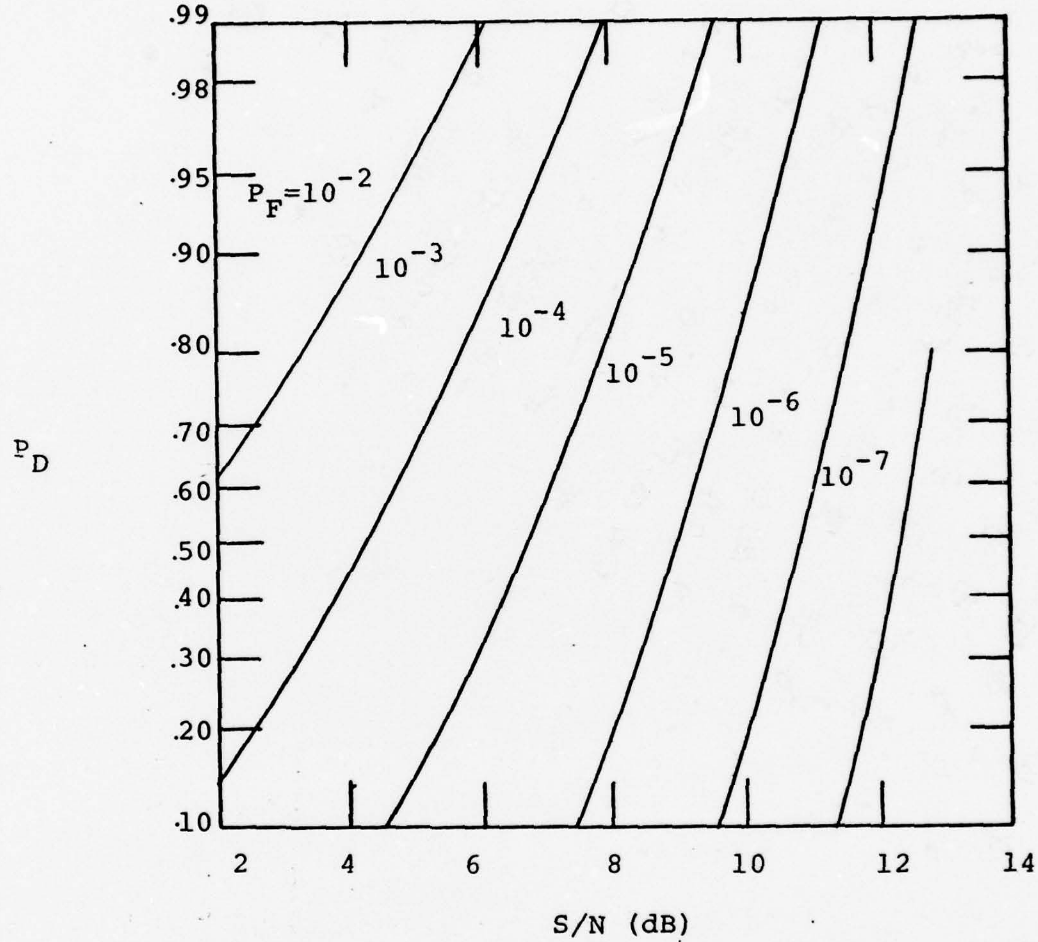
of detection at any fixed value of probability of false alarm in the case of lognormal is higher than that for contaminated normal, which in its turn is higher than the S/N in the Rayleigh case. This is due to the fact that the lognormal has the longest tail and the Rayleigh has the shortest tail.

Figures (5-8) show P_D verses S/N for a square law detector for the above distributions. Figures (9-12) compare the performance of the conventional square law-detector to the performance of the proposed M-detector. Examining figures (9) and (10) we notice that the M-detector always has a better performance than the square law detector. Notice also that the improvement in S/N becomes better as P_F decreases and achieves about 4 db at 10^{-6}. For the contaminated normal, figure (11) shows that this improvement is between (.5) and 1 db with the M-detector, better than the square law detector. It should be taken into account that the p-point family is not the best family to represent the contaminated normal and much more improvement can be achieved by the utilization of some other distribution classes e.g. 21. For the Rayleigh distribution, figure (12) shows that the M-detector is about 1 db worse than the square law detector. This loss should be expected since the square law detector is optimal in this case.



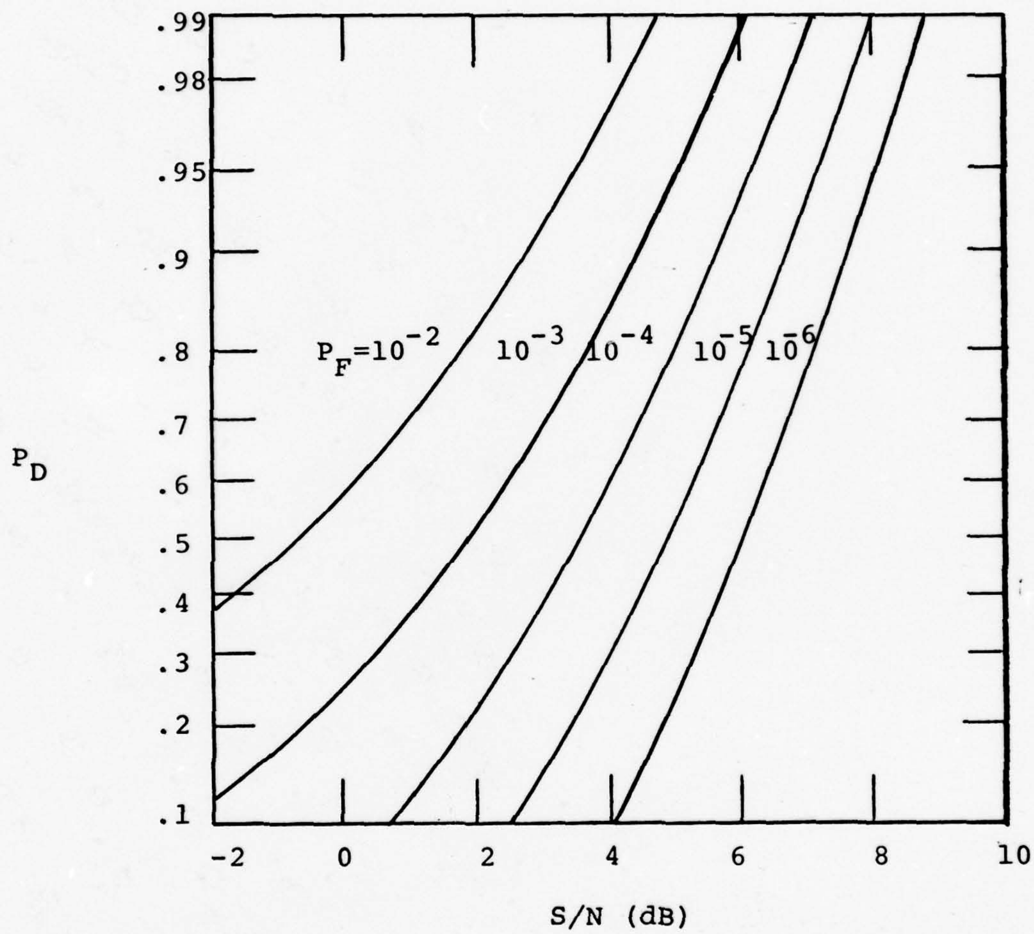
probability of detection vs S/N for the log-normal distribution (σ=6 dB) and the square law detector: N=30

Figure 5



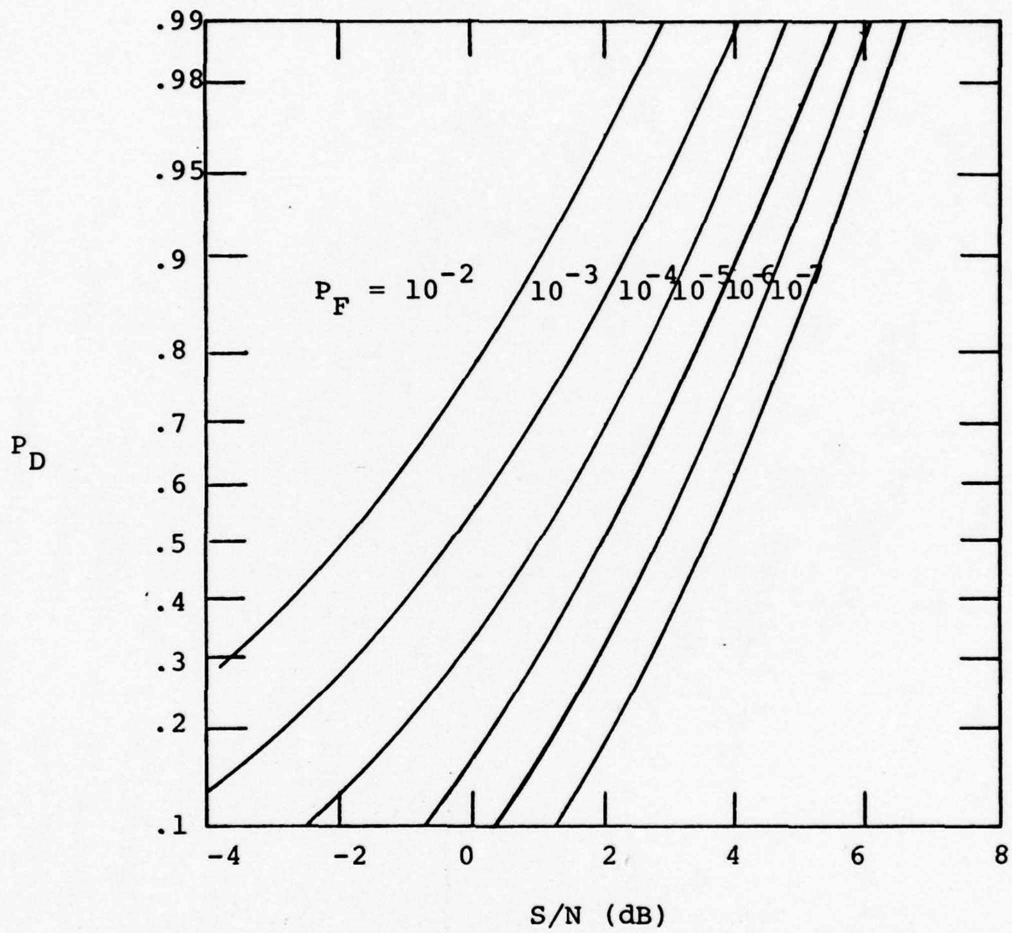
probability of detection vs S/N for the log-normal distribution (σ=6 dB) and the square law detector: N=10

Figure 6



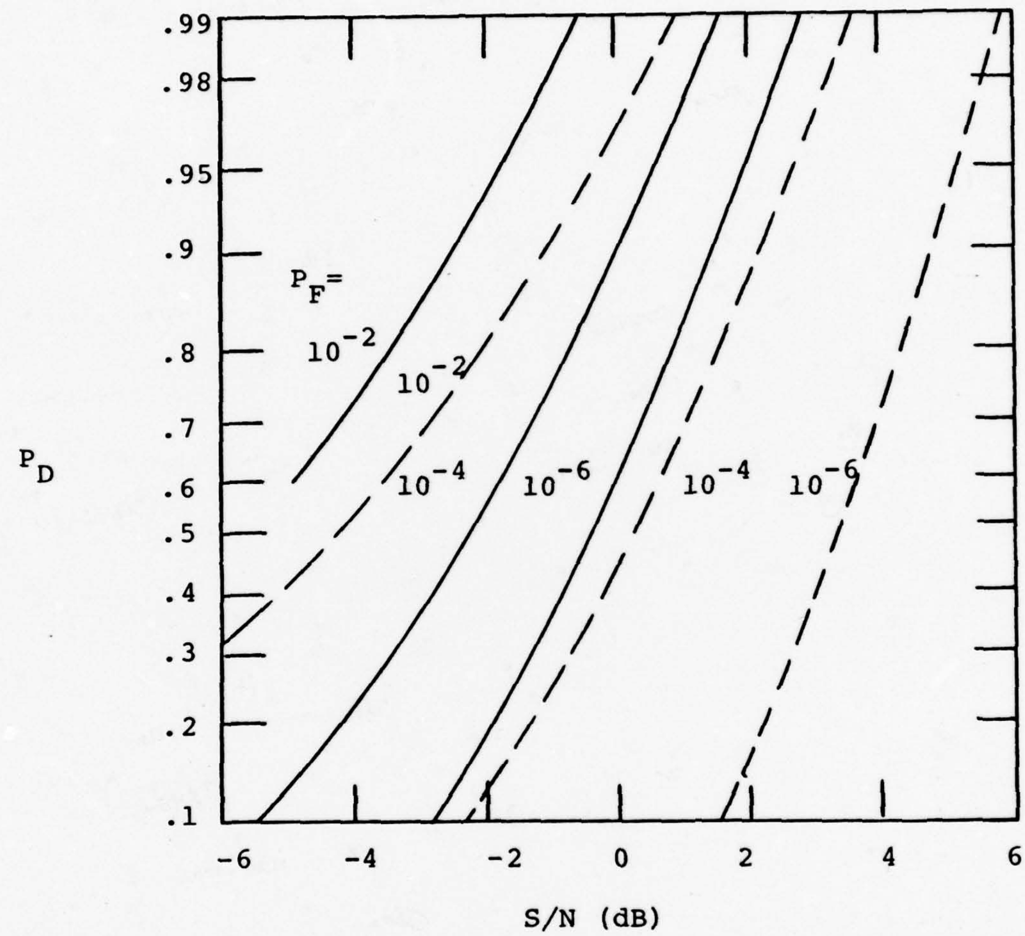
probability of detection vs S/N for the contaminated normal distribution (r=0.25,g=2.25) and the square law detector; N=10

Figure 7



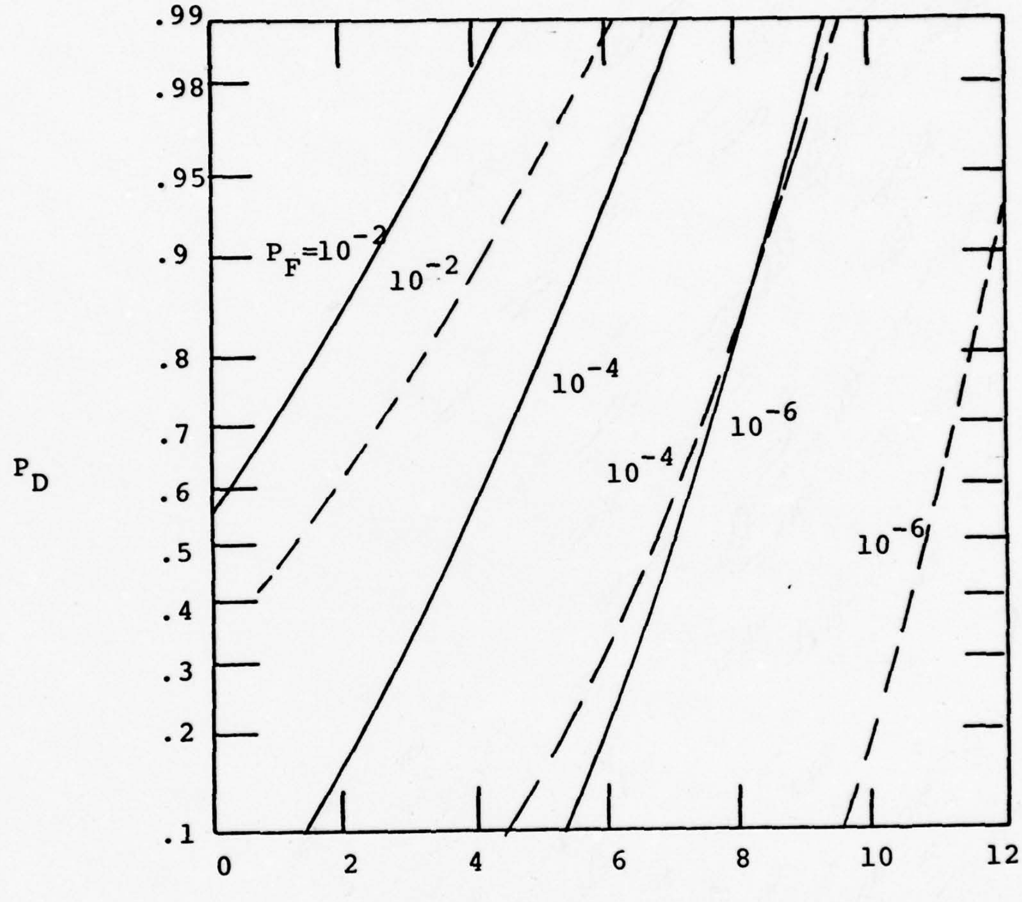
Probability of detection vs S/N for the Rayleigh distribution and the square law detector; N=10

Figure 8



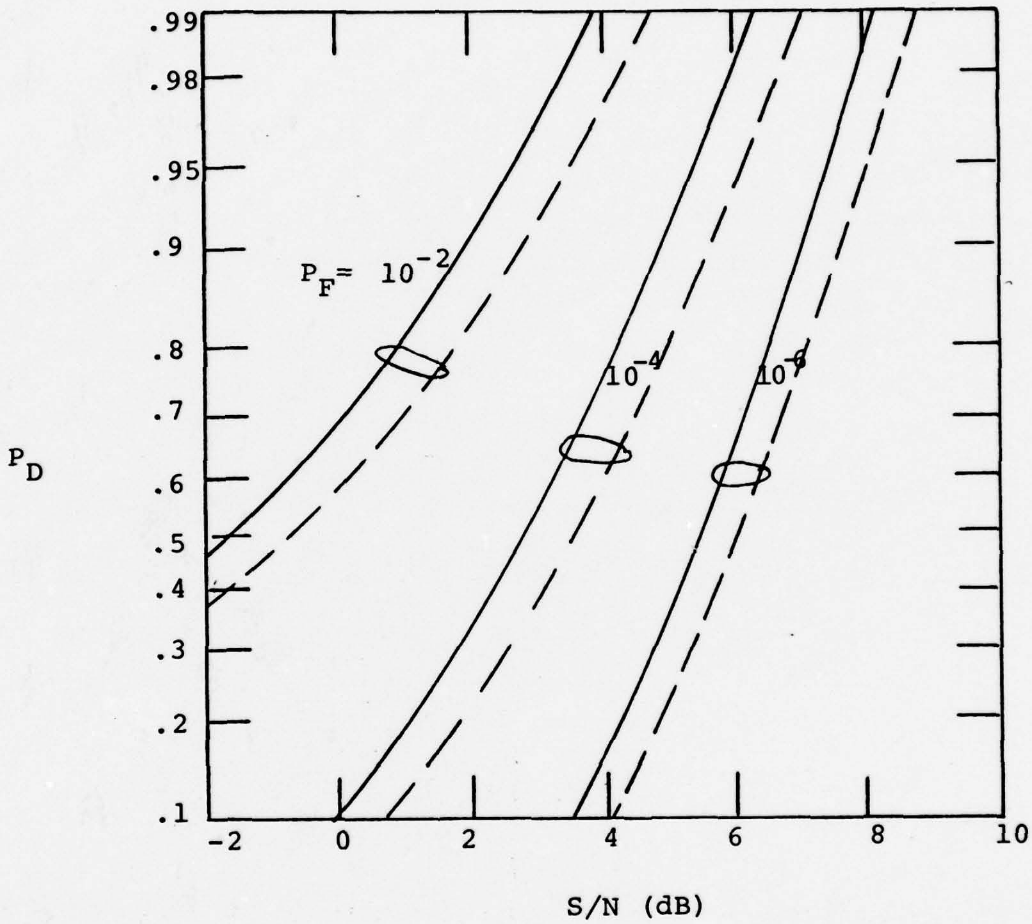
Comparison of square law detector (---) and M-detector on components (—) in log-normal (σ=6dB) interference; N=30

Figure 9



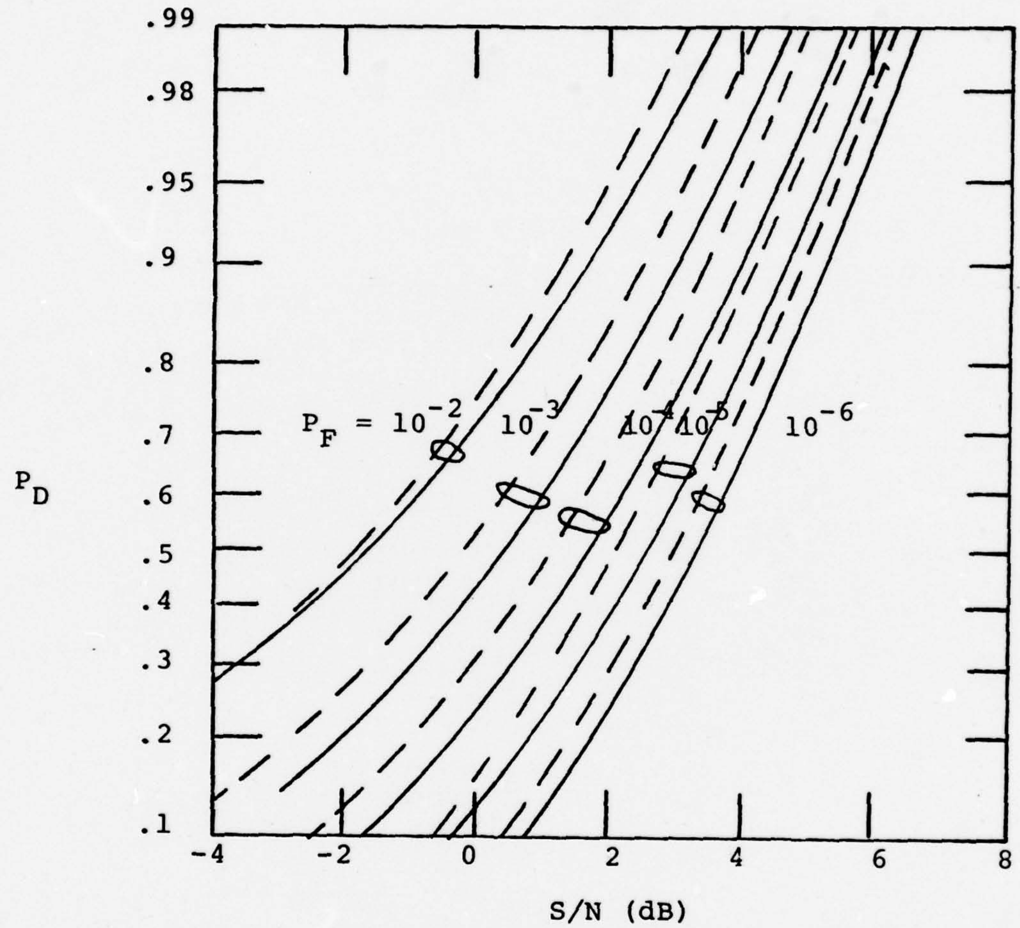
Comparison of square law detector (---) and M-detector on components (—) in log-normal (σ=6dB) interference; N=10

Figure 10



Comparison of square law detector (---) and M-detector on components (—) in contaminated normal interference (ϵ=.25); (g=2.25); N=10.

Figure 11



Comparison of square law detector (---) and M-detector on components (—) in Rayleigh interference; N=10

Figure 12

From the above results we can conclude that utilization of minimal information such as one quantile might lead to improvement in the performance of about 4 db or more. Comparing this to a loss of 1 db in the Rayleigh case, and taking into account that the longer tail distributions are much more likely than the shorter tail ones, we find the M-detector is highly preferable than the square law detector. One should also expect to get more improvement in the performance of the M-detector by utilization of more information. For example, if we have some idea about some of the partial moments, the family of p-point distributions would be replaced by the generalized moment constrained class defined by

$$F = \{f: \int S_n \, fdx \leq P_n \ , \ 1 \leq n \leq N$$

or

$$F = \{f: \int S_n \, fdx = P_n, \ 1 \leq n \leq M \text{ and}$$

$$\int S_n \, fdx \leq P_n, \ M+1 \leq n \leq N\}$$

Notice that, if $S_n(x) = 1$ for $-a \leq x < a$ and zero otherwise we get a quantiles and if $s_n(x) = x^r$ for $-b \leq x \leq b$ zero otherwise we get partial moments. More details about this class are in [22].

An important remark about these similation is that if we normalized the data with the appropriate value of a, then we can set one threshold to guarantee upper bound on the probability of false alarm even at small values of N as N=10. Also, a lower bound on the probability of detection exists. These bounds usually are the threshold and signal required to achieve certain P_F and P_D when the distribution is lognormal. In fact, it was noticed that the changes in P_F and P_D when the distribution of the noise changes from lognormal to contaminated normal are very small.

II.6 Summary

In this chapter we presented a procedure of the detection of signals with unknown phase in noise, under the assumption that the noise distribution is unknown member of a class of symmetric density functions F. It was shown that this method asymptotically guarantees an upper bound on both probability of false alarm and missing whenever the noise density function is element in the class F under consideration. Simulation results were given to show the finite sample size performance of this method and to compare it the conventional square law detector.

III. Detection of Partially Coherent Signals in Noise II

III.1 Introduction

In the previous chapter the problem of detection of partially coherent signals in noise was solved by introducing the M-detector on components. In this chapter, we continue investigating the same problem as defined in Section II.2. We introduce another approach which is called the stochastic approximation detector (SA). It may be noticed, that the M-detector depends on finding θ_{NH} and θ_{NV} as M-estimates of A_1 and A_2 respectively, and then conducting a threshold test using $\theta_N = \sqrt{\theta_{NH}^2+\theta_{NV}^2}$ as a test statistic. The SA-detector will be exactly the same as the M-detector except that the M-estimates (θ_{NH} and θ_{NV}) will be replaced by their stochastic approximation counterparts S_{NH} and S_{NV} respectively.

In the following section we shall give a brief description of the SA-estimates and their properties.

III.2 Stochastic Approximation estimation

Consider the sequence of observations $\{x_i\}$ given by

$$x_i = s + w_i \qquad i = 1,\ldots, N \tag{III-1}$$

where $\{w_i\}$ is a sequence of independent identically distributed random variables with a density function $f(w)$, which is an unknown member of a class of symmetric density functions F, and S is an additive signal. It is of interest to find an estimate S_N of the form

$$S_N = S_{N-1} + \frac{B}{N} g(x_N - S_{(N-1)}) \qquad \text{(III-2)}$$

where B is a constant and $g(\cdot)$ is some function such that S_N is a consistant estimate of S and asymptotically normal with variance $v_S^2(f,g)$. It would be of interest also to find a pair (f_0,g_0) such that

$$v_S^2(f_0,g_0) = \min_{g} \max_{f \varepsilon F} v_S^2(f,g) \qquad \text{(III-3)}$$

This problem has been investigated by several researchers and we shall summarize some of their results here.

Assume that $\mu(d) = E[g(x-d)]=0$ has a unique solution $d=s$. Consider also, the following set of assumptions:

Assumption (A_1): μ is a Borel-measurable function; and

$$(d-s)\mu(d) > 0$$

for all $d \neq s$.

Assumption (A_2): For some positive constants k and k_1, and for all d

$$k|d-s| \leq |\mu(d)| \leq k_1 |d-s|$$

Assumption (A_3): For all d

$$\mu(d) = \eta(d-s) + \delta(d,s)$$

where $\delta(d,s) = 0(|d-s|)$ as $(d-s) \to 0$ and $\eta \to 0$

Assumption (A_4): a) $\sup\limits_{d} E[g(x-d)-\mu(d)]^2 < \infty$

b) $\lim\limits_{d \to s} E[g(x-d)-\mu(d)]^2 = \sigma^2$

Assumption (A_5):

$$\lim_{R\to\infty} \lim_{\varepsilon\to 0^+} \sup_{|d-s|<\varepsilon} \int_{|z(d)|>R} z^2(d)\,dR = 0$$

where $z(d) = g(x-d) - \mu(d)$

Theorem (III-1) [23]:

Suppose that assumptions (A_1) through (A_5) are satisfied and $2Bk>1$. Then $N^{1/2}(S_N - S)$ is asymptotically normally distributed with mean 0 and variance $B^2\sigma^2(2B\eta-1)^{-1}$.

The proof will be omitted. Sacks also showed that the above theorem remains true (with K replaced by η) if assumption (A_2) is replaced by (A_2'):

Assumption (A_2'): For all d and some positive k_1

$$|\mu(d)| \leq k_1 |d-s|$$

and for every t_1, t_2 such that $0<t_1<t_2<\infty$

$$\inf_{t_1 \leq |d-s| \leq t_2} |\mu(d)| > 0$$

The above theorem gives conditions under which the SA-estimate is consistent and asymptotically normal. To achieve the minimax property, define the class of density functions F^* as

$$F^* = \{f \varepsilon F: I(f) < \infty\} . \qquad \text{(III-4)}$$

Assume that there exists a density $f_0(.) \varepsilon F^*$ such that $I(f_0) \leq I(f)$ for all f in F and define $g_0(.) = -\frac{f_0}{f_0}(.)$, then we have the following theorem from [22]

<u>Theorem (III-2)</u>:

Suppose that F is a convex set of probability density functions with F^* non-empty, $g_0(.)$ satisfies the conditions of theorem (III-1) and

$$m'_f(\alpha) = \frac{\partial}{\partial \alpha} E[-g_0(x-\alpha)] = -\int g_0(x) \; f'(x+\alpha)dx \qquad \text{(III-5)}$$

Then

$$v_S^2(f, g_0) \leq v_S^2(f_0, g_0) \leq v_S^2(f_0, g\ ) \qquad \text{(III-6)}$$

for all $f(.) \; \varepsilon \; F^*$.

Notice that the above conditions are very similar to the required conditions for the M-estimates. To see the difference in performance between the two estimators we use the following lemma from [22].

Lemma (III-1):

If the M-estimate and SA-estimate are both consistent and asymptotically normal at a density function f with positive finite variance then

$$v^2(f,L_0) \leq v_S^2(f,g_0) \qquad \text{(III-7)}$$

with equality iff $m_f'(\alpha)=E_{f_0}[g_0^2(x-\alpha)]$.

The above means that the M-estimate will always have a lower or equal variance. Another problem here is S_0, the initial value of S_N to be used in equation (III-1). Although S_N will not depend much on the specific value of S_0 asymptotically, this choice will effect S_N dramatically at finite sample sizes. Notice also that if $f=f_0$ and $g=g_0=-\frac{f_0'}{f_0}$, then $\sigma^2 = \frac{1}{I(f_0)}$ and if $B = \frac{1}{I(f_0)}$, then $v_S^2(f_0,g_0) = \frac{1}{I(f_0)}$ which is the Cramer-Rao lower bound and the variance of the estimate of S.

In the following section, we apply the above results to the detection problem in Section II.

III.3 Stochastic Approximation Detectors

Consider the problem of detection of partially coherent signals described in Section (II.2) where $\{x_i\}$ and $\{y_i\}$ are the inphase and quadrature components of the observations, respectively, given by

$$\left.\begin{aligned} x_i &= A_1 + w_i \\ y_i &= A_2 + v_i \end{aligned}\right\} \quad i=1,\ldots, N \qquad \text{(III-8)}$$

with the same assumptions and definitions in Section (II-2). Define S_N^H and S_N^V as the stochastic approximation estimates of A_1 and A_2 respectively given by

$$S_i^H = S_{(i-1)}^H + \frac{B}{i} g(x_i - S_{(i-1)}^H) \qquad \text{(III-9)}$$

and

$$S_i^V = S_{(i-1)}^V + \frac{B}{i} g(y_i - S_{(i-1)}^V) \qquad \text{(III-10)}$$

Define S_i as

$$S_i = \sqrt{(S_i^H)^2 + (S_i^V)^2} \qquad \text{(III-11)}$$

and the SA-test of hypotheses as

$$T_N = \sqrt{N}\, S_N \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \qquad \text{(III-12)}$$

where γ is the threshold of the test.

where $B_g(0|f)$ $[B_g(v|f)]$ refers to the probability of false alarm [detection] using the SA-detector with nonlinearity g when the actual noise distribution is f .

The above theorem shows that whenever the M-detector on components guarantees upper bounds on the two kinds of error probability, the SA-detector will guarantee the same bounds. On the other hand, for any density function for which the performance is better than these bounds the M-detector will have better performance. This means that in using the SA-detector in place of the M-detector we should expect a slight deterioration in performance in return for the computational convenience. In the next section we shall give some simulation results to evaluate the finite sample size performance of the SA-detector and to find out how much loss in efficiency we suffer by replacing the M-detector by the SA-detector.

III-4 Numerical Results

To evaluate the finite sample size performance of the SA-detector, simulations were done for three different distributions, lognormal, contaminated normal and Rayleigh. These distributions are the same utilized in the study of the M-detector. To conduct these simulations, each distribution was considered to be a member of a class of p-point distributions with p=0.5 as before. Then $f_0(.)$ is given by

$$f_0(x) = \begin{cases} b_1 \cos^2(c_1 x) & |x| < a \\ b_2 e^{-c_2 x} & |x| > a \end{cases} \qquad \text{(III-17)}$$

and

$$g_0(x) = \begin{cases} b \tan(c_1 x) & |x| < a \\ b \tan(c_1 a) \operatorname{sgn}(x) & |x| > a \end{cases} \qquad \text{(III-18)}$$

with b_1, b_2, b, c_1 and c_2 chosen such that f_0 is in F and $g_0(.)$ is continuous at $\pm a$. The constant B in equation (III-1) was chosen to be equal to $\frac{1}{I(f_0)}$. As a result, $v_S^2(f, g_0)$ should be asymptotically equal to B for all f in F. Three different choices for S_0, the initial value of the estimate, were utilized.

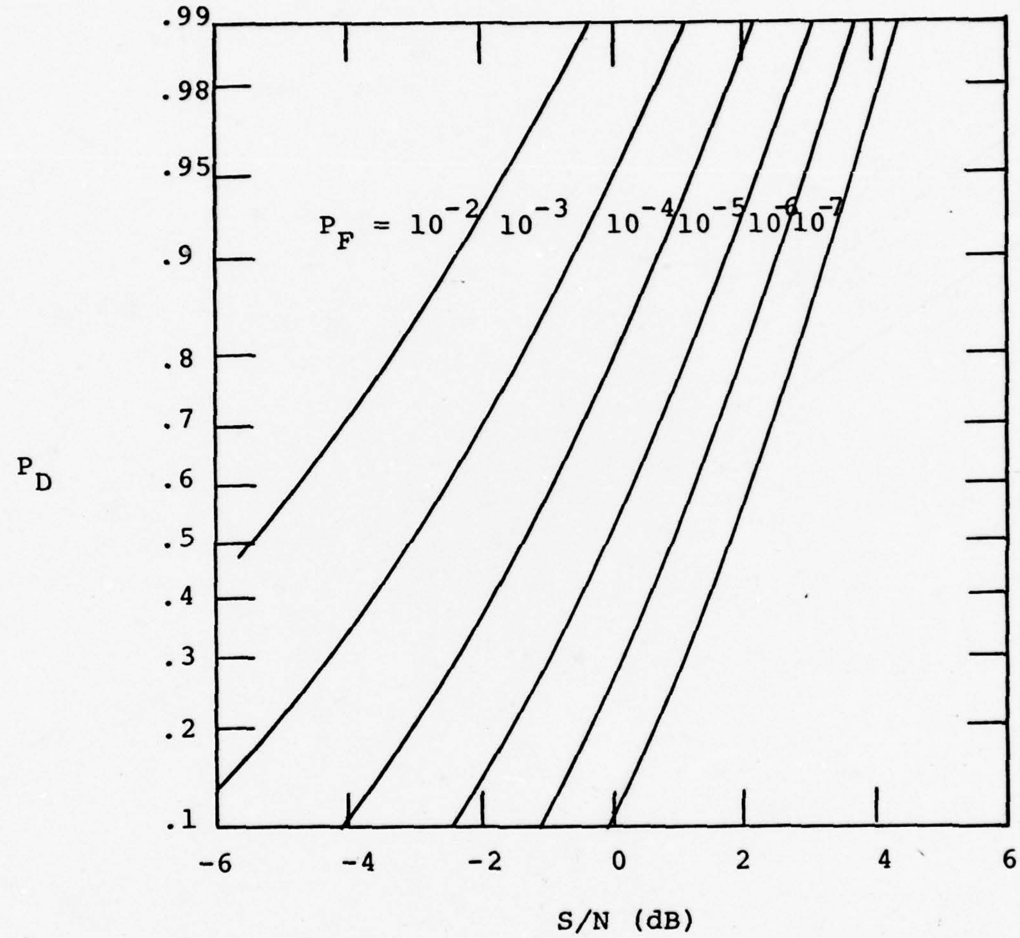
1) $S_0^H = x_1$ and $S_0^V = y_1$.

2) $S_0^H = S_0^V = 0$.

3) $S_0^H = \frac{\sum_{i=1}^{N} x_i}{N}$ and $S_0^V = \frac{\sum_{i=1}^{N} y_i}{N}$.

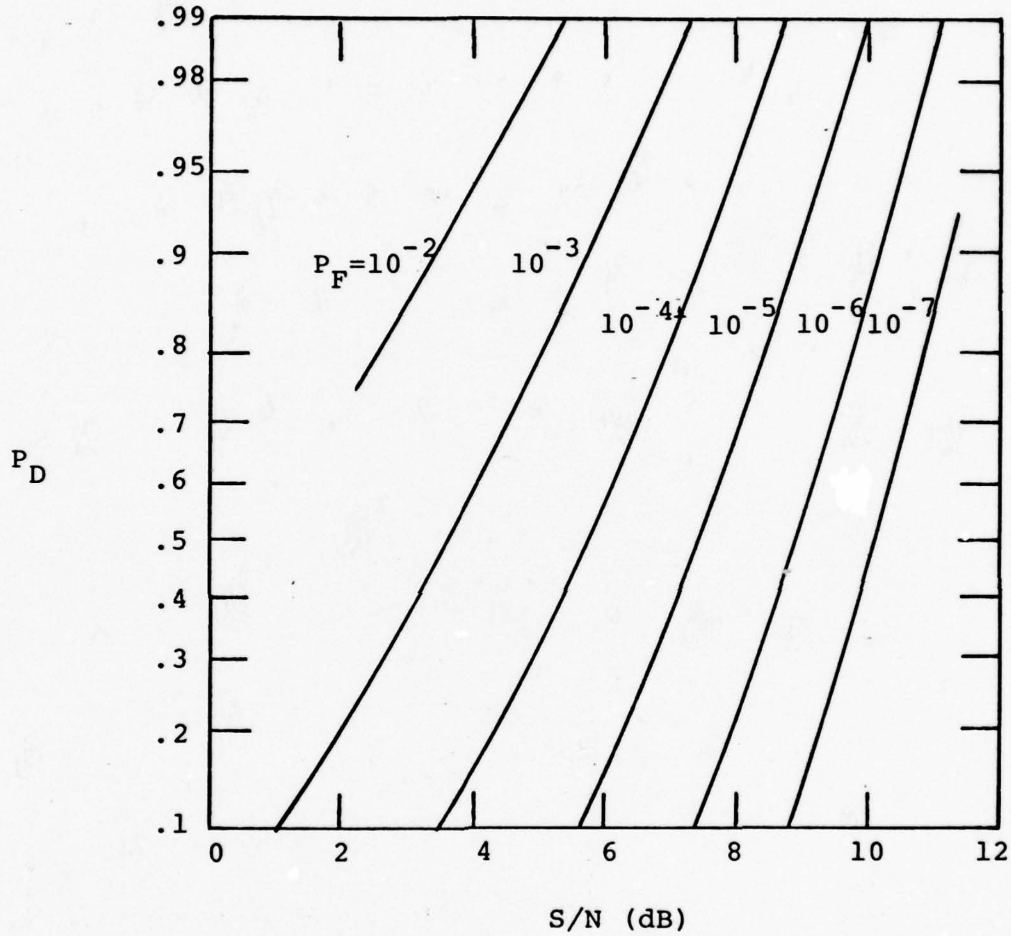
The third choice led to the best performance, so we report here on the results of these simulations only.

Figures 13,14,15 and 16 show the probability of detection for the SA-detector against the signal to noise ratio (S/N) for lognormal N=30, lognormal N=10, contaminated normal N=10 and Rayleigh N=10, respectively. Figures 17, 18 and 19 compare the performance of the SA-detector to the M-detector for lognormal and contaminated normal,



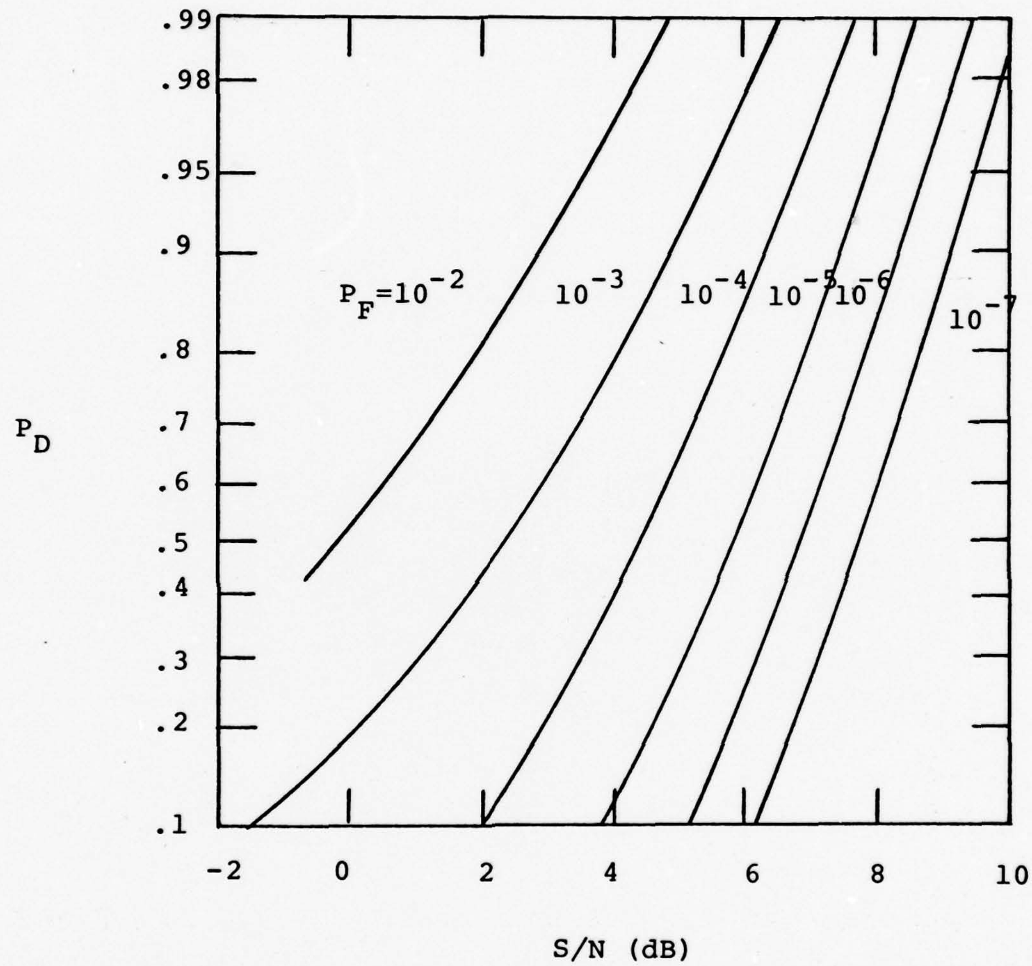
probability of detection vs S/N for log-normal distribution (σ=6dB) and the SA-detector; N=30

Figure 13



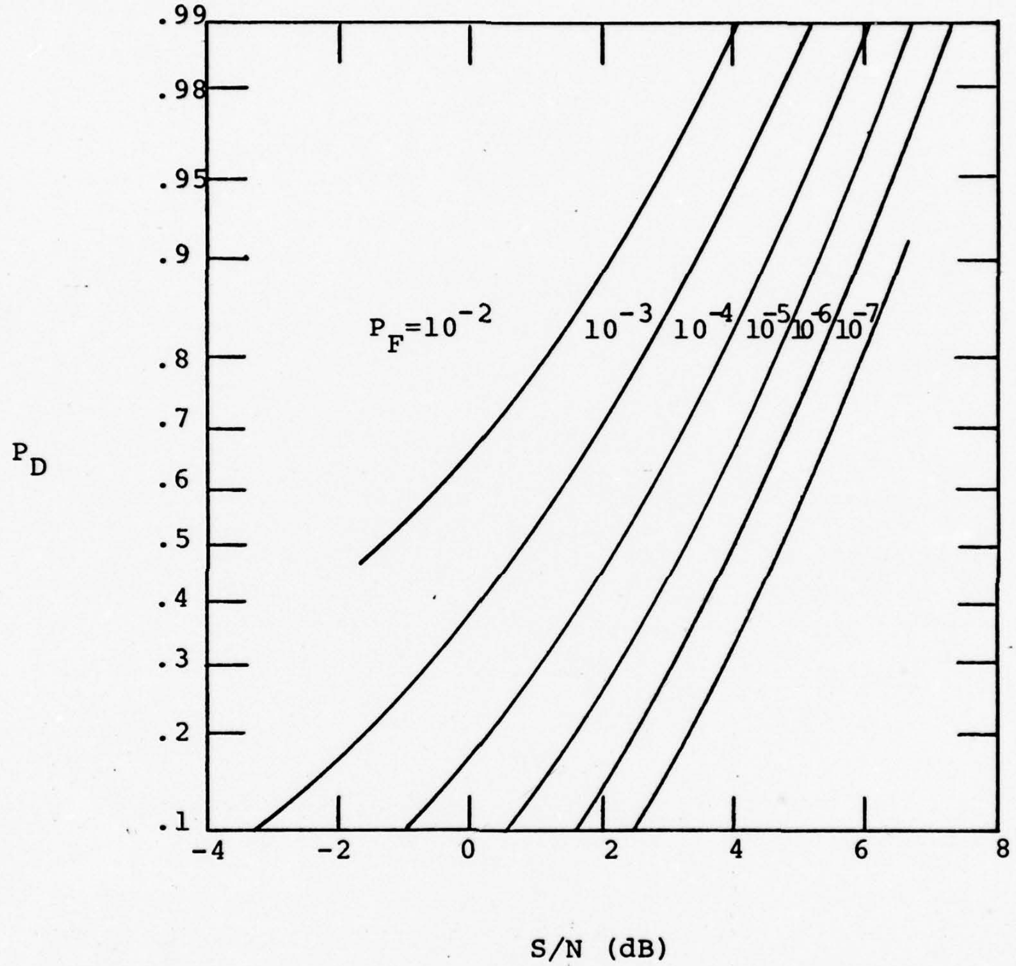
probability of detection vs S/N for log-normal distribution (σ=6dB) and the SA-detector; N=10

Figure 14



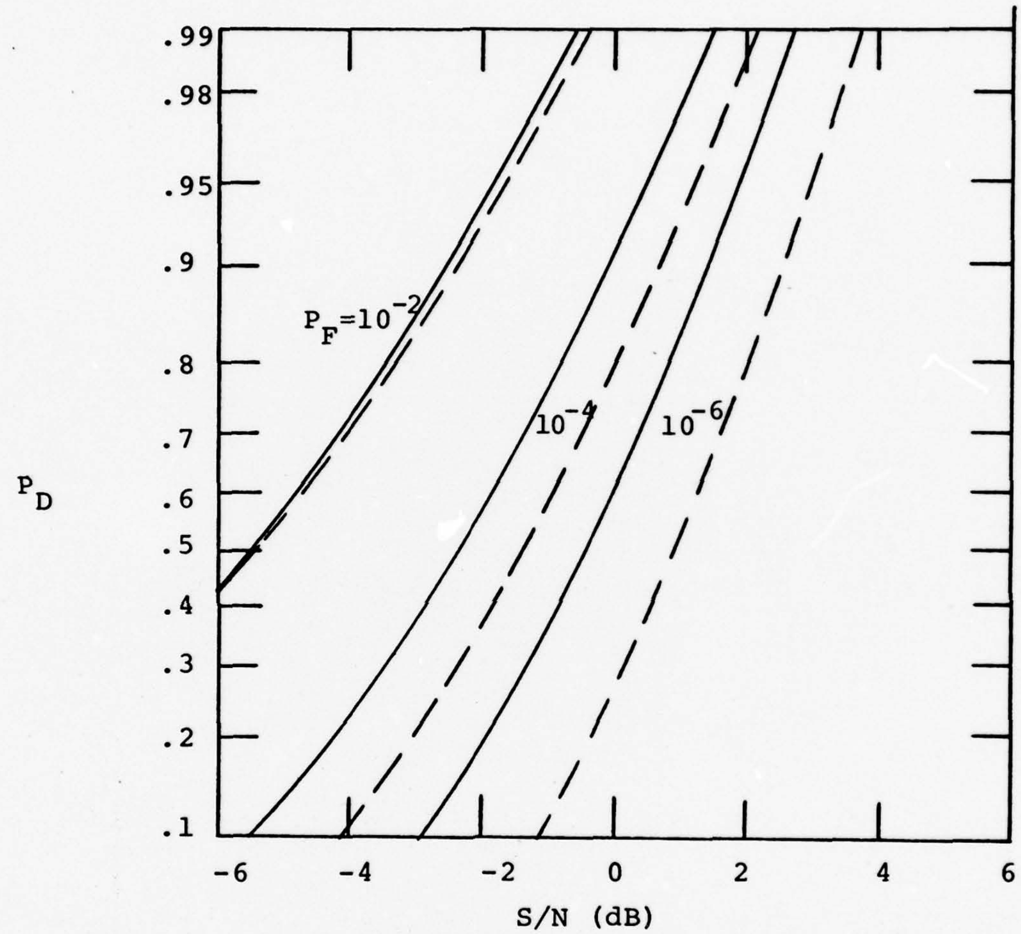
probability of detection vs S/N for contaminated normal distribution (r=0.25,g=2.25) and SA-detector; N=10

Figure 15



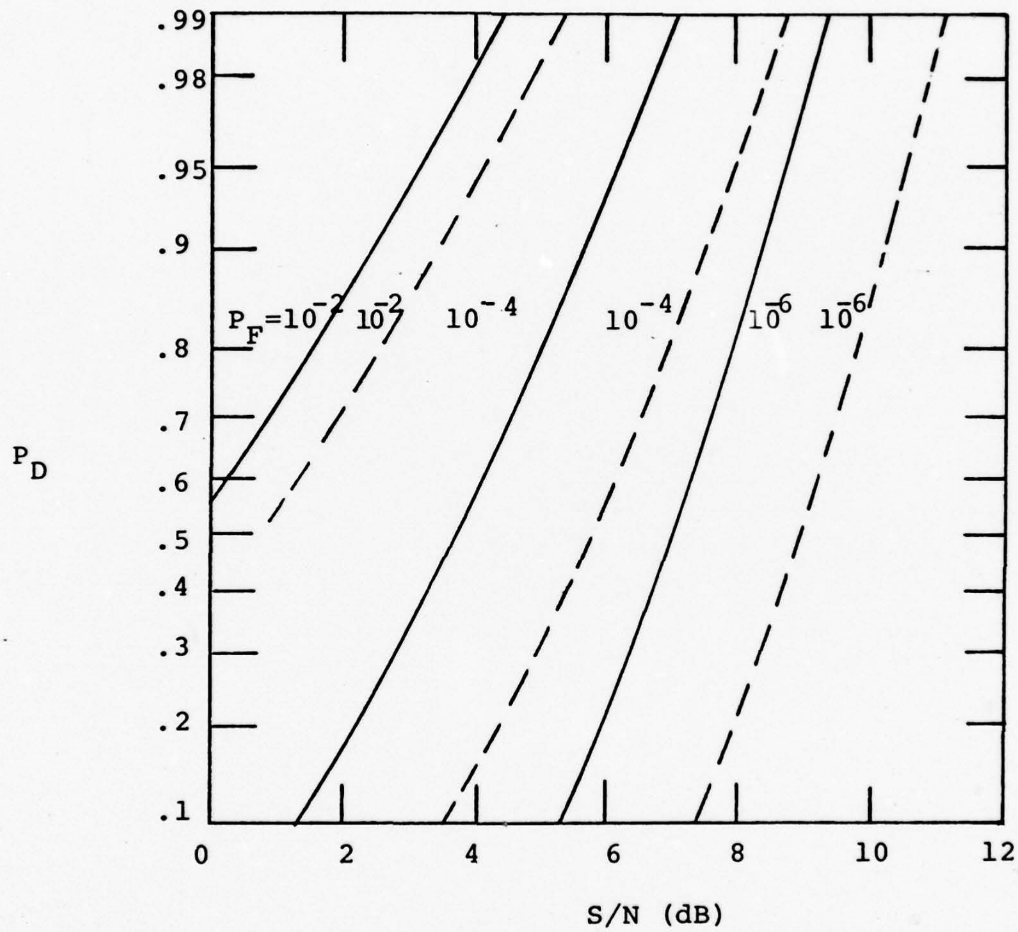
probability of detection vs S/N for Rayleigh distribution and SA-detector; N=10

Figure 16



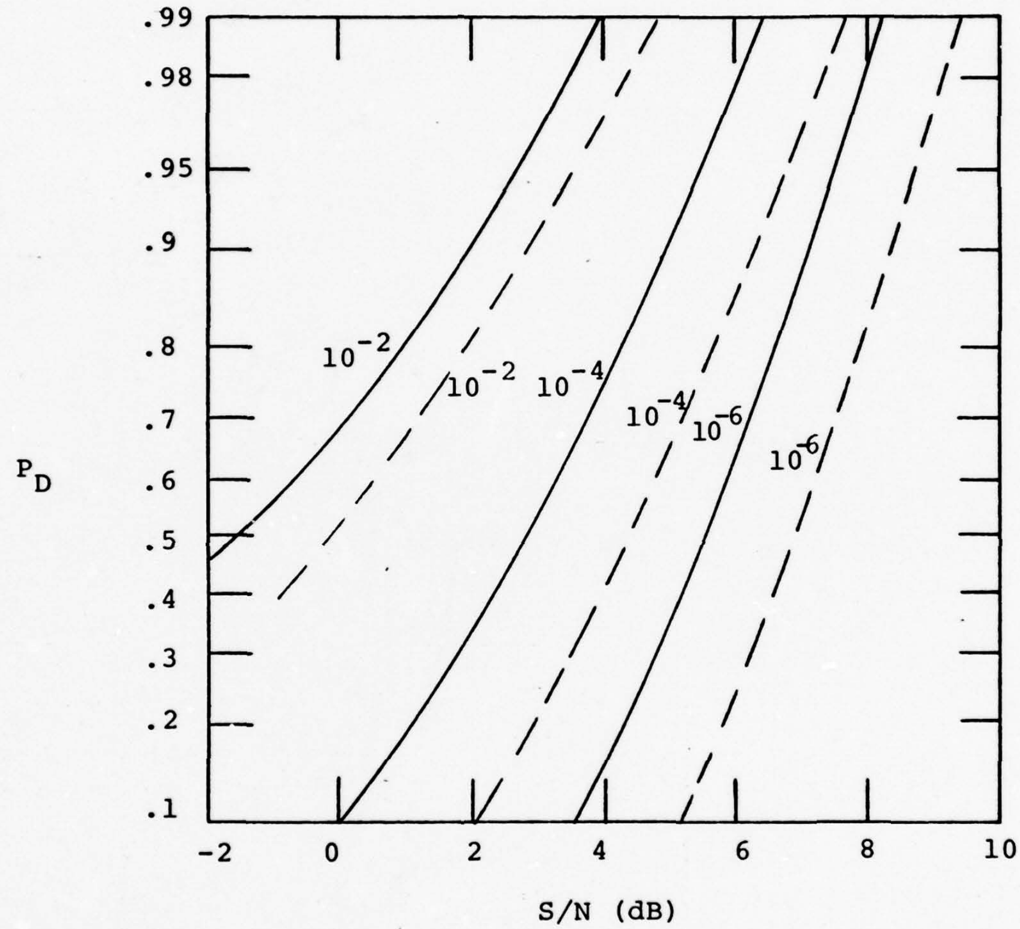
Comparison of SA-detector (---) and M-detector on components (—) in log-normal (σ=6 dB) interference; N=30

Figure 17



Comparison of SA-detector (---) and M-detector on components (—) in log-normal (σ=6 dB) interference; N=10

Figure 18



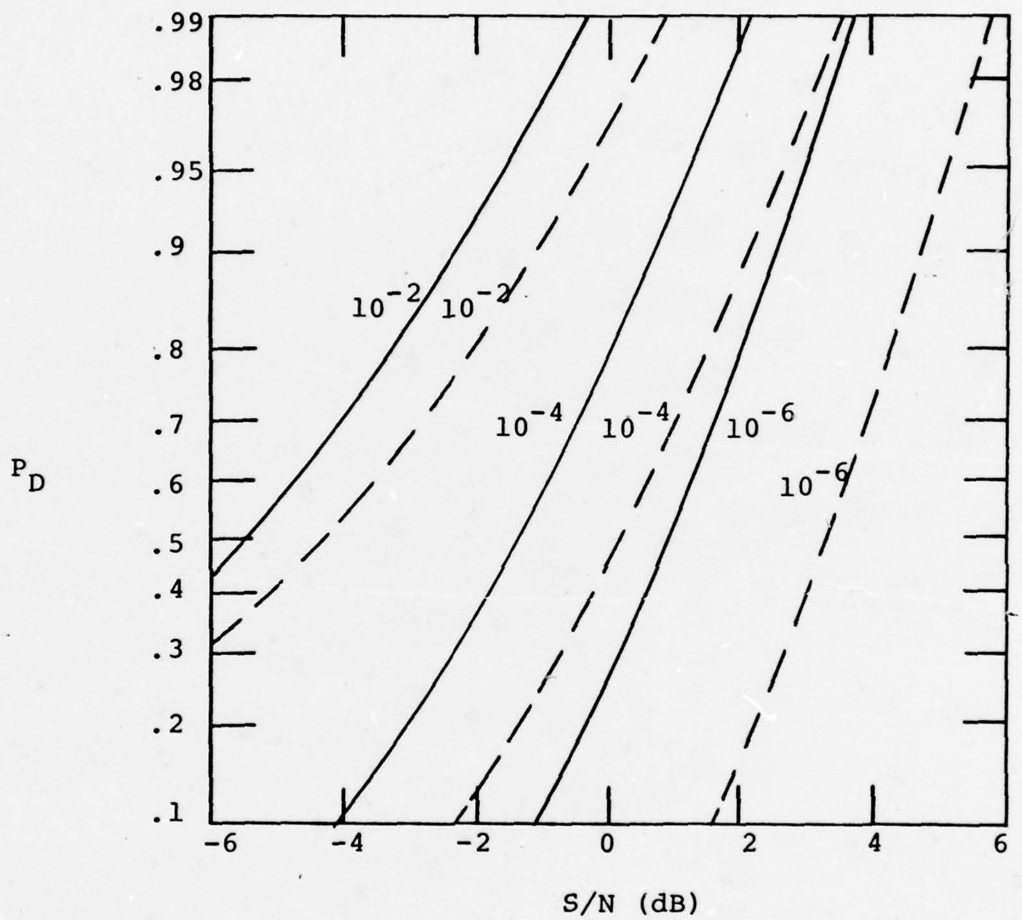
Comparison of SA-detector (---) and M-detector on components (—) in contaminated normal interference (r=0.25,g=2.24);N=10

Figure 19

and show a loss of up to about 2db in signal to noise ratio. Figures 20,21,22 compare the SA-detector to the square law detector. Although the SA-detector is better for the lognormal case it looses its superiority in the case of contaminated normal, and runs about 1db worse than the square law detector.

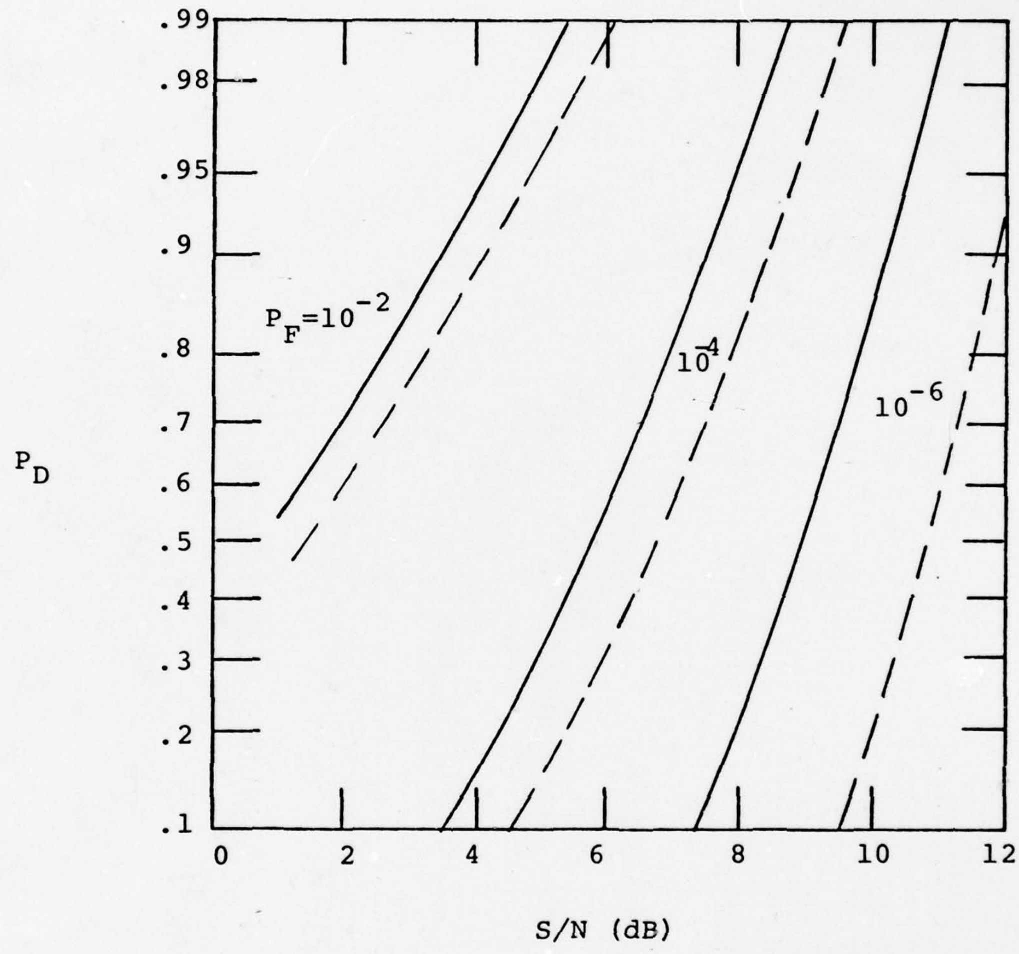
III-5 General Discussion:

The above simulations show that the deterioration in performance as a result of replacing the M-detector by the SA-detector is not slight but serious. It also shown that the performance of the SA-detector is affected not only by the choice of S_0, but also by the first few observations. To clarify this, consider N=10. Then if the first two observations are large with the same sign, then the chance of correcting this effect by the other eight observations is zero. On the other hand, the M-detector can be well-approximated by a one step Newton-Raphson version [20]. In addition, the M-detector may also be approximated by an on-line system if we are not interested in the specific value of θ_N but only in its relation to the threshold γ . By the monotonocity of the nonlinearity $l(.)$ in its argument, if $|\theta_{HN}|>d$ than either



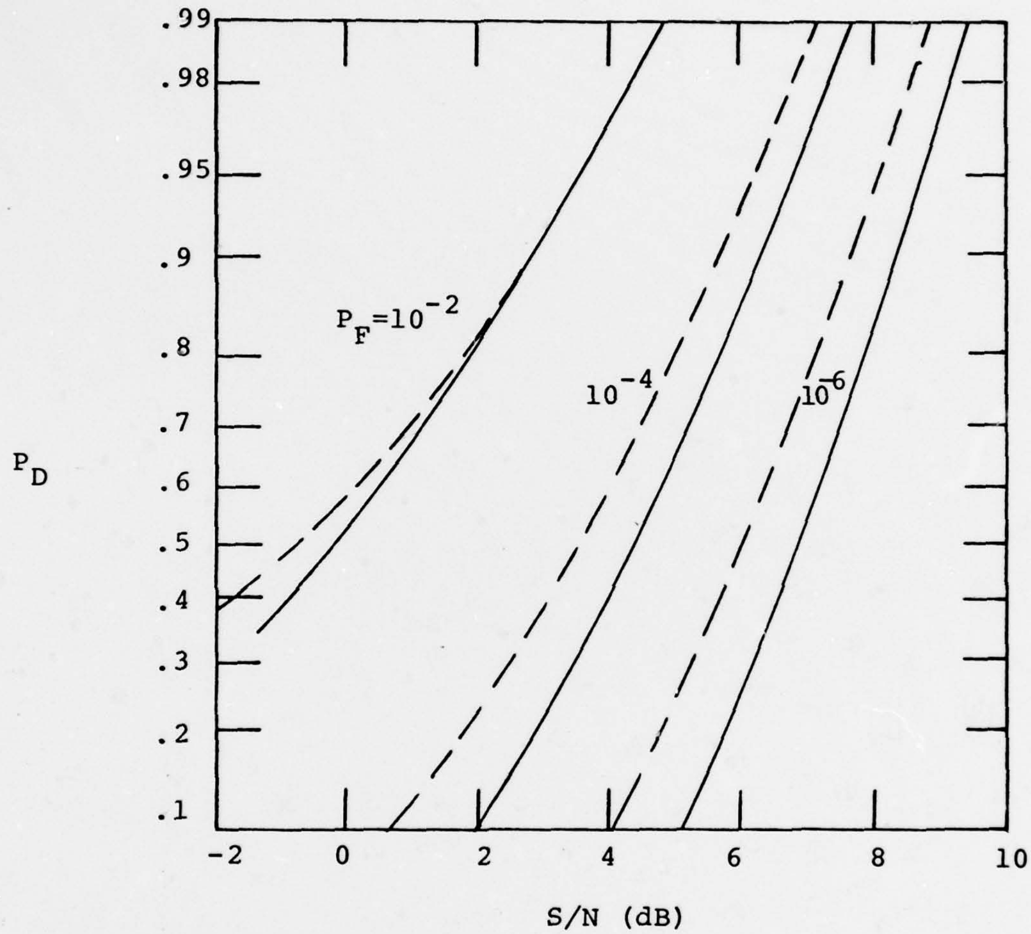
Comparison of SA-detector (—) and square law detector (---) in log-normal interference (σ=6 dB);N=30

Figure 20



comparison of SA-detector (—) and square law detector (---) in log-normal interference (σ=6 dB); N=10

Figure 21



Comparison of SA detector (—) and square law detector (---) in contaminated normal interference (r=0.25,g=2.25); N=10

Figure 22

$$\sum_{i=1}^{N} 1(x_i - d) > 0$$

or (III-19)

$$\sum_{i=1}^{N} 1(-x_i - d) > 0$$

and if $\theta_N > \gamma$ then there exists a number d such that $|\theta_{HN}| \geq d$ and $|\theta_{VN}| > \bar{d}$, where $\bar{d}^2 = \gamma^2 - d^2$. To construct the above approximate detector, we pick a set of numbers $\{\gamma_i\}$ between $\gamma/\sqrt{2}$ and γ and conduct the test as in equation (III-19). Note as the number of $\{\gamma_i\} \to \infty$ we get the exact M-detector. An example of the above implementation is shown in figure (23).

III.6 Summary

In this chapter we analysed another method for detection of signal with unknown phase in additive noise called stochastic approximation detector. It was shown that, although this method is very attractive and has good properties from the theoretical point of view, it fails completely in application at finite sample sizes. In addition, an approximate on line implementation for the M-detector on component was presented.

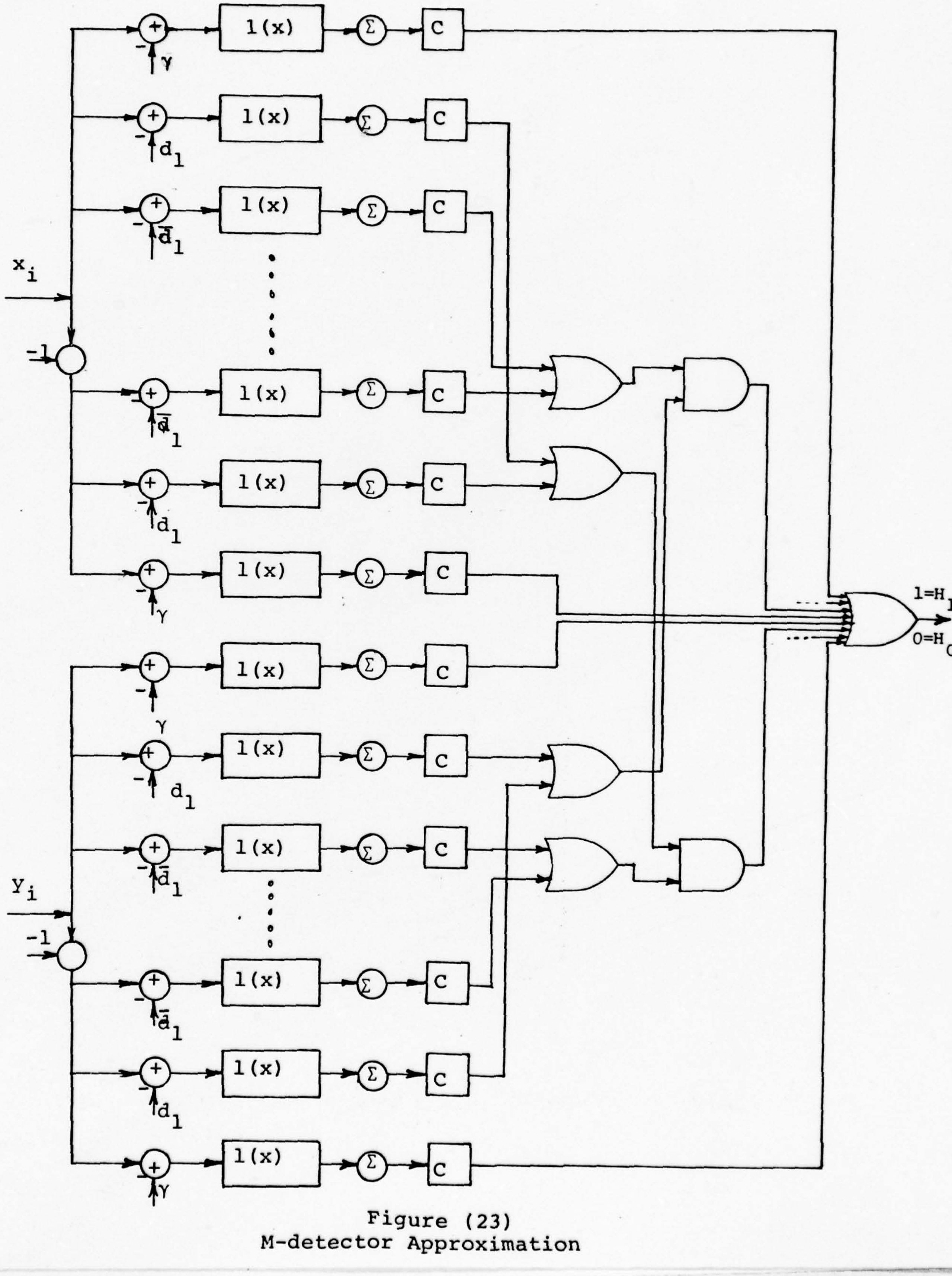


Figure (23)
M-detector Approximation

IV. Detection of Signal with Unknown Phase and Frequency

IV.1 Introduction

In the previous two chapters, the problem of detection of signals with unknown phase in the presence of additive noise was considered under the assumption that the noise distribution function is not completely known. It was assumed that all other signal parameters -- except perhaps the signal amplitude -- are known. Often in practice, in addition to the phase, the signal frequency is not completely known. The analog received signal in such cases can be modeled as

$$s_r(t) = B(t) \sin[(w_c+w)t + \phi] \tag{IV.1}$$

where w_c is a known carrier frequency. ϕ is an unknown phase taking any value in $[0,2\Pi]$ and w is an unknown frequency in the interval $[w',w'']$. Although this problem appears in many communication applications as a result of the uncertainty about the transmitter oscillator frequency, it is more frequently encountered in radar applications, especially in detection of moving targets [24]. In radar applications, this problem is usually referred to as "Doppler detection" and w is called the "Doppler Shift". Since our concern in this chapter will be mainly with the radar problem, it is in order here to discuss briefly the nature of this Doppler shift.

Assume that there exists a target which is moving with radial velocity v , then the range of this target as a function of time will be given by

$$R(t) = R_0 - v\ t \tag{IV-2}$$

where R_0 is some constant equal to the initial range at $t=0$. If the transmitted signal is

$$s_t(t) = B(t) \cos (w_c t) \ , \tag{IV-3}$$

then the signal returned from the target will be

$$s_r(t) = B(t- \tau(t)) \cos [w_c(t- \tau(t))] \tag{IV-4}$$

where $\tau(t)$ is the round-trip delay time

$$\tau(t) = \frac{2\ R[t - \tau(t)/2]}{c}$$

$$= \frac{2\ R_0/c}{1+v/c} - \frac{2(v/c)t}{1+v/c} \tag{IV-5}$$

where c is the velocity of propagation. But since $(v/c)<<1$, equation (IV-5) can be approximated by

$$\tau(t) = \frac{2R_0}{c} - \frac{2v}{c}\ t \tag{IV-6}$$

Substituting from equation (IV-6) in (IV-4), we get

$$s_r(t) = B[t-\tau+\frac{2v}{c}\ t] \cos [w_c(t-\tau)+\frac{2v}{c}\ w_c t] \tag{IV-7}$$

where $\tau = \frac{2R_0}{c}$ and is constant. Equation (IV-7) shows that the velocity of the moving target will have two main effects on the reflected signal:

a) A shift of the carrier frequency.

b) A compression or stretching of the time scale of the evelope.

The first effect is usually more serious than the second one [24]. In our study here we shall not take the second effect into consideration since the generalization of the solution will be straightforward.

Before proceeding to the solution it would be helpful to discuss the case in which the additive noise is known to be Gaussian.

IV.2 The Gaussian Noise Case

Consider the envelope of the transmitted signal to be a sequence of pulses each with duration D; the separation between pulses is T. Then if the pulse amplitude is A and the observations are corrupted by additive Gaussian noise, these observations will be in the form, under

$$H_1: \quad r_i = A\cos[(w_c+w)t + \phi] + n(t)$$

$$H_0: \quad r_i = n(t)$$

$$i=0,\ldots,(N-1) \qquad \text{(IV-8)}$$

But since the noise is Gaussian and independent, then the conditioned likelihood ratio would be expressed as

$$\lambda(r|w,\phi) = \prod_{i=0}^{(N-1)} \lambda(v_i|w\,,\phi) \qquad \text{(IV-9)}$$

where

$$\lambda(r_i|w,\phi) = \frac{\exp\{\frac{-1}{2\sigma^2}\int_{iT}^{iT+D}[r_i(t)-A\cos((w_c+w)t+\phi)]^2dt\}}{\exp\{\frac{-1}{2\sigma^2}\int_{iT}^{iT+D}[r_i(t)]^2\,dt\}}$$

$$= \exp\{\frac{+1}{\sigma^2}\int_{iT}^{iT+D} r_i(t)\;A\cos[(w_c+w)t+\phi]dt$$

$$-\frac{1}{2\sigma^2}\int_{iT}^{iT+D} A^2\cos^2[(w_c+w)t+\phi]dt \qquad \text{(IV-10)}$$

Substituting from (IV-9) in (IV-10)

$$\lambda(r|w,\phi) = F\exp\{\frac{+1}{2\sigma^2}\sum_{i=0}^{N-1}\int_{iT}^{iT+D} r_i(t)A\cos[(w_c+w)t+\phi]dt \qquad \text{(IV-11)}$$

where F corresponds to the received energy and is not a function of w or ϕ. Next, we break the remaining cosine terms up according to the trigonometric identity

$$\cos[w_0t+\phi] = \cos(w_0t)\cos(\phi) - \sin(w_0t)\sin(\phi)$$

and define

$$L_c = \sum_{i=0}^{N-1} \int_{iT}^{iT+D} r_i(t) \cos[(w_c+w)t]dt \tag{IV-12}$$

and

$$L_s = \sum_{i=0}^{N-1} \int_{iT}^{iT+D} r_i(t) \sin[(w_c+w)t]dt \tag{IV-13}$$

Thus, we can write equation (IV-11) as

$$\lambda(x|w,\phi) = F \exp\{ \frac{A}{\sigma^2} [L_c \cos \phi - L_s \sin \phi] \tag{IV-14}$$

If we assumed, as is usual in radar detector, that ϕ is uniformly distributed over the interval $[0,2\Pi]$, then

$$\lambda(r|w) = E_\phi[\lambda(r|w,\phi)]$$

$$= \int_0^{2\pi} \frac{1}{2\pi} F \exp\{ \frac{A}{\sigma^2} [L_c \cos \phi - L_s \sin \phi]\} d\phi$$

$$= F\, I_0[A(L_c^2+L_s^2)^{1/2}/\sigma^2] \tag{IV-15}$$

where $I_0(.)$ is the modified Bessel function of order 0. Notice that equation (IV-15) constitutes a complete solution to the detection problem when w is known [3]. If the frequency is unknown, as is the case here, we face one of two situations. Either we are interested in the specific value of w or not. If we are not interested in the value of w, we can average equation (IV-15) with respect to w to get $\lambda(r)$ and then compare it to a threshold [17]. If we are interested in w, then we choose a set of possible values for w, say $\{w_j\}_{j=1}^{M}$, such that this sequence forms a fine mesh on the interval [w',w"]. Next, we formulate a multiple hypotheses problem with a hypothesis for each of the M-values of w_j. That is,

$$H_0: \quad r_i = n(t)$$

$$H_1: \quad r_i = A\cos[(w_c+w_1)t + \phi] + n(t)$$

$$\vdots$$

$$H_M: \quad r_i = A\cos[(w_c+w_M)t + \phi] + n(t)$$

$$i = 0,\ldots,(N-1) \qquad \text{(IV-16)}$$

The likelihood ratio comparing the ith hypothesis to the null hypothesis is

$$\lambda_i = \lambda(x|w_i) \qquad i = 1,\ldots,M \qquad \text{(IV-17)}$$

If no λ_i is greater than the threshold, then the null hypothesis is accepted; otherwise, each w_i for which the threshold is exceeded is a potential detection. If only one target is present, the maximum likelihood estimate of w will be that w_i for which λ_i is the largest.

In the above analysis, it was assumed that the noise distribution is Gaussian with known parameters. If the variance of this Gaussian distribution is unknown, then an adaptive technique will be necessary to set the threshold. Many types of such techniques are possible [25]. One of these methods is the one-parameter threshold. In this procedure we utilize a set of reference cells, say K cells, then we integrate the pulses in each of these cells to get z_j, j=1,..., K. The threshold is then, with q constant,

$$T = q \sum_{j=1}^{k} \frac{z_j}{K} . \tag{IV-18}$$

Although this method is simple enough, it has two main disadvantages: a) it depends heavily on the assumption of normality which is doubtful in most radar application, b) if there is another target in one of the reference cells, this will increase the threshold and original target might be missed.

IV.3 M-Doppler detector

Examining the previous detection scheme, we notice that it consists mainly of two processes. First, assuming that the true frequency is w_j, we transform the original set of observations $r_i(t)$ into a new set of discrete in-phase and quadrature phase components x_{ij} and y_{ij}, respectively, given by

$$x_{ij} = \int_{iT}^{iT+D} r_i(t)\ A \cos\,[(w_c+w_j)t]dt \tag{IV-19}$$

and

$$y_{ij} = \int_{iT}^{iT+D} r_i(t)\ A \sin\left[(w_c+w_j)t\right]dt \qquad \text{(IV-20)}$$

Then, under H_0

$$x_{ij} = \int_{iT}^{iT+D} n(t)\ A \cos\left[(w_c+w_j)t\right]dt$$

$$= w_{ij} \qquad \text{(IV-21)}$$

and

$$y_{ij} = \int_{iT}^{iT+D} n(t)\ A \sin\left[(w_c+w_j)t\right]dt$$

$$= v_{ij}\ . \qquad \text{(IV-22)}$$

Under H_1

$$x_{ij} = \int_{iT}^{iT+D} A \cos[(w_c+w)t+\phi]A \cos[(w_c+w_j)t]dt+w_{ij}$$

$$= \frac{A^2}{2} \int_{iT}^{iT+D} \cos\left[(w_c-w_j)t+\phi\right]dt + w_{ij} \qquad \text{(IV-23)}$$

and since $w \ll \frac{1}{D}$, then

$$x_{ij} \simeq \frac{A^2}{2} \cos[(w-w_j)iT+\phi] + w_{ij} \qquad \text{(IV-24)}$$

Likewise,

$$y_{ij} \simeq \frac{A^2}{2} \sin[(w-w_j)iT+\phi] + v_{ij} \qquad \text{(IV-25)}$$

Noting that if $w=w_j$ the first term in both (IV-24) and (IV-25) reduces to $\frac{A^2}{2} \cos\phi$ and $\frac{A^2}{2} \sin\phi$, respectively, and that w_{ij} and v_{ij} are both Gaussian since $n(t)$ is Gaussian by assumption, the second step in the previous scheme was to utilize a square law detector to decide whether a signal is present at w_j.

As was discussed in Chapter II of this report, the performance of the square law detector depends on the assumption of normality of $n(t)$, and if $n(t)$ is not normal, as is the case of interest here, the performance of the square law detector might deteriorate seriously. To avoid this, we propose a more robust doppler detector which is called the M-Doppler detector. This detection scheme consists also of two steps. The first is exactly the same as in the previous scheme. The second utilizes the sequences $\{x_{ij}\}$ and $\{y_{ij}\}$ as inputs to a set of M-detectors on components presented in Chapter II. Since the mathematical work required for this procedure is the same as in Chapter II, we are going to give only a brief description of the procedure.

a) Find the nonlinearity $l(.)$ which is optimum in the minimax sense over the family of distributions F which contains $f(w_{ij})$ and $f(v_{ij})$.

b) Find $\theta_{NH}(j)$ and $\theta_{NV}(j)$ defined as

$$\sum_{i=0}^{N-1} l[x_{ij} - \theta_{NH}(j)] = 0 \qquad \text{(IV-26)}$$

and

$$\sum_{i=0}^{N-1} l[y_{ij} - \theta_{NV}(j)] = 0 \qquad \text{(IV-27)}$$

c) Find $\theta_N(j) = \sqrt{\theta_{NH}^2(j)+\theta_{NV}^2(j)}$.

d) Compare $\theta_N(j)$'s to a threshold γ .

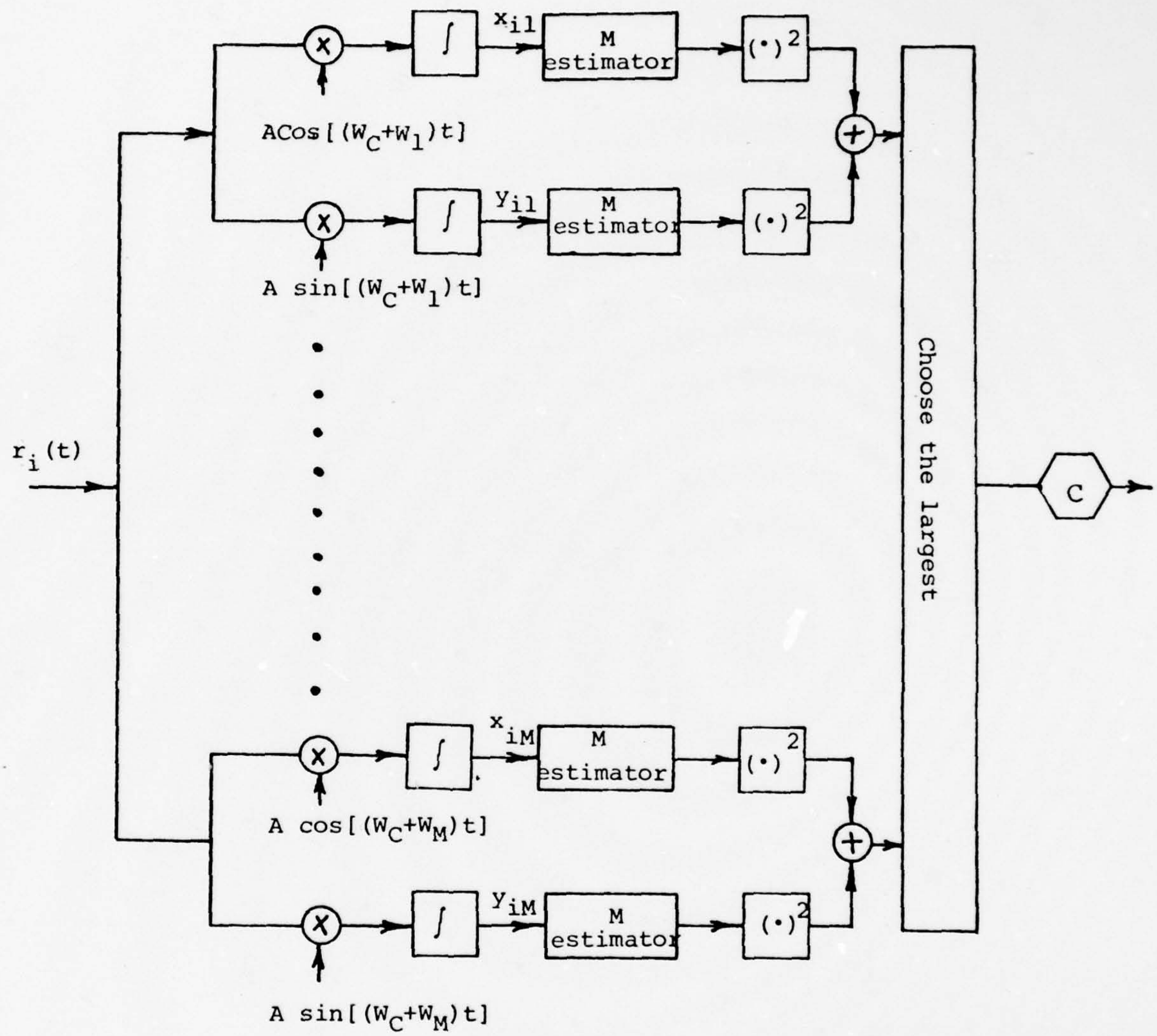
It should be noticed here that an advantage of this method is the stability of the threshold with respect to the distribution of $n(t)$, so an adaptive technique is not needed and there is no consequent loss in the power of the test. Figure (24) shows an M-Doppler detector.

IV-4 Numerical results

To test the performance of the proposed detector a set of simulations was conducted under the assumption that w_{ij} and v_{ij} are components of a lognormal distribution with $\sigma = 6$ db , and then repeated for the case when they are contaminated normal with $r = .25$ and $g = 2.25$. In both cases, $f(w_{ij})$ and $f(v_{ij})$ were considered to be members of a class of p-point distributions with $p=0.5$. Then $l(.)$ was found to be of the form

$$l(t) = \begin{cases} \tan(ct) & |t| \leq a \\ \tan(ca)\,\mathrm{sgn}(t) & |t| > a \end{cases}$$

for appropriate c such that $l(t)$ is a continuous function. The set w_j was chosen such that the interval $[w',w'']$ is divided into 16 equal parts. Using the above nonlinearity and values of w_j, the simulations were conducted and the



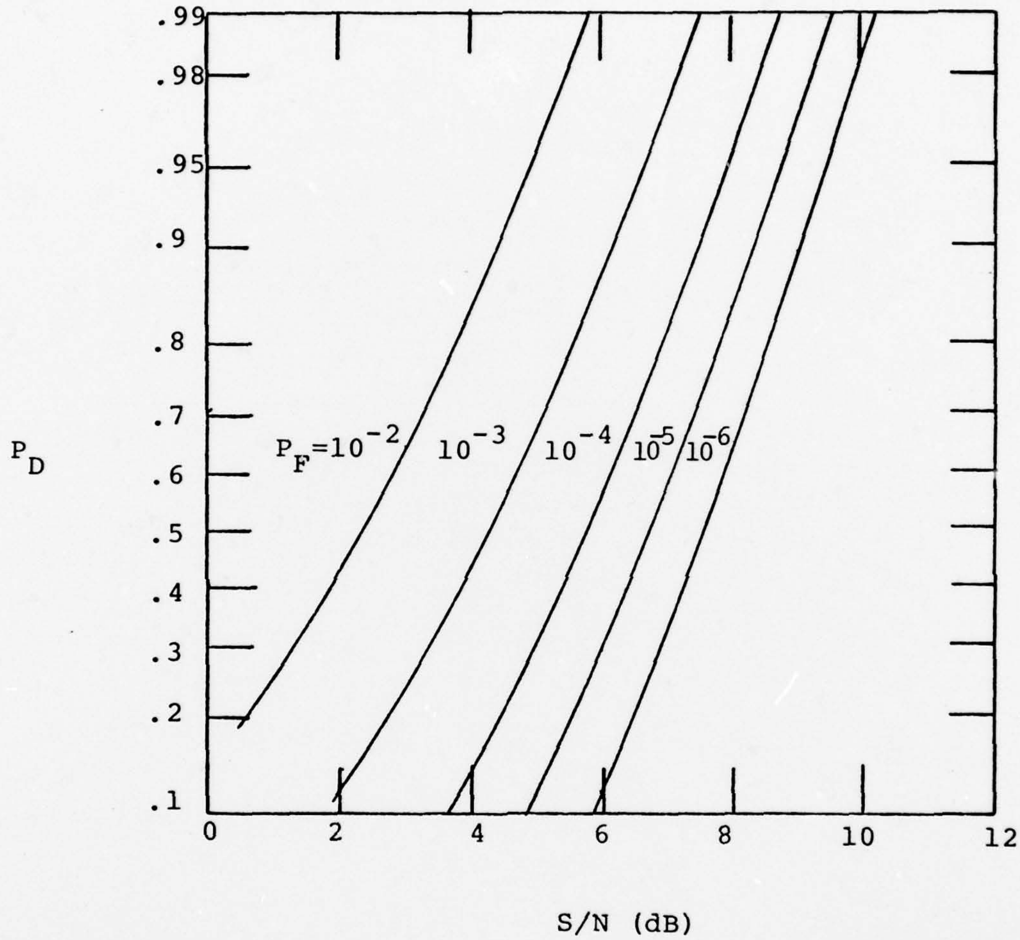
S/N (dB)

M-Doppler detector

Figure 24

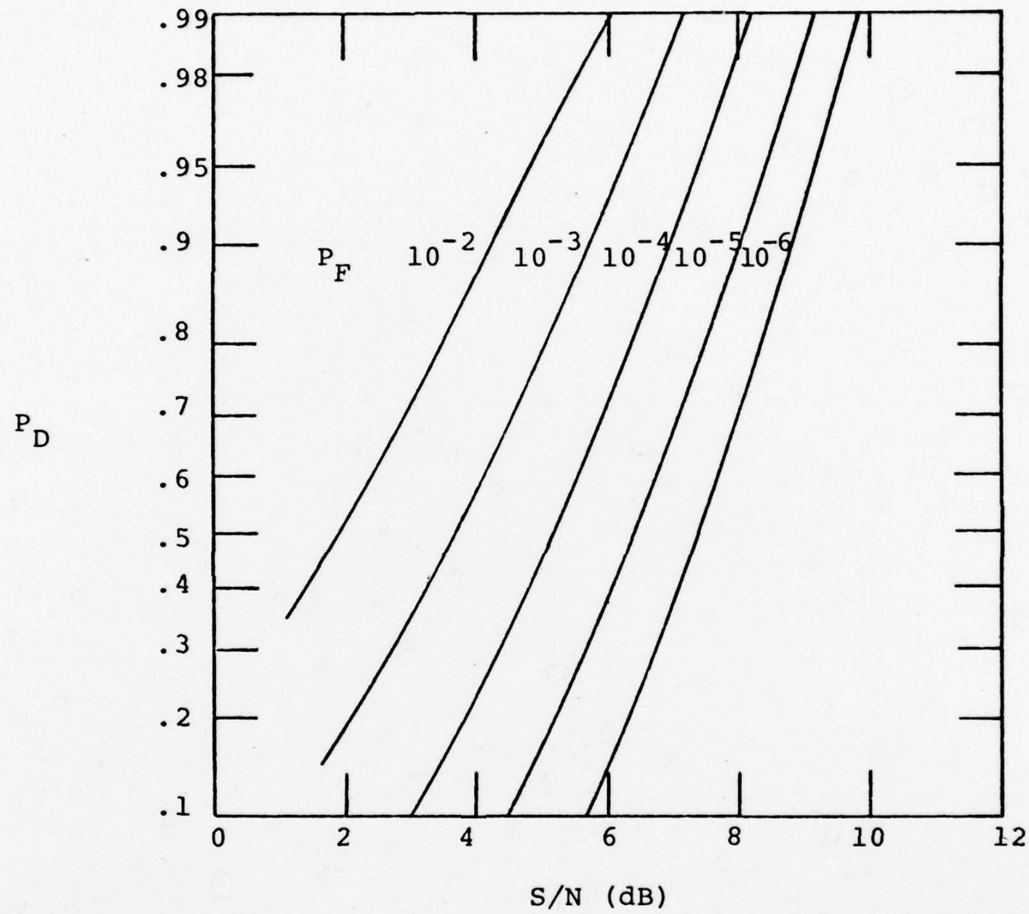
results are summarized in figures (25) and (26) which show the probability of detection verses the signal to noise ratio for different values of probability of false alarm for N=10. An important result which is not shown in these graphs is that the simulation confirmed again the stability of the threshold. It was found that if we normalized the thresholds in both cases by the value of a, we get almost the same results.

Figures (27) and (28) compare the M-Doppler detector to the one-parameter adaptive detector discussed in section (IV-2). It may be noticed that the superiority of the M-Doppler detector increases as P_F decreases and/or P_D increases and the difference in (S/N) comes to about 8 db at $P_F=10^{-6}$ and $P_D=.99$ for the lognormal case, and more than this for the contaminated normal case.



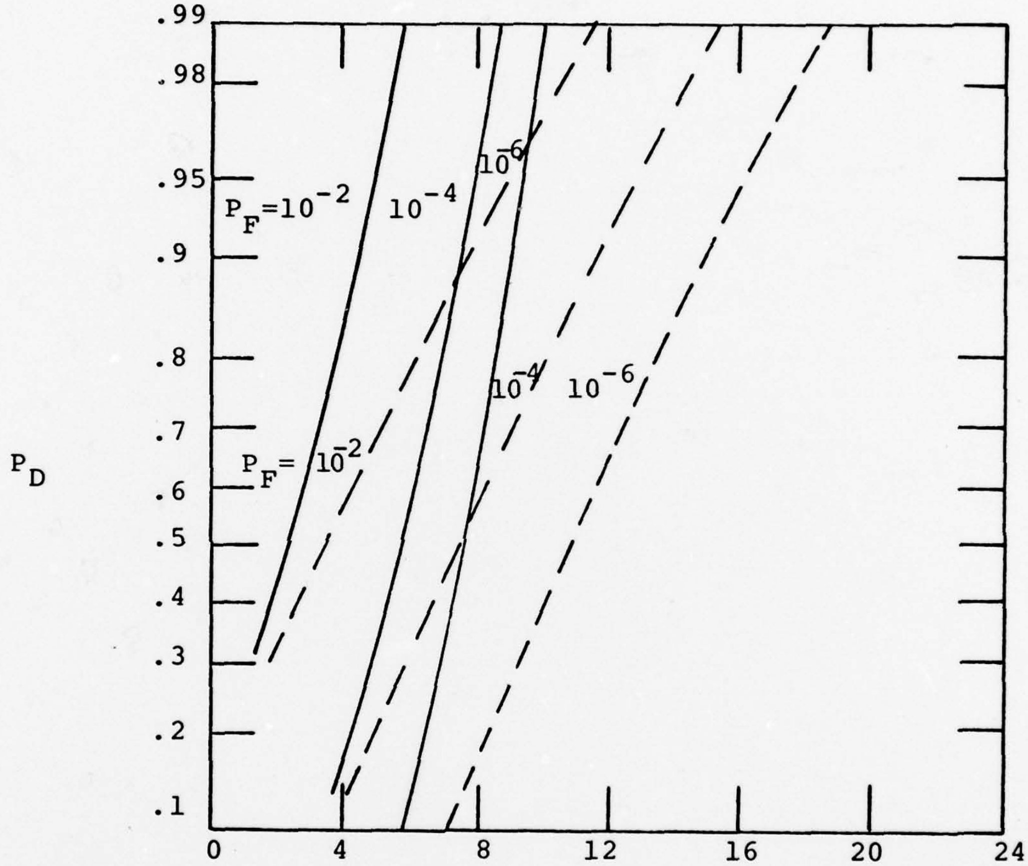
Probability of detection vs S/N for log-normal distribution (σ=6 dB) and M-Doppler detector; N=10

Figure 25



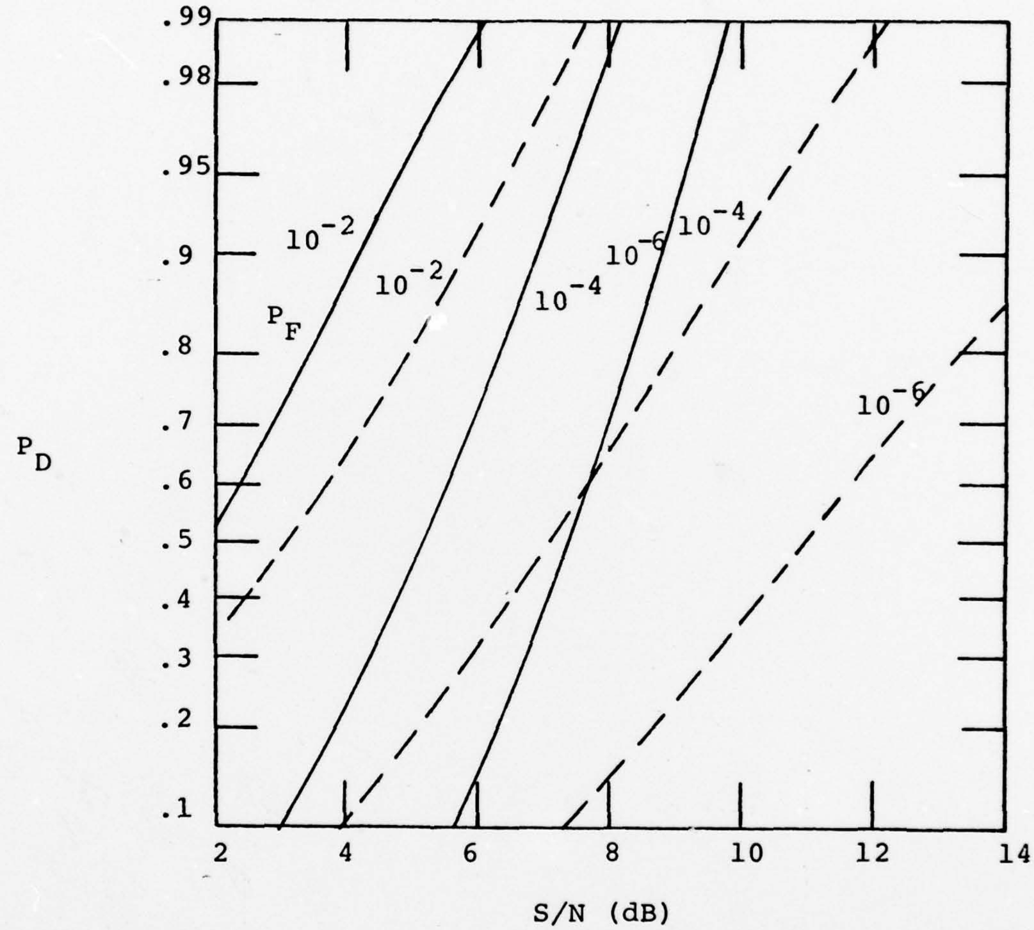
Probability of detection vs S/N for contaminated normal distribution (r=0.25,g=2.25) & M-Doppler detector; N=10

Figure 26



Comparison of M-Doppler detector (—) and one-parameter adaptive detector (---) for log-normal distribution (σ=6dB);N=10

Figure 27



Comparison of M-Doppler detector (—) and one-parameter adaptive detector (---) for contaminated normal distribution (r=0.25,g=2.25);N=10

Figure 28

V. Detection of Signals with random Phase in Sea Clutter

V-1 Introduction:

In the previous part of this report we assumed that the signal phase is an unknown nonrandom parameter over each set of observations but can change randomly from one set to another. In this chapter, we move one step ahead to assume that the signal phase changes randomly from one observation to another. The main difference between these two cases can be seen if we consider the distribution of the inphase and quadrature observation components. In the first case, the inphase and/or the quadrature components will have the same distribution under both hypotheses except for a location shift which is constant or almost constant under all observations, so it was suitable to utilize a test which was sensitive to location changes to distinguish between the two hypotheses. In the second case, the distribution of these components will also be the same except for a location shift, but this location shift will change randomly from one observation to another and will have a mean which is equal to zero under both hypotheses. This means that utilization of a hypothesis test which is sensitive to the means of the components should not be expected to lead to good performance, especially if this test is robust against the changes in assumptions as in the case of the M-detector on components presented in Chapter II. This detector was simulated under the assumptions of the second case and it was found that the signal to noise ratio had to be increased by at least 8 db to achieve the same probability of detection at the same probability of false alarm as in the first case.

On the other hand, it was noticed that the mean of the envelope of the observations is monotonically increasing with the signal in the case of random phase [26 Figures (4-7)]. As a result of the above observation, the main trend in treating the problem of detection of signals with random phase was generally to design detectors which use the envelope of the observations as input data instead of the components or the observations [17, 27 and 28]. In addition, if we examine the results in the previous chapters of this report we may notice that, for the types of noise under considerations, censoring the observations always leads to improvement in the performance. This conclusion was also supported by some previous results by Trunk [27] when he tested the trimmed-mean detector.

As a result of the above discussion, we are going to consider in this chapter only these detectors which utilize the envelope of the observations as input data, and which limit the effect of the tails of the observations distribution in some way. It should be emphasized here that we are looking for detectors which have good performance over a wide class of noise distributions when there is only partial information about these distributions. Also, we shall consider that the available information is always quantiles.

In Section (V-2) we present the "M-detector on envelope". In Sections (V-3) and (V-4) we discuss the "Robust Quantizer" detectors and the "Extreme Value Theorem" detectors respectively. Section (V-5) gives brief description of the "Trimmed Means" detector which will be utilized in Section (V-6) for comparisons.

(V-2) M-Detector on Envelope:

Let $z_1, z_2, \ldots, z_N$ be independent identically distributed samples taken from the probability density f(z) and let F(z) be the distribution function of z. Consider also that

$$z_i = \sqrt{x_i^2 + y_i^2} \qquad \text{(V-1)}$$

where

$$x_i = A \cos \phi_i + w_i \qquad \text{(V-2)}$$

and

$$y_i = A \sin \phi_i + v_i \qquad \text{(V-3)}$$

where v_i and w_i are noise components, ϕ_i is a random phase with uniform density function over the interval $[0,2\pi]$, and A is the signal amplitude which is equal to zero under H_0 and greater than zero under H_1. The exact value of A may be unknown. Assume also that there exist two known real numbers Z_1 and Z_2 such that $0<Z_1<Z_2<\infty$ and $F(Z_1|H_0)=P_1$ and $F(Z_2|H_0)=P_2$ where P_1 and P_2 are known and not equal to zero or one. It is required to design a detector which gives good performance against all distributions having the above properties, and which is insensitive to the variations of the actual distribution. As was discussed in the previous section this detector should be expected to limit the effect of large observations.

Before proceeding to describe the M-detector on envelope, we need the following definitions. Define the nonlinearity l(.) as

$$l(t) = \begin{cases} -b\ a & (t)\leq Z_1 \\ -b\ a + \dfrac{(b+1)a}{(Z_2-Z_1)}(t-z_1) & Z_1<t\leq Z_2 \\ a & Z_2<t \end{cases} \tag{V-4}$$

where a and b are two constants and Z_1,Z_2 are as defined before. Define θ_N as the value of θ such that

$$\sum_{i=1}^{N} g(z_i-\theta_N)=0 \tag{V-5}$$

$$g(z-\theta) = l(z-\theta) - E[l(z)|H_0] \tag{V-6}$$

Then the M-detector on envelope will be of the form

$$\sqrt{N} \quad \theta_N \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \gamma \tag{V-7}$$

where γ is the threshold.

Examining the above detection strategy, we find that the main idea is to use a fraction of the small observations (those less than Z_1) to get rid of the large observations (those larger than Z_2) without sorting the data. If b was chosen to be equal to $\frac{P_2}{P_1}$, then asymptotically this will be equivalent to the average sum of all observations in some intermediate interval of length (Z_2-Z_1). In this, the M-detector shares with the median detector the good property of deleting outlyers; however, it does not depend on one observation. So it is midway between the mean and median detectors.

To study the stability of the threshold we must first obtain an expression for the distribution of θ_N under H_0.

<u>Distribution of θ_N</u>

<u>Lemma (V.1)</u>

Under H_0, $\theta_N \to 0$ as $N\to\infty$.

Proof:

For any positive number ε the law of large numbers implies

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} g(z_i-\varepsilon) = E[l(z-\varepsilon)|H_0] - E[l(z)|H_0]$$

$$< 0$$

since $l(x)$ is monotonic in its argument. Also

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} g(z_i+\varepsilon) > 0$$

But since θ_N is defined such that $\sum_{i=1}^{N} g(z_i-\theta_N)=0$, then for any positive number ε, there esits a number N_0 such that for all $N \geq N_0$

$$-\varepsilon \leq \theta_N < \varepsilon$$

and

$$P\{|\theta_N| \leq \varepsilon\} \to 1 \quad \text{as } N \to \infty ,$$

Lemma (V.2):

Under H_0, $\sqrt{N}\,\theta_N$ is asymptotically distributed as normal with zero mean and variance

$$v^2(\theta_N) = \frac{E[g^2(z)]}{[E'(g(z-\theta))|_{\theta=0}]^2} \tag{V-8}$$

Proof:

$$P[\sqrt{N}\,\theta_N \leq k] = P[\theta_N \leq \frac{k}{\sqrt{N}}]$$

$$= P[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} g(z_i - \frac{k}{\sqrt{N}}) \leq 0] \tag{V-9}$$

Using the central limit theorem it can be shown that the sum in equation (V-9) tends to normal. Also,

$$\lim_{N\to\infty} E[g^2(z_i - \frac{k}{\sqrt{N}})] = E[g^2(z)]$$

and

$$\lim_{N\to\infty} \sqrt{N}\, E[g(z-\frac{k}{\sqrt{N}})] = \lim_{N\to\infty} \sqrt{N}\, E[1(z-\frac{k}{\sqrt{N}}) - 1(z)]$$

$$= -K \frac{\partial E}{\partial \theta} [1(z-\theta)]\Big|_{\theta=0}$$

$$= -k\ E' [g(z-\theta)|_{\theta=0}]$$

Then,

$$P\{\sqrt{N}\ \theta_N \leq k\} \to \Phi \left\{ \frac{k\ E'[g(z-\theta)|_{\theta=0}]}{E^{1/2}[g^2(z)]} \right\} \qquad \text{(V-10)}$$

where $\Phi(.)$ is the standard normal distribution function.

To evaluate the sensitivity of the threshold with respect to the distribution, we calculate the explicit value of the variance in equation (V-10).

$$E'\{g(z-\theta)|_{\theta=0} = E'\{1(z-\theta)|_{\theta=0}\}$$

$$= \frac{(b+1)a}{(Z_2-Z_1)} \int_{Z_1}^{Z_2} f(z)dz$$

$$= \frac{(b+1)a}{(Z_2-Z_1)} [P_2-P_1] \qquad \text{(V-11)}$$

i.e. it is independent of the specific shape of f(z).

$$E[g^2(z)] = E[l^2(z)] - E^2[l(z)]$$

$$= b^2 a^2 P_1 + a^2 [1-P_2] + [ba + \frac{(b+1)a}{Z_2 - Z_1} Z_1]^2 [P_2 - P_1][1+P_1-P_2]$$

$$+ [\frac{(b+1)a}{Z_2-Z_1}]^2 \{ \int_{Z_1}^{Z_2} z^2 f(z)dz - [\int_{Z_1}^{Z_2} z\, f(z)dz]^2 \}$$

$$-2\, [ba + \frac{(b+1)a}{Z_2-Z_1} Z_1][\frac{(b+1)a}{Z_2-Z_1}][1+P_1-P_2] \int_{Z_1}^{Z_2} z\, f(z)dz$$

(V-12)

It can be seen that the value of the threshold will change with changes in the true distribution only through the last two terms in equation (V-12). On the other hand, since we are dealing with detection in sea clutter, where all possible noise distributions range between Rayleigh and lognormal [29] a proper choice of Z_1 and Z_2 will minimize the effect of these two terms. For example, if we choose Z_1 and Z_2 such that P_1 and P_2 equal .1 and .5 respectively, the effect of these terms on the variance will be negligible. This choice also would help in calculating $E[l(z)|H_0]$ which is required to perform the test.

Receiver implementation

To perform the above test as it was described in equation (V-7), we must first calculate θ_N. This process may not be easy if the number of observations involved is large sine l(.) is a nonlinear function.

But, on the other hand, noting that g(.) is monotonic in its argument, for any number ε such that $\varepsilon > \theta_N$

$$\sum_{i=1}^{N} g(z_i-\varepsilon) < 0$$

and if $\varepsilon < \theta_N$, then

$$\sum_{i=1}^{N} g(z_i-\varepsilon) > 0$$

Accordingly if we determined the threshold to be γ, then the M-detector on envelope can be implemented in the form

$$\sum_{i=1}^{N} g(z_i-\gamma) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{V-13}$$

or, equivalently

$$\sum_{i=1}^{N} l(z_i-\gamma) \underset{H_0}{\overset{H_1}{\gtrless}} E[l(z)|H_0] \tag{V-14}$$

The implementation of (V-14) is shown in figure (28).

Numerical results:

To examine the finite sample size performance of the M-detector on envelope, simulation experiments were conducted at sample sizes of N=10 and 30 for the case when the noise has a lognormal distribution given by

$$f(z) = \frac{1}{\sqrt{2\pi} z} e^{-\frac{\ln^2 z}{2\sigma^2}} \tag{V-15}$$

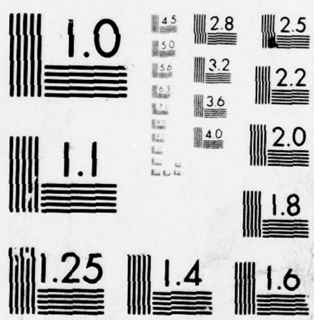
AD-A073 478 JOHNS HOPKINS UNIV BALTIMORE MD DEPT OF ELECTRICAL --ETC F/G 9/4
ROBUST NONCOHERENT DETECTION.(U)
FEB 79 H L WEINERT, A H EL-SAWY N00173-77-C-0189

JNCLASSIFIED NL

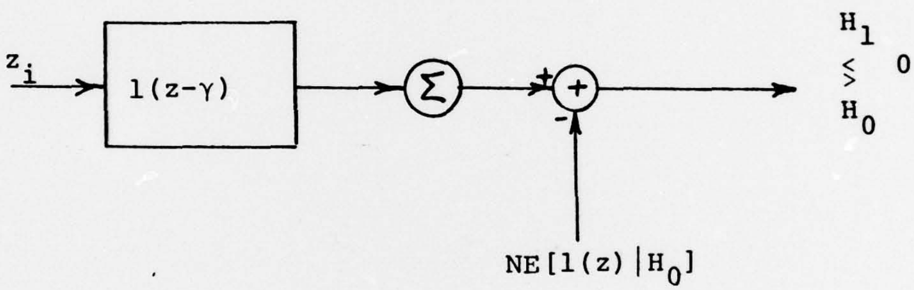
2 OF 2

AD
A073478

END
DATE
FILMED
10-79
DDC



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



M-dector on envelope

Figure 28'

where σ = 6 db, and at N=10 for the contaminated normal distribution given by

$$f(z) = (1-r^2)\, e^{\frac{-z^2}{2}} + \frac{r^2 z^2}{k^2}\, e^{\frac{-z^2}{2k^2}}$$

$$+ \frac{2\,(1-r)z}{k} \exp\left[\frac{-z^2(k^2+1)}{4k^2}\right] I_0\left[\frac{z^2(k^2-1)}{4k^2}\right] \qquad \text{(V-16)}$$

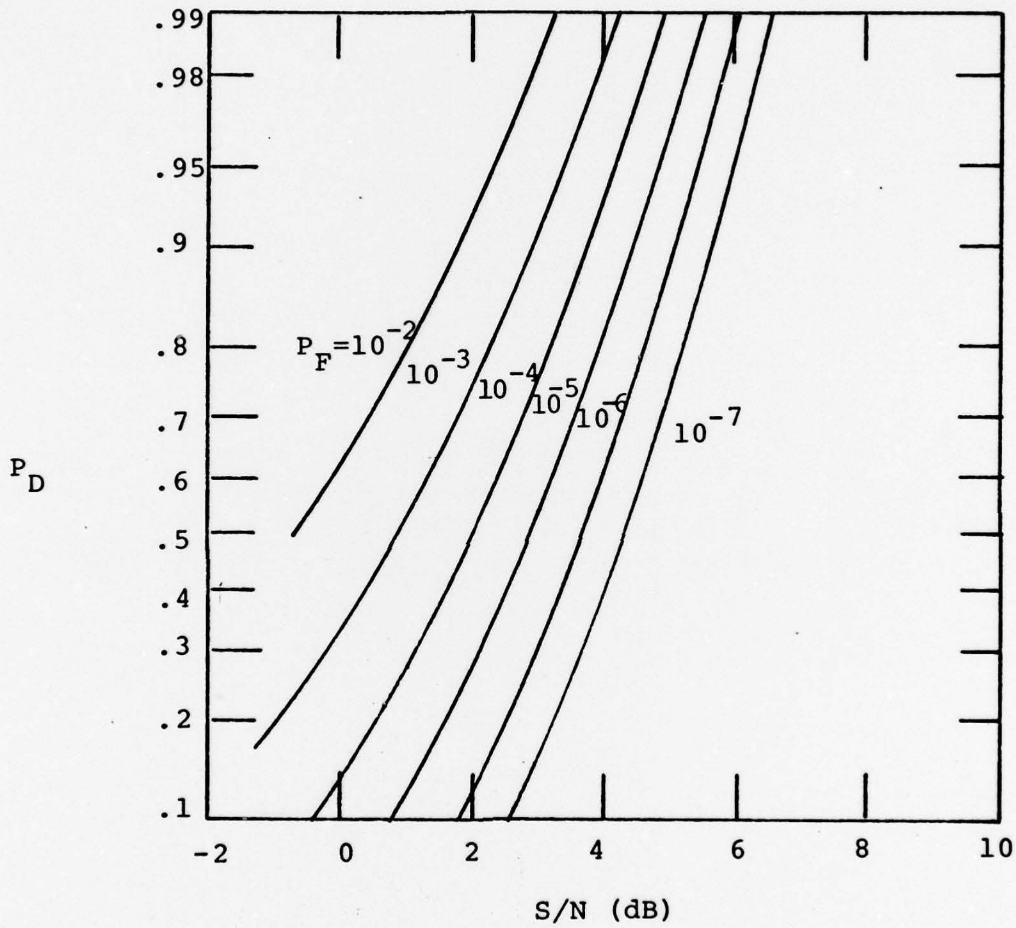
where $I_0(.)$ is the modified Bessel function, $r=0.25$ and $k=2.25$.

In each experiment, the values of Z_1 and Z_2 were found such that P_1 and P_2 equal 0.1 and 0.5 respectively. b was taken to be $\frac{P_2}{P_1} = 5$. To calculate $E[l(z)|H_0]$, a Rayleigh approximation of the form

$$f(z) = \frac{(z-m)}{\sigma^2}\, e^{\frac{-(z-m)^2}{2\sigma^2}} \qquad \text{(V-17)}$$

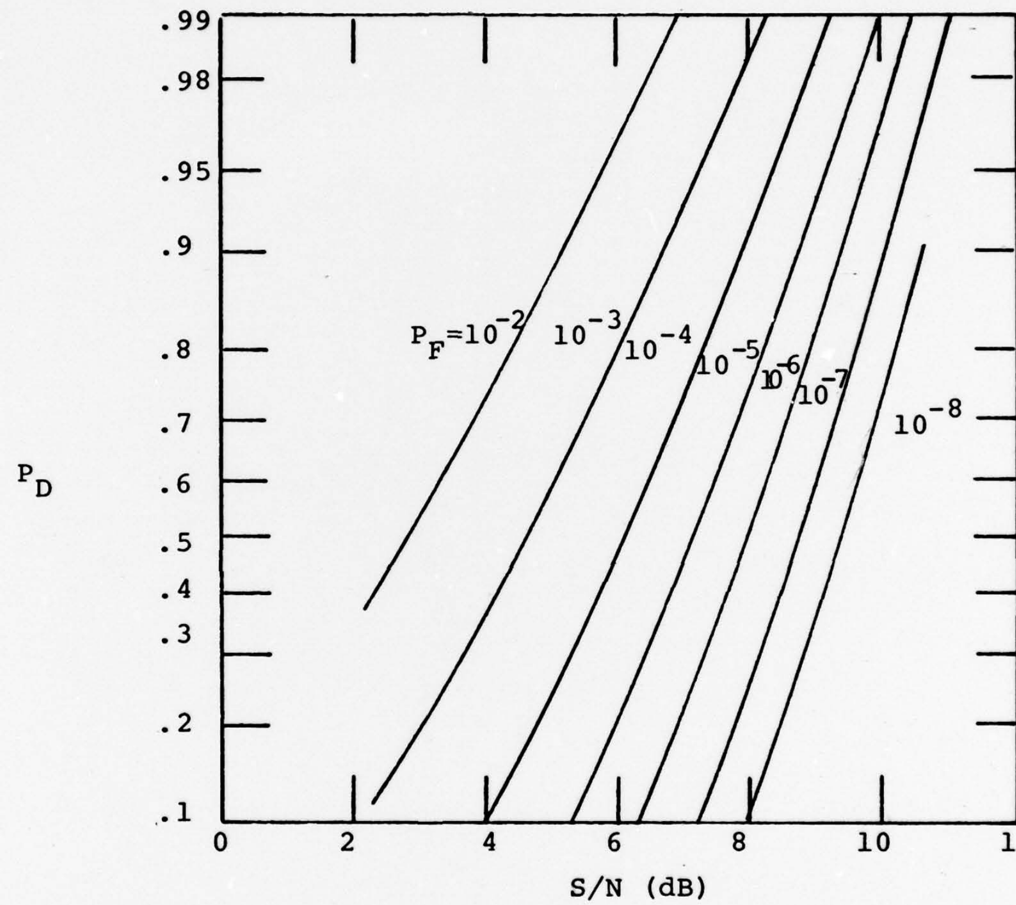
was utilized. First, we calculate m and σ such that this distribution satisfies the requirements at Z_1 and Z_2. Then, we use this approximation to calculate $E[l(z)|H_0]$. It was found that the maximum difference between the approximate and true values is about 5%. It was noticed also that the value of m is very small in both cases.

Figures (29-31) give the results of these simulations. The signal to noise ratio was taken as 20 $\log_{10}$(A/M) where A is the signal amplitude and M is the median of the distribution under consideration. Notice that the



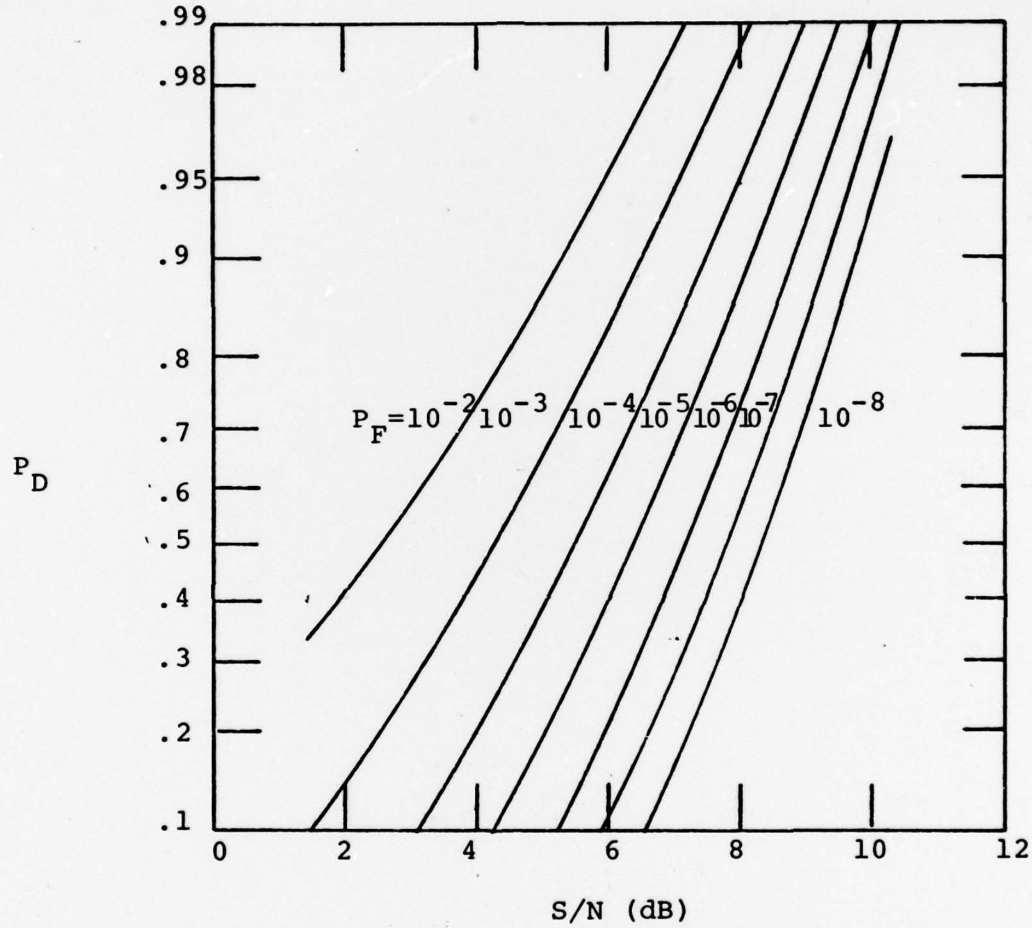
Probability of detection vs S/N (dB) for lognormal distribution (σ=6dB) and M-detector on envelope; N=30

Figure 29



Probability of detection vs S/N (dB) for lognormal distribution (σ=6dB) and M-detector on envelope; N=10

Figure 30



Probability of detection vs S/N (dB) for contaminated normal distribution (r=0.25,g=2.25) and M-detector on envelope; N=10

Figure 31

required (S/N) to achieve a certain P_D at the same P_F in the case of lognormal is slightly higher than that required in the contaminated normal case. The following remarks can be made:

1) $\frac{Z_1}{Z_2}$ is almost the same for the two distributions.

2) At N=10, if the thresholds were normalized by the appropriate value of Z_2, then the thresholds which would give a P_F of 10^{-2} and 10^{-6} when the noise is lognormal, would give a P_F of 10^{-2} and 10^{-7} when the noise is contaminated normal. This confirms the stability of the threshold.

3) At N=10, using the same normalization we found that at the same signal to noise ratio the probability of detection is also almost constant, independent of the true distribution.

Remarks 2 and 3 imply that the detector is robust or insensitive to changes in distribution, even at small sample sizes such as N=10. More about this detector will be presented later in this chapter.

<u>Summary</u>

In this section, the M-detector on envelope was presented as a solution to the problem of detection of signals with random phase in sea clutter. The main idea behind this detector was the use of some of the very small observations to delete the large ones without need for sorting the data. The robust properties of this detector were shown both theoretically and experimentally. In addition, a simple method for implementing this detector on line was presented.

(V-3) Robust Quantizer Detectors:

The main idea behind the set of detectors to be discussed in this section is in the utilization of a quantizer to transfer the input data to a set of preassigned values, then to utilize the quantizer outputs as an input to the decision rule. Assume that $\{z_i\}_{i=1}^{N}$ is the sequence of input data and $\{a_j\}_{j=0}^{m}$ is a sequence of real number such that $a_0=0$ and $a_m=\infty$. Then the output of the quantizer q_i will be given by

$$q_i = q(z_i) = l_j \qquad a_{j-1} \leq z_i < a_j \tag{V-18}$$

where $\{l_j\}_{j=1}^{m}$ is the set of preassigned values for the quantizer output. The test of hypotheses will be of the form

$$T_N = \frac{1}{N} \sum_{i=1}^{N} q_i \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{V-19}$$

where γ is the threshold. The above choice of T_N in fact does not in anyway limit the generality of the procedure since the q_i's are independent as long as the z_i's are, and we are still free to choose the l_j's to satisfy any optimality criterion we consider. Such quantizer detectors were presented in [30,31] under the assumption that the noise distribution is known and in [32] under the assumption that the noise distribution is partially known. The following remarks are in order.

1) With the proper choice of $\{a_j\}$ the detector will limit the effect of large observations to the desired degree.
2) The detector is insensitive to the specific shape of the distribution since it does not discriminate between the values of observations in the same interval.
3) If there is a family of distributions such that $[F(a_j)-F(a_{j-1})]$ under H_0 is the same for all members of this family for $j=1,\ldots, m$, then the distribution of T_N will be the same for any member of this family under H_0 and the test will be a constant false alarm rate test (CFAR) over this family.

From the above remarks and since we are concerned with the threshold stability, it seems that the most appropriate way to choose $\{a_j\}$ is in some way related to the quantiles of the distribution. This simplifies the implementation of the receiver in the case of detection in sea clutter since the designer usually has at least some idea about the quantiles of the distribution. In all detectors involved in this study, the a_j's were choosen such that $[F(a_j)-F(a_{j-1})] = \frac{1}{m}$ for all $j=1,\ldots, m$.

To choose the l_j's three different methods were used

a) The fixed step quantizer

In this method, l_j was chosen as $l_j=j$. So, if we assume that n_j observations fall between $a_{(j-1)}$ and a_j, the test statistic can be rewritten as

$$T_N = \frac{1}{N} \sum_{j=1}^{m} j\, n_j \qquad \text{(V-20)}$$

To find the distribution of T_N we first find the distribution of the n_j's which is given by

$$P_r[n_i=N_i,\ i=1,2,\ldots, m|H_0]=\frac{N!}{N_1!N_2!\ldots N_m!}\left(\frac{1}{m}\right)^m \qquad \text{(V-21)}$$

and

$$P_r[n_i=N_i,\ i=1,\ldots, m|H_1]=N!\left(\frac{P_1^{N_1}}{N_1!}\right)\ldots\left(\frac{P_m^{N_m}}{N_m!}\right) \qquad \text{(V-22)}$$

where P_j is the probability that one observation falls in the interval j under H_1. Using equation (V-20)-(V-22) we can get the distribution of T_N.

b) Coherent-Optimal quantizer:

In [32] it was shown that if the signal is a dc signal in additive noise then a locally most powerful quantizer detector of the above form can be constructed using

$$l_j = \frac{1}{m}\,[f_{(a_{j-1})} - f_{(a_j)}] \quad j=1,\ldots,m \tag{V-23}$$

Although this is not the case here, this quantizer was included in the study for the sake of completeness. If we assume the observations are distributed under H_0 as a lognormal with σ=6db then the values of the a_j's and l_j's would be given from table (V-1) below for m=10

j	1	2	3	4	5	6	7	8	9	10
a_j	.41	.556	.693	.841	1.0	1.10	1.44	1.8	2.43	∞
l_j	-.613	-.111	0	.059	.088	.107	.123	.125	.119	.104

Table (V-1) a_j's and l_j's for 10 level coherent optimal quantizer under lognormal distribution

Note that all the l_j's for j>5 are in the neighborhood of .11, and that the values of the l_j's for $j \geq 8$ are decreasing. This second remark is a setback in this method since it implies that as the signal increases T_N may decrease, which would be expected to lead to bad performance.

c) <u>Semi optimal-coherent quantizer</u>:

To correct for the descending part in the l_j's, l_9 and l_{10} were set equal to 0.126 and 0.127 respectively. Otherwise, this quantizer is the same as the previous one.

Numerical Results

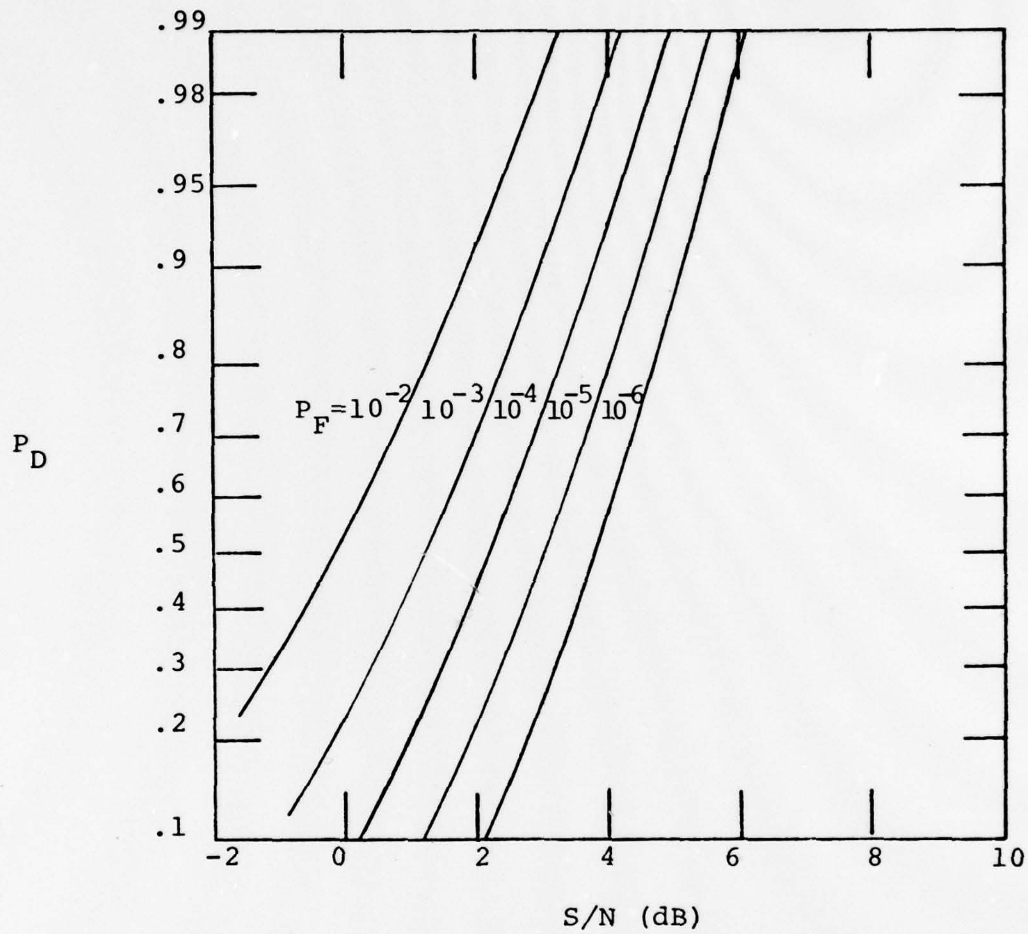
To evaluate the actual performance of the above three detectors, simulation experiments were conducted using lognormal and contaminated normal distributions as in Section (V-1). The results for the optimal coherent quantizer detector were as bad as expected especially for low values of P_F, so the results were not reported here.

Figures (32-34) show the probability of detection P_D versus the signal to noise ratio (S/N) at different probabilities of false alarms P_F, for the fixed step quantizer detector. Comparing figure (33) with (34) we can see that the signal to noise ratio required to achieve a certain P_d at the same P_F when the noise is lognormal is very close to that required when the noise is contaminated normal.

Figure (35) shows P_d versus (S/N) for the semioptimal coherent quantizer detector. Comparing figure (33) with (35) we notice that the fixed step quantizer detector is superior to the semioptimal coherent one especially at higher values of P_d.

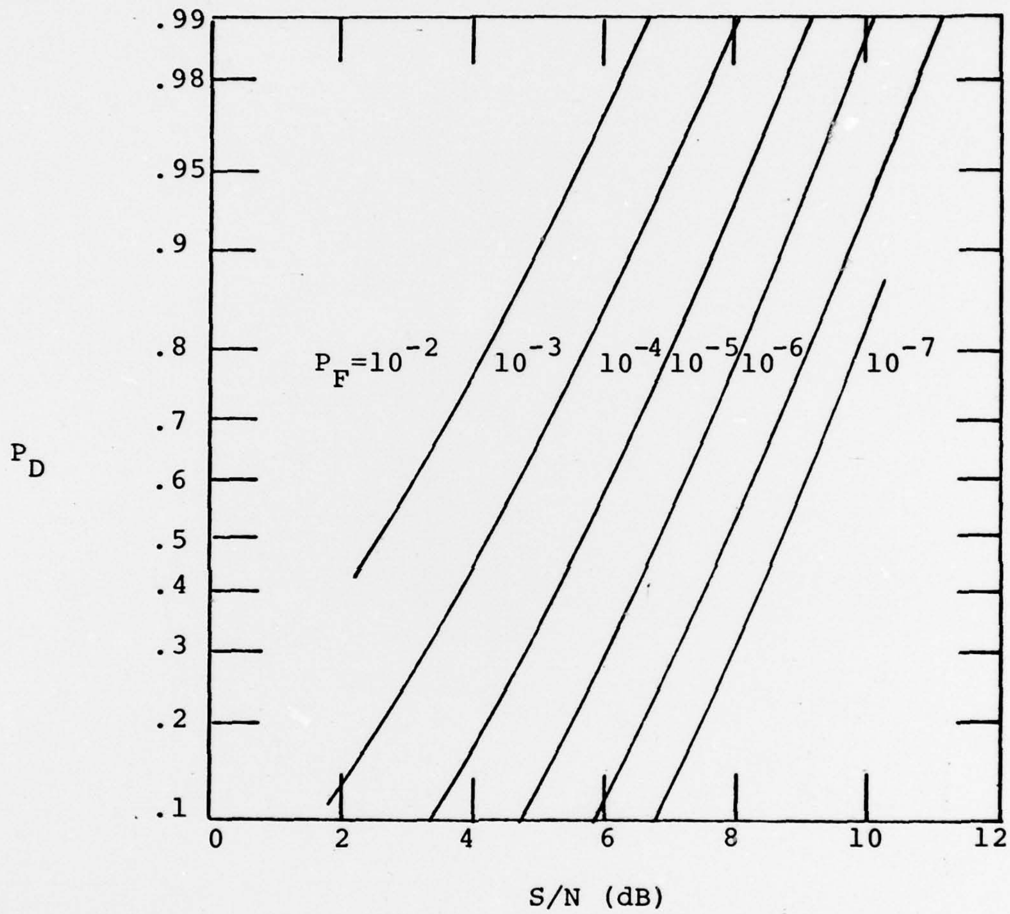
Summary

In this section, three quantizer detectors were investigated. It was shown that the simplest quantizer -- which is the fixed step quantizer -- is the best one. To implement this quantizer detector we need only to



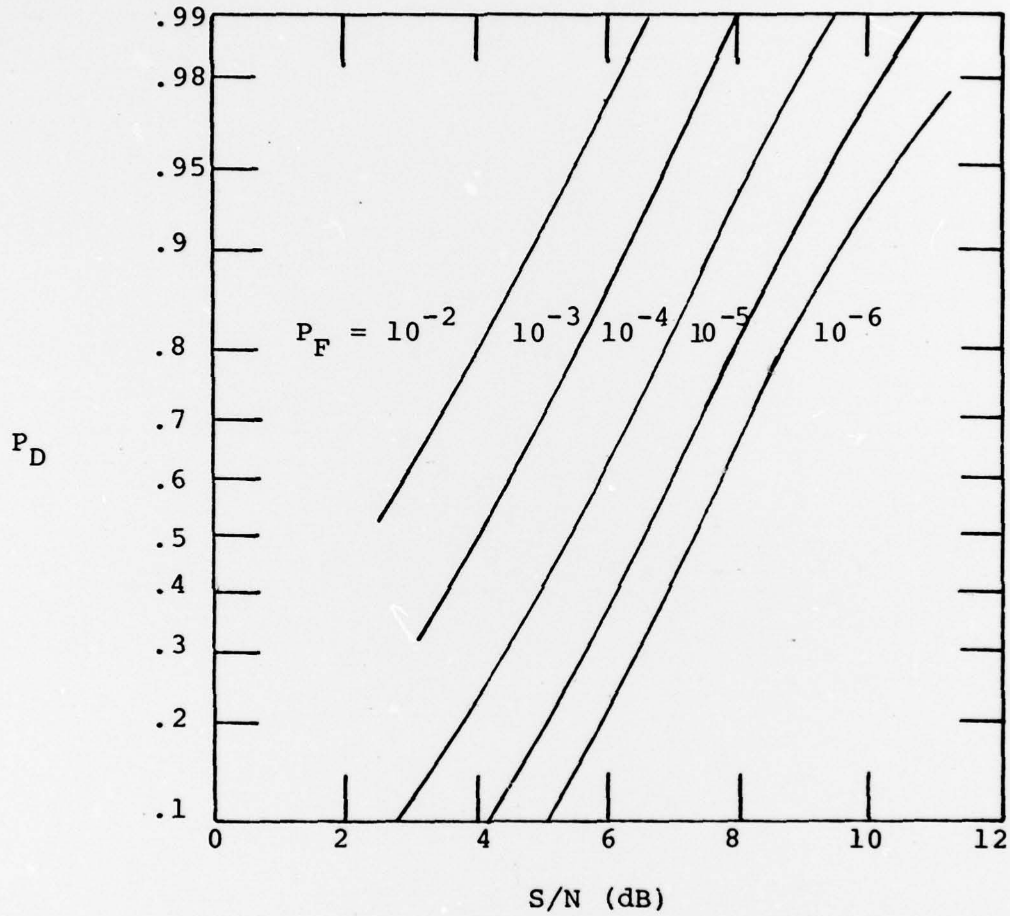
Probability of detection vs S/N (dB) for log-normal distribution (σ=6dB) and equal spaced quantizer detector, N=30

Figure 32



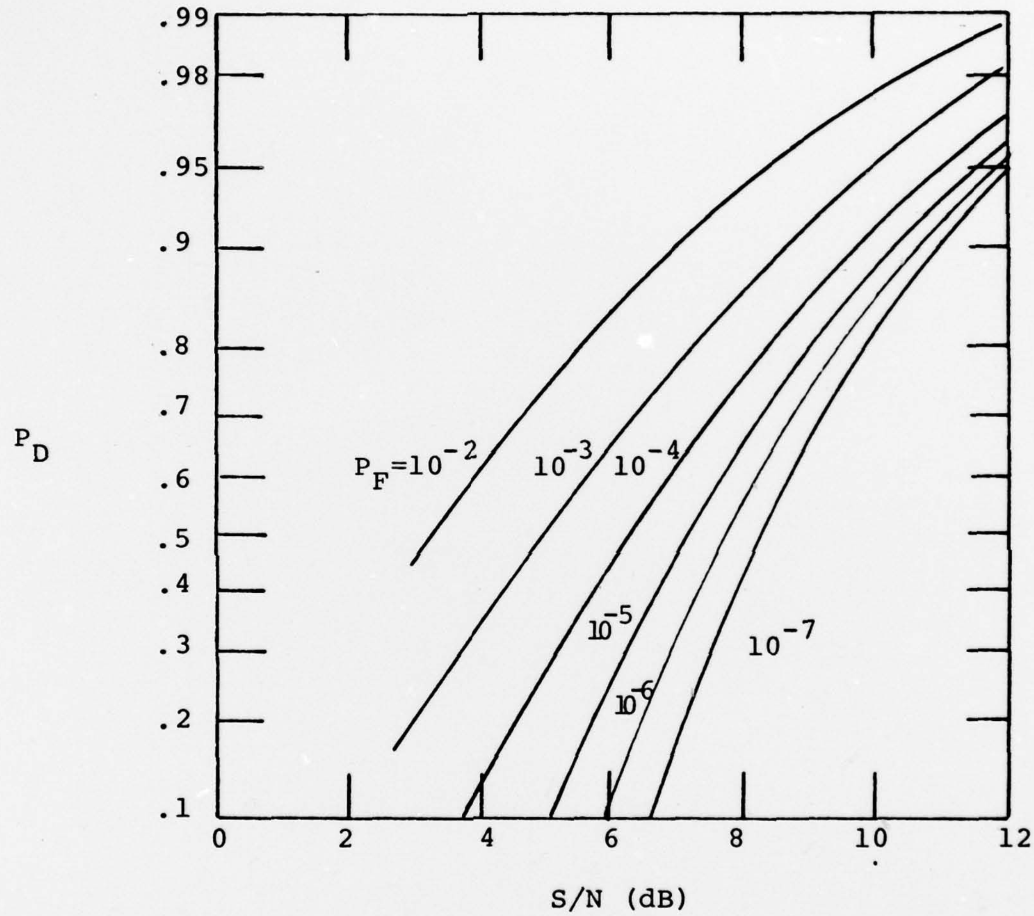
Probability of detection vs S/N (dB) for log-normal distribution (σ=6dB) and equal spaced quantizer detector; N=10

Figure 33



Probability of detection vs S/N (dB) for contaminated normal (r=0.25,g=2.25) and equal spaced quantizer detector; N=10

Figure 34



Probability of detection vs S/N (dB) for log-normal distribution (σ=6dB) and semi-optional coherent quantizer detector; N=10

Figure 35

know some the distribution's quantiles. The number of quantiles assumed to be known in the experimental study here was 9. Since this number is large, it is our opinion that some study should be devoted to the estimation of these quantiles. A first step in this direction was taken by the author in [33] for the case of detection coherent signals in symmetric noise. The fixed step quantizer detector will be compared to other detectors later in this chapter.

(V-4) The Extreme-Value Theory Detectors (EVT):

In [34], [35] Gumbel developed the extreme-value theory. It was shown that if the tails of an arbitrary distribution are exponential, then the right hand tail can be represented by

$$F(x) = 1 - \frac{1}{n} \exp[-\alpha_n (x-u_n)] \tag{V-24}$$

for x in an appropriately restricted region about the parameter u_n. Similarly, the left hand tail can be represented by

$$F(x) = \frac{1}{n} \exp[\alpha_1 (x-u_1)] \tag{V-25}$$

where the parameters u_n, u_1, α_n and α_1 are defined as

$$F(u_n) = 1 - \frac{1}{n} \tag{V-26}$$

$$\alpha_n = n\, f(u_n) \tag{V-27}$$

$$F(u_1) = \frac{1}{n} \tag{V-28}$$

$$\alpha_1 = n\, f(u_1) \tag{V-29}$$

and n is the number of samples from which the maximum or minimum sample is taken. In [35] Gumbel also gave a technique to estimate the parameters u_1, u_n, α_1 and α_n.

Using the above results Milstein et al [36] introduced the "Extreme Value Theory Detector" (EVT). The main idea in this detection scheme is that if the signal is fairly large, then the confusion results as to whether a certain observation is related to the lower tail of the distribution of signal plus noise or to the upper tail of the noise only distribution. To solve this problem they utilized an approximation of the likelihood ratio depending on the distribution approximation given in equations (V-24) and (V-25). So, the test for a single observation will be of the form

$$\Lambda(x) = \frac{\alpha_n \exp\,[-\alpha_n(x-u_n)]}{\alpha_1 \exp\,[-\alpha_1(x-u_1)]} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{V-30}$$

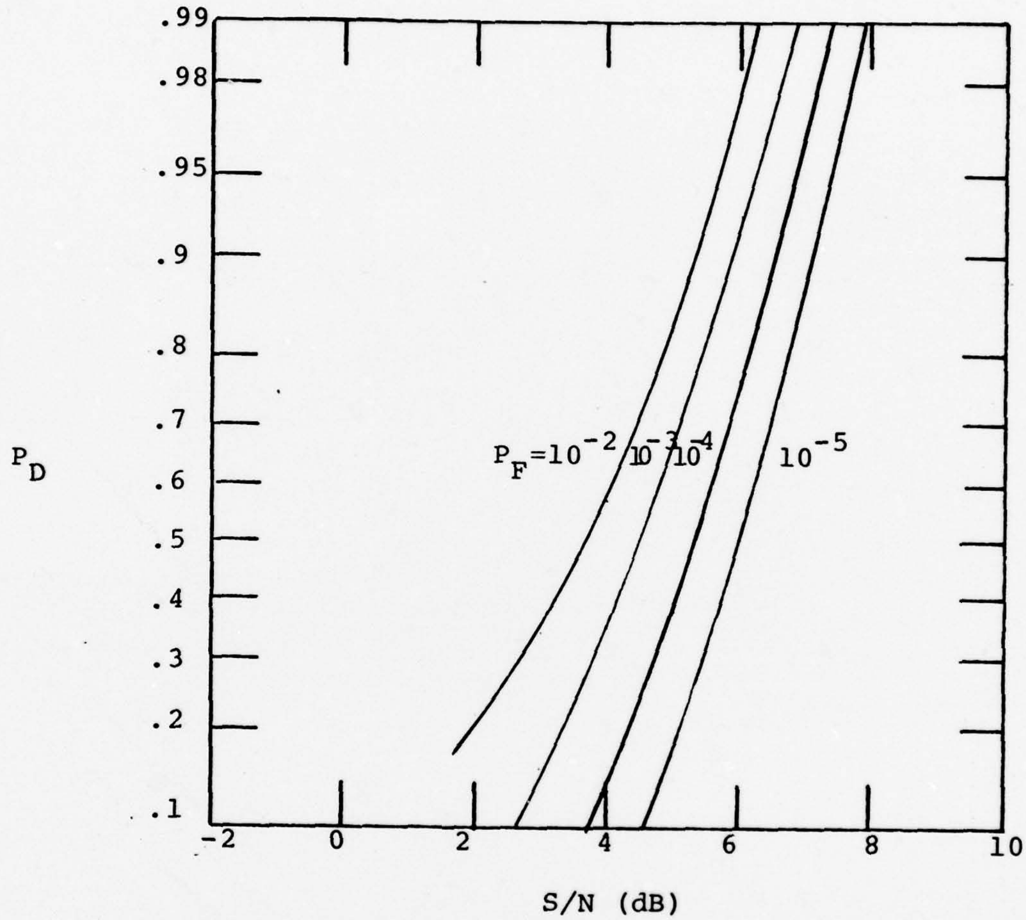
where γ is the threshold. Now, if there are N observations we repeat the test for each one of them and count how many observations out of these will exceed the threshold. If the probability of false alarm for a single observation is P_{F1}, and m observations exceeded the threshold, then the total P_F can be calculated using the binomial distribution. This detector can be implemented very simply by setting two thresholds. The first is γ , which would be exceeded under H_0 by any single observation with probability P_{F1}. The second is K, which is equal to the total number of observations exceeding γ , to get the total probability of false alarm P_F.

Notice that this detector is the same as the binary integrator detector presented in [29] for detection of signals in log normal clutter. It was also presented in [37] for detection of signals in Rayleigh noise channels. In general, it is another implementation of the rank detector.

A problem which might appear with such detectors is in the choice of the second threshold K. Test results show that the receiver performance depends on the choice of K and there is a different optimal choice for each distribution.

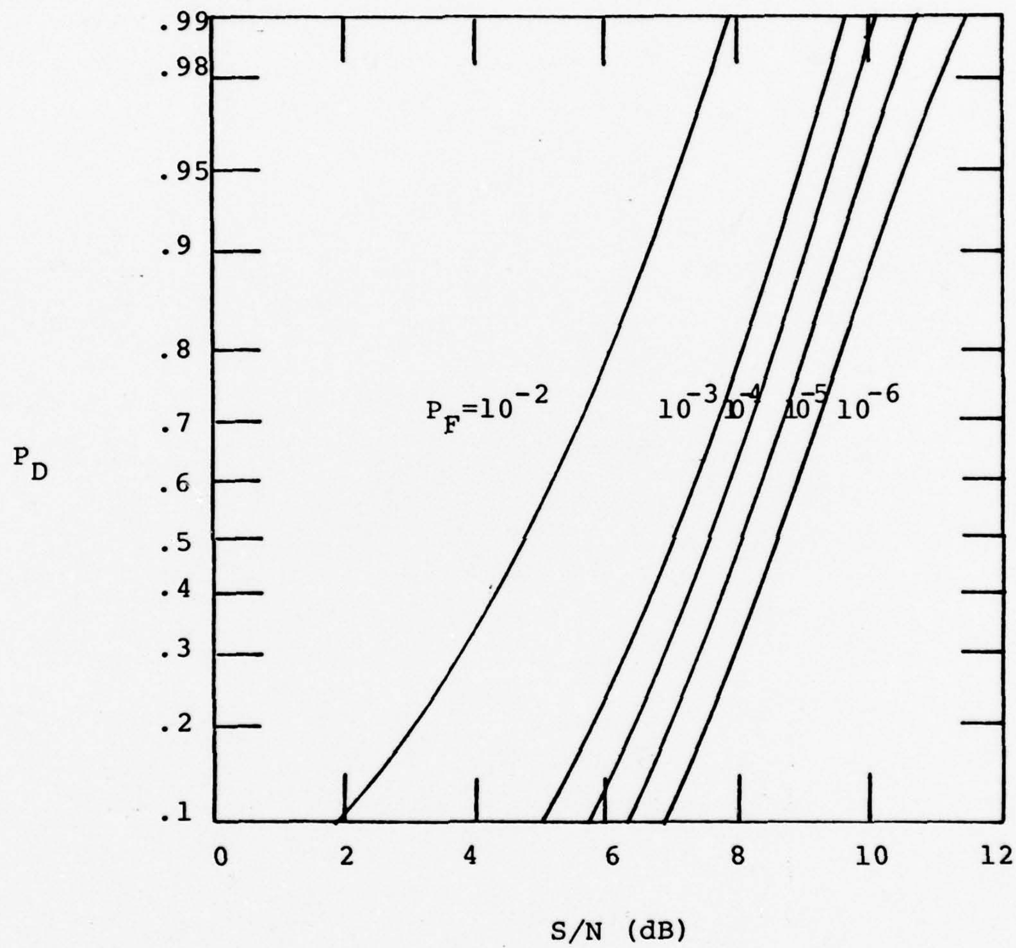
Numerical Results

We conducted simulation experiments for the above detector using lognormal and contaminated normal distributions as done before in the previous two sections. First we considered γ to be constant (it was changed in a very narrow region to guarantee P_F) and changed the second threshold K to get the required value of P_F. The results of this test are shown in figures (36) to (38).



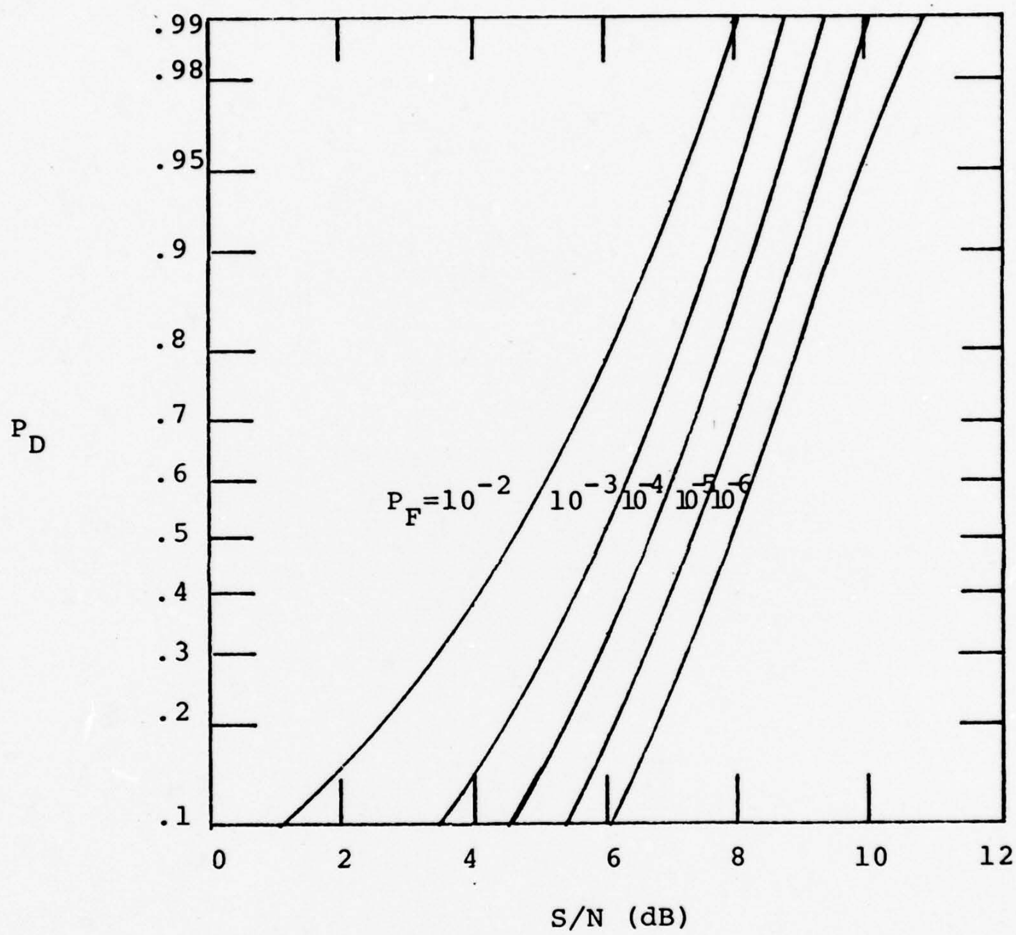
Probability of detection vs S/N (dB) for log-normal distribution (σ=6dB) and E.V.T. detector; N=30

Figure 36



Probability of detection vs S/N (dB) for log-normal distribution (σ=6dB) and E.V.T. detector, N=10

Figure 37



Probability of detection vs S/N (dB) for contaminated normal distribution and E.V.T. detector, N=10

Figure 38

It should be noticed that the required (S/N) for lognormal is slightly higher than that for contaminated normal to achieve the same P_d, especially at low P_F. The test was repeated for lognormal using K=7 and changing γ when N=10. Figure (39) shows the results of this test, which show improvement in the receiver performance, especially at $P_F = 10^{-2}, 10^{-3}$ and 10^{-4} where K was taken in the previous test to be equal to 4,5,6 respectively.

Summary

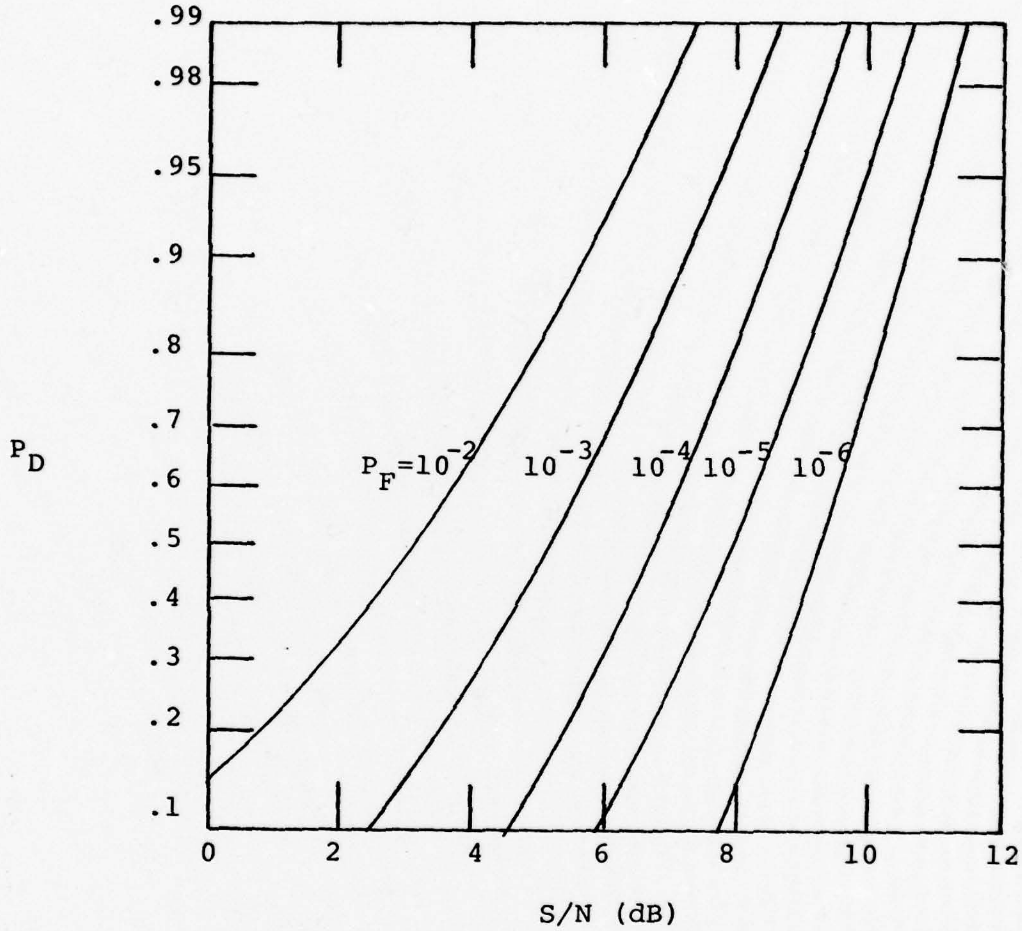
In this section, the EVT detector was described and it was shown that it leads to the same detection scheme as the binary integration. It was shown also that the test performance is sensitive to the second threshold K which cannot be chosen in an optimal way without knowing the exact distribution under consideration.

(V-5) The Trimmed-Mean detector:

In this section, we give a brief description of the trimmed-mean detectors which will be utilized in the next section for comparison. Assume there are N observations, then the designer chooses two numbers n_1 and n_2 such that $N \geq n_2 \geq n_1 \geq 0$, and the test statistics will be

$$T_N = \sum_{i=n_1}^{n_2} z_i$$

where z_i is the ith smallest observation, i.e. $z_1 \leq z_2 \leq \ldots \leq z_n$.



Probability of detection vs S/N (dB) for log-normal distribution and E.V.T. detector (m=7); N=10

Figure 39

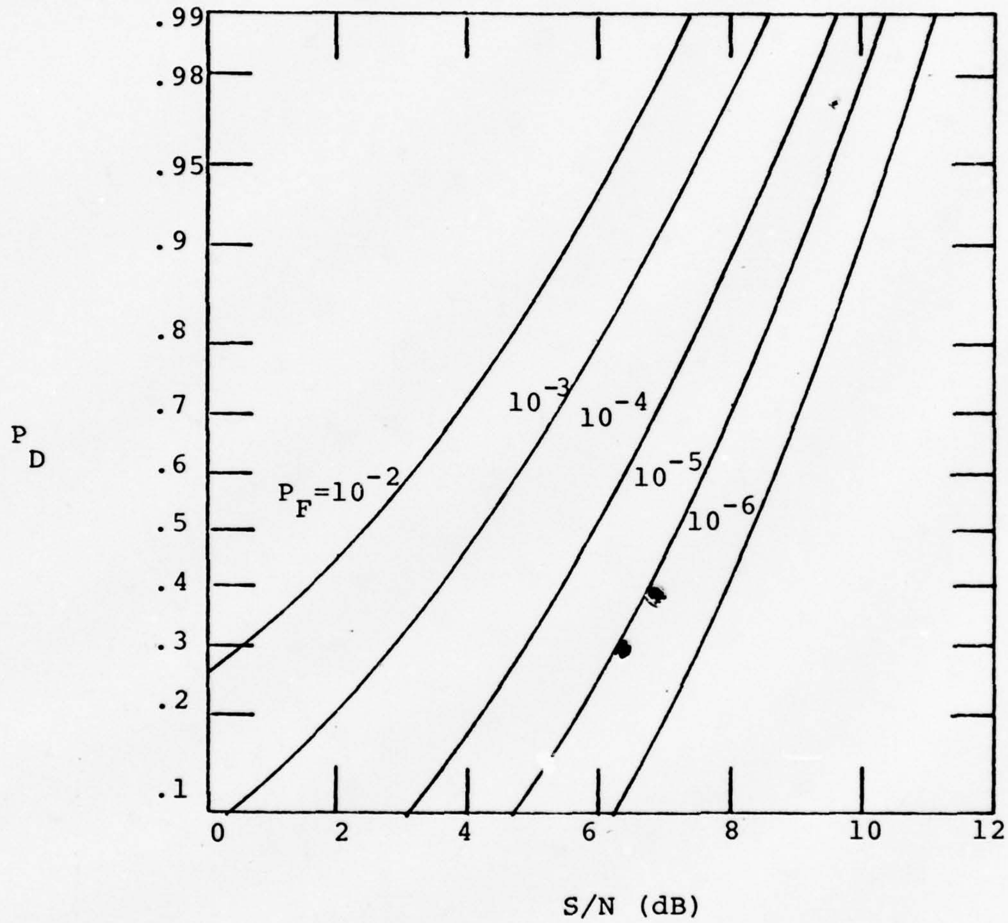
One difficulty with this test is that it requires sorting the data. The second difficulty is in setting the threshold since it depends on the distribution and there is no mathematical expression to be used. Notice also that if $n_1=n_2=[\frac{N}{2}]$ this turns out to be the median detector and if $n_1=1$ and $n_2=N$ it is the mean detector. On the other hand, it was shown in [27] that this detector performs better than most of the known detectors for both lognormal and contaminated normal distributions. More details about the trimmed-mean detector can be found in [38].

Figure (40) shows the performance of the trimmed mean detector against lognormal [σ=6db) when N=10, $n_1=2$ and $n_2=5$. The performance of this detector against contaminated normal is the same as the M-detector on envelope, so it will not be repeated here. It was found also that the M-detector is better for both distributions then the trimmed mean $n_1=4$, $n_2=7$.

(V-6) Detector Comparison

a) Performance

Comparing the performance results of the chosen detectors, it should be noted that for the lognormal distribution at N=10 the performance of the M-detector on envelope, the trimmed-mean detector ($n_1=2$, $n_2=5$) and the fixed step



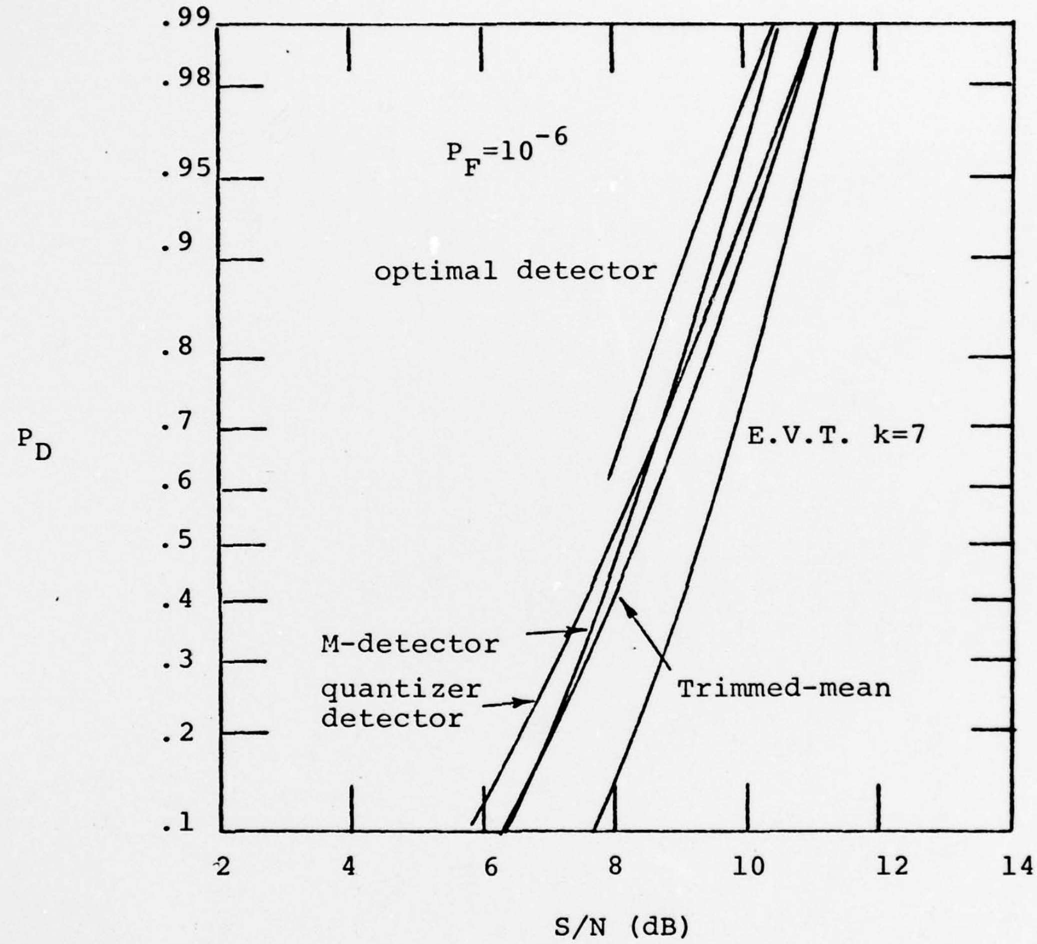
Probability of detection vs S/N (dB) for log-normal distribution (σ=6 dB) and Trimmed mean detector ($n_1=2, n_2=5$); N=10

Figure 40

quantizer detector are all very comparable. In fact the graphs for P_d against (S/N) overlap in most cases, and all of them are within less then 1db. Also, the performance of the M-detector gets bette than the others as P_F decreases and P_d increases. The E.V.T. detector always has poorer performance than the others, even for K=7 which was shown to be optimal for lognormal when N=10 [29]. Figure (41) shows an example of these comparisons at $P_F=10^{-6}$. It shows also the Chernoff lower bound as an approximation for the performance of the optimal detector. Both the M-detector and the fixed quantizer detectos are the closest to this bound, with the M-detector better at higher values of P_d.

At N=30, the performance of both the M-detector and the fixed step quantizer is almost the same with the M-detector better at low P_d. The E.V.T. is far behind, mainly because we did not utilize the optimal value for K.

For the contaminated normal, the performance of the trimmed mean and M-detector was the same as mentioned before. The E.V.T. was the worst detector in the group. The fixed quantizer is very slightly better than the M-detector at $P_F = 10^{-2}$ but the M-detector is much better at $P_F \leq 10^{-4}$.



Comparison of various detectors for log-normal distribution (σ=6dB) at $P_F=10^{-6}$; N=10

Figure 41

<u>General properties</u>

1) The performance of the E.V.T. detector is highly dependent on the value of the second threshold K and the true noise distribution, and since the main reason behind this study is the uncertainty about the noise distribution, this detector cannot be recommended for applications.

2) Although the trimmed-mean detector has comparable performance to the other two good detectors, implementation seems impractical since it requires sorting the data, and there is no known way to set the threshold for it.

3) The M-detector and the fixed step quantizer detector are the simplest to implement among all detectors studied, and they also have the best performance.

4) The M-detector has the advantage over the fixed step quantizer detector that it requires less information. It requires knowledge of only two quantiles, while the quantizer detector requires more quantiles.

5) Using the quantizer detector, P_F will be set only approximately and some values of P_F are unachievable for finite N.

6) If the designer is interested in the low P_F - high P_d region, the M-detector has better performance than the quantizer detector.

It is our opinion that the M-detector is the best of all the detectors involved in the study, for the above reasons. Meanwhile, we believe, as was mentioned earlier in this chapter, that some research efforts should be directed toward finding procedures for setting the quantizer parameters that depend on less information than is used here.

(V-7) Summary

The problem of detection of signals with random phase in sea clutter-like noise was treated in this chapter. Several detectors were developed and/or described. Comparing these detectors as to performance and implementation it was found that the M-detector on envelope is best.

VI. Performance of the M-Detector with Dependent Data

In all the above, it was assumed that the noise samples were independent. In this chapter, we will examine the performance of M-detectors in certain types of correlated noise.

Let $\{x_i\}$ be a dependent sequence of random variables given by

$$x_n = \sum_{i=1}^{m} \lambda_i \, x_{n-i} + w_n \ , \ n = 0, \pm 1, \pm 2, \ldots$$

where $\{w_i\}$ is a sequence of independent, identically distributed random variables with mean zero. Let the observations be

$$y_n = \theta + x_n$$

and let $\hat{\theta}$ satisfy

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} l(y_i - \hat{\theta}) = 0 \ .$$

We want to show that $\hat{\theta} \to \theta$ a.s. and that $\sqrt{N}(\hat{\theta} - \theta)$ is asymptotically normal.

Our principal tool is the following result due to Billingsley [39, P.184]. Let

$$\eta_n = f(\xi_n, \xi_{n-1}, \ldots) \ , \ n=1,2,\ldots$$

$$\eta_{\ell n} = f_\ell(\xi_n, \xi_{n-1}, \ldots, \xi_{n-\ell+1}) \ , \ n=1,2,\ldots$$

$$v(\ell) = E\{|\eta_0 - \eta_{\ell 0}|^2\} \ , \ \ell=1,2,\ldots$$

$$S_n = \eta_1+\eta_2+\ldots+\eta_n$$

where $f(\cdot)$ and $f_\ell(\cdot)$ are measurable functions, $\{\xi_n\}$ is a ϕ-mixing sequence with $\sum \phi_n^{1/2}<\infty$, and $\{\eta_n\}$ has zero mean and finite variance. Suppose also that $\sum[v(\ell)]^{1/2}<\infty$. Then the series

$$\sigma^2 = E(\eta_0^2) + 2 \sum_{j=1}^{\infty} E(\eta_0\eta_j)$$

converges absolutely; if $\sigma^2>0$, then $\frac{1}{\sqrt{N}} S_n$ converges in distribution to $N(0,\sigma^2)$.

To apply this result, we rewrite x_n as

$$x_n = \sum_{j=0}^{\infty} b_j w_{n-j}$$

$$b_j = \sum_{i=1}^{m} a_i r_i^j$$

where the r_i's are the roots of $r^m-\lambda_1 r^{m-1}-\ldots-\lambda_m$ and the a_i's are constants (see Sirvanci and Wolff [40]). Then

$$E(x_n^2) = \sum_{j=0}^{\infty} b_j^2 \sigma_w^2 < \infty .$$

Note that $\{w_i\}$ is ϕ-mixing since it is an independent sequence. It remains to show that

$$\sum_{\ell=1}^{\infty} \{E(1(x_0)-1(x_{0\ell}))^2\}^{1/2} < \infty , \qquad \text{(VI.1)}$$

where

$$x_{0\ell} = \sum_{j=0}^{\ell=1} b_j w_{-j} .$$

Lemma (VI.1):

If $|l(x_1)-l(x_2)| \le R\,|x_1-x_2|$ for all x_1, x_2, and for $0<R<\infty$, then (VI.1) holds.

Proof:

We have

$$|l(x_0)-l(x_{0\ell})| \le R|x_0-x_{0\ell}| = R\left|\sum_{j=\ell}^{\infty} b_j w_{-j}\right|$$

and thus

$$E\{l(x_0)-l(x_{0\ell})\}^2 \le R^2\, E\{\sum_{j=\ell}^{\infty} b_j w_{-j}\}^2 \le R^2 \sum_{j=\ell}^{\infty} b_j^2\, \sigma_w^2 .$$

Let a be the maximum of the $\{a_i\}$ and r be the maximum of the $\{r_i\}$. Then since $|b_j| \le mar^j$,

$$\sum_{j=\ell}^{\infty} b_j^2 \le m^2 \sum_{j=\ell}^{\infty} a^2 r^{2j} = m^2 a^2\, \frac{r^{2\ell}}{1-r^2}$$

and therefore

$$\sum_{\ell=1}^{\infty} \{E(l(x_0)-l(x_{0\ell}))^2\}^{1/2} \le \frac{Rm\sigma_w a}{\sqrt{1-r^2}} \sum_{\ell=1}^{\infty} r^{\ell} = \frac{Rm\sigma_w ar}{\sqrt{1-r^2}\,(1-r)}$$

since $0<r<1$.

Note that if $l(\cdot)$ is either of the following, it satisfies the requirements of the above lemma:

$$l(t) = \begin{cases} a , & t \ge a \\ t , & -a \le t \le a \\ -a , & t \le -a \end{cases}$$

or

$$l(t) = \begin{cases} \tan(ca) & , \quad t \geq a \\ \tan(ct) & , \quad -a \leq t \leq a \\ -\tan(ca) & , \quad t \leq -a \end{cases}$$

with $|ca| < \pi/2$.

To prove consistency of $\hat{\theta}$, we need a slight generalization of the above. Define

$$\eta_n = l(x_n - \theta) - E[l(x_n - \theta)]$$

$$\eta_{n\ell} = l(x_{n\ell} - \theta) - E[l(x_{n\ell} - \theta)] \ .$$

Then

$$\begin{aligned} v(\ell) &= E\{|\eta_0 - \eta_{0\ell}|^2\} \\ &= E\{|l(x_0-\theta) - l(x_{0\ell}-\theta)|^2\} - \{E(l(x_0-\theta) - l(x_{0\ell}-\theta))\}^2 \\ &\leq E\{l(x_0-\theta) - l(x_{0\ell}-\theta)\}^2 \leq R^2 E(x_0 - x_{0\ell})^2 \ . \end{aligned}$$

and

$$\sum_{\ell=1}^{\infty} [v(\ell)]^{1/2} < \infty \ .$$

Now from the above we have

$$\frac{1}{\sqrt{N}} \sum_{j=1}^{N} [l(x_i - \theta) - E[l(x_i - \theta)]] \rightarrow N(0, \sigma_1^2)$$

where

$$\sigma_1^2 = E(\eta_0^2) + 2 \sum_{j=1}^{\infty} E(\eta_0 \eta_j) \ .$$

That is, for any $\varepsilon>0$

$$\frac{1}{N}\sum_{j=1}^{N} l(x_j-(\theta_0+\varepsilon)) \to E[l(x_i-(\theta_0+\varepsilon))] > 0$$

$$\frac{1}{N}\sum_{j=1}^{N} l(x_j-(\theta_0-\varepsilon)) \to E[l(x_i-(\theta_0-\varepsilon))] < 0$$

where θ_0 satisfies

$$E[l(x_i-\theta_0)] = 0 .$$

Thus for any $\varepsilon_0>0$ there is a N_0 such that for $N > N_0$

$$\theta_0 - \varepsilon \leq \hat{\theta} \leq \theta_0 + \varepsilon$$

and therefore $\hat{\theta} \to \theta_0$ a.s.

Since $\sqrt{N}\,\hat{\theta} \leq \gamma$ if and only if $\frac{1}{\sqrt{N}} \sum_{j=1}^{N} l(x_j-\gamma/\sqrt{N}) \leq 0$, asymptotic normality follows from the proof of Thm. 2.1. in [39] if we take $\eta_i = l(x_i- \gamma/\sqrt{N})$.

To evaluate the finite sample size performance of the M-detector on components and on envelope in the presence of correlated noise, a set of simulations were conducted using correlated data. For the M-detector on components a correlated Rayleigh, log-normal, and contaminated normal noise were utilized. To generate the correlated Rayleigh (z) we utilized the following formula

$$x_i = 0.1\ x_{i-1} + u_i$$

$$y_i = 0.1\ y_{i-1} + v_i$$

$$z_i = (x_i^2 + y_i^2)^{1/2}$$

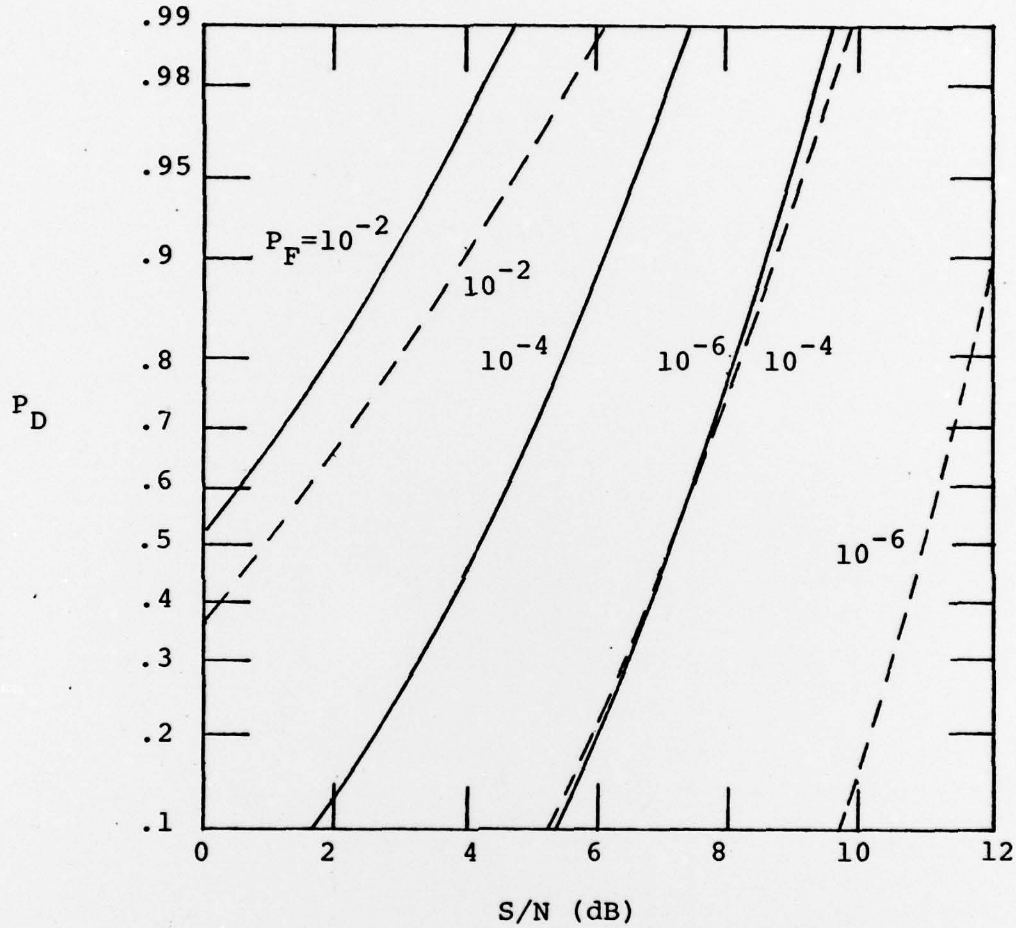
where $\{u_i\}$ and $\{v_i\}$ are independent samples from a zero mean normal distribution with variance 0.99. This value was chosen to make f(x) and f(y) members of the class of p-point distributions containing the standard normal distribution. This will simplify the comparison with the independent case treated in Chapter 2. The correlated log-normal was generated as

$$z_i = \exp(y_i)$$

$$y_i = 0.1\ y_{i-1} + u_i$$

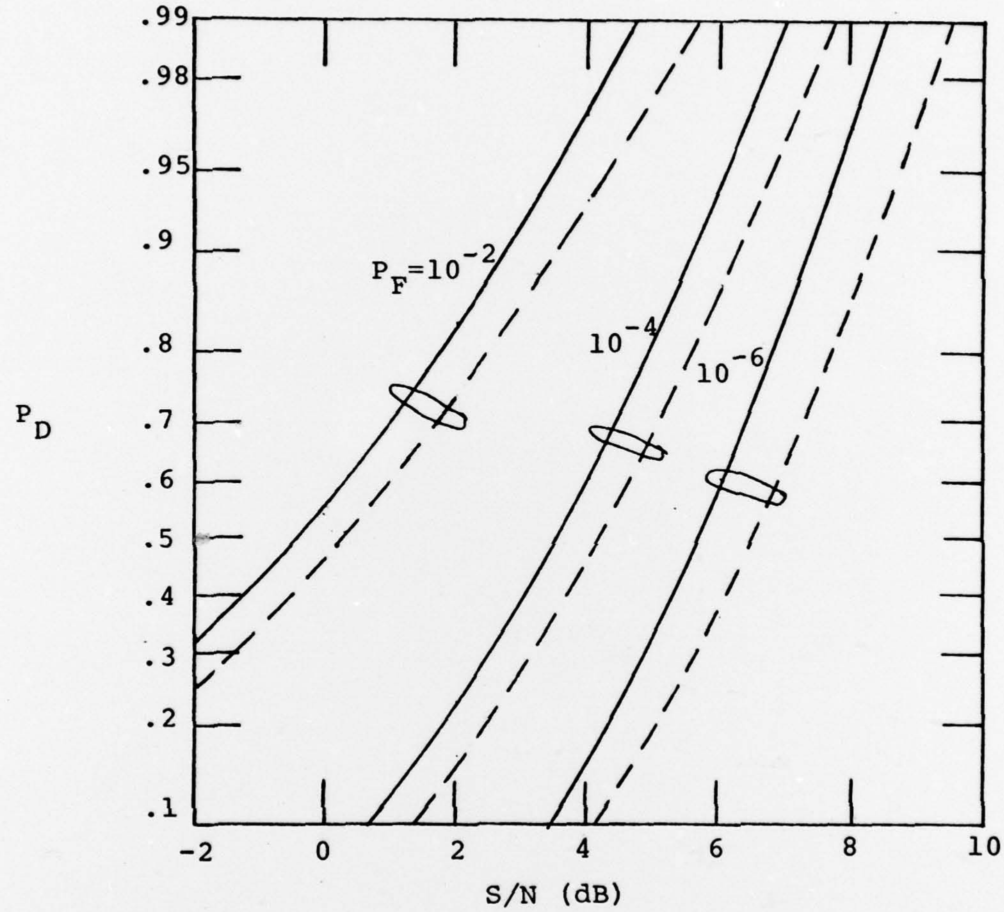
where $\{u_i\}$ is as before. The contaminated normal was generated as in the Rayleigh case except that $\{u_i\}$ and $\{v_i\}$ were independent samples from a contaminated normal distribution.

Figures 42-44 show performance of the M-detector on components (same set of detectors as in Section II.5) for log-normal, contaminated normal, and Rayleigh, respectively. Also shown is the performance of the conventional square-law detector. The M-detector is still superior to the square-law detector for both log-normal and contaminated normal cases, and it is less than 1dB worse than the square-law detector for



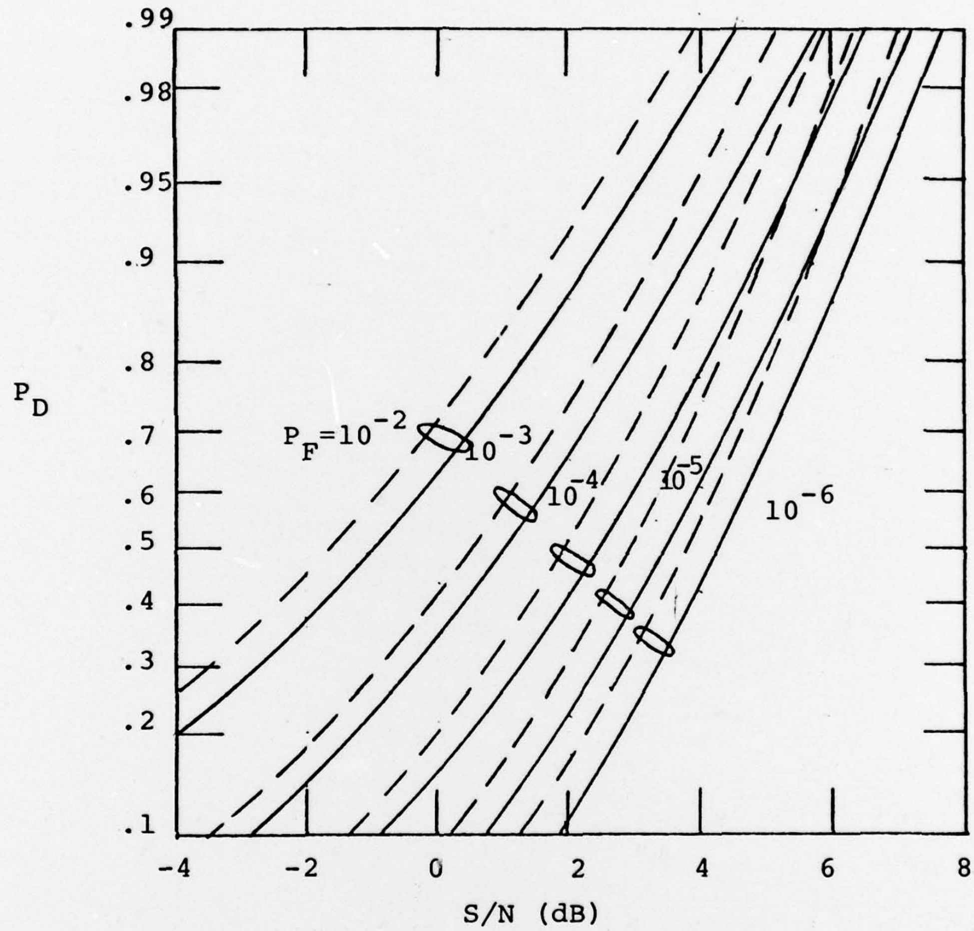
Probability of detection vs. S/N (dB) for correlated lognormal noise and M-detector on components (—) and square-law detector (---)

Figure 42



Probability of detection vs. S/N for correlated contaminated normal noise at N=10 for M-detector on components (—) and square-law detector (---)

Figure 43



Probability of detection vs. S/N (dB) for correlated Rayleigh noise for M-detector on components (—) and square-law detector (---)

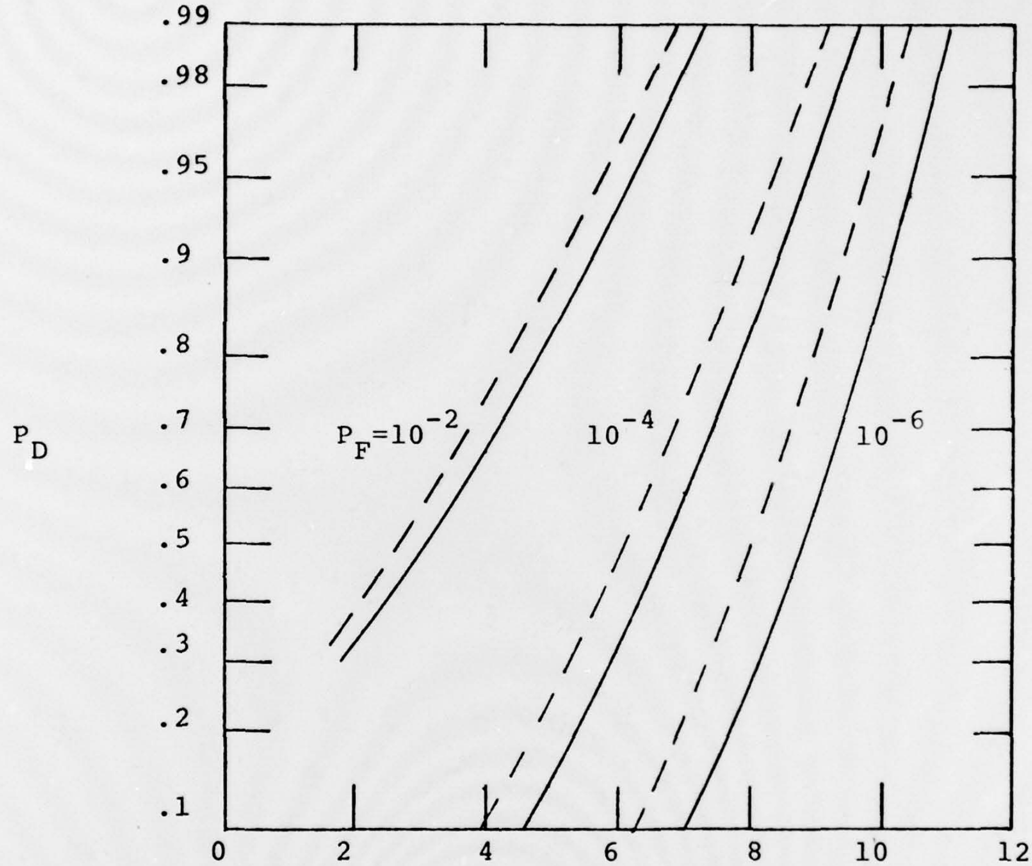
Figure 44

the Rayleigh distribution. Comparing these graphs with Figs. 2-4, it can be seen that there is a loss of power that achieves a maximum of about 1dB at $P_F = 10^{-6}$ when the noise is Rayleigh. The minimum power loss occurs in the log-normal case.

The graphs do not tell us how much power is lost if the threshold is set assuming correlated noise when in fact the noise is independent. This power loss was calculated and found to be less than 1.5dB, the maximum loss occuring in the Rayleigh case.

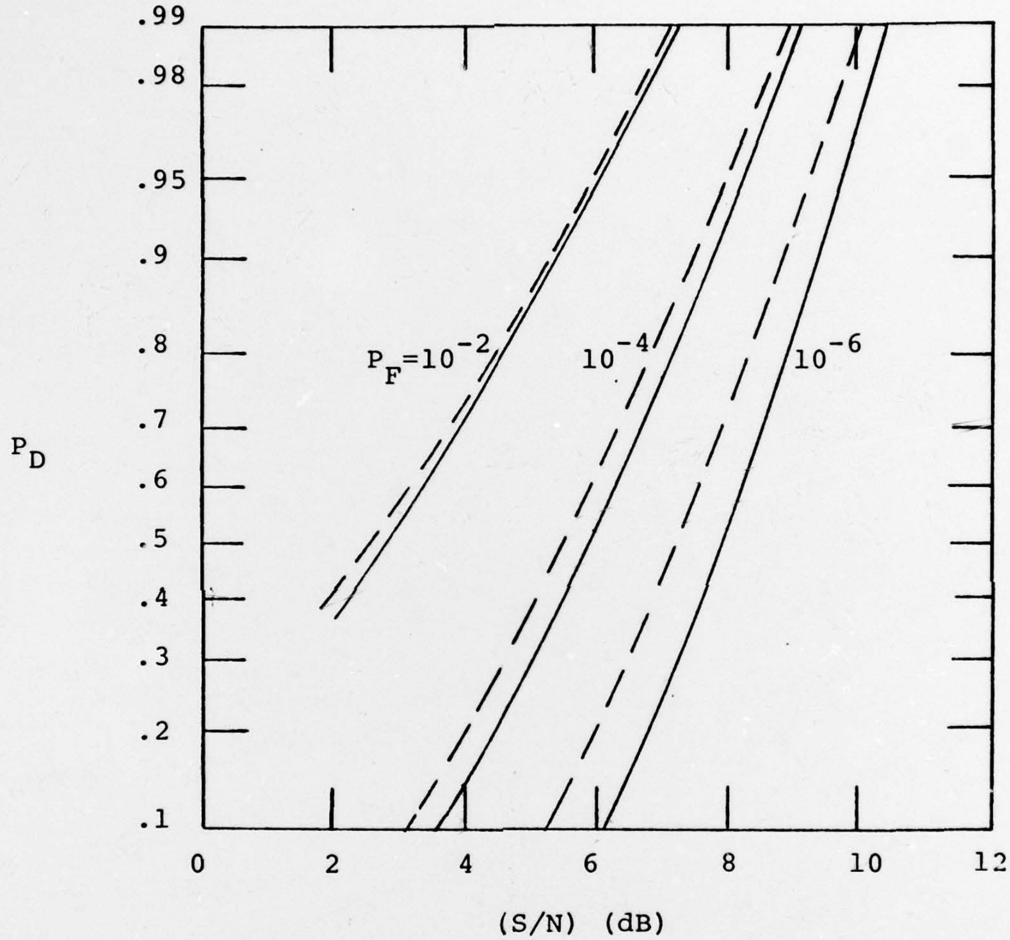
To test the performance of the M-detector on envelope for correlated noise the log-normal and contaminated normal cases were considered. Figs. 45-46 show the probability of detection vs. signal-to-noise ratio (S/N) for these cases at N=10. Also shown for comparison are the corresponding results for the independent case. The loss of power is always less than 1dB in both cases, increasing with decreasing P_F. If the threshold was set to give $P_F = 10^{-6}$ assuming correlated noise, then the actual P_F would be 10^{-7} if in fact the noise was uncorrelated. This is equivalent to a 1dB power loss. For larger values of P_F, the loss is smaller.

Finally, one may do better by choosing an $l(\cdot)$ other than the tangent function. However, an intelligent choice would require knowledge of the correlation, which is usually unknown _a priori_.



Probability of detection vs. (S/N) for M-detector on envelope and correlated lognormal noise (—) and independent lognormal noise (---) at σ=6dB, N=10

Figure 45



Probability of detection vs. (S/N) for M-detector on envelope and correlated contaminated normal noise (—) and independent contaminated normal (---), N=10.

Figure 46

VI. Summary and Conclusion

In this paper the problem of robust detection of non-coherent signals in noise is solved under the assumption that the noise distribution is unknown, but a member of a known class of distribution functions. This problem was divided into three main categories:

a) Partially-coherent signals

b) Unknown frequency and phase

c) Random phase

In the partially-coherent case, it was assumed that the unknown phase is constant over each observation period but changes randomly from one observation period to another. We have introduced two general methods for the design of detectors which are asymptotically most robust, in the sense of guaranteeing a non-trivial lower bound on the receiver performance over the class of distributions under consideration. It is shown that the design of any of these detectors requires the existence of a density function of minimum Fisher information number in the class of distributions. Two detectors were designed for the special case in which the distribution of the noise inphase and quadrature phase components are members of the class of p-point distributions. Simulation results for finite sample sizes were also given for different distributions. It was found that the M-detector performs better than both A-detector and the square law detector. The above results were extended also to the second case of unknown frequency and phase.

In the third case, it was assumed that the signal phase changes randomly from one observation to another. Three detectors were presented, the M-detector on envelope, the quantizer detector, and the extreme value theory (E.V.T.) detector. It was found that the M-detector is the best among the above three and the trimmed mean detector, judging from performance and ease of implementation in the presence of limited information about the noise distribution. All of the above detectors assume knowledge of some quantiles. The quantizer detector is the second best.

In general, we recommend the M-detector on component or on envelope when the observation distribution is not completely known, because of their threshold stability property and because they protect the decision rule against bad observations.

REFERENCES

[1] J. D. Gibbons, "Nonparametric Statistical Inference," McGraw-Hill, N.Y., 1971.

[2] E. L. Lemann, "Testing Statistical Hypotheses," John Wiley and Sons, N.Y., 1959.

[3] H. L. VanTrees, "Detection, Estimation and Modulation Theory," John Wiley and Sons, N.Y., 1968.

[4] S. S. Wilks, "Mathematical Statistics," John Wiley and Sons, N.Y., 1963.

[5] P. J. Huber, "A Robust Version of the Probability Ratio Test," Ann. Math. Statist., Vol. 36, pp. 1753-1758, December 1965.

[6] J. W. Tukey, "A Survey of Sampling from Contaminated Distributions," Contributions to Probability and Statistics, (ed Olkin), Stanford University Press, Stanford, Calif., pp. 448-485, 1960.

[7] F. R. Hample, "Contributions to the Theory of Robust Estimation," Ph.D. Dissertation, University of Calif., Berkeley, Calif., 1968.

[8] R. D. Martin and S. C. Schwartz, "Robust Detection of a Known Signal in Nearly Gaussian Noise," IEEE Trans. Inform. Theory, Vol. IT-17, pp. 50-56, January 1971.

[9] S. A. Kassam and J. B. Thomas, "Asymptotically Robust Detection of a Known Signal in Contaminated Non-Gaussian Noise," IEEE Trans. Inform. Theory, Vol. IT-22, pp. 22-26, January 1976.

[10] A. H. El-Sawy and V. D. VandeLinde, "Robust Detection of Known Signals," IEEE Trans. Inform. Theory, Vol. IT-23, pp. 722-727, Nov. 1977.

[11] D. R. Klose and L. Kurz, "A New Representation Theory and Detection Procedure for a class of Non-Gaussian Noise Channels," IEEE Trans. Comm. Tech., Vol. COM-17, pp. 225-234, April 1969.

[12] G. V. Trunk and S. F. George, "Detection of Targets in Non-Gaussian Sea Clutter," IEEE Trans. Aerosp. Electron. Syst., AES-6, pp. 620-628, Sep. 1970.

[13] R. L. Mitchell and J. F. Walker, "Recursive Methods for Computing Detection Probabilities," IEEE Trans. Aerosp. and Electron. Syst. AES-7, pp. 671-676, July 1971.

[14] G. V. Trunk, "Noncoherent Detection of Nonfluctuating Targets in Contaminated-Normal Clutter," NRL Rep. 6858, March 1969.

[15] P. J. Huber, "Robust Estimation of Location Parameter," Ann. Math. Statist., Vol. 35, pp. 73-101, March 1964.

[16] A. Papoulis, "Probability, Random Variables and Stochastic Processes," McGraw-Hill, N.Y., 1965.

[17] A. D. Whalen, "Detection of Signals in Noise," Acad. Press, N.Y.,1971.

[18] P. J. Huber, "Fisher Information and Spline Interpolation," Ann. Statist., Vol. 2, No. 5, pp. 1029-1033, Sep. 1974.

[19] E. L. Price and V. D. VandeLinde, "A Solution Technique for the Problem of Minimax Estimation of a Location Parameter," The Johns Hopkins University, Tech. Rep. EE 75-3, 1975.

[20] R. D. Martin and C. J. Masreliez, "Robust Estimation via Stochastic Approximation," IEEE Trans. Inform. Theory, Vol. IT-21, pp. 263-271, May 1975.

[21] A. H. El-Sawy and V. D. VandeLinde, "Detection of Signals in the Presence of Impulsive Noise," Conference on Inform. Sciences and Systems, Baltimore, MD, March 1978.

[22] E. L. Price and V. D. VandeLinde, "Robust Estimation Using the Robbins-Monro Stochastic Approximation Algorithm," Fifteenth Annual Allerton Conf. on Circuit and Syst. Theory, Monticello, Ill., Sep. 1977.

[23] J. Sacks, "Asymptotic Distribution of Stochastic Approximation Procedures," Ann. Math. Statist., Vol. 29, pp. 373-405, 1958.

[24] H. L. VanTrees, "Detection, Estimation and Modulation Theory," Vol. 3, John Wiley & Sons, N.Y., 1971.

[25] G. V. Trunk, "Range Resolution of Targets Using Automatic Detectors," NRL Rep. 8178, Nov. 1977.

[26] S. F. George, "The Detection of Nonfluctuating Targets in Log-Normal Clutter," NRL Rep. 6796, Oct. 1968.

[27] G. V. Trunk, "Non-Rayleigh Sea Clutter: Properties and Detection of Targets," NRL Rep. 7986, June 1976.

[28] G. V. Trunk, "Median Detector for Noncoherent Distributions," NRL Rep. 6898, May 1969.

[29] D. Schleher, "Radar Detection in Log-Normal Clutter," Ph.D. Dissertation, Polytech. Inst. of N.Y., June 1975.

[30] S. A. Kassam, "Optimum Quantization for Signal Detection," IEEE Trans. Comm., vol. COM-25, pp. 479-484, May 1977.

[31] H. V. Poor and J. B. Thomas, "Maximum-Distance Quantization for Detection," Proceedings of the Fourteenth Annual Allerton Conf. on Circuit and System Theory, pp. 925-934, Sep. 1976.

[32] Y. Ching and L. Kurz, "Nonparametric Detectors Based on m-Interval Partitioning," IEEE Trans. Inform. Theory, Vol. IT-18, pp. 251-257, March 1972.

[33] A. H. El-Sawy, "On the Design of Robust Quantizers For Detection," Proceedings of 1978 Conference on Information Sciences and Systems, Baltimore, MD, March 1978.

[34] E. J. Gumbel, "Statistical Theory of Extreme Values and Some Practical Applications," Washington, D.C., NBS, Appl. Math. Ser. 33, U.S. Government Printing Office.

[35] E. J. Gumbel, "Statistics of Extremes," Columbia University Press, N.Y., 1958.

[36] L. B. Milstein, et al., "Robust Detection Using Extreme-Value Theory," IEEE Trans. Inform. Theory, vol. IT-15, May 1969.

[37] M. Schwartz, "A Coincidence Procedure for Signal Detection," IEEE Trans. Inform. Theory, vol. IT-11, Jan. 1977.

[38] G. V. Trunk, "Trimmed-Mean Detector for Noncoherent Distributions," NRL Rep. 6997, Dec. 1969.

[39] P. Billingsley, "Convergence of Probability Measures," Wiley, N.Y., 1968.

[40] M. B. Sirvanci and S. S. Wolff, "Nonparametric Detection with Autoregressive Data," IEEE Trans. Info. Theory, Vol. IT-22, Nov. 1976.