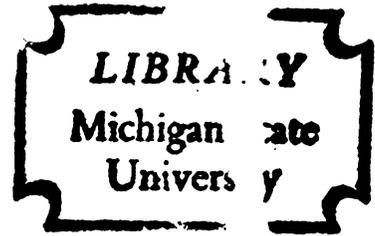


A ZERO-MEMORY NONLINEAR
FILTER FOR SIGNAL DETECTION

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
PHILIP WILLIAM ALLEN
1972



This is to certify that the
thesis entitled

A ZERO-MEMORY NONLINEAR
FILTER FOR SIGNAL DETECTION

presented by

PHILIP WILLIAM ALLEN

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in EESS

A handwritten signature in cursive script, appearing to read "Philip W. Allen".

Major professor

Date Dec. 29, 1971





ABSTRACT

A ZERO-MEMORY NONLINEAR FILTER FOR SIGNAL DETECTION

By

Philip William Allen

This thesis investigates the possibility of increasing the probability of detecting a constant signal in non-Gaussian noise by nonlinear processing of the receiver waveform. The approach taken uses the first-order statistics of the random observables as a starting point for the design of any optimal detection system, in contrast with other applications of nonlinear elements which are motivated by efforts to improve the signal-to-noise ratio in the receiver. The particular noise model admitted in this thesis is characterized by an amplitude probability density function having the form $f(x) = a \exp(-b|x|^\theta)$, $\theta \in [1,2]$. The noise combines additively with one of two transmitter symbols $\{0, \Delta\}$, where $\Delta > 0$, during each decision interval $[0, T]$. The detection problem is solved in an hypothesis testing framework.

Receiver design is constrained by a restriction on the maximum False Alarm Rate (FAR) which must be satisfied for every $\theta \in [1,2]$. This constraint, plus the mathematical intractability of the noise family under consideration, suggests the use of non-parametric detectors. The sign detector is very well suited because of its ease of implementation and its distribution-free



property. It is also the optimal device for weak signals in Laplacian noise ($\theta = 1$). Another detector considered is based on a sample mean. This detector is uniformly most powerful for a given FAR when $\theta = 2$. It is suboptimal for other values of θ but is often used when the underlying PDF's "appear" normal.

This thesis proposes a third detection method, one based on the inclusion of a nonlinear element. The nonlinear transfer characteristic is derived from a probability integral transformation which changes the specified input distribution to a standard normal distribution. This allows determination of a threshold value which guarantees satisfaction of the FAR restriction while avoiding any further randomization at the receiver. This is not possible, in practice, with a sign-detector because its test statistic has a discrete distribution.

The transfer characteristic chosen results from a study of the importance of Monotone Likelihood Ratios (MLR) in signal detection. A method for determining whether a single-parameter family of PDF's has MLR is given and applied to the noise model of interest. The MLR is shown to be invariant under a monotone nonlinear transformation, and this transformation is the basis for the nonlinear filter detector design. This detector is compared to the other detection systems under weak and strong signal conditions, and is shown to be a viable alternative to both.



A ZERO-MEMORY NONLINEAR
FILTER FOR SIGNAL DETECTION

By

Philip William Allen

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering
and
Systems Science

1972



G74662

TO MARSHA



ACKNOWLEDGEMENTS

I wish to thank the General Dynamics Corporation, Electrodynamics Division for the encouragement and financial support that made this thesis possible. Professor Richard Dubes, MSU Department of Computer Science, my teacher and thesis advisor, was a constant source of help, as was Professor Gerald Park, MSU Department of Electrical Engineering and Systems Science. Several helpful suggestions were made by Professors Dennis Gilliland, V. Mandrekar, and Robert Staudte, all of the MSU Department of Statistics and Probability.

Perhaps the greatest contribution was made by my wife whose patience, understanding, and encouragement contributed greatly to the success of this research. Therefore this thesis is dedicated to her.



TABLE OF CONTENTS

		Page
	List of Figures	vi
Chapter		
I	INTRODUCTION	1
	1.1 Requirements and Assumptions	3
	1.2 Parametric and Nonparametric Procedures	4
	1.3 Objections to Nonparametric Procedures	5
	1.4 Literature Review	6
	1.5 Thesis Noise Model	9
	1.6 Central Limit Theorem Errors	12
	1.7 Thesis Objective	15
	1.8 Organization of the Thesis	15
II	STATISTICAL BACKGROUND	16
	2.1 Simple Hypothesis Testing	17
	2.2 Composite Hypothesis Testing	19
	2.3 Signal-to-Noise Ratios (SNR)	23
	2.4 Optimal Detector of Constant Signal in White, Gaussian, Noise	24
	2.5 Measure Transformations and MLR	25
	2.6 Summary	29
III	ZERO-MEMORY NONLINEAR FILTER	30
	3.1 Clippers, Limiters and Integrators	31
	3.2 Measure Transformations	32
	3.3 Concavity	37
	3.4 NLF Output PDF under H_1	38
	3.5 Moment Approximations	41
	3.6 Summary	45
IV	COMPARISON OF DETECTORS	46
	4.1 Optimal Detector of Constant Signal in Laplacian Noise	47
	4.2 Weak Signal Detector Evaluation	49
	4.3 Strong Signal Detector Evaluation	57
	4.4 Simulation	61
	4.5 Summary	70



Chapter		Page
V	GENERAL CONCLUSIONS AND EXTENSIONS	72
	5.1 Conclusions	72
	5.2 Extensions	73
	BIBLIOGRAPHY	75
	APPENDIX I Derivation of $J(\theta:2)$ vs. θ	78
	APPENDIX II Realization of the Nonlinear Filter	82
	APPENDIX III Discontinuity of $g''(x)$ at the Origin	85
	APP NDIX IV Simulation Data	87

LIST OF FIGURES

Figure		Page
1.1	Multishot Communication System	2
1.2	ARE of Sign Test Relative to \bar{X} Test	11
1.3	$\frac{\mu_4}{\sigma^4}$ vs. θ	14
2.1	Invariance of MLR to Linear Transformation	27
2.2	Invariance of MLR to Nonlinear Transformation	28
3.1	"Gaussianizing" Transformation	33
3.2	Laplacian Transfer Characteristic	36
3.3	PDF of NLF-Output Under H_1	42
4.1	Probability Mass Function for S_N	50
4.2	Nonlinear Filter Detector	54
4.3	Simulation Block Diagram	62
4.4	P_D vs. SNR, Laplacian Noise, Sample Size = 4 ...	64
4.5	P_D vs. SNR, Gaussian Noise, Sample Size = 4	65
4.6	P_D vs. SNR, Laplacian Noise, Sample Size = 12 ..	66
4.7	P_D vs. SNR, Gaussian Noise, Sample Size = 12 ...	67
4.8	P_D vs. SNR, Laplacian Noise, Sample Size = 20 ..	68
4.9	P_D vs. SNR, Gaussian Noise, Sample Size = 20 ...	69
A.1	Divergence vs. θ	81
A.2	Piece-Wise Approximation	83
A.3	Piece-Wise Approximation Flow Chart	84

CHAPTER I
INTRODUCTION

This thesis deals with a decision-theoretic approach to a problem commonly encountered in the context of radio communications, namely, the detection of a constant signal in additive, stationary, white noise. When the noise is assumed Gaussian the design and performance of optimal detectors have been developed by Helstrom [H1], Van Trees [V1], Hancock and Wintz [H2] and Root [R3], among others. When the Gaussian noise approximation cannot be made, application of nonparametric testing methods has been proposed by Helstrom [H1], Hancock and Lainiotis [H3], Daly and Rushforth [D1], and many others.

This thesis will point out two criticisms of nonparametric approaches and propose a viable alternative based on a probability integral transformation.

The detection problem of interest is shown in Figure 1.1. The requirements and assumptions which determine the detector design are listed in Section 1.1. Section 1.2 reviews parametric and nonparametric detectors as applied to the detection of constant signals in noise while Section 1.3 points out some criticisms of nonparametric detectors as alternatives to parametric methods. Section 1.4 presents some approaches advanced by others for the problem, while Section 1.5 discusses the applicability of various

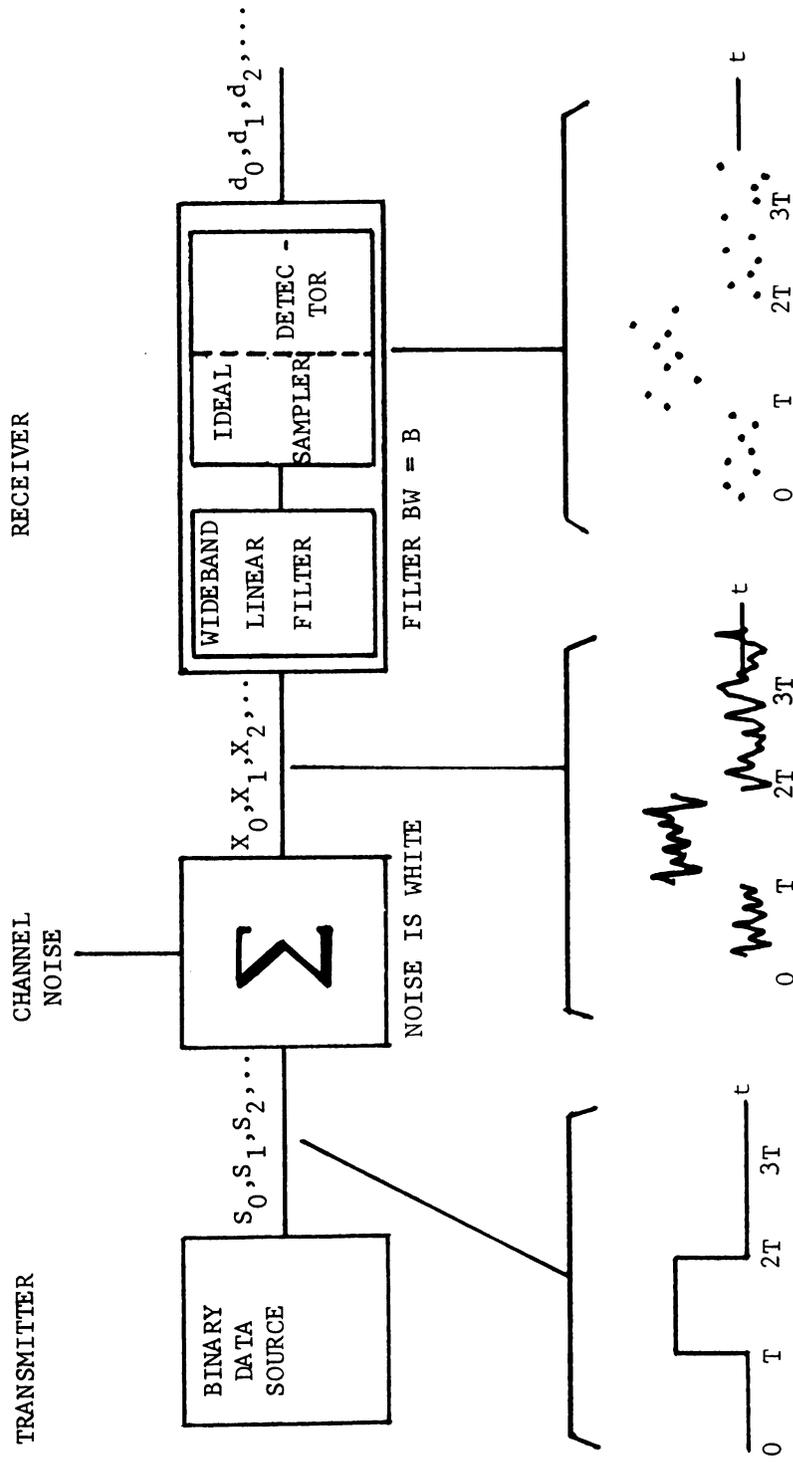


Figure 1.1 Multishot Communication System

parametric and nonparametric approaches to the problem. Section 1.6 investigates central limit theorem errors resulting from tests based on a sample sum. Section 1.7 contains the thesis objective and Section 1.8 explains the organization of the thesis.

1.1 Requirements and Assumptions

All detectors for the problem considered in this thesis are to be designed in accordance with the following assumptions.

- A1 The N independent, identically distributed samples taken at the receiver are to be classified as coming from either the density $f(X)$ or, alternatively, the density $f(X - \Delta)$, $\Delta > 0$, where the form of the probability density function (PDF) is in the family $a \exp(-b|X|^\theta)$.
- A2 The design criterion is Neyman-Pearson; i.e., the probability of signal detection (P_D) is to be maximized subject to a constraint on the False Alarm Rate (FAR)¹.
- A3 The data received in each signaling interval are statistically independent of the data received in all other intervals.
- A4 The source symbol selected for transmission during each signaling interval is statistically independent of the symbol selected during any other signaling interval.
- A5 The transmitter and receiver are synchronized. A receiver interval of T seconds [or N samples]

¹ The statistical literature refers to power and size rather than detection probability and False Alarm Rate, respectively.

corresponds to exactly one transmitter symbol.

A6 The wideband linear filter serves to bound the noise power at the A/D converter (ideal sampler) input without producing sample-to-sample dependence.

Assumptions A3 and A4 guarantee that the analysis of a one-shot receiver, wherein the decision about the data received in the interval $(0, T)$ does not depend on data received in any other interval, is applicable to the multishot system [H2] of Figure 1.1.

The noise model was introduced by Kanefsky and Thomas [K2] as a reasonable representation of impulse noise, one of the most common and troublesome non-Gaussian noise processes. If an impulse process is modeled as a train of independently occurring pulses with small overlapping, then several simple pulse shapes result in amplitude densities having the form specified in A1 [G1].

1.2 Parametric and Nonparametric Procedures

An extensive collection of results has been obtained and published for problems of detecting signals in additive noise. The vast majority of these results are based on the assumption that the underlying noise density is Gaussian and white; thus it can be completely characterized by two parameters. In many cases the Gaussian noise assumption is dictated by physical considerations; thermal noise, for example. Often one of the Central Limit Theorems is invoked to justify consideration of the asymptotic convergence in distribution of a suitably normalized sum of random variables to a standard normal distribution. The mathematical

tractability of the Gaussian noise model has provided an abundance of theoretical results. "There is undoubtedly a temptation to regard distributions as normal, unless otherwise proven, and to use the standard normal theory wherever possible." [K1].

A great amount of interest has centered on the development of a different approach to this problem. This nonparametric approach, formulated in the late 1930's made few assumptions about the probabilistic structure of the noise and used simple and unsophisticated methods for detector design. Because these detectors were based on very general assumptions concerning the noise density, they displayed relative insensitivity to departures from the assumptions, a feature called "robustness."

It has been noted [C1] that whereas parametric procedures provide exact solutions to problems stated approximately, their nonparametric counterparts provide approximate solutions to problems stated exactly. It is in this sense that a nonparametric test may prove to be superior to an "optimum" test designed under invalid assumptions.

The consideration of alternatives to optimal detectors may be prompted by the complexity, both analytic and practical, of the optimal design. Nonparametric detectors are intended to be simple to implement at a cost of some deterioration of performance relative to the optimal detector [T1].

1.3 Objections to Nonparametric Procedures

Asymptotic Relative Efficiency (ARE) provides a convenient measure of relative performance of two detectors which are

designed for the same problem. It allows comparison between a nonparametric detector and an "optimal" detector and measures the effect of departures from the assumptions on which the "optimum" detector is designed. Because ARE is essentially a large sample, small signal performance measure, its validity has been criticized for practical receivers [T1], [N1].

Another criticism of small sample nonparametric detectors is that the distribution of the test statistic is discrete which may necessitate an additional randomization to achieve the requirement of A2. This complication is avoided by the method proposed in this thesis and introduced in Section 1.7.

1.4 Literature Review

Rappaport and Kurz [R1] developed an optimal nonlinear detector for digital data transmission through non Gaussian channels. They noted the practical difficulty of evaluating the PDF of their test statistic and suggested Monte Carlo simulation for any specified noise PDF and sample size. To avoid the very large amounts of computer time involved in any such simulation, the authors investigated the asymptotic performance of the nonlinear detector which was shown to depend on a single SNR parameter. Several signaling waveforms were considered in a Cauchy noise environment and asymptotic system performance was shown to be relatively independent of the signaling waveform.

A suboptimal nonparametric alternative for constant-signal detection was described by Hancock and Lainiotis [H3]. Their detector, based on the Median, or Sign, Test was robust, a

desirable property. It used a nonlinear element (ideal clipper) to reduce the input samples to binary random variables. The distribution of the test statistic under the hypothesis H_0 was binomial. By evaluating the detector's performance on an asymptotic basis through the Central Limit Theorem, the randomization required for small sample problems was not encountered.

An application of a nonlinear filter (NLF) for the improvement of detection reliability for signals in non-Gaussian noise was also investigated by Richard and Gore [R2]. The NLF transfer characteristic was determined from the first-order statistics of the interfering noise and the expected signal amplitude. A narrow bandwidth, linear, integrating, low pass filter (LPF) followed the NLF. The long integration time of the LPF and the Central Limit Theorem was used to justify the assumption that the Gaussian model "approximately" described the LPF output. This permitted the calculation of detection probability and false alarm rate as functions of the NLF output signal-to-noise ratio (SNR). For a very special class of non-Gaussian noises, resulting from passing Gaussian noise through a piece-wise linear nonlinearity, the computed and experimental SNR improvement factors at the NLF output indicated that a considerable improvement in detection reliability could be obtained.

Hancock and Wade [H7] investigated the problem of attaining near-optimum reception of known binary signals over wide-band channels in the presence of narrow-band interference. They noted that the purpose of the receiver was to divide the space of received waveforms into two disjoint regions - the "mark" and

"space" regions. The linear receiver which they described reduced the dimensionality of the received waveform to a univariate statistic which was compared to a decision threshold. The decision surface generated by the linear receiver was determined to be a plane in the received signal space while it was noted that the Bayes optimum decision surface had the form $p_n(\mathbf{X} - \mathbf{M}) = p_n(\mathbf{X} + \mathbf{M})$, with $p_n(\cdot)$ the interference PDF, \mathbf{X} the received waveform vector, and \mathbf{M} the Mark waveform vector. Optimum reception with a linear receiver could only be possible under certain conditions, Gaussian interference being one of them. For more general interference, the authors described a method by which a nonlinear coordinate transformation could transform the decision surface into a hyperplane. Subsequent processing of the transformed coordinates with a linear receiver was shown to result in optimum reception.

This thesis departs from these approaches in the following way. The signal detection problem is viewed at the outset as one to which the existing literature of statistical decision theory can be applied. The property of Monotone Likelihood Ratio is investigated fully and its importance in the detection problem of this thesis is discussed. Then a particular coordinate transformation is chosen within the framework of the decision-theoretic approach, rather than from a heuristic notion of signal-to-noise ratio improvement.

Based on decision-theoretic arguments, this thesis presents a practical alternative to nonparametric and "approximately" optimum parametric detectors.

1.5 Thesis Noise Model

Attention is here focused on the noise model to be considered in this thesis. The noise will be characterized by its first-order statistics, specifically, its PDF

$$f_X(x) = a \exp(-b|x|^\theta), \text{ where } \theta \in [1,2], \quad (1)$$

and a , b and θ are related through:

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (2)$$

$$\int_{-\infty}^{\infty} x^2 f_X(x) dx = \sigma_X^2. \quad (3)$$

An Optimal detector of a constant signal in additive stationary noise having the PDF (1) can be determined, at least in theory. However, the complexity of implementation leads to the consideration of other (perhaps) suboptimal procedures. Furthermore, the detector designed to be optimal for one value of θ may prove unsatisfactory for some other value of θ . The fact that θ is an unknown parameter having a range $[1,2]$ must be included in the design of any suboptimal procedure.

Previous work for this detection problem has concentrated on certain nonparametric detectors. Their application is motivated by simplicity and robustness. Their utility is often justified by the criterion of ARE, a small-signal, large sample performance measure. Comparison between robust detectors and those based on an assumed Gaussian noise model show that "optimal" detection can be misleading when the underlying assumptions are not valid.

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every entry, no matter how small, should be recorded to ensure the integrity of the financial statements. This includes not only sales and purchases but also expenses and income. The document provides a detailed list of items that should be tracked, such as inventory levels, accounts payable, and accounts receivable. It also outlines the proper procedures for recording these transactions, including the use of double-entry bookkeeping and the importance of regular reconciliations.

The second part of the document focuses on the analysis of the financial data. It explains how to interpret the various components of the financial statements, such as the balance sheet, income statement, and cash flow statement. It provides a step-by-step guide to calculating key financial ratios and metrics, such as the current ratio, debt-to-equity ratio, and return on assets. The document also discusses the significance of these ratios and how they can be used to assess the financial health and performance of the organization.

The final part of the document addresses the reporting requirements for the financial statements. It outlines the format and content of the financial statements, including the required disclosures and the role of the auditor. It also discusses the importance of transparency and accountability in financial reporting, and provides guidance on how to communicate the results of the financial analysis to stakeholders.

What is very interesting about the noise model admitted is that the optimal tests for the extreme values, $\theta = 1$ and $\theta = 2$, are well known. When $\theta = 1$, the test based on signs (S_N) is locally most powerful (LMP) while the test based on the sample mean (\bar{X}) is uniformly most powerful (UMP) for $\theta = 2$. Because of the optimal property of the sign detector, other non-parametric tests (e.g., the Wilcoxon or van der Waerden tests) will not be considered in this thesis. Both the sign and sample mean tests are easily implemented although the former has the disadvantage of a discrete PDF for its detector statistic.

The ARE of the sign test, S_N , relative to that based on the sample mean, \bar{X} , can be shown to be $ARE_{S_N, \bar{X}} = \frac{\theta^2 \Gamma(\frac{3}{\theta})}{\Gamma^3(\frac{1}{\theta})}$, which is plotted in Figure 1.2. If $1 \leq \theta \leq 1.4$, the Sign Test would be preferred to the \bar{X} test if ARE were the sole selection criterion. The range of θ for which the sign test is preferred must be reduced if the following are also considered:

- a) ARE may not be a valid comparison measure in small sample problems.
- b) The discrete PDF of the test statistic may require an additional randomization at the receiver.
- c) The optimality property of the sign test is a local property, i.e., it is based on a vanishingly small SNR.

For these reasons the test based on \bar{X} seems preferable to the sign test for a large range of θ , but there is a major disadvantage to its use. For the small sample sizes considered here central limit theorem errors must be taken into account

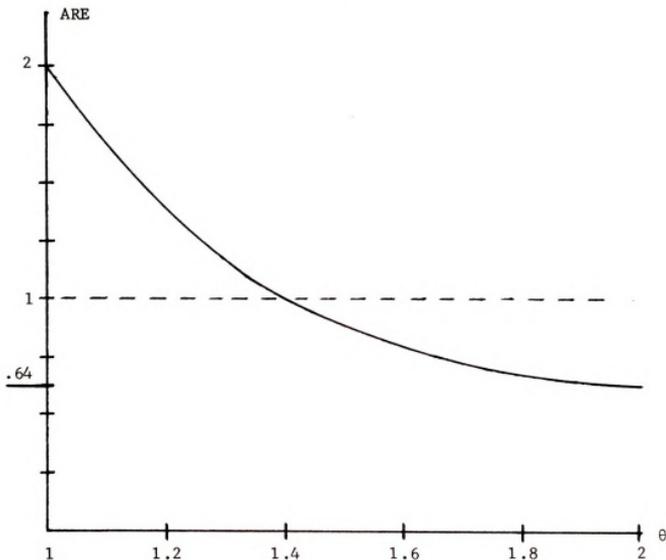


FIGURE 1.2 ARE of Sign Test Relative to \bar{X} Test

during any receiver design. Quite obviously, the threshold value determined by consideration of the design specification on FAR and assuming that \bar{X} is Gaussian will be optimal only for $\theta = 2$.

It is also obvious that no single detector structure can be optimal for every member of the family of noise distributions and all possible Signal to Noise Ratios (SNR), defined as the ratio $\frac{A}{\sigma_X}$, i.e., the rms signal power divided by the rms noise power. Because of this, the design criterion A2 must be modified (adaptive receivers are of no concern in this thesis

although they do offer an alternative approach) to state specifically the terms under which optimization is to be accomplished.

For the purposes of this thesis, the FAR will be assumed as the primary criterion of design; i.e., any detector design must guarantee a FAR which is less than, or equal, to the design goal, for all $\theta \in [1,2]$.

As previously mentioned, the sign test is applicable to this type problem. Because of its distribution-free properties, a specified value of FAR can be achieved and maintained for all possible θ , although it may require an additional randomization in practical problems.

There has been no consideration of alternatives to the nonparametric detector because of the mathematical intractability of the family of noise densities. However it can be, and often is, argued that the Central Limit Theorem can be applied to this type of problem directly because the assumptions made ensure the convergence in distribution of any properly normalized statistic based on the sample sum. For practical sample sizes there will be some error in this approximation, and this situation is now investigated.

1.6 Central Limit Theorem Errors

The central limit theorem errors are more severe for those density functions whose shape reduces the speed of convergence. This section deals with the problem of determining which member of a family "differs" the most from a standard normal, when the

family is characterized by the PDF of equation (1).

Let $F_n(x)$ be the distribution function of the standardized sample sum. Then the error using a standard normal approximation may be written [W1, p. 265]

$$\begin{aligned}
 F_n(x) - \Phi(x) = & -\frac{1}{\sqrt{n}} \left(\frac{\alpha_3}{3!}\right) \Phi^{(3)}(x) \\
 & + \frac{1}{n} \left[\frac{1}{4!} (\alpha_4 - 3) \Phi^{(4)}(x) + \frac{10}{6!} \alpha_3^2 \Phi^{(6)}(x) \right] \\
 & - \frac{1}{3/2} \left[\frac{1}{5!} (\alpha_5 - 10\alpha_3) \Phi^{(5)}(x) + \frac{35}{7!} \alpha_3 (\alpha_4 - 3) \Phi^{(7)}(x) \right. \\
 & \left. + \frac{280}{9!} \alpha_3^3 \Phi^{(9)}(x) \right] + o\left(\frac{1}{n^{3/2}}\right), \quad (4)
 \end{aligned}$$

where $\alpha_j \triangleq \frac{\mu_j}{\sigma^j}$, μ_j the j -th central moment, and σ the standard deviation of the X_i . By the symmetry of (1), this becomes

$$F_n(x) - \Phi(x) = \frac{(\alpha_4 - 3) \Phi^{(4)}(x)}{4! n} + o\left(\frac{1}{n^{3/2}}\right). \quad (5)$$

The central limit theorem error will be largest for that member $\theta \in [1, 2]$ having the largest fourth central moment.

$$\mu_4 = \int_{-\infty}^{\infty} a X^4 e^{-b|X|^\theta} dx = 2a \int_0^{\infty} X^4 e^{-bX^\theta} dx. \quad (6)$$

Making a change of variables and integrating, yields

$$\mu_4 = \frac{2a}{\theta b^{5/\theta}} \Gamma(5/\theta). \quad (7)$$

Using (A4) and (A7) [Appendix 1] in (7), results in

$$\mu_4 = \frac{\sigma^4 \Gamma(5/\theta) \Gamma(1/\theta)}{\Gamma^2(3/\theta)}.$$

The graph of $\frac{\mu_4}{\sigma^4}$ is shown in Figure 1.3. Thus, when $\theta = 1$, indexing the Laplacian family, the central limit theorem errors are largest. This same result is found in Appendix I using the J-divergence as a measure of difference between members of the family (1).

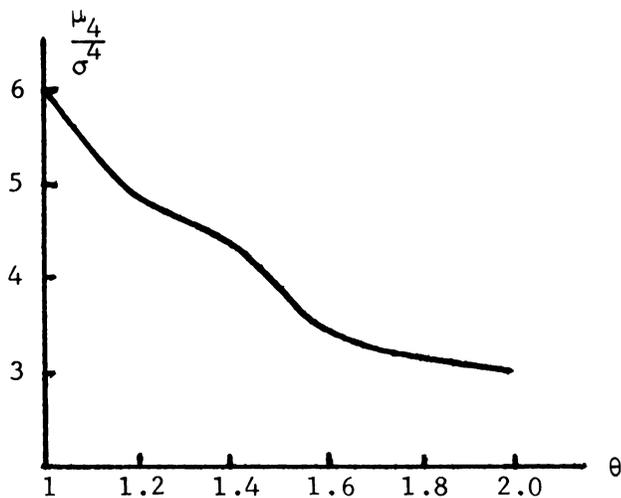


Figure 1.3 $\frac{\mu_4}{\sigma^4}$ vs. θ

The well-known fact that the Laplacian density has "fatter" tails than a normal density of the same mean and variance [H6, p. 43] ensures that the FAR determined by a Gaussian assumption will be optimistic; i.e., it will underestimate the true FAR and is therefore not a viable alternative.

Previous efforts [R1], [R2] with nonlinear filters have demonstrated their capability of improving SNR and detection reliability. As noted by Richard and Gore, SNR is not in itself a significant performance parameter [R2, p. 440]. Of greater importance is the exact density of the test statistic from which the FAR and detection probability can be calculated. This thesis will investigate one particular way of obtaining this important information.



1.7 Thesis Objective

This thesis will examine the effect of zero-memory non-linear filters on the detection problem stated in the introduction. The transfer characteristic of the NLF will be determined by a probability integral transformation. Generally speaking, the transformation will insure that the test statistic has a continuous PDF. The type of noise considered at the receiver input and the monotonicity of the transfer characteristic of the NLF guarantee the optimality of a threshold test. Taken together, they enable satisfaction of any specified false alarm rate without randomization through an easily implemented threshold detector. The continuity of the test statistic PDF will not depend on "large" samples.

1.8 Organization of the Thesis

Chapter II presents a relevant background of statistical decision theory for this communications problem. Chapter III concentrates on the development of the applicable nonlinear filter (NLF) and examines its properties. Chapter IV compares the NLF detector with other alternative approaches, while Chapter V summarizes the thesis and suggests some directions for further research.

CHAPTER II

STATISTICAL BACKGROUND

This chapter summarizes the relevant background in statistical decision theory. The definitions and theorems are in large part taken from the books of Lehmann [L1], Fraser [F1], and Ferguson [F2], but are adapted to directly apply to the detection problems encountered in the context of radio communications. The primary purpose of this chapter is to present a set of conditions which guarantee that the optimal test for location, where optimality is in the sense of UMP, is a threshold test. To that end Section 2.1 develops the Best Critical Region (BCR) for Binary Hypothesis Testing problems. The notion of Monotone Likelihood Ratio (MLR) is used to define the BCR in terms of a decision threshold. Section 2.2 extends threshold detection to UMP tests, and presents a sufficient condition by which a location parameter family can be tested for MLR. Section 2.3 defines some measures of SNR and ARE from parametric and nonparametric theory, respectively, and shows how they are related. Section 2.4 develops the optimal detector of a constant signal in white, Gaussian noise, and Section 2.5 investigates the effect of zero memory filters, linear and nonlinear, on the MLR property. Section 2.6 summarizes the main results of the chapter.

2.1 Simple Hypothesis Testing

The following definitions are intended to make precise the idea of location shifts of a real valued random variable X , whose cumulative distribution function (CDF) depends on a single real parameter Δ . Let $F(x|\Delta)$ denote the distribution function of the random variable X when Δ is the true parameter value; F is assumed to be absolutely continuous with probability density function (PDF) $f(x|\Delta)$.

Definition. The parameter Δ is a location parameter for the distribution of X if, and only if, $f(x|\Delta) = f(x - \Delta)$ for some density $f(x)$.

Definition. A single-parameter family of distributions is said to have monotone likelihood ratio if whenever $\Delta_1 < \Delta_2$ the likelihood ratio $\frac{f(x|\Delta_2)}{f(x|\Delta_1)}$ is a non-decreasing [or non-increasing] function of x .

A non-randomized decision rule for the binary hypothesis testing problem

$$H_0: \Delta = \Delta_0$$

$$H_1: \Delta = \Delta_1 > \Delta_0$$

may be defined by a measurable subset, Ω_0 , of the sample space \mathcal{X} , $\Omega_0 \subset \mathcal{X}$, with the understanding that if the observable variable X falls in Ω_0 , announce H_0 is true; otherwise say H_1 is true. The set $\Omega_1 \triangleq \mathcal{X} - \Omega_0$ is called the critical region. The false alarm rate (FAR) of the test is given by $\int_{\Omega_1} p(X|H_0) dX$

while the probability of detection (P_D) of the test is given

$$\text{by } \int_{\Omega_1} p(X|H_1) dX.$$

An optimum test, in the Neyman-Pearson sense, is defined by a critical region Ω_1 satisfying $\int_{\Omega_1} p(X|H_0) dX \leq P_F$ and $\int_{\Omega_1} p(X|H_1) dX = P_D = \text{maximum}$. This region will be called a Best Critical Region (BCR) and the test defined by a BCR is called a most powerful (MP) test at FAR P_F .

Definition. A region $\Omega_1 \subset \mathcal{X}$ is said to be best at FAR P_F for testing H_0 against H_1 if $\int_{\Omega_1} p(X|H_0) dX = P_F$ and if for every other region Ω'_1 for which $\int_{\Omega'_1} p(X|H_0) dX = P_F$, $\int_{\Omega_1} p(X|H_1) dX \geq \int_{\Omega'_1} p(X|H_1) dX$, i.e. $P_D(\Omega_1) \geq P_D(\Omega'_1)$.

In other words, a region Ω_1 is best at FAR P_F if, out of all regions having the same FAR, its measure under the alternative (P_D) is largest. A general method for finding the best test of a simple hypothesis against a simple alternative can be found in the Neyman-Pearson lemma.

Lemma: In the test of a simple hypothesis H_0 against a simple alternative H_1 , the region $\Omega_1 \triangleq \{X: \Lambda(X) \geq k\}$ where

$$\Lambda(X) \triangleq \frac{f(X|H_1)}{f(X|H_0)} \text{ and } k \text{ satisfies } \int_{\Omega_1} p(X|H_0) dX = P_F \text{ where}$$

$0 < P_F < 1$, is the BCR at the given FAR.

The proof appears in Hogg and Craig [H4] p. 274 and is not repeated here.

The preceding lemma and the definition of MLR allow statement and proof of the next theorem.

Theorem: The most powerful test of a simple hypothesis H_0 against a simple alternative H_1 when the likelihood ratio is monotone (increasing) is a threshold test defined by the critical region $\Omega_1 = \{X: X > K\}$. If K satisfies $\int_K^{\infty} p(X|H_0) dX = P_F$ then Ω_1 is a BCR at FAR P_F .

Proof: By the Neyman-Pearson lemma the region $\Omega_1 \triangleq \{X: \Lambda(X) > k\}$ is best at its FAR where k is determined by $\int_k^{\infty} p(\Lambda|H_0) d\Lambda = P_F$.

If $\Lambda(X)$ is monotone in X then $\forall X \geq K, \Lambda(X) \geq \Lambda(K)$. Choose $K \ni \int_{\{X: X \geq K\}} p(X|H_0) dX = P_F = \int_{\Lambda(X_0)}^{+\infty} p(\Lambda|H_0) d\Lambda = \int_k^{+\infty} p(\Lambda|H_0) d\Lambda$. Then $\Lambda(K) = k$ and the regions $\{X: X \geq K\}, \{X: \Lambda(X) \geq k\}$ are equal. Because the latter region is MP at its FAR, then the threshold test is also MP.

2.2 Composite Hypothesis Testing

The larger detection problem concerns composite hypotheses; i.e., when the distributions under H_0 and H_1 depend on a parameter which is assumed to take on values in some (real) space R .

Let $\omega \subset R$ be the subset of parameter values such that

$\{P_{\Delta} | \Delta \in R\}$ denotes the family of distributions for the random

variable X when H_0 is true. Then $R - \omega$ corresponds to the possible parameter values when H_1 is true. The hypothesis

testing problem can be restated as $H_0: \Delta \in \omega$

$H_1: \Delta \in R - \omega$.

As before, the problem is to dichotomize the observation space \mathcal{X} so as to satisfy some criterion of "goodness". For this situation the notion of best test at FAR P_F must be redefined.

Definition. In the test of a composite hypothesis H_0 against an alternative (simple or composite) the region Ω_1 is said to be at FAR P_F if $\sup_{\Delta \in \omega} \int_{\Omega_1} p(X|\Delta) dX = P_F$.

When the alternative H_1 is also composite the notion of best test must also be generalized.

Definition. A region Ω_1 is said to be Uniformly Most Powerful (UMP) at FAR P_F for testing $H_0: \Delta \in \omega$ against $H_1: \Delta \in R - \omega$ if Ω_1 has FAR P_F and if, for any other region Ω'_1 at FAR P_F , it is true that $\int_{\Omega_1} p(X|\Delta) dX \geq \int_{\Omega'_1} p(X|\Delta) dX$ for each $\Delta \in R - \omega$.

The Neyman-Pearson lemma guarantees that, in the class of regions having FAR P_F , there exists one whose P_D is maximum at any fixed element $\Delta_1 \in R - \omega$. There is no reason why this region should also maximize the P_D for some other element $\Delta_2 \in R - \omega$. It is, therefore, not surprising that UMP tests exist only in special circumstances. Fortunately, one of these circumstances is widely found in radio communications problems and will be investigated in this thesis.

Theorem: If the distribution of X has non-decreasing monotone likelihood ratio, then the region $\Omega_1 = \{X: X \geq K\}$ where K is chosen to satisfy $\int_{\Omega_1} p(X|H_0) dX = P_F$, $0 < P_F < 1$, is UMP of FAR P_F for the hypothesis testing problem $H_0: \Delta \leq \Delta_0$ against $H_1: \Delta > \Delta_0$.



Proof: See Ferguson [F2], page 210.

This theorem and its counterpart for binary hypothesis testing illustrate the effect of the MLR property. Tests for location simplify to threshold tests when the underlying $f(x|\Delta) = f(x - \Delta)$ have MLR. Because of the importance of the MLR property in this thesis, the sufficient conditions under which a density may have a MLR with respect to a location parameter are now investigated.

Definition. A density function $f(X)$ is called strongly unimodal if $-\ln f(X)$ is a convex function of X .

Lehmann [L1, p. 330] shows that strong unimodality is a sufficient condition for a density to have MLR.

Theorem: Let $f(x)$ be a density on R . A sufficient condition that $f(x|\Delta) = f(x - \Delta)$ have MLR is that $f(x)$ be strongly unimodal.

Proof: Let $f(X)$ be strongly unimodal and suppose $X < X'$,

$\Delta < \Delta'$. The condition of MLR in X , namely, $\frac{f(X - \Delta')}{f(X - \Delta)} \leq \frac{f(X' - \Delta')}{f(X' - \Delta)}$, is equivalent to

$$\log f(X' - \Delta) + \log f(X - \Delta') \leq \log f(X - \Delta) + \log f(X' - \Delta') .$$

The four differences can be arranged, from smallest to largest, as $(X - \Delta')$, $(X - \Delta)$, $(X' - \Delta')$, $(X' - \Delta)$ or $(X - \Delta')$, $(X' - \Delta')$, $(X - \Delta)$, $(X' - \Delta)$. In either case by defining

$t \triangleq (X' - X)/(X' - X + \Delta' - \Delta)$ it is true that

$$X - \Delta = t(X - \Delta') + (1 - t)(X' - \Delta)$$

$$X' - \Delta' = (1 - t)(X - \Delta') + t(X' - \Delta) .$$

Define $h(X) \triangleq -\log f(X)$; by hypothesis, $h(X)$ is convex. Then for every t_1, t_2 and $0 \leq \lambda \leq 1$

$$h[\lambda t_1 + (1 - \lambda)t_2] \leq \lambda h(t_1) + (1 - \lambda)h(t_2)$$

$$\log f[\lambda t_1 + (1 - \lambda)t_2] \geq \lambda \log f(t_1) + (1 - \lambda)\log f(t_2) .$$

Letting $\lambda = t$, $t_1 = X - \Delta'$, $t_2 = X' - \Delta$,

$$\log f(X - \Delta) \geq t \log f(X - \Delta') + (1 - t) \log f(X' - \Delta) .$$

Letting $\lambda = 1 - t$,

$$\log f(X' - \Delta') \geq (1 - t) \log f(X - \Delta') + t \log f(X' - \Delta) .$$

Adding these two equations yields

$$\log f(X - \Delta) + \log f(X' - \Delta') \geq \log f(X - \Delta') + \log f(X' - \Delta),$$

which is the condition for $f(X)$ to have MLR.

Q.E.D.'

This theorem provides a method of ascertaining whether the family of densities

$$f(x|\Delta) = a \exp(-b|x - \Delta|^\theta), \quad \theta \text{ fixed,}$$

has a MLR. This is seen to be the case since $-\ln f(x) = -a + b|x|^\theta$ is convex in x for each fixed $\theta \in [1, 2]$.

2.3 Signal-to-Noise Ratios (SNR)

In Section 1.5, the measure of signal-to-noise ratio (SNR) applicable to this thesis was given as the rms signal power divided by the rms noise power. Because the detector of this thesis is to be compared to others in Chapter IV, this section introduces performance measures which have previously been defined for parametric and nonparametric detection. The general Gaussian problem is the hypothesis testing problem in which the densities under both classes are Gaussian, $i = 0, 1$. When the covariance matrices under the hypotheses are equal but the means are different, the quantity

$$d^2 \triangleq \frac{[\mathcal{S}(\ell|H_1) - \mathcal{S}(\ell|H_0)]^2}{[\text{Var}(\ell|H_1)]}$$

with ℓ the test statistic, has been used as a SNR [V1, p. 99]. This same quantity was specified by Hancock and Lainiotis [H3] for the problem of detecting an unknown d.c. signal having magnitude $\theta > 0$ in zero-mean noise with a symmetric, but otherwise largely unspecified, PDF.

The physical justification for this concept of SNR rests on the idea that the two densities of the received random variable, conditional on signal present and signal absent, become more distinguishable as either the distance between their central locations (means) becomes greater or the concentration of their values around the central location becomes greater (smaller variances).

A theorem of Capon in 1961 [C2] introduced the concept of efficacy in determining the asymptotic relative efficiency (ARE) of a test U_n relative to the test U_n^* , where both tests satisfied certain regularity conditions. The ARE of the U_n -detector with respect to the U_n^* -detector was given by

$$E_{u, u^*} \triangleq \frac{e(U_n)}{e(U_n^*)}, \text{ where } e(U_n),$$

the efficacy of the U_n -detector, was defined as

$$e(U_n) \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \left[\frac{\frac{\partial}{\partial \theta} E_{\theta}(U_n)}{\sigma_0(U_n)} \right]_{\theta=0}^2.$$

The sample size of the U_n -detector is n , and $E_{\theta}(U_n)$ is the expected value of the detector statistic under the alternative hypothesis H_1 , when the signal strength is θ . The following shows that efficacy depends directly on the previous measure of SNR and the relationship is derived using the regularity conditions of Capon.

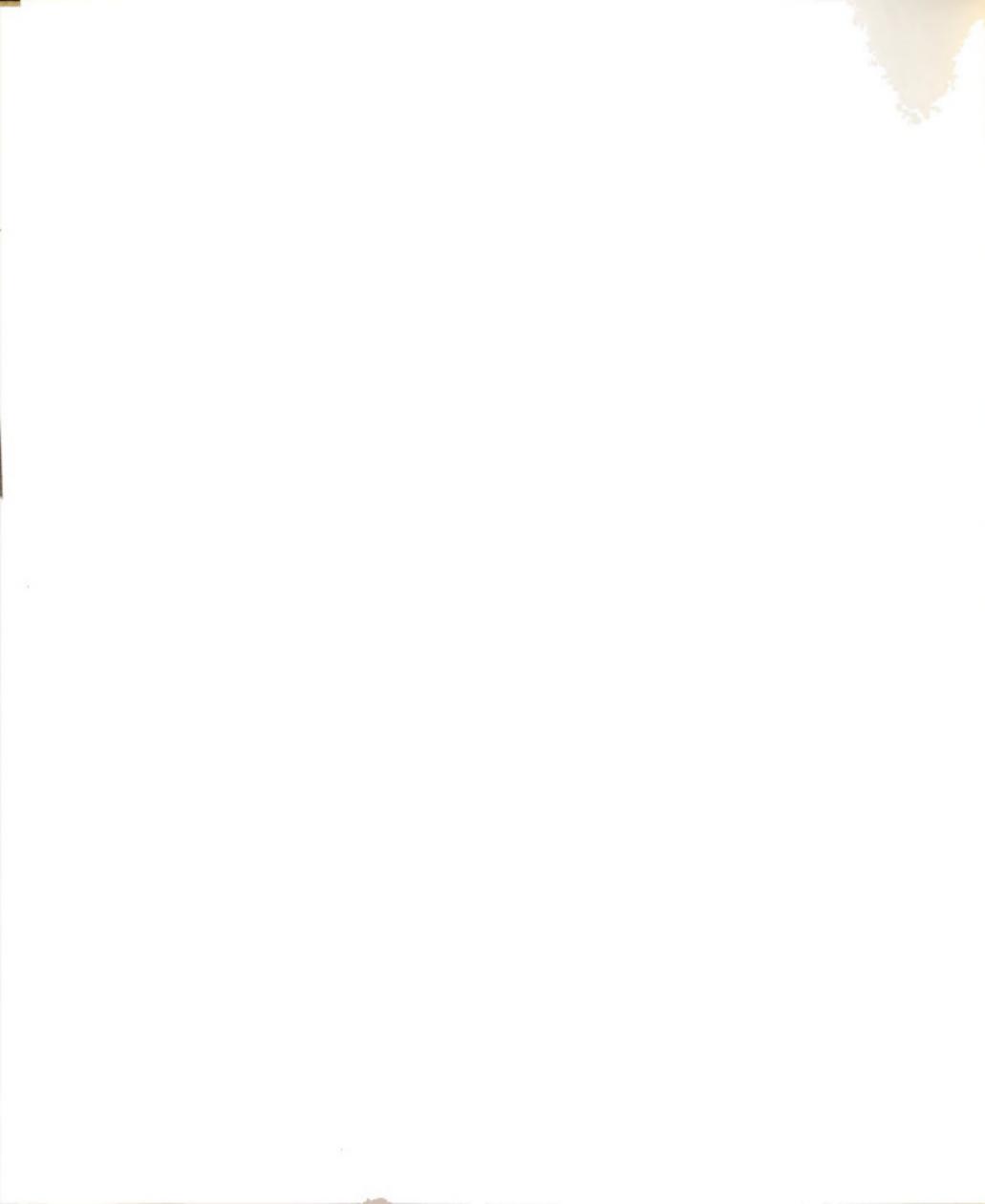
Given that $\frac{\partial}{\partial \theta} E_{\theta}(U_n)$ exists for all $\theta \in (0, a)$, and is continuous at $\theta = 0$, a Taylor Series expansion of $E_{\theta}(U_n)$ about the origin yields $E_{\theta}(U_n) - E_0(U_n) = \theta_n \frac{\partial}{\partial \theta} E_{\theta}(U_n) \Big|_{\theta=0} + o(\theta_n^2)$ as $\theta_n \rightarrow 0$ [H5, p. 54]. Then

$$e(U_n) \doteq \lim_{n \rightarrow \infty} \left[\frac{E_{\theta}(U_n) - E_0(U_n)}{\sqrt{n} \theta_n \sigma_0(U_n)} \right]^2. \text{ The relationship between}$$

efficacy and SNR is obvious.

2.4 Optimal Detector of Constant Signal in White, Gaussian, Noise

This section develops the optimal detector of a constant signal in white, Gaussian, noise. According to well-known results



on sufficiency, the optimal test can be based on the sufficient statistic $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$. The distribution of \bar{Y} under H_0 is $\eta(0, \frac{\sigma^2}{N})$ and under H_1 , \bar{Y} is $\eta(\Delta, \frac{\sigma^2}{N})$. The normal family has increasing MLR with respect to location (variance fixed), and therefore the best test having a False Alarm Rate of P_F is

$$\begin{array}{c} H_1 \\ \bar{Y} > K \\ H_0 \end{array}$$

where K is defined to satisfy

$$P_F = \int_K^{+\infty} \left(\frac{2\pi\sigma^2}{N}\right)^{-\frac{1}{2}} e^{-\frac{N}{2}\left(\frac{y}{\sigma}\right)^2} dy .$$

Thus $K = \frac{\sigma}{\sqrt{N}} \operatorname{erfc}_*^{-1} P_F$, where the complementary error function is

defined by $\operatorname{erfc}_* z \triangleq \int_z^{+\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du$. Hence, the BCR is defined as

$C \triangleq \{\bar{Y}: \bar{Y} > \frac{\sigma}{\sqrt{N}} \operatorname{erfc}_*^{-1} P_F\}$. The probability of detection for this critical region is defined as

$$P_D \triangleq \int_K^{+\infty} \left(\frac{2\pi\sigma^2}{N}\right)^{-\frac{1}{2}} e^{-\frac{N}{2}\left(\frac{y-\Delta}{\sigma}\right)^2} dy .$$

$$P_D = \operatorname{erfc}_* \left[\operatorname{erfc}_*^{-1} P_F - \frac{\sqrt{N} \Delta}{\sigma} \right] .$$

This equation relates the probability of detection directly to False Alarm Rate, and is the basis of the receiver operating characteristic (ROC) [V1, p. 38].

2.5 Measure Transformations and MLR

The purpose of this section is to investigate the effect that a zero memory filter has on its input with respect to the

desirable property of MLR. In Section 2.5.1 measure transformations are defined while in Section 2.5.2 the MLR property is shown to be preserved under monotone measure transformation.

2.5.1 Measure Transformations

Let $f_X(X)$ be the density function for a continuous random variable X , and suppose a real-valued point function $g(\cdot)$ is specified. The random variable $Y = g(X)$ is defined and its density function $f_Y(y)$ is to be determined. Assuming that $g(\cdot)$ is monotone increasing wherever $f_X(X) \neq 0$ the density of Y is given by

$$f_Y(v) = f_X(g^{-1}(v)) \left| \frac{d}{dv} g^{-1}(v) \right|. \quad (1)$$

Proof: See Dubes [D2], p. 243.

2.5.2 Monotone Measure Transformations

If X has PDF $f_X(x|\Delta)$ which has increasing MLR, and $Y = g(X)$ where $g(\cdot)$ is monotone increasing and satisfies the regularity conditions which guarantee (1), then the family of PDF's for Y ,

$$f_Y(y|\Delta) \stackrel{\Delta}{=} f_X(g^{-1}(y)|\Delta) \left| \frac{d}{dy} g^{-1}(y) \right|,$$

has increasing MLR. This is immediate since if $\Delta_1 > \Delta_2$,

$$\frac{f_Y(y|\Delta_1)}{f_Y(y|\Delta_2)} = \frac{f_X(g^{-1}(y)|\Delta_1)}{f_X(g^{-1}(y)|\Delta_2)},$$

$\frac{f_X(x|\Delta_1)}{f_X(x|\Delta_2)}$ is monotone increasing, and $g^{-1}(\cdot)$ is monotone



increasing. The result is now illustrated with examples of linear and non-linear monotone transformations on location parameter families.

Example 1. . Let X under H_0 be uniform on $[-1,1]$. Under H_1 , X will be uniform on $[-1 + \alpha, 1 + \alpha]$, $\alpha > 0$, and suppose the linear filter has the transfer characteristic $aX + b$, where both a and b are known positive constants [for the sake of simplicity]. Thus $f(Y|H_1) = \frac{1}{2a}$ for $a(\alpha-1) + b \leq y \leq a(\alpha+1) + b$, $f(Y|H_0) = \frac{1}{2a}$ for $b-a \leq y \leq b+a$. These densities are shown in Figure 2.1 and it is easily verified that Y has the MLR property.

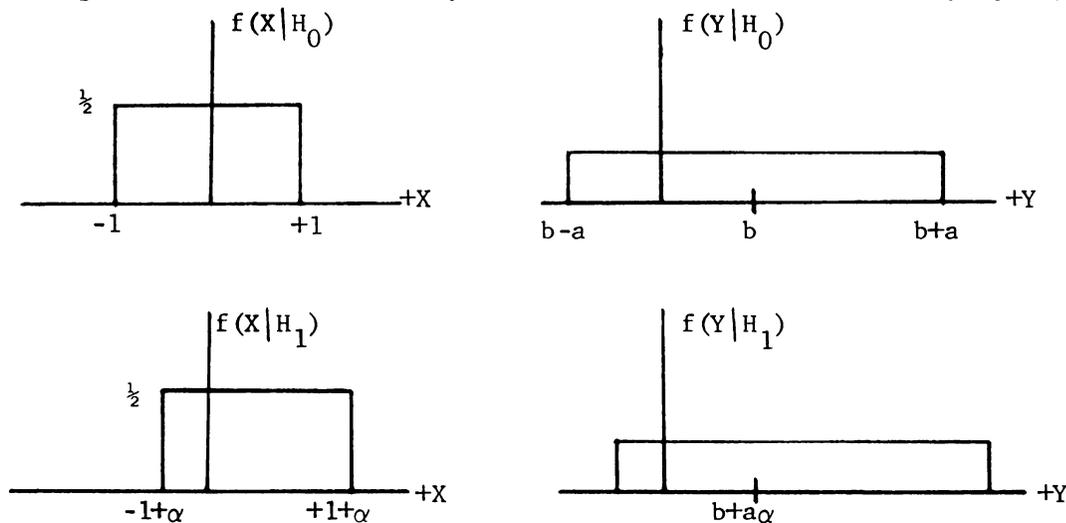


Figure 2.1 Invariance of MLR to Linear Transformation

Example 2. Let the input random variable X be uniform on $(-\frac{1}{2}, \frac{1}{2})$ under H_0 and uniform $(\Delta, \Delta+1)$ under H_1 with $-\frac{1}{2} < \Delta \leq 0$. Let $g(X)$ be defined as $g(X) = \sin X$ for $-\frac{\pi}{2} \leq X \leq +\frac{\pi}{2}$, $+1$ for $X > \frac{\pi}{2}$, -1 for $X < -\frac{\pi}{2}$. The situation is illustrated in Figure 2.2.

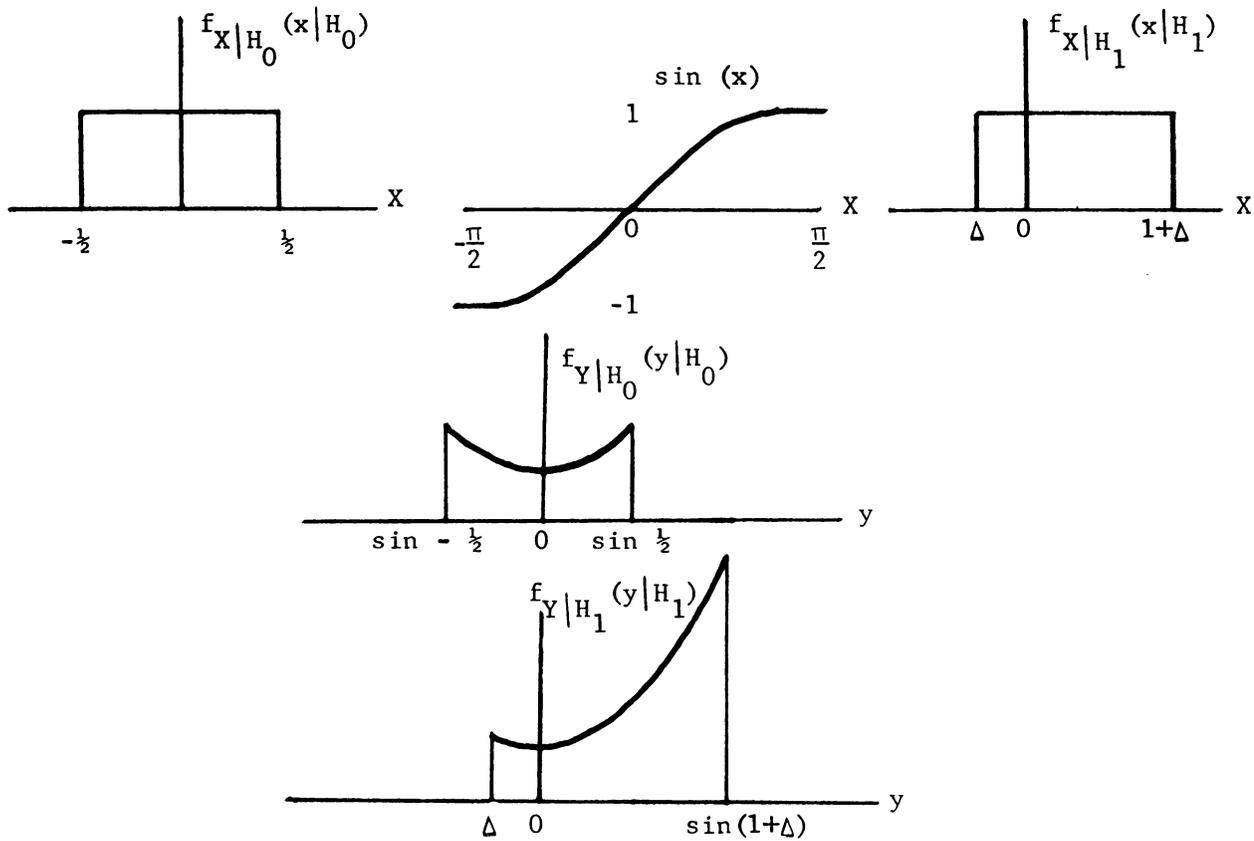


Figure 2.2 Invariance of MLR to Nonlinear Transformation

2.5.3 A Transformation to Achieve a Specified Distribution

Let X be a random variable having CDF F . Subject to certain regularity conditions [H6, p. 33], the transformation

$$G^{-1}(F(X))$$

will yield a random variable Y , having CDF G .

2.6 Summary

This chapter presented some concepts from statistical decision theory which are applicable to the communications problem introduced in Chapter I. The property of Monotone Likelihood Ratio was shown to be a sufficient condition for a test of location (known direction) to be implemented by a threshold test. The equivalence of the Efficacy of Capon and the SNR of Van Trees was shown and Monotone Likelihood Ratio was shown to be an invariant property when a family having it was passed through a nonlinear filter having an odd-symmetric, monotone increasing transfer characteristic. A sufficient condition for MLR was derived and applied to show that the family of noise densities introduced in Chapter I has the MLR property.



CHAPTER III

ZERO-MEMORY NONLINEAR FILTER

This chapter discusses the use of a nonlinear filter between the sampler and detector of Figure 1.1. Section 3.1, which is concerned with integrators and clippers, provides motivation for investigating the measure transformation of Section 3.2. This section outlines the procedure which "Gaussianizes" random variables; that is, the transfer characteristic of a zero-memory nonlinear filter (NLF) is determined through a probability integral transformation, which produces independent Gaussian random variables at the NLF output, when the filter is driven by a finite sequence of independent random variables having a Laplacian probability density function. Section 3.3 then discusses a special property of the NLF transfer characteristic while Section 3.4 derives the exact distribution of the NLF output when the input is noise plus signal. Section 3.5 develops first and second moment approximations for the NLF output under H_1 , while Section 3.6 summarizes the main results of the chapter.



3.1 Clippers, Limiters and Integrators

In many practical communications receivers, the input process is first filtered to remove any noise spectral components which may lay outside of the signal bandwidth. Cascaded linear, tuned networks, whose overall bandwidth is wide enough to pass with considerable amplification the significant signal components while rejecting those frequencies containing noise alone, are widely used. If the input to a filter is a quasi-monochromatic¹ signal, its output at any time is a weighted time-average of its past input. When the filter bandwidth is much narrower than the noise bandwidth at its input, the central limit theorem can be invoked to justify assuming that the amplitude distribution of the filtered output is Gaussian.

In other receiver types the input process is first sampled, thereby reducing the continuous-time input to a finite-length time series. Many applications of nonparametric statistical methods have been prompted by the fact that the sampled data process often has a large variance and consequently may not satisfy the conditions of a central limit theorem [R1]. Therefore, so-called "optimum" tests based on the Gaussian assumption may be highly unreliable. The inapplicability of Gaussian methods is also prevalent in underwater sound receivers which pick up sporadic biological noise and telecommunications networks which are subject to randomly occurring impulses due to lightning, sparks in ignition systems and faulty connections or switches [H1, p. 317].

¹ Narrow Band



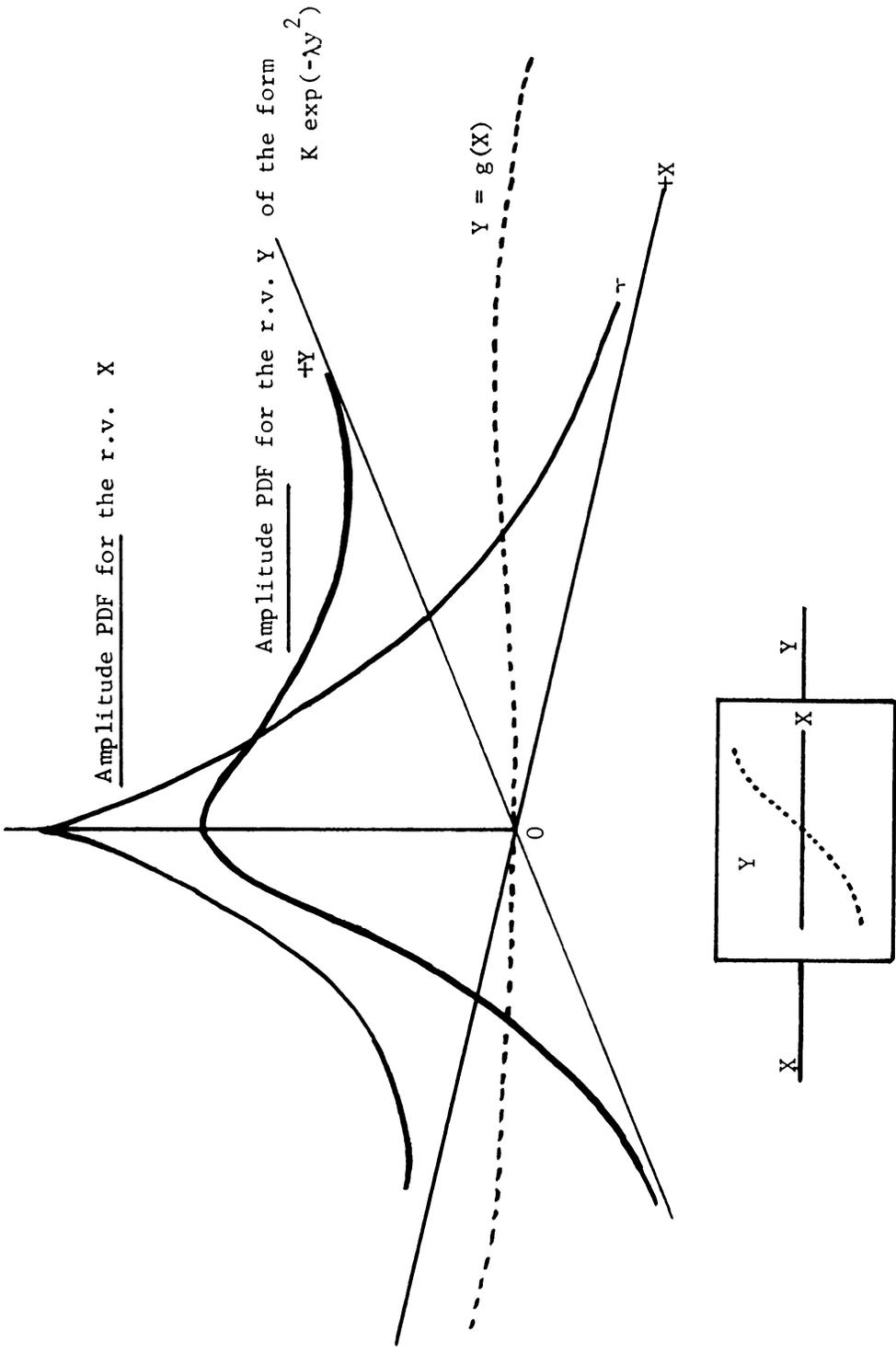
A large variance for the input random variable may preclude application of a central limit theorem without some form of noise suppression. One method is the clipping, or limiting, of the received signal, the usual justification being that this increases the SNR. Although the introduction of a clipper before processing the received signal seems intuitively a good idea, it is not at all clear that a simple clipping level is in some sense the best system design that can be achieved. Moreover, when severe models of impulsive noise amplitudes are considered, commonly defined SNR's may not be useful criteria of system performance. The structure of the nonlinear device for suppression should be determined by whatever knowledge is available about the noise PDF so that a minimal amount of received information pertinent to the signal decision is destroyed [R1], [R2].

3.2 Measure Transformations

The essential problem is sketched in Figure 3.1 and can be stated as follows: Let X be a random variable having a PDF which is known completely and which is absolutely continuous with respect to the Gaussian PDF. Determine the transfer characteristic of a device such that with the r.v. X applied at the input, the output random variable, Y , has a Gaussian PDF.

This problem is different from the more common one involved in functions of random variables. There, the transfer characteristic of a device is specified and the input is a r.v. having a known PDF. The problem is to then determine the distribution of the output. This was discussed in Section 2.5.1. The results of





Nonlinear Zero-memory Filter

Figure 3.1 "Gaussianizing" Transformation



Section 1.6 indicate that the PDF of interest in this thesis is Laplacian.

Problem Statement: Find a transfer characteristic $g(x)$ which transforms the random variable X having the Laplacian PDF

$f_X(x) = a \exp(-b|x|)$, $-\infty < x < +\infty$, with a and b defined by

$$\int_{-\infty}^{\infty} f_X(x) dx = 1, \int_{-\infty}^{\infty} x^2 f_X(x) dx \triangleq \sigma_X^2 = \eta_B \text{ to the Gaussian random}$$

variable Y , which has 0 mean and variance σ_0^2 .

Problem Solution: Constrain $g(x)$ to be a monotone, strictly increasing function on the real line having odd symmetry and require $Y = g(X)$ a.e. Then

$$P[Y \leq y] = P[X \leq x], \text{ with } x = g^{-1}(y). \quad (1)$$

Case #1: $x < 0$

$$\int_{-\infty}^{y/\sigma_0} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}} du = \int_{-\infty}^x a e^{+b\omega} d\omega \quad (2)$$

$$\text{or, } \Phi\left[\frac{y}{\sigma_0}\right] = \frac{a}{b} e^{bx} \quad \begin{array}{l} x < 0 \\ y < 0 \end{array} \quad (3)$$

Case #2: $x > 0$

$$\int_{-\infty}^{y/\sigma_0} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}} du = \frac{1}{2} + \int_0^x a e^{-b\omega} d\omega \quad (4)$$

$$\text{or, } \Phi\left[\frac{y}{\sigma_0}\right] = \left(\frac{1}{2} + \frac{a}{b}\right) - \frac{a}{b} e^{-bx} \quad (5)$$

Using the first condition on X , $\int_{-\infty}^{\infty} f_X(x) dx = 1$, the ratio $\frac{a}{b}$ is determined.



$$\int_{-\infty}^{\infty} a e^{-b|x|} dx = 1 \quad (6)$$

$$\int_0^{+\infty} a e^{-bx} dx = \frac{1}{2} \quad (7)$$

$$-\frac{a}{b} e^{-bx} \Big|_0^{+\infty} = \frac{1}{2} \Rightarrow \frac{a}{b} = \frac{1}{2} \quad (8)$$

Combining (3), (5) and (8) yields

$$\Phi \left[\frac{y}{\sigma_0} \right] = \frac{1}{2} e^{bx} \quad x < 0 \quad (9)$$

$$= 1 - \frac{1}{2} e^{-bx} \quad x > 0 \quad (10)$$

Solving for y in (9) and (10) gives the form of $g(x)$ required:

$$y = \sigma_0 \Phi^{-1} \left[\frac{e^{bx}}{2} \right] \quad x < 0 \quad (11)$$

$$y = \sigma_0 \Phi^{-1} \left[1 - \frac{e^{-bx}}{2} \right] \quad x > 0 \quad (12)$$

Figure 3.2 shows the function $g(x)$ of (12).

Using the second condition on the PDF $f_X(x)$ determines the value of the coefficient b .

$$\int_{-\infty}^{\infty} x^2 f_X(x) dx \stackrel{\Delta}{=} \sigma_X^2 \Rightarrow 2 \int_0^{\infty} a x^2 e^{-bx} dx = \sigma_X^2 \quad (13)$$

Equation (13) is easily put in the form

$$\int_0^{\infty} u^2 e^{-u} du = \frac{b^3 \sigma_X^2}{2a}$$

where the left-hand-side is immediately recognized as the Gamma



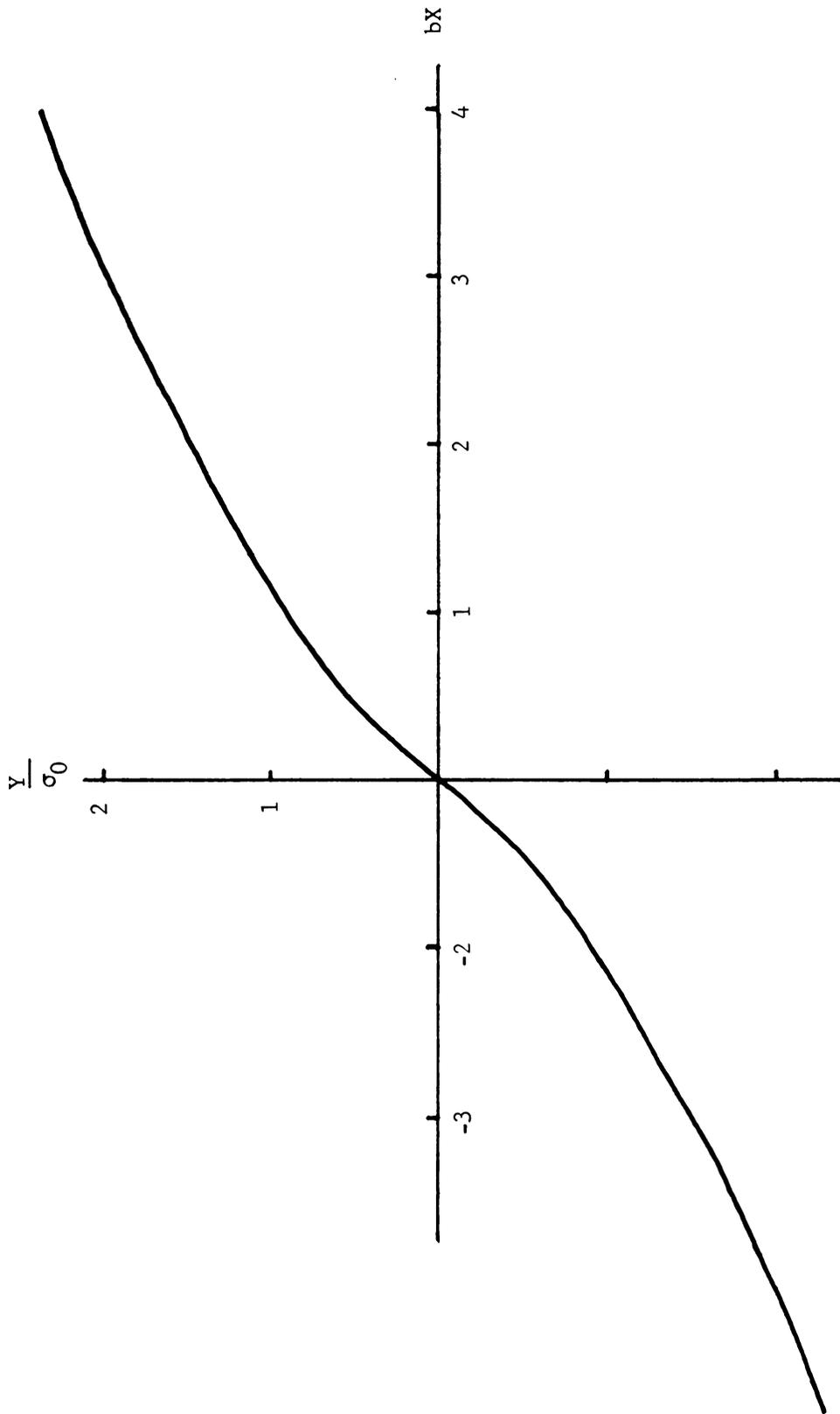


Figure 3.2 Laplacian Transfer Characteristic

integral; therefore

$$\frac{b^3 \sigma_X^2}{2a} = \Gamma(3) = 2! = 2. \quad (14)$$

Finally, $\frac{b^3 \sigma_X^2}{2a} = 1 \Rightarrow \left(\frac{b}{a}\right) b^2 \sigma_X^2 = 4$. Using (8) in (14) yields

$$b^2 = \frac{2}{\sigma_X^2}, \quad \text{or}$$

$$b = \frac{\sqrt{2}}{\sigma_X}. \quad (15)$$

The given PDF can more easily be written in the form

$$f_X(x) = \frac{b}{2} \exp(-b|x|), \quad -\infty < x < +\infty. \quad (16)$$

3.3 Concavity

This section will prove that the transfer characteristic determined in the previous section is convex on the region $x > 0$.

Given

$$g(x) = \sigma_0 \Phi^{-1}\left(1 - \frac{e^{-bx}}{2}\right), \quad x > 0 \quad (17)$$

write it as

$$\frac{e^{-bx}}{2} = \int_{g/\sigma_0}^{+\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du. \quad (18)$$

Take derivatives on both sides with respect to x twice

$$\frac{-b^2 e^{-bx}}{2} = \frac{1}{\sigma_0} \left[g'' - (g')^2 \frac{g}{\sigma_0} \right] \frac{e^{-\frac{1}{2}\left(\frac{g}{\sigma_0}\right)^2}}{\sqrt{2\pi}} \quad (19)$$

$$g'' = \frac{(g')^2}{\sigma_0} [gg' - b\sigma_0^2]. \quad (20)$$

Thus to show that g is concave on $x > 0$ it is sufficient to show that $gg' - b\sigma_0^2 < 0$, for $x > 0$.

$$g' = \sigma_0^2 b e^{\frac{1}{2} \left(\frac{g}{\sigma_0}\right)^2} \int_{g/\sigma_0}^{+\infty} e^{-\frac{u^2}{2}} du \quad (21)$$

Integrating (21) by parts yields

$$g' = \frac{\sigma_0^2 b}{g} - \sigma_0^2 b e^{\frac{1}{2} \left(\frac{g}{\sigma_0}\right)^2} \int_{(g/\sigma_0)}^{+\infty} \frac{e^{-\frac{1}{2} u^2}}{u^2} du \quad (22)$$

Hence

$$g'g - \sigma_0^2 b = -\sigma_0^2 b g e^{\frac{1}{2} \left(\frac{g}{\sigma_0}\right)^2} \int_{(g/\sigma_0)}^{+\infty} \frac{e^{-\frac{1}{2} u^2}}{u^2} du \quad (23)$$

All terms on the right hand side are positive for $x > 0$, therefore $gg' - \sigma_0^2 b < 0$ and $g''(x) < 0$ for $x > 0$.

3.4 NLF Output PDF under H_1

This section assumes the NLF has the transfer characteristic specified in (11) and (12) and that the NLF input has the Laplacian amplitude PDF of (16) translated by the amount $\Delta > 0$. The PDF of the NLF output will be derived using the methods of Section 2.5.1. This will show the dependence of the form of the output PDF on the input SNR, and will be used as justification for certain approximations made in Chapter IV.

Write

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| \quad (24)$$



where

$$\begin{aligned} y &= \sigma_0 \Phi^{-1}\left[1 - \frac{e^{-bx}}{2}\right] & x > 0 \\ &= \sigma_0 \Phi^{-1}\left[\frac{e^{bx}}{2}\right] & x < 0 \end{aligned} \quad (25)$$

and

$$f_X(x) = \frac{b}{2} e^{-b|x-\Delta|}, \quad -\infty < x < +\infty. \quad (26)$$

It will be shown that the case for $\Delta = 0$, i.e., x is distributed according to H_0 , can be considered as a special case of this computation, thus demonstrating the "Gaussianizing" property of the NLF.

Case 1: $y < 0 \Rightarrow x < 0$

$$y = g(x) = \sigma_0 \Phi^{-1}\left(\frac{e^{bx}}{2}\right). \quad (27)$$

Then
$$g^{-1}(y) = \frac{1}{b} \ln\left(2\Phi\left(\frac{y}{\sigma_0}\right)\right) \quad (28)$$

$$\frac{dg^{-1}(y)}{dy} = \frac{1}{b} \frac{\frac{2}{\sigma_0} \Phi'(y/\sigma_0)}{2\Phi(y/\sigma_0)} \quad (29)$$

$$f_Y(y) = \frac{b}{2} e^{b\left[\frac{1}{b} \ln 2\Phi(y/\sigma_0) - \Delta\right]} \frac{1}{b\sigma_0} \frac{2\Phi'(y/\sigma_0)}{2\Phi(y/\sigma_0)}$$

$$f_Y(y) = \frac{e^{-b\Delta}}{\sigma_0} \Phi'(y/\sigma_0) = e^{-b\Delta} \frac{e^{-\frac{1}{2}(y/\sigma_0)^2}}{\sqrt{2\pi} \sigma_0} \quad y < 0 \quad (30)$$

Case 2: $y > 0 \Rightarrow x > 0$

$$f_Y(y) = \frac{b}{2} e^{-b|g^{-1}(y) - \Delta|} \left| \frac{dg^{-1}(y)}{dy} \right| \quad (31)$$



Combining the results (30), (39), and (41) gives

$$f_Y(y) = \begin{cases} e^{-b\Delta} \frac{e^{-\frac{1}{2}(y/\sigma_0)^2}}{\sqrt{2\pi} \sigma_0} & y < 0 \\ e^{b\Delta} \cdot \frac{e^{-\frac{1}{2}(y/\sigma_0)^2}}{\sqrt{2\pi} \sigma_0} & y > g(\Delta) \\ \frac{e^{-b\Delta}}{4[1-\Phi(y/\sigma_0)]^2} \cdot \frac{e^{-\frac{1}{2}(y/\sigma_0)^2}}{\sqrt{2\pi} \sigma_0} & 0 < y < g(\Delta) \end{cases} \quad (42)$$

When $\Delta = 0$, $g(\Delta) = 0$ and (42) reduces to a zero-mean Gaussian PDF with variance σ_0^2 . The graph of $f_Y(y)$ is shown in Figures 3.3 (a), (b), (c), and (d), for increasing values of SNR. The assumption of weak signal normality of Y appears well justified when $b\Delta \leq \frac{1}{10}$, approximately -29 db.

3.5 Moment Approximations

Under the alternative (H_1) the PDF for the random variable X , the NLF input, is specified exactly. It is a replica of the PDF of X under H_0 , translated by $\Delta > 0$. Assumptions A1 and A6¹ guarantee that under H_1 , the random variable X has expectation $\Delta > 0$, and a known variance, σ_X^2 . The problem of this section is to find an approximation for the mean and variance of Y , the NLF output random variable, when X contains signal plus noise.

Although the exact form of the PDF of Y is given in the previous section, the complexity of determining the moments

¹ Chapter I



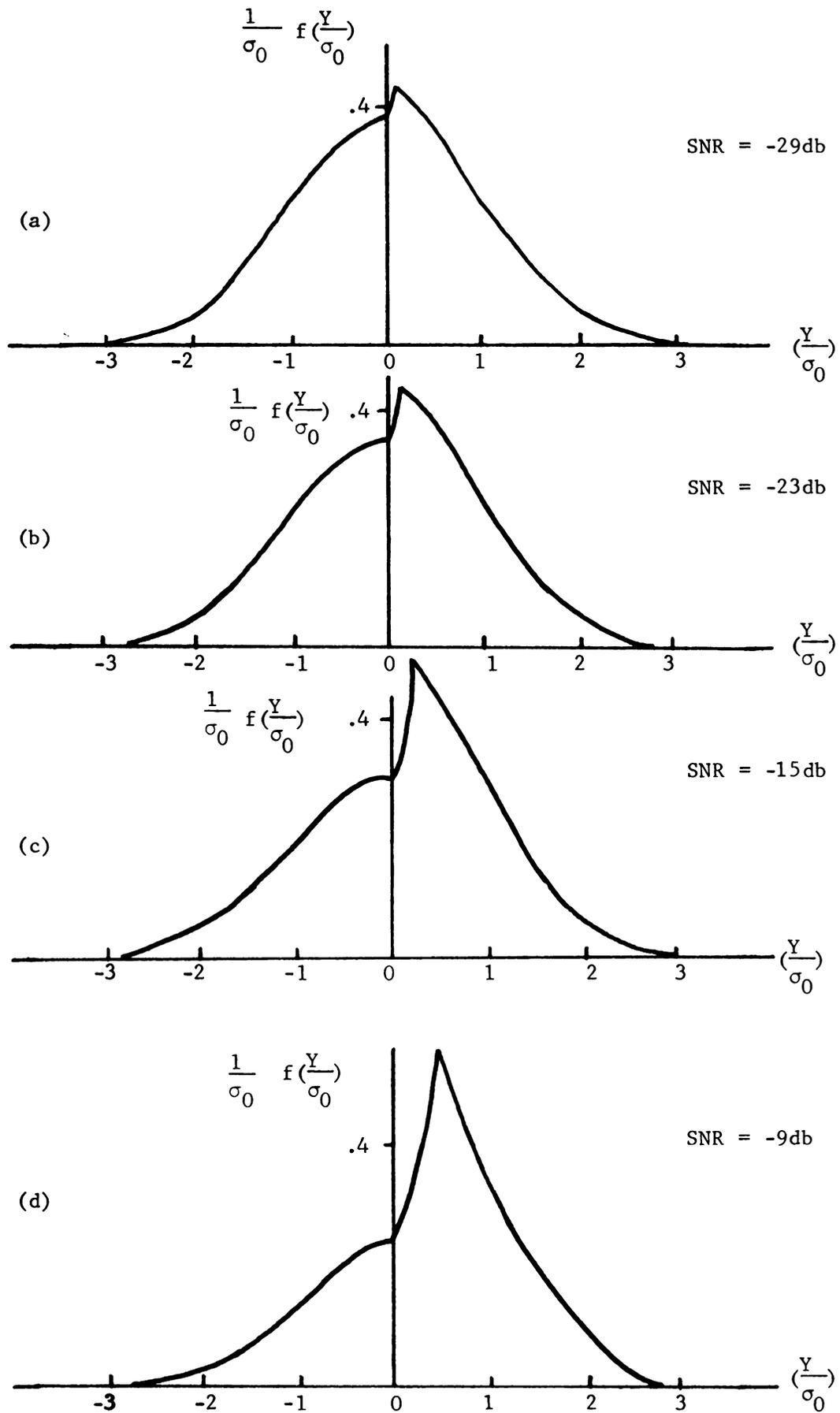
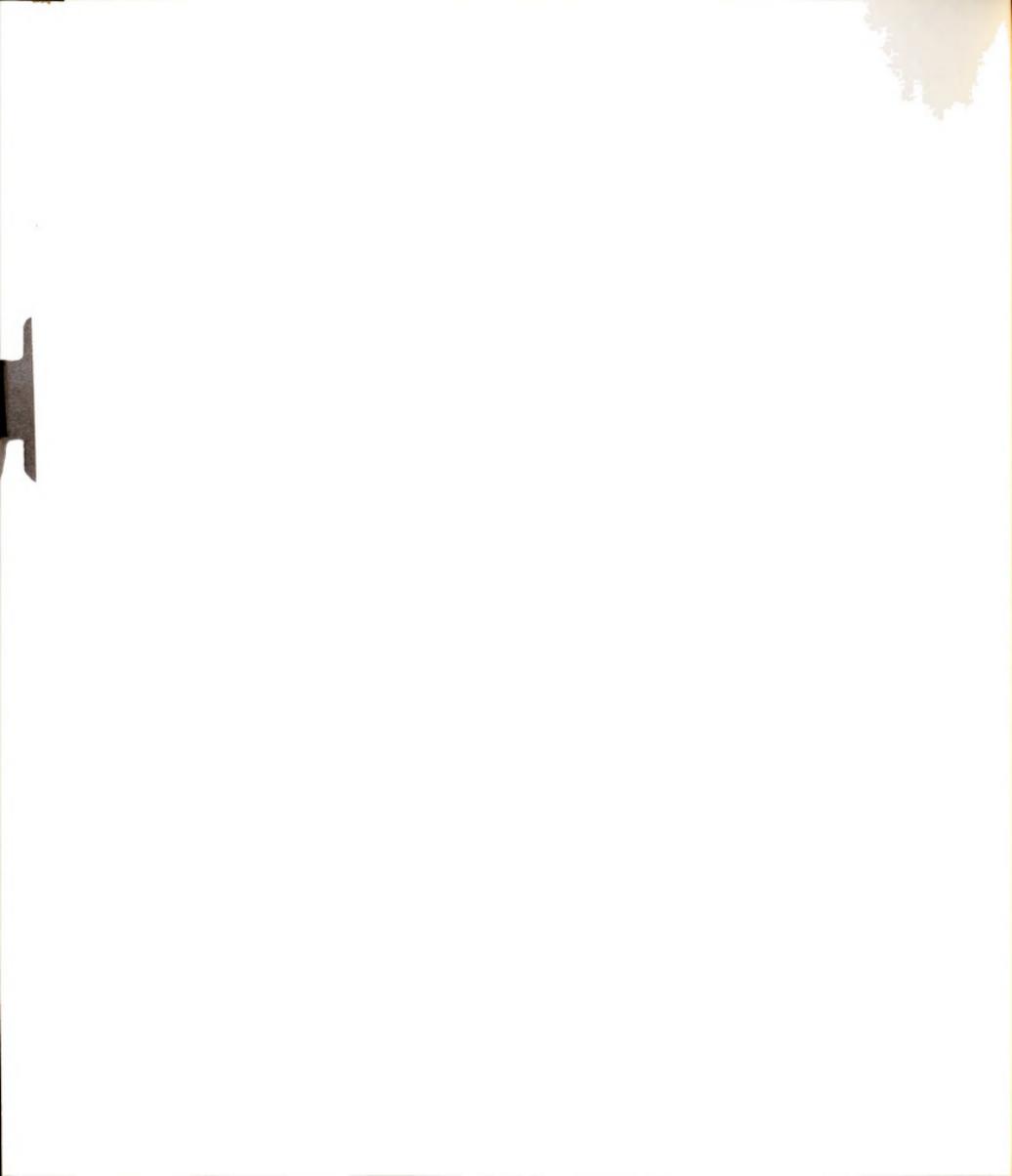


Figure 3.3 PDF of NLF-Output Under H_1



encourages the introduction of simplifying approximations. The expected value and variance of Y , the NLF output, are derived below.

Let \bar{Y}_i , $i = 0, 1$ indicate the expected value of the NLF output random variable when hypothesis H_i is true.

$$\bar{Y}_1 = \int_{-\infty}^{\infty} y f_Y(y|H_1) dy = \int_{-\infty}^{\infty} g(x) f(x|H_1) dx \quad (43)$$

Expanding $g(x)$ in a Taylor series about Δ yields (subject to a few regularity conditions), for some $\xi \in (0, \Delta)$,

$$g(x) = g(\Delta) + \sum_{j=1}^{r-1} \frac{(x-\Delta)^j}{j!} g^{(j)}(\Delta) + \frac{(x-\Delta)^r}{r!} g^{(r)}(\xi) \quad (44)$$

Substituting (44) into (43) and interchanging summation and integration operations,

$$\bar{Y}_1 = \int_{-\infty}^{\infty} \left[g(\Delta) + \sum_{j=1}^{r-1} \frac{(x-\Delta)^j}{j!} g^{(j)}(\Delta) + \frac{(x-\Delta)^r}{r!} g^{(r)}(\xi) \right] \cdot f(x|H_1) dx \quad (45)$$

$$\bar{Y}_1 = g(\Delta) + \frac{g''(\Delta)\sigma_X^2}{2} + \frac{g^{(iv)}(\Delta)}{4!} \mu_{X,4} + \dots \quad (46)$$

Assuming $g(\cdot)$ is sufficiently smooth to allow consideration of only the first two terms, gives

$$\bar{Y}_1 \doteq g(\Delta) + \frac{g''(\Delta)\sigma_X^2}{2} \quad (47)$$

The second central moment is:

$$\bar{Y}_1^2 = \int_{-\infty}^{\infty} g^2(x) f(x|H_1) dx \quad .$$

Let $h(x) = g^2(x)$. Then, using (47),

$$\bar{Y}_1^2 \doteq h(\Delta) + \frac{h''(\Delta)}{2} \sigma_X^2 \quad (48)$$

$$h''(\Delta) = 2[(g'(\Delta))^2 + g(\Delta)g''(\Delta)] \quad (49)$$

$$\bar{Y}_1^2 = g^2(\Delta) + [(g'(\Delta))^2 + g(\Delta)g''(\Delta)]\sigma_X^2 \quad (50)$$

Noting that $\sigma_1^2 = \bar{Y}_1^2 - (\bar{Y}_1)^2$, (47) and (56) give the desired result, namely,

$$\sigma_1^2 \doteq [g'(\Delta)]^2 \sigma_X^2 \quad (51)$$

It must be noted that (47) and (51) are of questionable validity in the weak signal case. This is because the function $g(x)$ has a discontinuous second-derivative at the origin.¹ Although the function is sufficiently smooth away from the origin, and therefore allows (47) and (51) to be useful in the large SNR case, other methods of approximation must be sought for the weak signal situation.

Referring directly to the PDF, Section 4, for very small Δ , (42) becomes

$$f_Y(y) = e^{-b\Delta} \frac{e^{-\frac{1}{2}\left(\frac{y}{\sigma_0}\right)^2}}{\sqrt{2\pi} \sigma_0} \quad \text{if } y < 0$$

$$f_Y(y) \doteq e^{b\Delta} \frac{e^{-\frac{1}{2}\left(\frac{y}{\sigma_0}\right)^2}}{\sqrt{2\pi} \sigma_0} \quad \text{if } y > 0 \quad (52)$$

Then

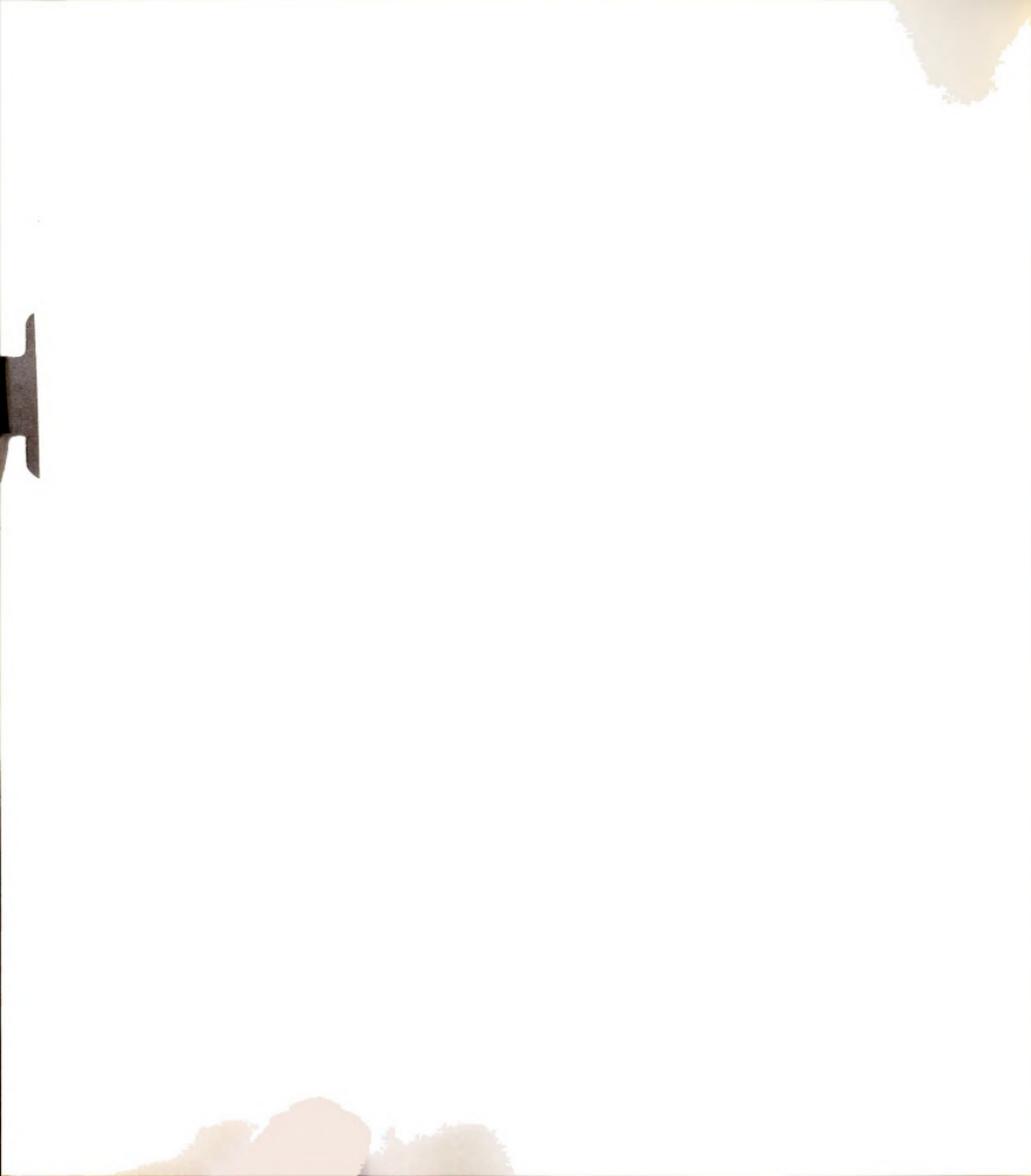
¹ See Appendix III.

$$\begin{aligned}
E(Y) &\doteq e^{-b\Delta} \sigma_0 \int_{-\infty}^0 y \frac{e^{-\frac{1}{2}\left(\frac{y}{\sigma_0}\right)^2}}{\sqrt{2\pi} \sigma_0^2} dy + e^{b\Delta} \sigma_0 \int_0^{\infty} y \frac{e^{-\frac{1}{2}\left(\frac{y}{\sigma_0}\right)^2}}{\sqrt{2\pi} \sigma_0^2} dy \quad (53) \\
&= \sqrt{\frac{2}{\pi}} \sigma_0 \sinh(b\Delta) \\
&= \sqrt{\frac{2}{\pi}} \sigma_0 b\Delta + o(\Delta^2) \\
&\doteq \frac{2}{\sqrt{\pi}} \left(\frac{\sigma_0}{\sigma_X}\right) \Delta \quad (54)
\end{aligned}$$

The variance of Y under H_1 and weak signal conditions will be assumed identical to the variance of Y under the null hypothesis. By assuming that the variances under the two hypotheses are equal, an optimum decision rule can be defined. However, no claim will be made for optimality; the important feature will be the improvement in performance which may be attained by using a nonlinear element in the detector.

3.6 Summary

This chapter discussed reasons for using nonlinear elements in the detection process. Because it is not clear that a simple clipper is in any sense the best nonlinear element that can be used, the "Gaussianizing" nonlinear filter was introduced. The nonlinear filter (NLF) that transforms Laplacian variables to Gaussian random variables was derived. The concavity of the transfer characteristic on the positive real axis was shown and the output PDF under H_1 was found. Approximations for the mean and variance of the NLF output were derived and their regions of validity were discussed.



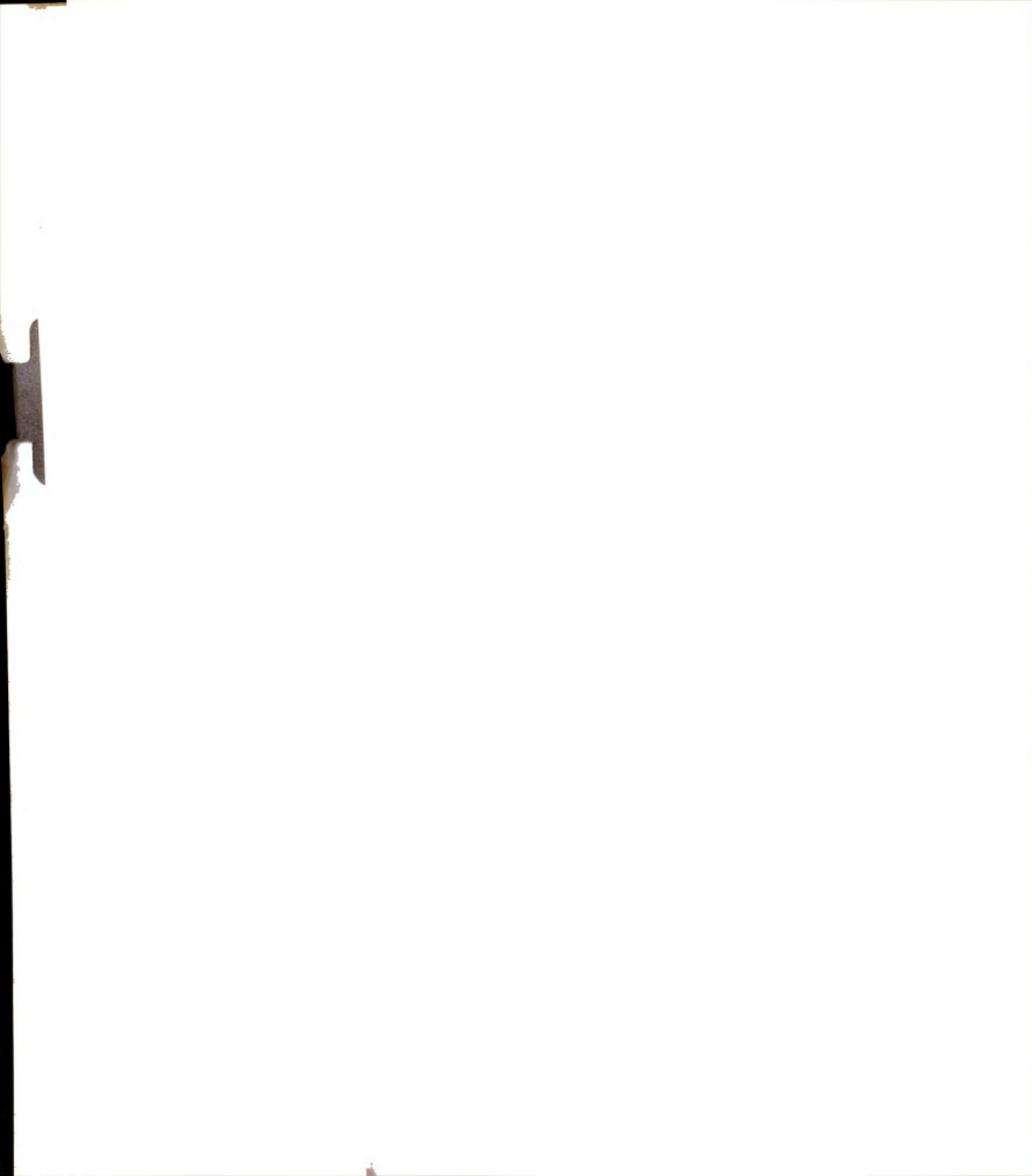
CHAPTER IV
COMPARISON OF DETECTORS

This chapter evaluates the performance of a detector which includes the NLF derived in Chapter III. It investigates the possibility of improving the detection reliability. To do so, the probability of detection (P_D) as a function of false-alarm rate (FAR) is found. For the weak signal detection problem the NLF-detector performance is compared to the optimal detector (the nonparametric sign test) and also to a suboptimal detector which is based on the assumption of a Gaussian noise model. The comparison is repeated for the case of strong signal.

The measure of performance of one detector relative to another is based on the following definition [C1, p. 89].

Definition: Let T_1 and T_2 represent two tests that test the same H_0 against the same H_1 , with the critical regions size α , and with the same values of β . The relative efficiency of T_1 to T_2 (or "efficiency of T_1 relative to T_2 ") is the ratio n_2/n_1 , where n_1 and n_2 are the sample sizes of the tests T_1 and T_2 respectively.

This is a reasonable and useful measure of the relative efficiency of T_1 with respect to T_2 . An asymptotic relative efficiency (ARE) of T_1 with respect to T_2 is obtained by passing to the weak-signal limit (when it exists) as H_1 approaches H_0 in a properly restricted sense [C5, p. 296].



Section 4.1 develops the optimal detector for location alternatives for a sample of size N from the Laplacian density $f(x) = \frac{b}{2} \exp(-b|x-\Delta|)$. Section 4.2 focuses on the weak-signal detection properties of the sign test, the "optimal" test based on an assumed Gaussian noise model, and the NLF-detector. Section 4.2 concludes with a comparison between the sign detector and the NLF detector. Section 4.3 investigates the large-signal detection properties of the three detectors. In Section 4.4 a Monte Carlo simulation is performed to investigate the performance of the three detection systems for small sample sizes in both Laplacian and Gaussian noise. Section 4.5 summarizes the main results.

4.1 Optimal Detector of Constant Signal in Laplacian Noise

The optimal detector, as was discussed in Chapter II, is based on the likelihood ratio, where $\Lambda(x^n) \stackrel{\Delta}{=} \frac{f(x_1 x_2 \cdots x_n | H_1)}{f(x_1 x_2 \cdots x_n | H_0)}$, and $f(x_1 \cdots x_n | H_i)$ is the likelihood function of the sample $x^n \stackrel{\Delta}{=} (x_1 x_2 \cdots x_n)$, when hypothesis H_i is true. Let the signal-plus-noise PDF be of the form $f(u) = \frac{b}{2} \exp(-b|u-\Delta|)$, for $-\infty < u < +\infty$, with b determined from the relationship

$$\int_{-\infty}^{\infty} x^2 f(x) dx = \sigma_X^2 \quad (1)$$

(See Chapter III, Section 3.2 for this development.)

$$\Lambda(x^n) = \left(\prod_{i=1}^n a e^{-b|x_i-\Delta|} \right) / \left(\prod_{i=1}^n a e^{-b|x_i|} \right) \quad (2)$$

$$= \exp \left(-b \sum_{i=1}^n (|x_i-\Delta| - |x_i|) \right) \quad (3)$$

$$-\frac{1}{b} \ln \Lambda(x^n) = \sum_{i=1}^n (|x_i-\Delta| - |x_i|) \quad (4)$$

To proceed further with this analysis the real line is broken up into 3 disjoint intervals, as follows:

Let $I_1 = \{x | x \geq \Delta > 0\}$, $I_2 = \{x | x < 0\}$ and $I_3 = \{x | \Delta > x > 0\}$. Now for each X_i which falls in I_1 , the right-hand-side of (4) reduces to $|X_i - \Delta| - |X_i| = -\Delta$. For each X_i in I_2 , the right-hand-side of (4) becomes $|X_i - \Delta| - |X_i| = \Delta$; and for each X_i in I_3 , the right-hand-side of (4) becomes $\Delta - 2X_i$. If among the N observations, N_i are assumed to fall in interval I_i , the right-hand-side of (4) may be rewritten as:

$$\Delta(N_2 - N_1) + \sum_{N_3} (\Delta - 2X_i). \quad (5)$$

This can also be written as

$$\Delta(N_2 - N_1) + N_3\Delta - \sum_{N_3} (\Delta - 2X_i). \quad (6)$$

Using the fact that $N_3 = N - (N_1 + N_2)$, (6) may be written

$$\Delta(N_2 - N_1) + [N - (N_1 + N_2)]\Delta - 2\sum_{N_3} X_i \quad (7)$$

$$= \Delta[N - 2N_1] - 2\sum_{N_3} X_j. \quad (8)$$

The optimum processor consists of three-level threshold circuitry followed by a counter and summing unit. If the threshold circuitry detects a sample in I_1 , Δ units are subtracted from the counter while a positive Δ units are added to the counter when a detected sample falls in I_2 . Each sample occurring in I_3 is summed and stored. The storage unit and counter indication are summed after the N samples are processed and this test statistic is compared to the decision threshold. The exact distribution of this test statistic can, theoretically, be obtained although actual analytic evaluation would be cumbersome.



4.2 Weak Signal Detector Evaluation

If interest is focused on the vanishingly small SNR case ($\Delta \rightarrow 0$), a result from classical hypothesis testing literature becomes immediately applicable. If X_1, X_2, \dots, X_N is a sample from the Laplacian distribution with density $\frac{1}{2} \exp(-|X - \Delta|)$, with Δ a location parameter, then the Locally Most Powerful (LMP) test of $\Delta \leq 0$ against $\Delta > 0$ is the sign test [L1, p. 344]. The concept of LMP tests is applicable to communication problems because of the desirability of achieving optimum detection for weak signals. Strong signals will be detected even if the detector is well below optimum, as was noted by Capon [C2, p. 67].

As $\Delta \rightarrow 0$, the test statistic of (8) becomes $\Delta(N - 2N_1)$, where N_1 is the number of X_i greater than Δ . Thus an equivalent test can be stated in the form $S_N(X^N) = \sum_{i=1}^N u(X_i)$ where

$$u(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The test based on S_N has been investigated by Capon [C3], Kanefsky and Thomas [K2], and others [C4], [H6], [C5].

It is a well known result [F2] that the statistic S_N is sufficient for p , the probability that $X > 0$. The distribution of S_N is binomial with parameters N and $\frac{1}{2}$ under H_0 , and therefore the threshold value for any specified FAR can be determined. The problem of randomization can best be explained by an example.



Example 1: Determine the threshold value that will guarantee a FAR of 10% when the test statistic is S_N and $N = 8$. The probability mass function for S_N is shown in Figure 4.1. The randomization problem becomes obvious unless the FAR is specified as one of the values (0.0, 0.0039, 0.0351, 0.1445,...).

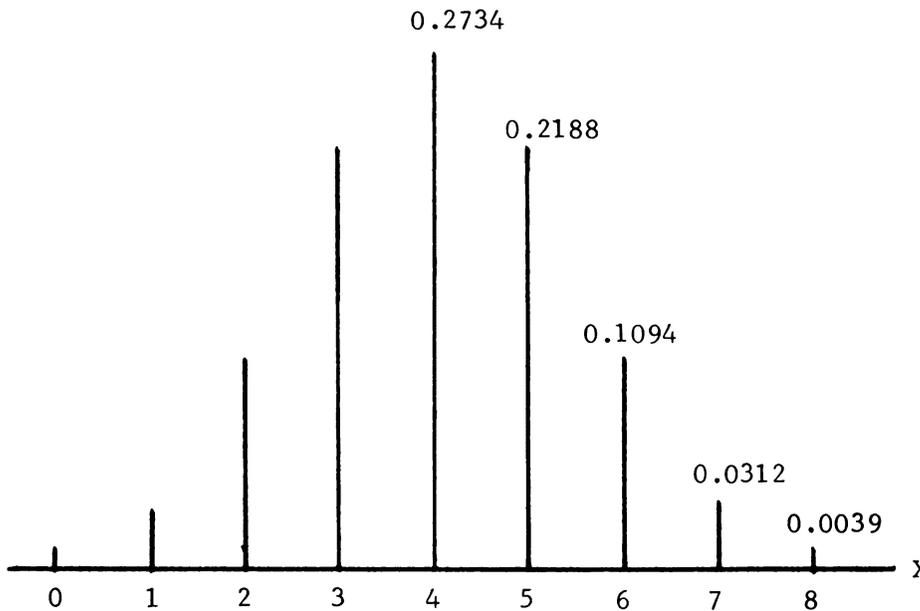


Figure 4.1 Probability Mass Function for S_N

To achieve the specified FAR of 10%, randomization of the threshold value of $X = 6$ is required.

$$0.10 = (.0039) + (0.0312) + \beta(0.1094)$$

$$\beta = \frac{.0649}{.1094} \doteq .594 .$$

The test is based on S_N rejects the hypothesis H_0 if $X = 7$ or 8 and with probability .594 rejects H_0 if $X = 6$.



This section continues with the derivation of the asymptotic distributions of the sign detector under the hypothesis and alternative and its probability of detection as a function of FAR.

Write

$$P[S_N \leq n] = \sum_{j=0}^n \binom{N}{j} p_i^j (1 - p_i)^{N-j}, \quad n = 0, 1, 2, \dots, N \quad (9)$$

where

$$p_i \triangleq P[X \geq 0 | H_i] . \quad (10)$$

With $p_0 = \frac{1}{2}$ (by Assumption A2, Chapter I), define $K \ni$

$$\sum_{j=K}^N \binom{N}{j} = 2^N P_F . \quad (11)$$

This expression assumes that randomization is unnecessary. Because asymptotic results ($N \rightarrow \infty$) are of interest in this chapter, this will be of no consequence. On the basis of (11), write

$$P_D = \sum_{j=K}^N \binom{N}{j} p_1^j (1 - p_1)^{N-j} \quad (12)$$

with p_1 given by (10). For large sample sizes, the DeMoivre-Laplace limit theorem can be invoked to simplify computation of the probabilities (11) and (12), [F4].

Defining

$$S_{N,i}^* \triangleq \frac{S_N - Np_i}{\sqrt{Np_i(1-p_i)}} \quad (13)$$

the theorem states that

$$P[z_1 \leq S_{N,i}^* \leq z_2] \xrightarrow{N} \eta(z_2) - \eta(z_1) . \quad (14)$$

Then (11) may be written

$$P_F = \frac{\int_{K - \frac{N}{2}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du}{\sqrt{\frac{N}{4}}} \quad (15)$$

$$P_D = \frac{\int_{K - Np_1}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du}{\sqrt{Np_1(1-p_1)}} \quad (16)$$

From (15)

$$K = \sqrt{\frac{N}{4}} \operatorname{erfc}_*^{-1} P_F + \frac{N}{2} . \quad (17)$$

Using (17) in (16) yields the receiver operating characteristic of the sign detector, namely,

$$P_D^s = \operatorname{erfc}_* \left[\frac{\operatorname{erfc}_*^{-1} P_F - \sqrt{N} (2p_1 - 1)}{\sqrt{4p_1(1-p_1)}} \right] \quad (18)$$

where the s superscript indexes the sign detector ROC.

When X has the Laplacian PDF

$$f(X) = \frac{b}{2} \exp(-b|X|), \quad -\infty < X < +\infty \quad (19)$$

then

$$p_1 = 1 - \frac{e^{-b\Delta}}{2}$$

$$1 - p_1 = \frac{e^{-b\Delta}}{2} \quad (20)$$

Substituting these expressions into (18) yields



$$P_D^S = \operatorname{erfc}_* \left[\frac{\operatorname{erfc}_*^{-1} P_F - \sqrt{N} (1 - e^{-b\Delta})}{\left[(2 - e^{-b\Delta}) e^{-b\Delta} \right]^{\frac{1}{2}}} \right]. \quad (21)$$

Substituting $b = \frac{\sqrt{2}}{\sigma_X}$ into (21) yields

$$P_D^S = \operatorname{erfc}_* \left[\frac{\operatorname{erfc}_*^{-1} P_F - \sqrt{N} \left(1 - e^{-\sqrt{2} \frac{\Delta}{\sigma_X}} \right)}{\left[\left(2 - e^{-\sqrt{2} \frac{\Delta}{\sigma_X}} \right) e^{-\sqrt{2} \frac{\Delta}{\sigma_X}} \right]^{\frac{1}{2}}} \right]. \quad (22)$$

Equation (22) can be used to plot P_D as a function of P_F with N and $\frac{\Delta}{\sigma_X}$ as parameters. The ROC of (22) can be compared to that derived in Chapter II, Section 2.5 for the "optimal" detector based on an assumed Gaussian noise model, namely,

$$P_D^0 = \operatorname{erfc}_* \left[\operatorname{erfc}_*^{-1} P_F - \sqrt{N} \frac{\Delta}{\sigma_X} \right]. \quad (23)$$

For the case of small SNR, i.e., $\Delta \sim 0$, (22) becomes

$$P_D^S = \operatorname{erfc}_* \left[\operatorname{erfc}_*^{-1} P_F - \sqrt{2N} \frac{\Delta}{\sigma_X} \right]. \quad (24)$$

Suppose that the numbers of samples N in (23) and (24) are not necessarily equal, i.e., in (23) let $N = N_0$ while in (24), let $N = N_s$.

If the two tests have the same FAR and P_D , then equating (23) and (24) yields

$$\operatorname{erfc}_*^{-1} P_F - \sqrt{N_0} \frac{\Delta}{\sigma_X} = \operatorname{erfc}_*^{-1} P_F - \sqrt{2N_s} \frac{\Delta}{\sigma_X} \quad (25)$$

or simply $\frac{N_0}{N_s} = 2$ which is precisely the Pitman ARE for the nonparametric sign detector relative to the "optimal" Gaussian detector.



The weak signal performance of the NLF detector is now presented. The results of Chapter III are summarized in Figure 4.2.

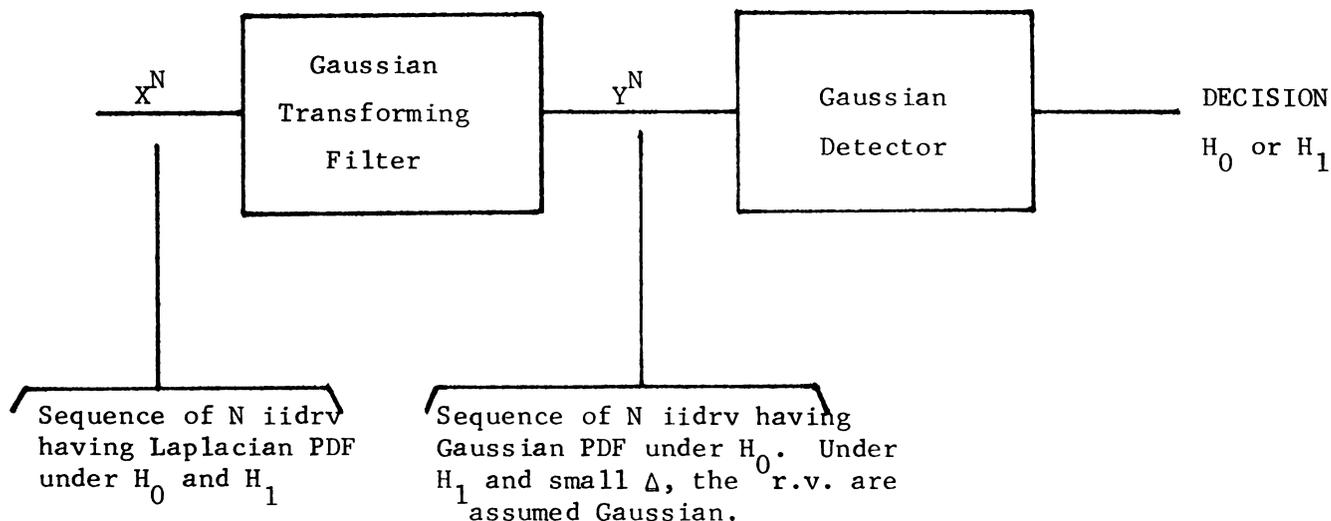


Figure 4.2 Nonlinear filter detector

Well known results concerning the sufficiency of the sample mean can be invoked for the problem described in Figure 4.2. The detector based on the sample (Y^N) mean is easily formulated. Under H_0 , $\bar{Y}_N \sim \eta(0, \frac{\sigma_0^2}{N})$ while under H_1 , $\bar{Y}_N \sim \eta(\frac{2}{\sqrt{\pi}} (\frac{\sigma_0}{\sigma_X}) \Delta, \frac{\sigma_0^2}{N})$. The expression for P_D as a function of FAR is then given by

$$P_D^N = \text{erfc}_* \left[\text{erfc}_*^{-1} P_F - \frac{\sqrt{N}}{\sigma_0} \left(\frac{\sqrt{4}}{\sqrt{\pi}} \left(\frac{\sigma_0}{\sigma_X} \right) \Delta \right) \right] \quad (26)$$

the validity of which is based on $\Delta \sim 0$. The N-superscript indexes the NLF detector ROC.

Suppose that the numbers of samples N in (24) and (26) are not necessarily equal, i.e., in (24) let $N = N_s$ and in (26), let $N = N_N$. If the two tests have the same FAR and P_D , then equating (24) and (26) yields

$$\operatorname{erfc}_*^{-1} P_F - \sqrt{2N_s} \frac{\Delta}{\sigma_X} = \operatorname{erfc}_*^{-1} P_F - \frac{\sqrt{N}}{\sigma_0} \left(\frac{4}{\pi} \left(\frac{\sigma_0}{\sigma_X} \right) \right) \Delta$$

$$\frac{N_s}{N_N} = \frac{2}{\pi} \doteq 0.636 . \quad (27)$$

Using (25) and (27) the ARE of the NLF-detector relative to the "optimal" Gaussian detector is found to be

$$\frac{N_0}{N_N} = \frac{N_0}{N_s} \cdot \frac{N_s}{N_N} = \frac{4}{\pi} \doteq 1.27 .$$

On the basis of ARE, the NLF-detector is better than the "optimal" Gaussian detector, but not as good as the sign detector.¹ This result could have been anticipated under the assumptions on which it was based. The following theorem makes this situation clear.

Theorem: Let X and Y be univariate random variables related by $Y = g(X)$, $g(\cdot)$ a monotone increasing function wherever $f_X(X)$ is nonzero. Then $J_X(1:2) = J_Y(1:2)$.

Proof: $J_X(1:2) \triangleq \int_{-\infty}^{\infty} (f_{X1}(X) - f_{X2}(X)) \ln \frac{f_{X1}(X)}{f_{X2}(X)} dX$. Make the change of variable $X = g^{-1}(y)$, which is monotone increasing and defined uniquely for each y . Then,

¹ Note that the NLF does improve SNR without a corresponding improvement in detection probability.



$$\begin{aligned}
J_X(1:2) &= \int_{-\infty}^{\infty} (f_{X1}(g^{-1}(y)) - f_{X2}(g^{-1}(y))) \left(\frac{d}{dy} g^{-1}(y)\right) \ln \frac{f_{X1}(g^{-1}(y))}{f_{X2}(g^{-1}(y))} dy \\
&= \int_{-\infty}^{\infty} (f_{Y1}(y) - f_{Y2}(y)) \ln \frac{f_{Y1}(y)}{f_{Y2}(y)} dy \\
&= J_Y(1:2) .
\end{aligned}$$

This formalizes the notion that a nonlinearity whose characteristic is monotonic does not affect P_D and FAR [R2, p. 438]. This is because the probability that the output exceeds a given threshold is exactly the same as the probability that the input exceeds a corresponding threshold, there being a one-to-one relationship between input and output amplitudes.

The optimal test for location was shown in Chapter II to be of the form $\Lambda(X^N) \underset{H_0}{\overset{H_1}{>}} K$, where $\Lambda(X^N)$ was the likelihood ratio and K was picked so that $P[\Lambda(X^N) > K | H_0] = P_F$. By the monotonicity of the logarithm, an equivalent optimal test is of the form $\ln \Lambda(X^N) \underset{H_0}{\overset{H_1}{>}} K^*$, with K^* satisfying

$$P[\ln \Lambda(X^N) > K^* | H_0] = P_F .$$

Then
$$P[\ln \Lambda(X^N) > K^* | H_1] = P_D .$$

If $\ln \Lambda(X^N)$ has a symmetric PDF under H_0 and H_1 , and if the variance of $\ln \Lambda(X^N)$ is the same under both H_0 and H_1 , then P_D depends on the separation of means of $\ln \Lambda(X^N)$, the larger the separation the larger the detection probability will be. Then $E[\ln \Lambda(X^N) | H_1] - E[\ln \Lambda(X^N) | H_0]$ is a useful measure of the distance between distributions. When normalized by the common

variance, the result corresponds to the SNR measure of Van Trees [Section 2.3]. But

$$E[\ln \Lambda(X^N) | H_1] = \int_{-\infty}^{\infty} f_1(x) \ln \frac{f_1(x)}{f_0(x)} dx ,$$

and

$$E[\ln \Lambda(X^N) | H_0] = \int_{-\infty}^{\infty} f_0(x) \ln \frac{f_1(x)}{f_0(x)} dx$$

which yields

$$\begin{aligned} & E[\ln \Lambda(X^N) | H_1] - E[\ln \Lambda(X^N) | H_0] \\ &= \int_{-\infty}^{\infty} [f_1(x) - f_0(x)] \ln \frac{f_1(x)}{f_0(x)} dx \triangleq J(1:0) \end{aligned}$$

which is the J-divergence distance measure introduced in Appendix I. This shows that the optimal test for small signals at the nonlinear filter output is not based on the sample mean. However, as pointed out in Section 3.5, no claim for optimality was intended; the NLF-detector is being investigated as a suboptimal detector which provides a continuous test statistic of known form under the null hypothesis.

4.3 Strong Signal Detector Evaluation

Letting the sample sizes of the sign detector and optimal Gaussian detector be N_s and N_0 , respectively, and equating (22) and (23) yields

$$\frac{\text{erfc}_*^{-1} P_F - \sqrt{N_s} (1 - e^{-\beta})}{[(2 - e^{-\beta})e^{-\beta}]^{\frac{1}{2}}} = \text{erfc}_*^{-1} P_F - \sqrt{\frac{N_0}{2}} \beta \quad (28)$$

where $\beta \triangleq \sqrt{2} \Delta / \sigma_x$. Define $h(\beta) \triangleq 1 - [(2 - e^{-\beta})e^{-\beta}]^{-\frac{1}{2}}$, and write



$$\frac{N_0}{N_s} = \frac{2}{\beta^2} \left[\frac{\text{erfc}^{-1} P_F}{\sqrt{N_s}} h(\beta) + (1 - e^{-\beta}) \right]^2 \quad (29)$$

In the region $0 < \beta < .7$, $|h(\beta)| < .15$, and for sufficiently large N_s , (29) becomes

$$\frac{N_0}{N_s} \doteq 2 \left(\frac{1 - e^{-\beta}}{\beta} \right)^2 > 1. \quad (30)$$

The weak-signal limit ($\beta \rightarrow 0$) is easily found by applying L'Hospital's rule twice, resulting in

$$\frac{N_0}{N_s} = 2 \quad (31)$$

in agreement with the conclusion of Section 4.2. Thus, the large-sample efficiency improvement of the sign detector relative to the optimal Gaussian detector in Laplacian noise is seen to be more than a local property; i.e., valid at vanishingly small SNR's.

This result, although developed here for the first time, is not entirely unexpected. Recall that the variance of the sign detector under H_1 was given by $Np_1(1 - p_1)$, with $p_1 \triangleq P[X > 0 | H_1]$. Note that as p_1 increases from $\frac{1}{2}$ to 1, the variance decreases thus concentrating more of the probability mass about its average. With a fixed threshold, the probability of detection will dramatically increase.

The large signal mean and variance of the NLF detector were found in Chapter III, Section 3.5. Invoking the central limit theorem to evaluate P_D and FAR, and using those moment approximations with a sample size of N_N yields the ROC

$$P_D^N = \operatorname{erfc}_*^{-1} \left[\frac{\sigma_0 \operatorname{erfc}_*^{-1} P_F}{\sigma_X g'(\Delta)} - \sqrt{N_N} \left(\frac{g(\Delta) + \frac{1}{2} g''(\Delta) \sigma_X^2}{\sigma_X g'(\Delta)} \right) \right] . \quad (32)$$

Equating this to (29) yields

$$\begin{aligned} \sigma_0 \operatorname{erfc}_*^{-1} P_F - \sqrt{N_N} \left(g(\Delta) + \frac{1}{2} g''(\Delta) \sigma_X^2 \right) \\ = \sigma_X g'(\Delta) \left[\operatorname{erfc}_*^{-1} P_F - \sqrt{N_0} \left(\frac{\Delta}{\sigma_X} \right) \right] . \end{aligned} \quad (33)$$

Assuming N_N and N_0 are sufficiently large, (33) becomes

$$\frac{N_0}{N_N} \doteq \left[\frac{g(\Delta) + \frac{1}{2} g''(\Delta) \sigma_X^2}{\Delta g'(\Delta)} \right]^2 . \quad (34)$$

By the mean-value theorem (34) may be written

$$\begin{aligned} \frac{N_0}{N_N} &\doteq \left(\frac{g(0) + \Delta g'(\xi) + \frac{1}{2} g''(\Delta) \sigma_X^2}{\Delta g'(\Delta)} \right)^2 \quad \text{for some } \xi \in (0, \Delta) . \\ &\cong \left(\frac{g'(\xi)}{g'(\Delta)} \right)^2 \geq 1 \quad \text{by the convexity of } g(\cdot) . \end{aligned}$$

This requires that $|\frac{1}{2} g''(\Delta) \sigma_X^2| \ll |\Delta g'(\xi)|$ or,

$$\frac{|g''(\Delta)|}{g'(\xi)} \ll \sqrt{2} \left(\frac{\Delta}{\sigma_X} \right) \frac{\sqrt{2}}{\sigma_X} .$$

Now $\frac{|g''(\Delta)|}{g'(\xi)} \cong \frac{|g''(\Delta)|}{g'(\Delta)}$ which can be written (using equation (20) of Section 3.3) as

$$\frac{|g''(\Delta)|}{g'(\Delta)} = \frac{b\sigma_0^2 - g(\Delta)g'(\Delta)}{\sigma_0^2} \leq b = \frac{\sqrt{2}}{\sigma_X} .$$

But for sufficiently large signal-to-noise ratio,



$$\frac{|g''(\Delta)|}{g'(\xi)} \leq \frac{|g''(\Delta)|}{g'(\Delta)} \leq b = \frac{\sqrt{2}}{\sigma_X} \ll \sqrt{2} \left(\frac{\Delta}{\sigma_X}\right) \frac{\sqrt{2}}{\sigma_X} .$$

Thus, the NLF detector has greater efficiency than the "optimal" Gaussian detector. This result could also be anticipated because the variance of the test statistic, given by $[g'(\Delta)]^2 \sigma_X^2$, asymptotically approaches zero as the signal amplitude increases.

Lemma: The variance of the NLF output under H_1 becomes vanishingly small as the input signal amplitude increases. That is

$$\lim_{\Delta \rightarrow \infty} \text{Var} (Y|H_1) = 0.$$

Proof: Recall that $\text{Var} (Y|H_1) \doteq [g'(\Delta)]^2 \sigma_X^2$. Because σ_X^2 is a fixed, known constant it is sufficient to show that $\lim_{\Delta \rightarrow \infty} g'(\Delta) = 0$.

In Section 3.3 it was shown that $b\sigma_0^2 - gg' \geq 0$. Then because $g > 0$, $\forall x > 0$, and $b\sigma_0^2$ is finite,

$$\frac{b\sigma_0^2}{g} - g' \geq 0 .$$

Therefore

$$g' \leq \frac{b\sigma_0^2}{g} .$$

But by the definition of g as a monotone increasing function, the limit as Δ approaches infinity for the right side is zero. Because $g' \geq 0$ when $x > 0$, this holds also for the left side. Q.E.D.



4.4 Simulation

This section provides a Monte Carlo evaluation of the three detection systems for small sample sizes in both Laplacian ($\theta = 1$) and Gaussian ($\theta = 2$) noise. It is expected that the performance of the detectors for values of $\theta \in (1,2)$ will fall somewhere between their performance at the extremes. Figure 4.3 is a block diagram of the overall simulation procedure. Threshold values yielding a False Alarm Rate of 10% were read into the program, along with the randomization strategy for the sign detector. Signal-to-Noise ratios were made dependent on a single parameter, Δ , by choosing $\sigma_0 = \sigma_x = 1$. The NLF was synthesized as discussed in Appendix II and the number of Monte Carlo experiments performed during each pass through the program was set at 250. This yields the interval (.07, .14) as a 95% confidence interval for the FAR, [C1, p. 99] a suitable range for the purposes of this simulation. The average FAR for each sample size and each detector is shown in Table 4.1.

Laplacian Noise					
	4	8	12	16	20
SIGN	.089	.104	.115	.101	.087
XBAR	.089	.103	.095	.094	.095
NLF	.100	.103	.114	.095	.103
Gaussian Noise					
	4	8	12	16	20
SIGN	.098	.088	.120	.108	.086
XBAR	.103	.108	.112	.093	.109
NLF	.129	.128	.105	.104	.111

Table 4.1 False Alarm Rates for Detectors

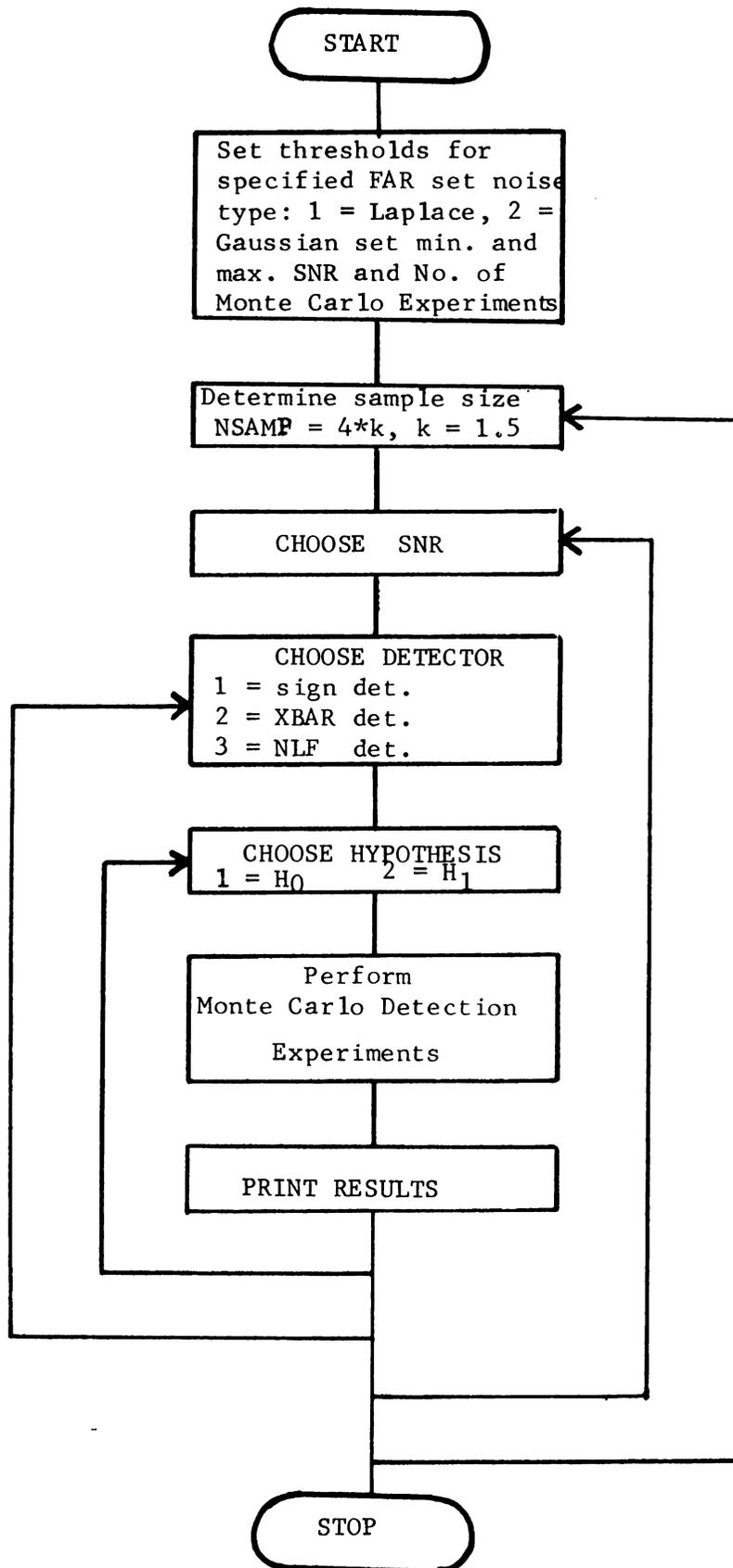


Figure 4.3 Simulation Block Diagram



The data are presented in Figures 4.4 through 4.9 for sample sizes 4, 12 and 20. Appendix IV contains a complete tabulation of simulation data.

Before the data can be adequately evaluated and conclusions drawn as to the relative performance of each of the detector systems, it is important that the simulation procedure be examined so that any biases appearing in the data be clearly defined as to source and overall effect on the simulation.

The nonlinear filter implementation is discussed in Appendix II. As mentioned there, a piece-wise linear approximation to the continuous transfer characteristic of Section 3.2 was made in ten uniformly spaced intervals extending from x equals zero, to x equals three. A linear approximation to $g(x)$ for x greater than three was based on the slope of the chord between x equals 3 and x equals 4. Because the transfer characteristic is concave, this approximation tends to amplify outliers and results in an increased FAR and P_D . However, the effect was not substantial as evidenced by Table 4.1. It was enough to cause the detection probability of the NLF-detector to be greater than that of the optimal detector in Gaussian noise, a theoretical impossibility. This is particularly evident at small sample sizes. This was the only noticeable bias in the simulation procedure and therefore the following conclusions can be made.

1. In Laplacian noise the sign detector is inferior to both the XBAR detector and the NLF-detector over the greater part of the range of experimental SNR's for small samples [See Figures 4.4, 4.6 and 4.8]. As the sample size increases the performance

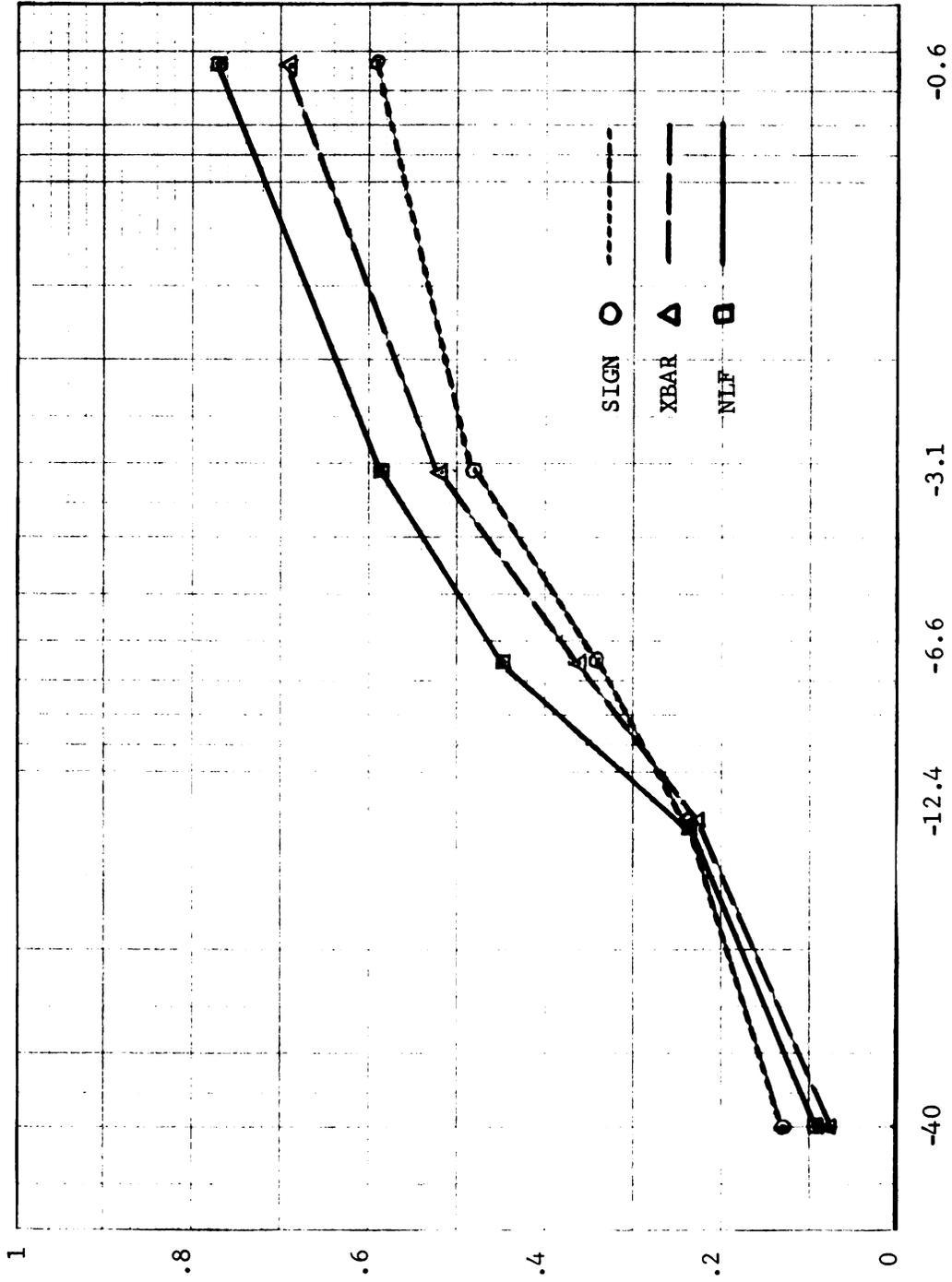


Figure 4.4 P_D vs. SNR, Laplacian Noise, Sample Size = 4

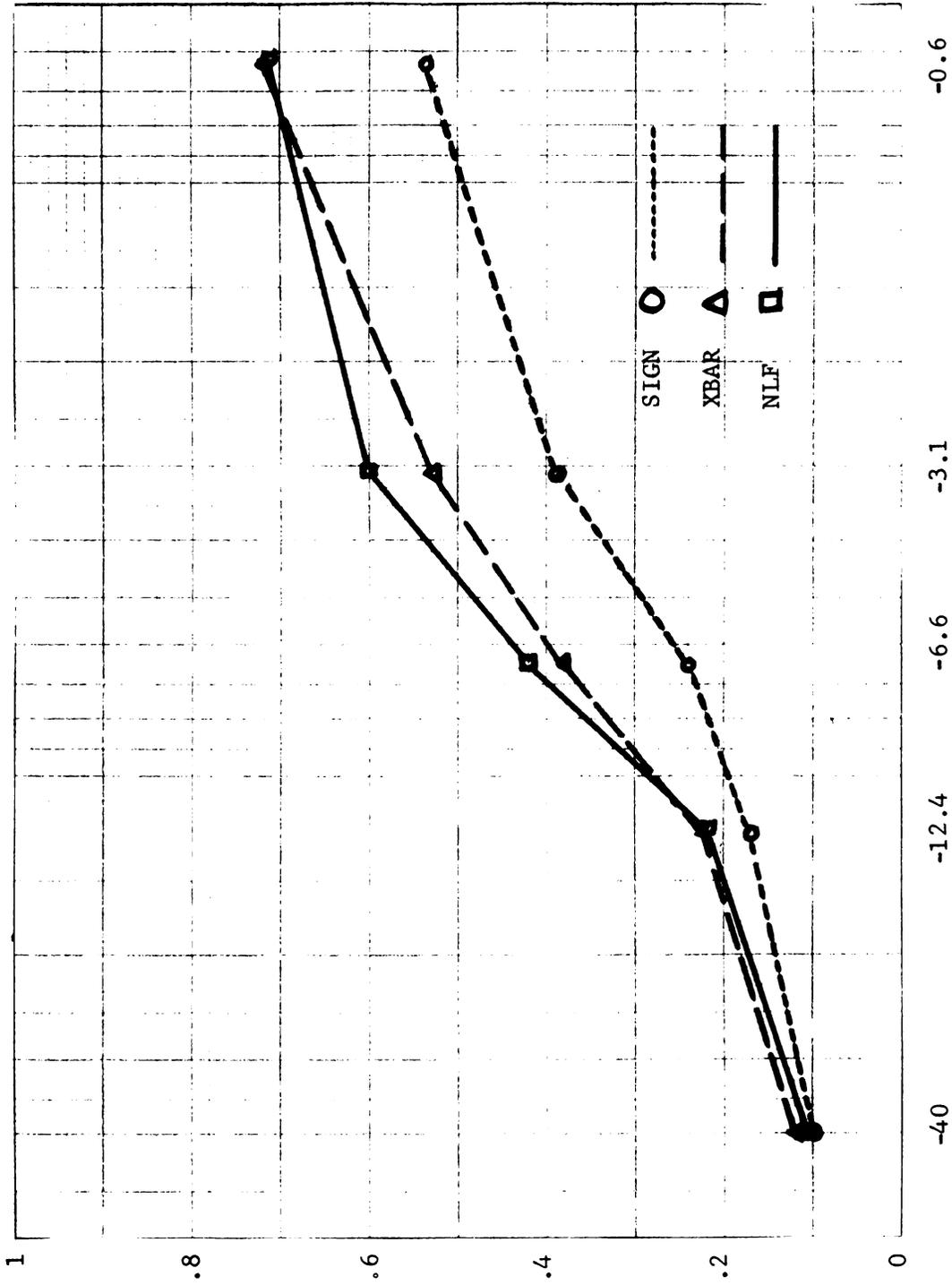


Figure 4.5 P_D vs. SNR, Gaussian Noise, Sample Size = 4



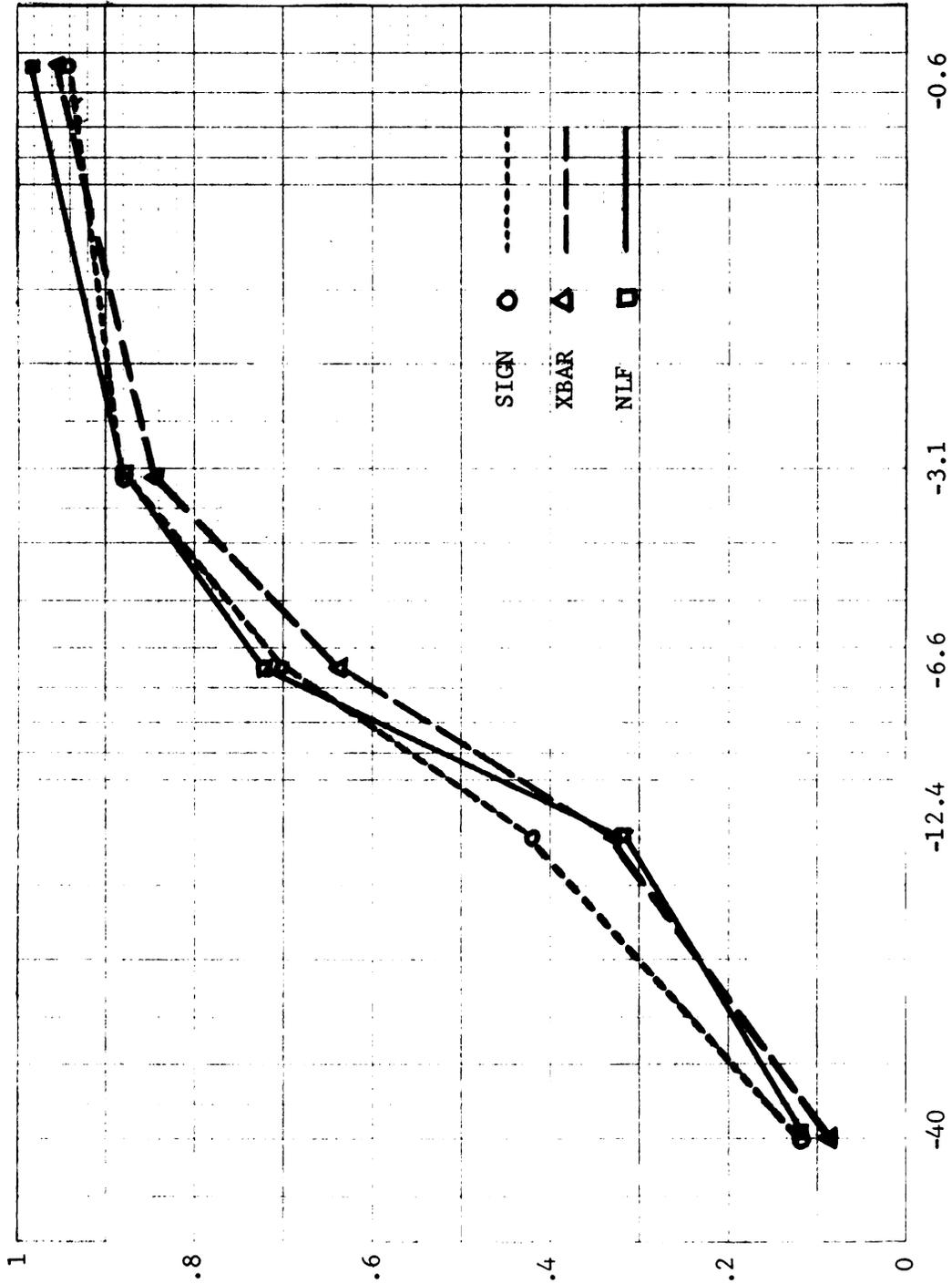


Figure 4.6 P_D vs. SNR, Laplacian Noise, Sample Size = 12



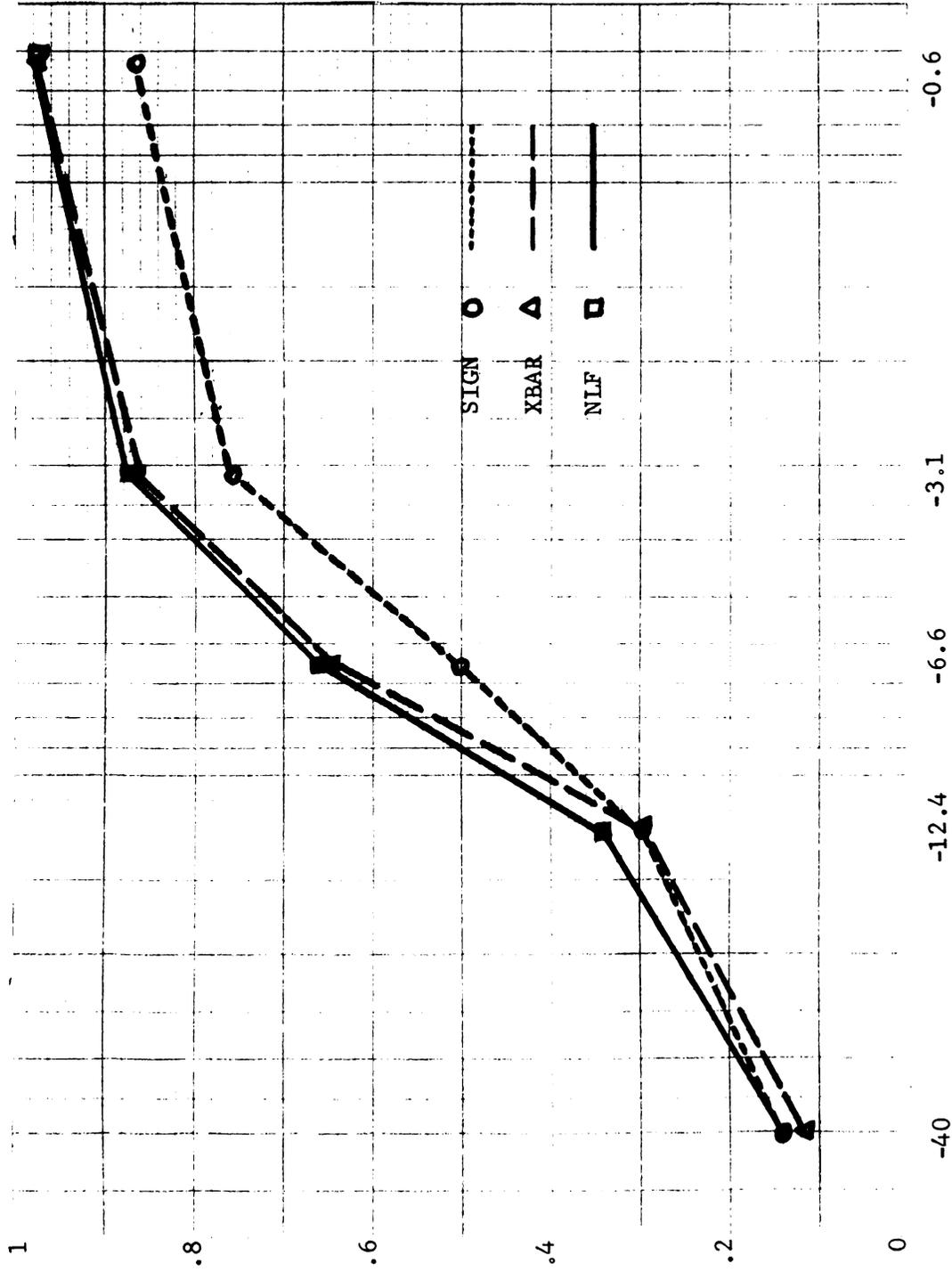


Figure 4.7 P_D vs. SNR, Gaussian Noise, Sample Size = 12



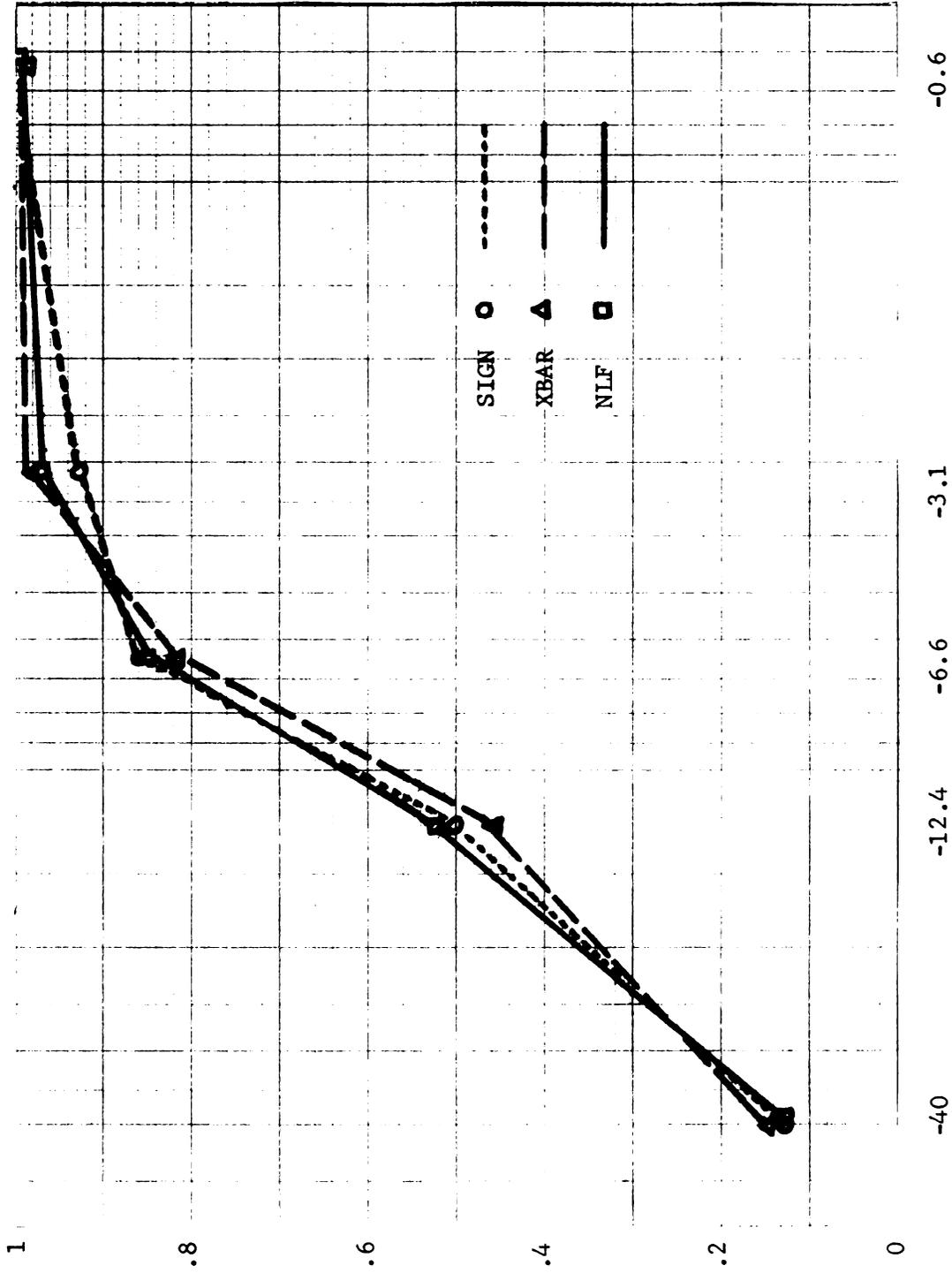


Figure 4.8 P_D vs. SNR, Laplacian Noise, Sample Size = 20



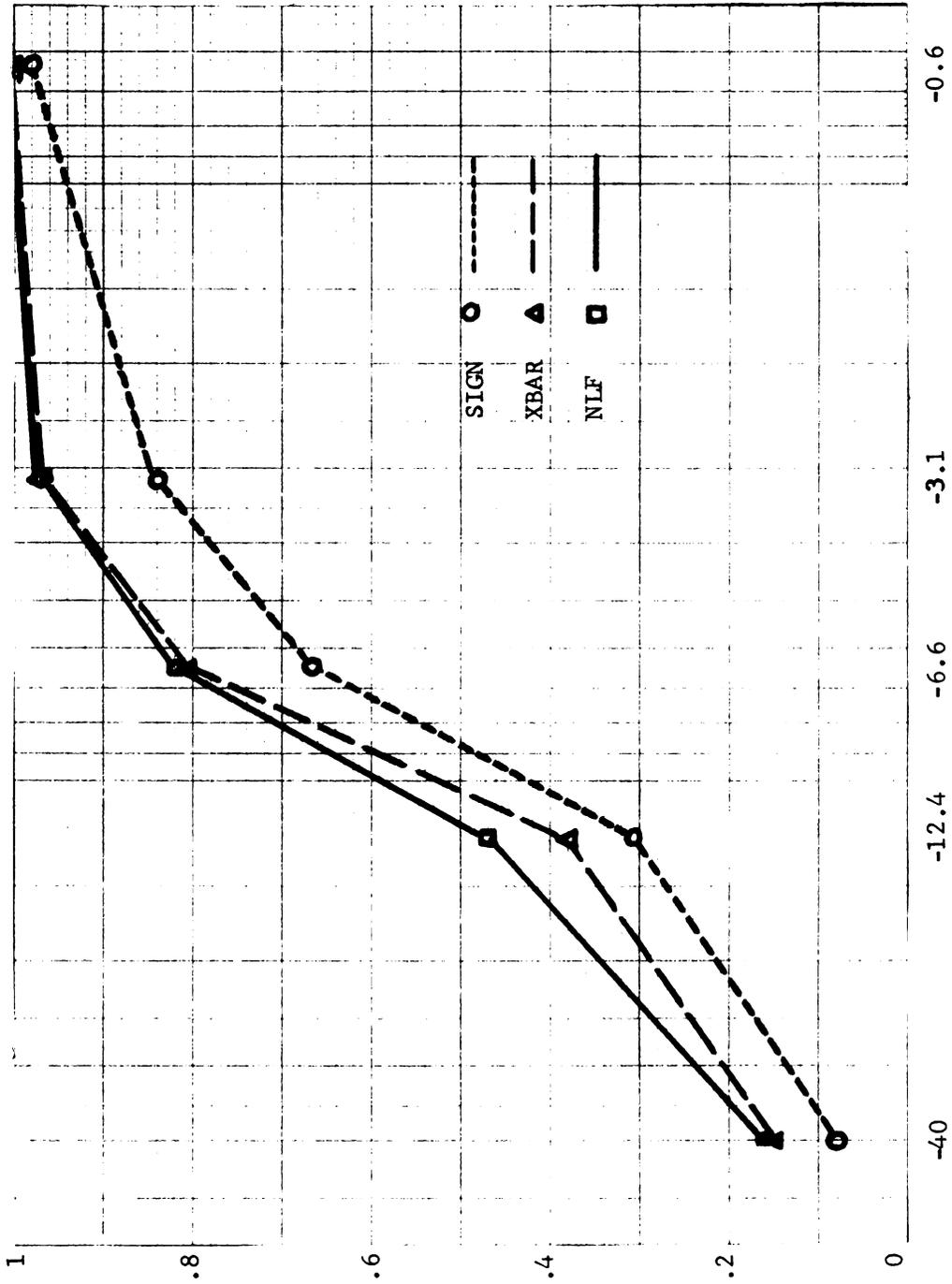


Figure 4.9 P_D vs. SNR, Gaussian Noise, Sample Size = 20



of the sign detector improves to the point of being equivalent to the other two, in agreement with the result of Section 4.3 [Figure 4.8].

2. In Laplacian noise, the NLF detector is superior to the XBAR detector for small sample sizes but as sample size increases the performance of the XBAR detector improves to the point of being equivalent to that of the NLF-detector. This lends support to the conclusions of Sections 4.2 and 4.3 [See Figures 4.5, 4.7, 4.9].

3. In Gaussian noise, the XBAR detector provides a substantial improvement in detection probability compared to the sign detector for all sample sizes and SNR's [See Figures 4.5, 4.7, 4.9]. However, the NLF-detector and the XBAR detector appear equivalent for all SNR's and sample sizes.

4. The PDF of the NLF output can theoretically be obtained for all values of $\theta \in [1,2]$, although this thesis concentrated on the special case $\theta = 1$. For the case $\theta = 2$, the method of Section 2.5.1 was invoked; the resultant PDF was mathematically intractable. For this reason, the NLF-detector FAR cannot be computed for values of $\theta \in (1,2]$. It therefore suffers the same disadvantage as the XBAR detector for $\theta \in [1,2)$. [See Section 1.5.]

4.5 Summary

The optimal detector of constant signals in Laplacian noise was derived and the difficulty of evaluating its performance was discussed. It was shown that for vanishingly small signals a locally most powerful detector could be defined and easily implemented. Its performance could be easily evaluated for all sample sizes. Because the test statistic had a discrete PDF, it

was pointed out that randomization of threshold values might be necessary to achieve a specified FAR and this was illustrated by an example. The weak-signal detection property of the "optimal" detector, i.e., the detector designed on the basis of an assumed Gaussian noise model, was compared to that of the LMP detector on the basis of ARE and shown to be markedly inferior. The weak-signal detection property of the NLF detector was compared to the LMP detector, and while still inferior, was an improvement over the "optimal" detector. That this result follows from the type of nonlinear transformation assumed in this thesis was shown in a theorem. The strong signal detection properties of all three detectors were compared and again the NLF detector was shown to be a viable alternative to the others.

A Monte Carlo simulation of the detection problem was implemented and the three detector systems were evaluated in both Laplacian and Gaussian noise for a large range of SNR and small sample sizes. It was demonstrated that the sign detector was inefficient relative to the XBAR detector in Laplacian noise and small sample sizes unless the signal level was vanishingly small. The NLF detector was shown to be equivalent to the XBAR detector and a definite improvement over the sign detector for a wide range of SNR's and sample sizes in both Laplacian and Gaussian noise.

CHAPTER V

GENERAL CONCLUSIONS AND EXTENSIONS

The major conclusions of the thesis are reviewed in this chapter and possibilities for future research are discussed.

5.1 Conclusions

This thesis has been concerned with the design of a zero-memory, nonlinear filter, and its application to the problem of detecting a constant signal in non Gaussian noise. It was shown in Chapter III that a nonlinear filter could be designed which transformed a sequence of independent, identically distributed, Laplacian random variables to a sequence of independent, identically distributed, Gaussian random variables. This allowed satisfaction of a specified false alarm rate without randomization at the receiver, a disadvantage when the intractability of the input distribution suggests application of non-parametric procedures.

The transfer characteristic of the nonlinear filter (NLF) was restricted to be monotonic increasing. In Chapter II, the property of monotone likelihood ratio (MLR) was shown to be invariant under this transformation. Thus, optimal tests for location consisted of threshold tests.

In Chapter IV, the NLF-detector was compared to the optimal detector and another based on an assumed Gaussian noise model, for



vanishingly small signals. On the basis of ARE, its performance was an improvement over the detector based on the Gaussian noise model, but it was not as good as the locally most powerful sign detector. A theorem showed that the NLF transfer characteristic did not affect the discriminability between the hypotheses. It was then concluded that the NLF detector was not the optimal detector for location alternatives. For strong signals, both the sign detector and the NLF detector were shown to be superior to that based on an assumed Gaussian noise model. It was shown that the variance of the test statistic under the alternative for the sign detector and the NLF detector asymptotically decreased with increasing signal-to-noise ratio.

5.2 Extensions

"Little has ever been done in the matter of finding to what extent this form of transformation might be utilized in deriving distribution laws for empirical data, but it has been used now and then to throw unusual types of distribution curves into forms that conformed better to established types, especially to the normal curve." [F3, pp. 273-274]. It was this statement that provided the direction of this thesis. Although a first step has been taken in investigating the utility of transforming distributions, there remains a great deal more to be done. The choice of a transformation yielding a normal curve is not necessarily best; perhaps another PDF could be more easily realized than the method of Appendix II indicates is required for "Gaussianizing". Further, transforming to another PDF could

possibly improve the efficiency of a NLF-detector relative to the optimal sign detector.

Another direction which should prove fruitful is the investigation of the improvement afforded by a nonlinear filter when the optimality criterion is minimum average probability of error. This is most often the optimality criterion in a communications network.



BIBLIOGRAPHY



BIBLIOGRAPHY

- [C1] Conover, W.J., Practical Nonparametric Statistics, Wiley, New York, 1971.
- [C2] Capon, J., "On the Asymptotic Efficiency of Locally Optimum Detectors," IRE Trans. on Information Theory, Vol. IT-7, No. 2.
- [C3] Capon, J., "Optimum Coincidence Procedures for Detecting Weak Signals in Noise," IRE International Convention Record, Pt. 4, 1960.
- [C4] Chadwick, H.D. and Kurz, L., "Two Sequential Nonparametric Detection Procedures," NYU Technical Report 400-155, March 1967.
- [C5] Carlyle, J.W., "Nonparametric Methods in Detection Theory," Chapter 8 of Communication Theory, A.V. Balakrishnan, Ed., McGraw-Hill, New York, 1968.
- [D1] Daly, R.F. and Rushforth, C.K., "Nonparametric Detection of a Signal of Known Form in Additive Noise," IEEE Trans. on Information Theory, Vol. IT-11, No. 1.
- [D2] Dubes, R.C., The Theory of Applied Probability, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1968.
- [F1] Fraser, D.A.S., Nonparametric Methods in Statistics, Wiley, New York, 1957.
- [F2] Ferguson, T.S., Mathematical Statistics: A Decision Theoretic Approach, Academic Press, New York, 1967.
- [F3] Fry, T.C., Probability and Its Engineering Uses, Van Nostrand Co., Princeton, N.J., 1965.
- [F4] Feller, W., An Introduction to Probability Theory and Its Applications, 3rd Ed., Wiley, New York, 1957.
- [G1] Gilbert, E.N. and Pollack, H.O., "Amplitude Distribution of Shot Noise," Bell System Technical Journal, Vol. 39, No. 2, pp. 333-350.
- [H1] Helstrom, C.W., Statistical Theory of Signal Detection, Pergamon Press, Oxford, 1968.

- [H2] Hancock, J.C. and Wintz, P.A., Signal Detection Theory, McGraw-Hill, New York, 1966.
- [H3] Hancock, J.C. and Lainiotis, D.G., "Distribution-Free Detection Procedures," Purdue University Technical Report TR-EE-64-8, 1964.
- [H4] Hogg, R.V. and Craig, A.T., Introduction to Mathematical Statistics, 3rd Ed., Macmillan, London, 1970.
- [H5] Heathcote, C.R., Probability: Elements of the Mathematical Theory, Wiley, New York, 1971.
- [H6] Hájek, J., Nonparametric Statistics, Holden-Day, San Francisco, 1969.
- [H7] Hancock, J.C. and Wade, W.D., "A Digital Wide-Band Non-linear Receiver Capable of Near Optimum Reception in the Presence of Narrow-Band Interference," IEEE Trans. on Communication Systems, Vol. CS-11, No. 3.
- [K1] Kendel, M.G. and Stuart, A., The Advanced Theory of Statistics, Vol. 2, Hafner Publishing Co., New York, 1947.
- [K2] Kanefsky, M. and Thomas, J.B., "On Polarity Detection Schemes with Non-Gaussian Inputs," Journal of the Franklin Institute, Vol. 280, No. 2.
- [K3] Kullback, S., Information Theory and Statistics, Dover, New York, 1968.
- [L1] Lehmann, E.L., Testing Statistical Hypotheses, Wiley, New York, 1959.
- [N1] Noether, G.E., Elements of Nonparametric Statistics, Wiley, New York, 1967.
- [P1] Parzen, E., Modern Probability Theory and Its Applications, Wiley, New York, 1960.
- [R1] Rappaport, S.S. and Kurz, L., "An Optimal Nonlinear Detector for Digital Data Transmission Through Non-Gaussian Channels," IEEE Trans. on Communications Technology, Vol. COM-14, No. 3.
- [R2] Richard, R.H. and Gore, W.C., "A Nonlinear Filter for Non-Gaussian Interference," IEEE Transactions on Communication Systems, Vol. CS-11, No. 4.
- [R3] Root, W.L., "The Detection of Signals in Gaussian Noise," NASA, NSG-2-59, September 1965.



- [S1] Siegel, S., Nonparametric Statistics, McGraw-Hill, New York, 1956.
- [T1] Thomas, J.B., "Nonparametric Detection," Proceedings of the IEEE, Vol. 58, No. 5.
- [V1] Van Trees, H.L., Detection, Estimation, and Modulation Theory, Part I, Wiley, New York, 1968.
- [W1] Wilks, S.S., Mathematical Statistics, Wiley, New York, 1962.



APPENDICES



APPENDIX I

DERIVATION OF $J(\theta:2)$ VS. θ

The noise amplitude model can be written in the form $f_Y(y) = a e^{-b|y|^\theta}$, $-\infty < y < +\infty$, where the relationship between the parameters a , b and θ can be derived from the following equations:

$$\int_{-\infty}^{\infty} f_Y(y) dy = 1 \quad \text{and} \quad (A1)$$

$$\int_{-\infty}^{\infty} y^2 f_Y(y) dy = \sigma^2 . \quad (A2)$$

This form of noise was chosen because it realistically represents the amplitude distribution of a noise source consisting of an additive combination of Gaussian and impulse noise.

By the symmetry of $f_Y(y)$, equations (A1) and (A2) can be written

$$\int_0^{\infty} f_Y(y) dy = \frac{1}{2} \quad \text{and} \quad (A1)'$$

$$\int_0^{\infty} y^2 f_Y(y) dy = \frac{\sigma^2}{2} . \quad (A2)'$$

In equations (A1)' and (A2)' make the substitution $u = by^\theta$.

$$\int_0^{\infty} u^{\frac{1}{\theta}-1} e^{-u} du = \frac{1}{2a} \frac{\theta b}{\theta} . \quad (A3)$$

The left side is recognized as $\Gamma(\frac{1}{\theta})$. Therefore,

$$\Gamma\left(\frac{1}{\theta}\right) = \frac{\theta b^{\frac{1}{\theta}}}{2a} . \quad (\text{A4})$$

Equation (A2)' results in

$$\int_0^{\infty} u^{\frac{3}{\theta}-1} e^{-u} du = \frac{b^{\frac{3}{\theta}} \theta \sigma^2}{2a} . \quad (\text{A5})$$

Therefore,

$$\Gamma\left(\frac{3}{\theta}\right) = \frac{\theta b^{\frac{3}{\theta}} \sigma^2}{2a} . \quad (\text{A6})$$

Eliminating the variable a from equations (A4) and (A6) yields

$$b = \left(\frac{\Gamma\left(\frac{3}{\theta}\right)}{\sigma^2 \Gamma\left(\frac{1}{\theta}\right)} \right)^{\frac{\theta}{2}} . \quad (\text{A7})$$

The J-divergence can be used in Chapter I as a measure of discriminability between densities in the family $a \exp(-b|x|^{\theta})$, with $\theta \in [1,2)$ and a standard normal density ($\theta = 2$) [K3]. Let $f_{\theta}(x)$ represent the PDF of any member of the family, and $f_2(x)$ represent the PDF of a standard normal density. By definition

$$J(\theta:2) = \int_{-\infty}^{\infty} (f_{\theta}(x) - f_2(x)) \log \frac{f_{\theta}(x)}{f_2(x)} dx \quad (\text{A8})$$

$$\begin{aligned} J(\theta:2) &= \int f_{\theta}(x) \log a \sqrt{2\pi} \exp\left(\frac{x^2}{2} - b|x|^{\theta}\right) dx - \int f_2(x) \log a \sqrt{2\pi} \\ &\quad \exp\left(\frac{x^2}{2} - b|x|^{\theta}\right) dx \\ &= \log a \sqrt{2\pi} \int f_{\theta}(x) dx - \log a \sqrt{2\pi} \int f_2(x) dx + \int \left(\frac{x^2}{2} - b|x|^{\theta}\right) f_{\theta}(x) dx \\ &\quad - \int \left(\frac{x^2}{2} - b|x|^{\theta}\right) f_2(x) dx \\ &= b \int |x|^{\theta} (f_2(x) - f_{\theta}(x)) dx + \frac{1}{2} \left[\int x^2 f_{\theta}(x) dx - \int x^2 f_2(x) dx \right] . \end{aligned}$$



The latter term is zero because of the assumed equality of the second moment (unity) for all members $\theta \in [1,2]$.

$$J(\theta:2) = b \int |x|^\theta (f_2(x) - f_\theta(x)) dx . \quad (A9)$$

By symmetry of the PDF's, this is written

$$J(\theta:2) = \frac{2b}{\sqrt{2\pi}} \int_0^\infty x^\theta e^{-\frac{x^2}{2}} dx - 2ab \int_0^\infty x^\theta e^{-bx^\theta} dx .$$

By changing variables, this may be reduced to

$$\begin{aligned} J(\theta:2) &= \frac{2b}{\sqrt{2\pi}} \int_0^\infty (2y)^{\frac{\theta-1}{2}} e^{-y} dy - \frac{2a}{\theta} \int_0^\infty \left(\frac{y}{b}\right)^{\frac{1}{\theta}} e^{-y} dy \\ &= \sqrt{\frac{2}{\pi}} b 2^{\frac{\theta-1}{2}} \Gamma\left(\frac{\theta+1}{2}\right) - \frac{2a}{\theta} \left(\frac{1}{b}\right)^{\frac{1}{\theta}} \Gamma\left(\frac{1}{\theta} + 1\right) . \end{aligned} \quad (A10)$$

Substituting from equation (A7) for $b = \left[\frac{\Gamma(3/\theta)}{\Gamma(1/\theta)} \right]^{\theta/2}$ and using

$$\frac{2a}{\theta b} = \frac{1}{\Gamma(1/\theta)} \quad \text{from equation (A4) yields}$$

$$\begin{aligned} J(\theta:2) &= \frac{1}{\sqrt{\pi}} \left(\frac{2\Gamma(3/\theta)}{\Gamma(1/\theta)} \right)^{\frac{\theta}{2}} \Gamma\left(\frac{\theta+1}{2}\right) - \frac{\Gamma(1/\theta + 1)}{\Gamma(1/\theta)} \\ &= \frac{1}{\sqrt{\pi}} \left(\frac{2\Gamma(3/\theta)}{\Gamma(1/\theta)} \right)^{\frac{\theta}{2}} \Gamma\left(\frac{\theta+1}{2}\right) - \frac{1}{\theta} . \end{aligned} \quad (A11)$$

The value of θ that maximizes (A11) indexes the member of the family (1), Chapter I, which most differs from a standard normal, and which can be expected to produce the largest central limit theorem error.

The graph of $J(1:2)$ vs. θ is shown in Figure A1, and the maximum divergence occurs when $\theta = 1$, which indexes the Laplacian distribution.

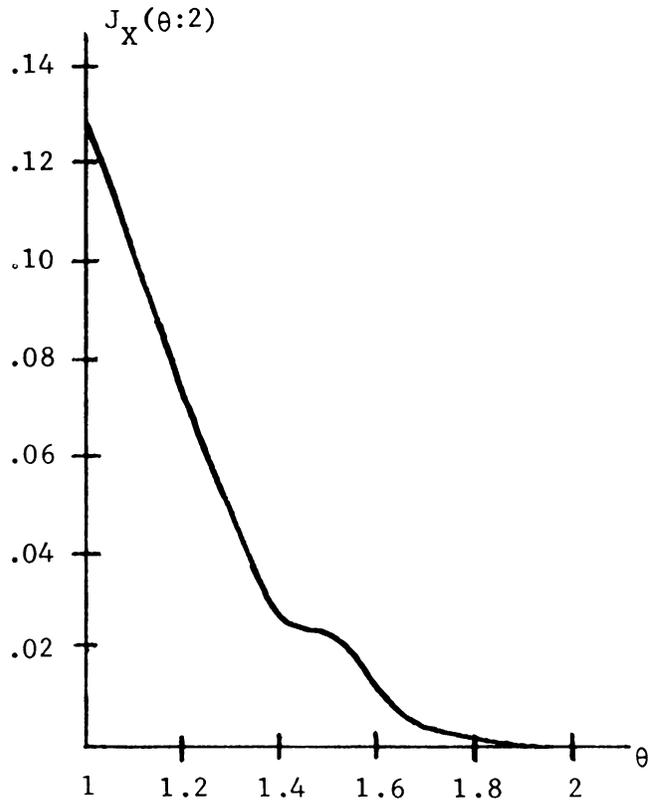


FIGURE A.1 Divergence vs. θ



APPENDIX II
REALIZATION OF THE NLF

The NLF of Section 3.2 was implemented on a digital computer using well-known techniques of piece-wise linear approximation. The odd symmetry property of the transfer characteristic considerably simplified the implementation. In 10 uniformly spaced intervals [See Figure A2] extending from $x = 0$ to $x = 3$, the equation of the chord connecting adjacent ordinate values was determined by the computer. This required the reading in of ordinate values at the 10 abscissa points. An asymptotic expression for values of $x > 3$ was based on a linear approximation using the slope of the chord between $x = 3$ and $x = 4$. Because of the concavity of the transfer characteristic, $g(x)$, this approximation tended to bias the NLF-detector as discussed in Section 4.4. A flow-chart is shown in Figure A2.

Let $F_n(x)$ be the piece-wise linear approximation to the continuous odd symmetric curve $g(x)$ derived in Section 3.2.

Define

- NTC: Number of piece-wise linear segments (10)
- TC(I): I-th positive ordinate of the curve $g(x)$, $I = 1, 2, \dots, \text{NTC}$.
- RANGE: Largest abscissa value (3)
- APSL : Slope and intercept of approximating straight line for
: abscissa values greater than RANGE.
- APIN : abscissa values greater than RANGE.
- AINC: Abscissa increment (.3)

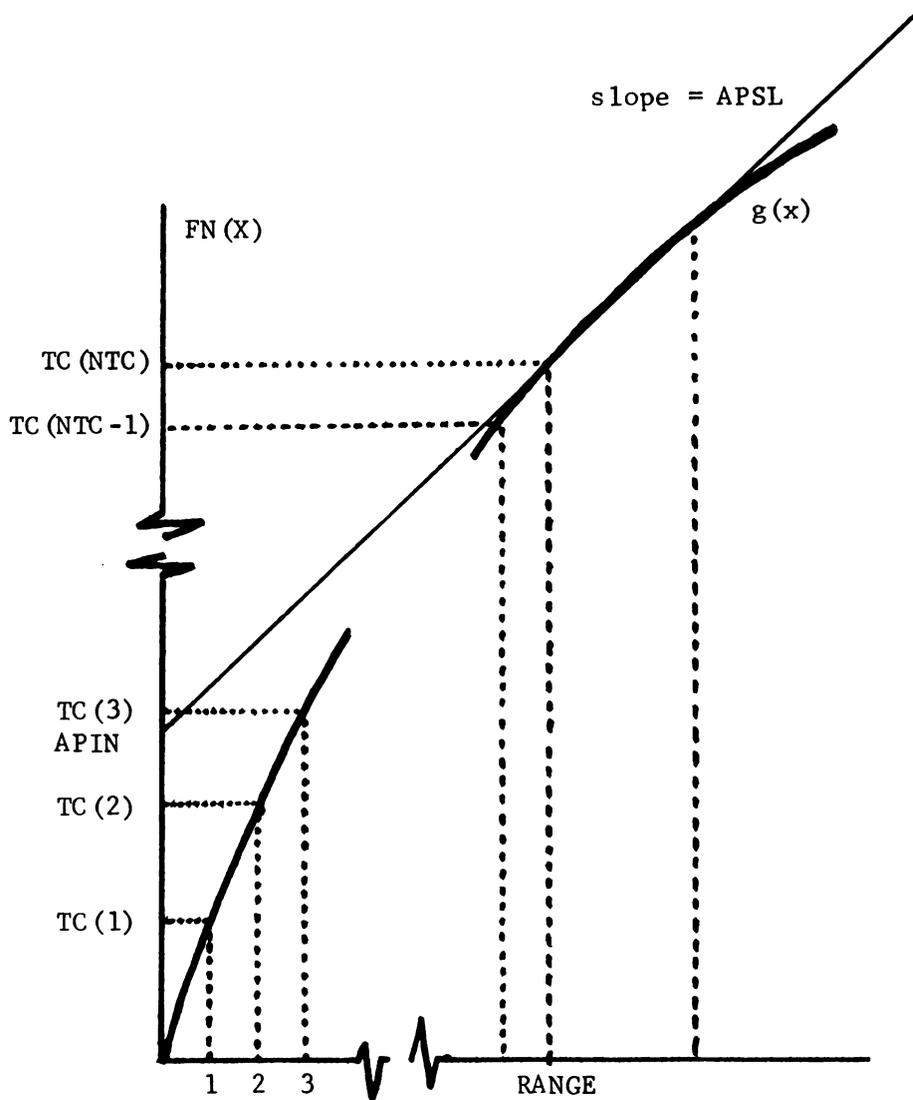


Figure A.2 Piece-Wise Approximation

The numbers in parenthesis are the values used during the simulation (See Section 4.4).

A flow chart is shown in Figure A3.

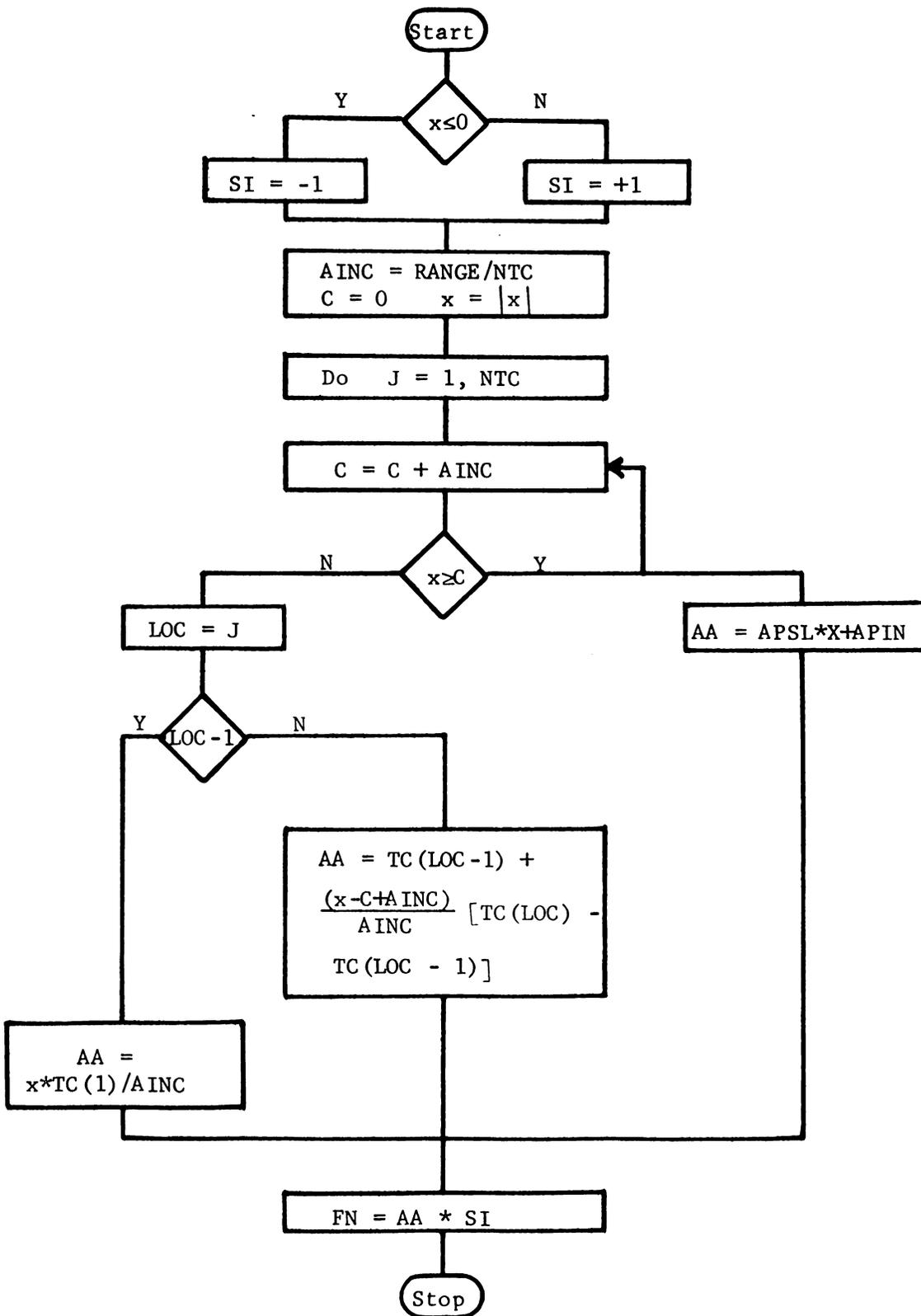


Figure A.3 Piece-Wise Approximation Flow Chart



APPENDIX III

DISCONTINUITY OF $g''(x)$ AT THE ORIGIN

The function $g(x)$ is shown in Chapter III, Figure 3.2. That the second derivative of $g(x)$, evaluated at the origin is discontinuous is shown here.

Lemma: Let $g(x)$ be an absolutely continuous function satisfying $g(x) = -g(-x)$. If $g(x)$ is convex for $x > 0$, then $g(-x)$ is concave on $x < 0$. That is, if $g''(x) \leq 0$ for $x > 0$, then $g''(-x) \geq 0$ for $x < 0$.

Proof:
 $g(-x) = -g(x)$
 $g'(-x) = +g'(x)$
 $g''(-x) = -g''(x)$

Then $g''(-x) = -(g''(x))$ and $g''(x)$ is ≤ 0 by hypothesis.

Then $g''(-x) \geq 0$. Q.E.D.

Corollary: All even derivatives, defined and continuous at $x = 0$, must vanish at $x = 0$.

Equation (20) of Chapter III gives an explicit expression for $g''(x)$, namely

$$g''(x) = \frac{g'(x)}{\sigma_0} (g(x)g'(x) - \sigma_0^2 b), \quad x > 0.$$

Thus,

$$g''(0+) = -b g'(0).$$

The term $g'(0)$ is easily found from (18) to be

$$g'(0) = \frac{\pi}{2} b .$$

Therefore

$$g''(0+) = -b^2 \frac{\pi}{2} \neq 0 .$$

However, by the above lemma

$$g''(0-) = b^2 \frac{\pi}{2} .$$

The second derivative approaches different limits at $x = 0$ from the left and right. Therefore, the second derivative is discontinuous at the origin.



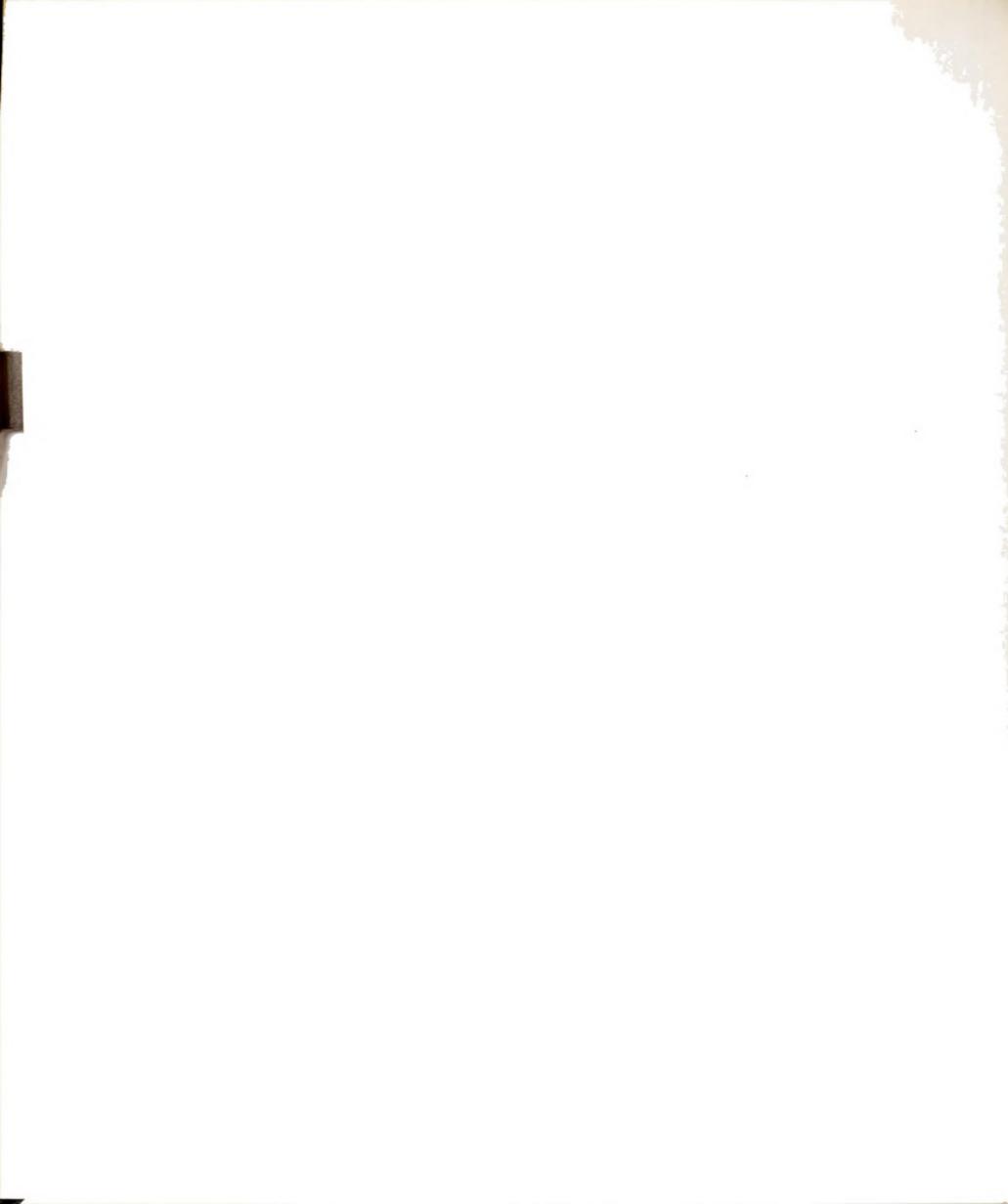
APPENDIX IV

Simulation Data

	Laplacian						Gaussian						
	SIGN		XBAR		NLF		SIGN		XBAR		NLF		
	FAR	P _D	FAR	P _D	FAR	P _D	FAR	P _D	FAR	P _D	FAR	P _D	
4	-40.	.092	.132	.080	.080	.120	.096	.1075	.1	.0975	.1125	.148	.100
	-12.4	.060	.244	.080	.232	.084	.236	.09	.1775	.0975	.225	.104	.220
	-6.6	.100	.336	.108	.356	.072	.452	.1	.24	.12	.38	.140	.420
	-3.1	.092	.480	.080	.520	.128	.584	.105	.39	.12	.5325	.108	.600
	-0.6	.092	.592	.088	.692	.080	.768	.11	.5375	.0875	.7225	.144	.716
	+1.3	.100	.740	.112	.904	.116	.892	.075	.6725	.095	.85	.128	.888
	-40.	.132	.088	.104	.144	.068	.076	.085	.1075	.1	.0875	.128	.144
	-12.4	.100	.292	.084	.248	.104	.328	.075	.195	.08	.2625	.124	.304
	-6.6	.128	.572	.128	.552	.096	.528	.1	.3975	.115	.475	.140	.580
	-3.1	.088	.664	.112	.780	.128	.852	.0825	.5125	.11	.75	.096	.784
	-0.6	.080	.792	.108	.892	.132	.928	.095	.755	.1325	.9275	.136	.912
	+1.3	.096	.900	.084	.968	.092	.996	.075	.85	.11	.9775	.144	.976
	-40.0	.104	.116	.100	.092	.136	.120	.145	.15	.1	.1175	.088	.156
	-12.4	.108	.420	.068	.332	.124	.316	.1	.305	.15	.3025	.092	.340
	-6.6	.144	.704	.120	.636	.104	.720	.0925	.5125	.13	.64	.112	.660
	-3.1	.112	.884	.096	.852	.088	.884	.13	.76	.1075	.8575	.136	.884
	-0.6	.116	.948	.100	.952	.104	.984	.13	.8975	.095	.9825	.108	.980
	+1.3	.108	.968	.084	.992	.128	.996	.12	.965	.09	.9925	.092	1.000

-40.0	.088	.072	.084	.100	.092	.124	.09	.0975	.1125	.105	.116	.120
-12.4	.104	.492	.116	.372	.096	.408	.0875	.28	.1075	.355	.092	.440
-6.6	.108	.712	.092	.700	.096	.772	.1125	.6	.06	.74	.112	.732
16	-3.1	.124	.084	.952	.116	.956	.1275	.805	.105	.94	.112	.944
-0.6	.088	.984	.084	.996	.076	.992	.1075	.9275	.075	.9875	.088	.988
+1.3	.092	1.000	.096	1.000	.092	1.000	.1175	.9825	.1	1.0	.104	1.000
-40.0	.084	.128	.088	.148	.108	.132	.068	.084	.12	.136	.080	.160
-12.4	.076	.504	.088	.464	.092	.574	.088	.308	.104	.384	.136	.468
-6.6	.080	.864	.108	.816	.124	.856	.1	.668	.144	.804	.108	.816
20	-3.1	.116	.104	.988	.124	.976	.072	.844	.088	.968	.112	.976
-0.6	.076	1.000	.112	.988	.092	1.000	.076	.98	.088	1.0	.100	1.000
+1.3	.108	1.000	.072	1.000	.076	1.000	.112	.996	.108	1.0	.128	1.000





MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03037 9683