ESC-TR-93-238

(1)

Technical Report
937

AD-A279 581

Adaptive Detection in Stationary and Nonstationary Noise Environments

P. Monticciolo

24 February 1994

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

Prepared for the Department of the Air Force under Contract F19628-90-C-0002.

This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Department of the Air Force under Contract F19628-90-C-0002.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Gary Tutungian
Administrative Contracting Officer
Directorate of Contracted Support Management

PAGES _____

ARE

MISSING

IN

ORIGINAL

DOCUMENT

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

ADAPTIVE DETECTION IN STATIONARY AND NONSTATIONARY NOISE ENVIRONMENTS

P. MONTICCIOLO
Group 44

TECHNICAL REPORT 937

24 FEBRUARY 1994

LEXINGTON MASSACHUSETTS

ABSTRACT

This report describes the statistical performance of several radar-based adaptive detection schemes in both stationary and nonstationary noise and interference environments. The detectors under study must be able to correctly determine the presence of a target in a range gate with a high degree of probability given that the probability of misclassification is a fixed small value. The hostile noise environment is assumed to consist of possibly time-varying, spatially correlated interference along with Gaussian background noise. In a typical radar environment, the mean value of the returned radar signal and the noise covariance matrix are unknown parameters; therefore, generalized likelihood ratio test procedures were used to develop decision rules that meet the Neyman-Pearson criterion. Three major cases of interest were examined. First, the single-pulse test developed by Kelly[1] is reviewed. The multiple-pulse return test case is extremely complicated and was divided into distinct detector forms: noncoherent and coherent. The performance of each detector is a function of the signal-to-noise ratio, the number of radar pulse returns used in the decision rule, and the quality of the covariance estimate.

[1]E.J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron Syst.*, **AES-22** (1986).

ACKNOWLEDGMENTS

I would like to acknowledge the financial support and computing facilities provided by MIT Lincoln Laboratory under the auspices of Ken Senne. I would like to thank John Proakis for his guidance and the opportunity to be associated with both Northeastern University and Lincoln Laboratory over the last several years. Finally, I would like to thank Ed Kelly who was always willing to lend his support, guidance, and friendship.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

This section establishes the basic premises of this investigation. The intent of this study is to quantify the statistical performance of several detectors derived from a generalized likelihood ratio (GLR) test. A general overview of the detection process is presented, and the terminology and assumptions used in the report are defined. Four separate cases are considered, and the general problem definition, solution techniques, and associated results are briefly discussed for each.

1.1 General Overview

This report addresses the development and analysis of radar-based adaptive decision rules. The particular concern is to determine whether the returned radar signal indicates the presence of a target in the range gate currently under observation. If detector output exceeds a threshold, certain actions such as tracking can be initiated. Because the returned signal includes a possible echo amplitude proportional to target radar cross-sectional area, directional interference, and Gaussian sky noise, a statistical approach to decision making is taken to minimize the probability of an incorrect decision.

It is important to note that the term "adaptive" used in the current context refers to the time dependency of the *actual parameter values* used in the decision rule. Adaptive does not refer to the time dependency of *parameter value updates* that occur in algorithms such as the least mean square and recursive least squares [1]. To expand on this terminology further, the scalar output of an adaptive array processor is defined as

$$y_i = \mathbf{w}_i^H \mathbf{x}_i \quad , \tag{1}$$

where $\mathbf{w}_i$ is the weight vector and $\mathbf{x}_i$ is the data vector.

At each sample instant the adaptive algorithms update the weight vector, the value of which is determined according to a specified cost function. This study deals purely with raw input data and does not form the adaptive array output. Estimates of the parameters found in the subsequent decision rules are time varying. Also, the decision rule contains forms similar to the optimum weight vector for the known covariance matrix case.

In almost all cases, this report is occupied with the presence of noise and interference that causes errors in the receiver output. For a system in which the signal is corrupted by noise, it is most desirable that the receiver minimizes the probability of error. Statistical decision theory can be used to obtain expressions for the probability of error, i.e., the probability of misclassifying the received binary signal, and the probability of detection (P_D) [2]. The misclassification error that is of most interest to radar system analysts is the probability of false alarm (P_{FA}). This error occurs when the decision rule indicates that a target is present when actually it is not. In practice, radar systems typically require a P_{FA} on the order of 10^{-6}.

Several additional levels of complication can be added to this scenario. Assuming white Gaussian receiver noise and a multiple-element receiver array, the probability density function (pdf) of the received data is uniquely specified by two parameters—the mean value vector and the covariance matrix. The mean and covariance matrix are also referred to as the "first- and "second-order" noise statistics. If one or both parameters are unknown and/or time varying, the analysis can be become extremely difficult, if not intractable. This report focuses on examining conditions where the noise environment is nonstationary, i.e., where the first- and/or second-order noise statistics are time varying. In addition, the true values of the parameters are unknown. Section 1.2 describes the general radar system model used in this investigation.

1.2 Radar System Model

A comprehensive review of basic radar principles and operations can be found in sources such as the *Radar Handbook* [3]. A radar system scans a three-dimensional region of space, the coordinates of which are azimuth, elevation, and range. The search volume is divided into segments called range bins or range gates. To detect a target, a sequence of modulated pulses is emitted and the system then listens for the returned echo. The time between transmission and echo reception corresponds to the total propagation delay that is a function of distance. The received data are processed and input to a decision rule to determine whether a target is present in the range bin under observation. If the output of the decision rule is greater than some predefined threshold with a value that is proportional to the desired false alarm rate, a "hit" is declared. In general, if a sufficient number of hits occurs in a range gate, further action is taken.

1.2.1 Array and Receiver Structure

The radar system under examination is an array with N elements that are connected to separate receiver channels. The array configuration is not specified but a generalized steering vector is denoted as the N-vector $\mathbf{p}$. The steering vector defines the directional orientation of a possible target and is determined from target position and array geometry. By not assuming any antenna characteristics or special array configurations, the problem can be analyzed in its most general form; however, for subsequent simulation studies a particular array geometry will be designated. The receiver section of the radar system is assumed ideally to demodulate, match filter, and sample the incoming planar narrowband waveform. The current model is idealized to concentrate solely on first-order effects.

If the target is present in the range bin currently under observation and designated as hypothesis H_1, the received signal consists of N_p modulated pulses plus noise. Under the null hypothesis,

designated as H_0, the received signal consists solely of noise. A sample of the matched filter outputs at time instant i, $i = 1, \ldots, N_p$, is denoted as the complex[2] N-vector $\mathbf{z}_i$. The vector $\mathbf{z}_i$ is referred to as a primary data vector. The samples of the matched filter outputs are assumed to be uncorrelated with respect to time, but there can be time-varying spatial correlations among the multiple receiver channels.

In radar applications the matched filter output, which directly corresponds to the complex returned signal amplitude, is not known a priori and as such must be treated as an unknown parameter. The analysis is restricted to the case where the returned signal amplitude is a nonrandom variable, i.e., the target is nonfluctuating; therefore, Swerling fluctuating target distribution models [4] are not included in this analysis. The complex returned signal amplitude is assumed constant over the N_p observation interval, implying a constant radar cross-sectional area. The results obtained here can be extended to the fluctuating case using the Swerling models at the expense of increased complexity.

1.2.2 Noise Environment

The statistical characterization of the received signal is now examined. The noise environment is assumed to be composed of Gaussian background noise and/or directional interference such as jamming. Therefore, the individual matched filter outputs are independent complex Gaussian random variables (r.v.'s). A complex Gaussian r.v. possesses a circularly symmetric pdf with the real and imaginary parts being independent and each part having a variance equal to $\sigma^2/2$. Thus a complex r.v. has variance σ^2 given the preceding definition. The N-element primary data vector $\mathbf{z}_i$ is defined to be complex multivariate Gaussian-distributed with its $N \times N$ covariance matrix denoted as $\mathbf{M}_i = \mathrm{E}\{(\mathbf{z}_i - \bar{\mathbf{z}}_i)(\mathbf{z}_i - \bar{\mathbf{z}}_i)^H\}$[5]. The expected value of $\mathbf{z}_i$ is $\mathbf{0}$ under H_0 and $b\,\mathbf{p}$ under H_1 for all i. The term b is the complex matched filter amplitude given that a target is present in the range bin, and $\mathbf{p}$ is the steering vector. As stated previously, the true value of b is unknown. The true covariance matrix is also unknown and possibly time varying. Hence, the notation $\mathbf{M}_i$ indicates the time-varying nature of the noise environment.

The time-varying nature of the second-order statistic implies nonstationarity. As a result, techniques that are dependent upon wide-sense stationarity assumptions cannot be applied [6]. In particular, ensemble averaging is not a valid technique if the noise environment is nonstationary. A nonstationary environment can be best illustrated by an example. The power associated with a directional interference source can fluctuate throughout the observation interval, and averages of the sample covariance matrix would not approach the true covariance matrix in the limit. Only over sampling intervals where the source intensity is constant will an estimate based upon averaging

---

[2]The matched filter output has in-phase and quadrature components that give rise to the complex representation.

be valid. An extreme case of a nonstationary environment is that of a blinking jammer where the jamming power is randomly switched ON and OFF. (This scenario is investigated in Section 4.)

In the sequel no assumptions are made about the structure of the covariance matrix. By not imposing any predefined covariance structure, e.g., the sum of dyads plus the identity matrix, an unconstrained maximization of the likelihood function can be performed with respect to this unknown covariance matrix. The lack of constraints will be shown to lead to the sample covariance matrix as the estimate of the true covariance matrix at time instant i.

1.2.3 The Decision Rule

Suppose a binary hypothesis test is to be performed on a "data" vector $\mathbf{x}$. The general form of the pdf of $\mathbf{x}$ is known but one or more of the true parameter values of the distribution are unknown. To determine whether the data belong to H_0 or H_1 , they are processed and the resulting output is compared to a threshold. The GLR test procedure is a reasonable solution to a problem of this form. Although not guaranteed to be optimal, the GLR test provides a general formulation for problems that contain unknown parameters [2].

The GLR is defined as the quotient of the joint pdf of the data vector $\mathbf{x}$, assuming hypothesis H_1 ; and the joint pdf of the data, assuming hypothesis H_0. The terms Θ_1 and Θ_0 are defined as the vectors of unknown parameters under the respective hypotheses. The numerators and denominators of the ratio (the likelihood functions) are then maximized independently with respect to the unknown parameters:

$$\mathrm{GLR} = \frac{\sup\limits_{\Theta_1} \; p_1(\mathbf{x}\,;\Theta_1|H_1)}{\sup\limits_{\Theta_0} \; p_0(\mathbf{x}\,;\Theta_0|H_0)} \quad . \tag{2}$$

The maximum likelihood estimates (MLEs) are the terms that maximize the respective joint pdf's in Equation (2). For example, the MLE of the mean assuming a Gaussian pdf is simply the sample mean. The MLE of each unknown parameter is then substituted into the likelihood ratio in place of the true parameter value. Thus a decision rule is obtained. Using $\mathbf{x}$ as the input, the output of the decision rule is compared to a threshold. If the output exceeds the threshold, $H_1$is chosen, otherwise the null hypothesis is chosen.

The statistical performance of the GLR test is measured in terms of P_D and P_{FA}; in many cases it is quite difficult to obtain their exact expressions, and one may resort to approximations or simulations to obtain these measures of performance.

1.3 Case Studies

This section previews the four cases to be examined. For each case, a brief overview is presented and pertinent assumptions or approximations are stated. The general form of the solution

and important results are discussed. This information can be found in Table 1 and complete details can be found in the appropriate sections.

TABLE 1

Case Study Summary

Case	No. of Pulses	Noise Environment	Comments
1	1	N/A	P_D and P_{FA} expressed in terms of closed-form expressions
2	N_p	Stationary	Can be transformed to Case 1 via unitary transformation
3	N_p	Nonstationary	Noncoherent structure, exact P_{FA} expression, P_D via simulation
4	N_p	Nonstationary	Coherent structure, P_D and P_{FA} via simulation

1.3.1 Case 1: The Single Pulse

This section examines the GLR test-based decision rule for one data sample. This case has been extensively studied by Kelly [7,8]. These results will be completely reproduced in Section 2 to provide a framework for subsequent analyses. Exact expressions for the P_D and the P_{FA} are obtained for this case. The P_D is conveniently expressed as a function of signal-to-noise ratio (SNR) and the threshold value corresponding to a specified P_{FA} is evaluated using the incomplete gamma function [9].

Some basic terminology is now defined. A sample of the N matched filter outputs corresponding to single return is denoted as the N-vector $\mathbf{z}$. The expected value of the complex multivariate Gaussian r.v. $\mathbf{z}$ under H_1 is $b\mathbf{p}$ and $\mathbf{0}$ under H_0. The complex returned signal amplitude b is unknown. In addition, the covariance matrix under each hypothesis is unknown. The true covariance matrices are designated as $\mathbf{M}_1$ and $\mathbf{M}_0$, respectively. These unknown parameters must be estimated according to GLR test methodology.

From basic detection and estimation theory, the MLE of an unknown parameter asymptotically approaches the true parameter value, i.e., it is a consistent estimator. For the multivariate Gaussian distribution, the MLE of the covariance matrix is the sample covariance matrix that is a consistent estimator. In many cases, the decision rule obtained from GLR test procedures is identical in form to a decision rule obtained when the covariance matrix is assumed known except that the sample covariance matrix replaces the true covariance matrix in the decision rule. Therefore, the "quality" of the covariance estimate will play a key role in detector performance in that a

"better" estimate will enable the GLR-based detector performance to approach that of the known covariance case.

In an attempt to improve detector performance in cases where the true covariance matrix is unknown, Kelly proposed augmenting the primary data vector with K additional data samples that are assumed to be free of signal returns. These additional samples are taken from K range gates that are in close proximity of the range gate currently under observation. These supplementary vectors are defined as secondary vectors and are denoted as $\mathbf{z}(k)$, $k = 1, \ldots, K$. There are now $K + 1$ terms available for covariance matrix estimation at a particular time instant.

Given sufficient range and spatial resolution, it is not unreasonable to assume that multiple targets would be sparsely distributed. As a result, each N-element secondary vector is assumed to be zero-mean and possess the same covariance matrix as the primary vector, and each secondary vector is statistically independent of the primary vector and the other $K - 1$ secondaries. These assumptions are realistic and significantly reduce the complexity of the subsequent analysis.

The GLR test for the single-pulse case is now formulated. The respective likelihood functions are the joint pdf of the primary vector and the K secondary vectors under under each hypothesis. Under H_1, the joint pdf is the product of the primary vector nonzero-mean multivariate Gaussian pdf and K identical zero-mean multivariate Gaussian pdf's. Under H_0, the joint pdf is simply the product of $K + 1$ zero-mean multivariate Gaussian pdf's.

The GLR test procedure requires that the unknown parameters be replaced by their MLEs. First, each likelihood function is maximized with respect to the unknown covariance matrix while holding the returned signal amplitude fixed. The resulting MLEs are the sample covariance matrices. Note that the GLR test procedure leads directly to the sample covariance matrix form, which is an unbiased estimator and asymptotically approaches the true covariance. Therefore, the inclusion of the K additional secondary vectors yields a more accurate estimate of the covariance.

Estimates of the covariance matrices are substituted into the likelihood functions in place of the true covariance matrices. Given the multivariate Gaussian distribution assumption, the likelihood ratio at this point is the quotient of the determinant of the MLE of the covariance under H_0 and the determinant of the MLE of the covariance matrix under H_1. The MLE of the covariance matrix under H_1 is still a function of the unknown returned signal amplitude. This problem is now equivalent to the Wilks' Lambda test commonly found in statistical literature [10]. For this case, the MLE of b is equivalent in form to the minimum variance unbiased estimator of the returned amplitude. Substituting the MLE of b into the test yields the test form that will be the basis of the single-pulse analysis.

Under H_1, the final form of the single-pulse test is statistically equivalent to the inverse of a noncentral β r.v. conditioned on the β-distributed loss factor ρ. This loss factor arises from the need to estimate the covariance matrix. Reed, Mallet, and Brennan have extensively studied this r.v. and have shown that estimating the covariance matrix effectively reduces the available SNR to the detector, resulting in performance degradation [11]. Section 2 shows that as the number of secondary vectors used for covariance estimation increases, estimation loss decreases

and detector performance approaches that of the known covariance (matched filter) case. Under the null hypothesis, the test is statistically equivalent to a central β r.v. Also, exact expressions for the P_{FA} and P_D are derived in Section 2.

1.3.2 Case 2: Multiple Pulses in a Stationary Environment

The test for this case uses N_p observation intervals to make a decision. By making the following assumptions, this problem can be transformed into a structure identical to that described in case 1. First, the complex returned amplitude b is assumed constant for the duration of the test. Second, for this case a wide-sense stationary noise environment is assumed over the entire observation interval; therefore, the second-order noise statistic, the true covariance, does not vary as a function of time. The data are assumed to be statistically independent of the noise and interference.

Suppose an $N \times N_p$ matrix is formed, consisting of N_p statistically independent primary vectors and collect K secondary data vectors during the observation intervals. It can be shown that postmultiplication of this primary data matrix by an appropriate data-dependent unitary transformation results in the rotation of the mean value component of each N_p primary vector into the primary data matrix first column position [12]. The remaining $N_p - 1$ primary vectors are now signal-free or equivalently, they are zero-mean vectors. Thus the $N_p$pulse test has been transformed into the single-pulse test form but with $K + N_p - 1$ zero-mean secondary vectors. The same P_D and P_{FA} formulas as derived for case 1 can be applied.

For large N_p, the additional $N_p - 1$ vectors used for covariance estimation will significantly reduce estimation loss. Of course, secondaries obtained from each observation interval ($N_p \times K$) can be used for the covariance estimate, but the additional memory and computational requirements could be expensive. The stationarity assumption allows the primary data vectors themselves to be used to improve the covariance estimate.

Because the case 2 test can be transformed into a previously solved problem, the analysis is included in Section 2. In either case, improved covariance estimation results in improved detector performance.

1.3.3 Multiple Pulses in a Nonstationary Environment

When considering the multiple-pulse test, a nonstationary noise environment rules out the solution technique used for case 2. A unitary transformation cannot be applied to the primary data matrix because each vector is no longer statistically identical; therefore, the discussion continues with the general formulation of the GLR for multiple-pulse conditions where there are numerous unknown parameters.

The sampled data are assumed statistically independent. Due to this assumption, the likelihood functions used to form the GLR are simply the product of the N_p likelihood functions associated with each data sample. If the noise is Gaussian, each of the N_p likelihood functions

takes on the form of the multivariate Gaussian distribution. GLR test procedures indicate that the likelihood function for the complete data set under each hypothesis must then be maximized with respect to each unknown parameter. In the case where each unknown parameter varies as a function of time, maximization of the overall likelihood function is equivalent to maximizing each of the N_p likelihood functions independently. If one or more of the unknown parameters remains constant, however, the maximization process can become quite difficult or nearly impossible.

For the problem under investigation, the true covariance matrix is assumed to be unknown and time varying. As previously discussed, the signal amplitude is constant over the complete observation interval. The MLE of each of the N_p covariance matrices can be evaluated by maximizing the likelihood function under each hypothesis at each time instant. The MLE of each covariance matrix is the sample covariance matrix obtained using the primary vector and the K secondary vectors; however, because the signal amplitude is constant, the maximization of the GLR with respect to this unknown parameter requires the solution of a $2N_p$-degree polynomial. For even a relatively small N_p, the problem is unyielding.

Two possible techniques were developed to deal with this problem. First, allow the model for the returned amplitudes be time varying. This modeling assumption enables the likelihood ratio to be expressed as the product of N_p independent ratios, which results in the same form as the ratio found in Equation (2). This reformulation of the problem is the basis of the case 3 analysis. Second, do not allow the returned amplitude model to be time varying (the actual condition for a constant radar cross-sectional area). After maximizing the test with respect to the N_p unknown covariances, the test can be expressed in terms of a series expansion using an algebraic identity. The series expansion approach is the basis of the case 4 analysis. These two cases are now discussed in greater detail.

Case 3 : Noncoherent Detection Using Multiple Pulses. By allowing the model for the signal amplitude to be time varying, the decision rule can be expressed as the product of N_p of the likelihood ratios described in case 1 [13]. This assumption enables independent maximization of each N_p likelihood function with respect to the unknown parameters. The GLR for this case is statistically equivalent to the product of N_p independent β-distributed r.v.'s. Section 3 will show that the test is equivalent to a noncoherent integration process. Performance of this noncoherent detector is expected to approach the matched filter performance level as $K \rightarrow \infty$.

To evaluate noncoherent detector performance, the threshold value corresponding to a specified P_{FA} must be determined. Fortunately, using the transformation of variable technique [6], the product of independent central β r.v.'s with a special set of degrees-of-freedom parameters can be transformed into a central χ^2 r.v. with N_p degrees of freedom. The threshold value is obtained using the incomplete gamma function [14].

The P_D for this test presents considerable difficulties. For $N_p > 2$, it is virtually impossible to obtain an exact expression for the distribution of the product of noncentral β r.v.'s. A Chernoff bound for this distribution was obtained and provided an analytical means for evaluating performance. This approach is quite different from the approximate χ^2 techniques preferred by

statisticians when analyzing Wilks' Lambda tests [15]. A Monte Carlo simulation was used to obtain experimental results. As expected, the analysis revealed that noncoherent detector performance improved as a function of N_p and K.

Case 4: Coherent Detection Using Multiple Pulses. Unlike case 3, the model for the returned signal amplitude is held constant over the observation interval. The GLR test, as per case 3, can be maximized with respect to the N_p unknown covariance matrices; however, the maximization of the GLR with respect to b is almost impossible. This problem can be solved using an approximation. The GLR is expressed in terms of a truncated power series expansion. The approximate GLR can then be maximized with respect to b. After some manipulation, the final approximate GLR can be expressed as the product of a term that is a function of the amplitude estimate and the GLR obtained for case 3. This result is defined as a coherent decision rule because the estimate of b incorporates data from the complete N_p pulse observation interval. The statistical analysis of this decision rule, however, remains daunting and analytical attempts were abandoned in favor of Monte Carlo simulations.

Several conditions were examined via simulation. As a performance benchmark, the coherent detector was tested in the less severe stationary noise environment. Given a reasonable number of secondary vectors, detector performance exceeded that of the noncoherent, known covariance detector. Under nonstationary conditions, the coherent detector also exceeded the performance of the noncoherent detector. In fact, coherent detector performance decreased only when the number of covariance estimates during an N_p pulse observation interval was reduced. Among the noise environments used for test purposes were the blinking jammer and multiple jammer scenarios.

1.4 Report Outline

The remainder of this report is organized as follows: Section 2 presents the single-pulse analysis based upon the GLR test. Because the multiple-pulse stationary case can be reduced to the form of case 1, its analysis is incorporated here. Section 3 explores the performance of the noncoherent detector, including a complete derivation of the Chernoff bound. The coherent detector that is based upon a series expansion is examined in Section 4. The final section summarizes the results and conclusions for all cases (see Table 1).

2. SINGLE-PULSE ADAPTIVE DETECTOR

In the course of deriving the GLR test, basic terminology is presented. The resulting test is shown to be statistically equivalent to 1 plus a conditional F-distributed r.v. Using a relationship between F and β r.v.'s, the inverse of the test statistic is shown to possess a conditional β distribution. Using the cumulative distribution function of the β distribution, exact expressions for the P_D and P_{FA} are obtained, and the effect of covariance estimation on detector performance is discussed [7].

2.1 Single-Pulse GLR Test

The analysis begins by defining the complex multivariate Gaussian pdf that characterizes the primary and secondary data vectors. As stated in Section 1, the real and imaginary parts of each element of the data vectors are independent. Under the signal plus noise hypothesis, the mean of the primary data vector is $b\mathbf{p}$; and under the noise-only hypothesis the mean vector is $\mathbf{0}$. The pdf of the primary vector under H_1 is

$$p_1(\mathbf{z}|H_1) = \frac{1}{\pi^N \|\mathbf{M}_1\|} e^{-(\mathbf{z}-b\mathbf{p})^H \mathbf{M}_1^{-1} (\mathbf{z}-b\mathbf{p})} \quad , \tag{3}$$

where

$$\mathbf{M}_1 = \left\{ E(\mathbf{z} - b\mathbf{p})(\mathbf{z} - b\mathbf{p})^H \right\} \quad ,$$

b is the complex returned signal amplitude, and $\mathbf{p}$ is the steering vector. Without loss of generality, the steering vector is defined to have unit length.

The primary vector pdf under H_0 is

$$p_0(\mathbf{z}|H_0) = \frac{1}{\pi^N \|\mathbf{M}_0\|} e^{-\mathbf{z}^H \mathbf{M}_0^{-1} \mathbf{z}} \quad , \tag{4}$$

where

$$\mathbf{M}_0 = E\left\{ \mathbf{z}\mathbf{z}^H \right\} \quad .$$

The secondary vectors, $\mathbf{z}(k), k = 1, \ldots, K$, are also zero-mean and possess the same pdf as the primary vector under H_0; $\mathbf{M}_1$ and $\mathbf{M}_0$ are assumed unknown.

From previous studies where covariance matrices are assumed unknown, detector performance was shown to be directly related to the quality of the covariance estimate [7,11]. Because the sample covariance matrix is the estimator form obtained from the GLR test procedure, this estimate will asymptotically approach the true value as the number of sample vectors used to form the estimate increases. Therefore, the covariance matrix estimator should use the maximum number of available samples.

The joint pdf of the primary and K secondary vectors forms the basis for the GLR procedure. Because nearby range gates are assumed to share the same covariance matrix, $K + 1$ data vectors possessing the same covariance are now available to form a decision rule. The primary and secondary vectors are statistically independent; hence, the joint pdf of the primary vector and the K secondary vectors is simply the product of the individual pdf's. The joint pdf is denoted as

$$\mathrm{p}_i\{\mathbf{z}, \mathbf{z}(1), \ldots, \mathbf{z}(K)|\mathrm{H}_i\}, i = 0, 1 \quad . \tag{5}$$

Forming the ratio of the respective joint pdf's, the single-pulse likelihood ratio is

$$\Lambda = \frac{\mathrm{p}_1\{\mathbf{z}, \mathbf{z}(1), \ldots, \mathbf{z}(K)|\mathrm{H}_1\}}{\mathrm{p}_0\{\mathbf{z}, \mathbf{z}(1), \ldots, \mathbf{z}(K)|\mathrm{H}_0\}} \quad . \tag{6}$$

The numerator and denominator pdf's are referred to as the "likelihood functions."

Equation (6) is a function of the unknown parameters b, $\mathbf{M}_1$, and $\mathbf{M}_0$. Because of the presence of unknown parameters, the standard GLR test procedure is used to arrive at a decision rule. The GLR expressed in terms of the multivariate Gaussian pdf is defined as:

$$\Lambda(\mathbf{M}_1, \mathbf{M}_0, b) = \frac{\sup\limits_{\mathbf{M}_1, b} C_1^{K+1} e^{-\left\{(\mathbf{z}-b\mathbf{p})^H \mathbf{M}_1^{-1}(\mathbf{z}-b\mathbf{p})+\mathbf{Q}\right\}}}{\sup\limits_{\mathbf{M}_0} C_0^{K+1} e^{-\left\{\mathbf{z}^H \mathbf{M}_0^{-1}\mathbf{z}+\mathbf{Q}\right\}}} \quad , \tag{7}$$

where

$$C_i = \frac{1}{\pi^N \|\mathbf{M}_i\|}, i = 0, 1 \quad ,$$

and

$$Q = \sum_{i=1}^{K} \mathbf{z}^H(k)\mathbf{M}_0^{-1}\mathbf{z}(k) \quad .$$

For mathematical simplicity, apply the inner product identity $\mathbf{x}^H\mathbf{Y}^{-1}\mathbf{x} = \mathrm{Tr}\{\mathbf{Y}^{-1}\mathbf{x}\mathbf{x}^H\}$ to the $(K+1)$th root of Equation (7). Using this identity yields

$$\Lambda^{\frac{1}{K+1}} = \frac{\sup_{\mathbf{M}_1 b} C_1 \; e^{-\mathrm{Tr}\left[\mathbf{M}_1^{-1}\frac{1}{K+1}\{(\mathbf{z}-b\mathbf{p})(\mathbf{z}-b\mathbf{p})^H+\mathbf{S}\}\right]}}{\sup_{\mathbf{M}_1} C_0 \; e^{-\mathrm{Tr}\left[\mathbf{M}_0^{-1}\frac{1}{K+1}\left\{\mathbf{z}\mathbf{z}^H+\mathbf{S}\right\}\right]}} \quad , \tag{8}$$

where

$$\mathbf{S} = \sum_{i=1}^{K} \mathbf{z}(k)\mathbf{z}^H(k) \quad .$$

Λ is now redefined as the $(K+1)$th root of the ratio.

The GLR test procedure requires that the MLEs of the unknown parameters under each hypothesis be evaluated and then substituted into the test in place of the true parameter values. The preceding expression is solved to obtain the MLEs of the unknown covariance matrix. The MLEs are denoted as $\hat{\mathbf{M}}_1$ and $\hat{\mathbf{M}}_0$, respectively. From Equation (8), the MLEs of the covariance matrix under each hypothesis are

$$\hat{\mathbf{M}}_1(b) = \frac{1}{K+1}\left\{(\mathbf{z}-b\mathbf{p})(\mathbf{z}-b\mathbf{p})^H + \mathbf{S}\right\} \tag{9}$$

and

$$\hat{\mathbf{M}}_0 = \frac{1}{K+1}\left\{\mathbf{z}\,\mathbf{z}^H + \mathbf{S}\right\} \quad . \tag{10}$$

As stated in Section 1, the sample covariance form found in Equations (9) and (10) arises from the problem formulation, not from a priori assumption.

The MLE $\hat{\mathbf{M}}_1$ is still a function of the unknown returned amplitude that must be subsequently estimated. After substituting the MLEs of the covariance matrices into their respective likelihood functions, the GLR reduces to the following ratio of determinants:

$$\Lambda(b) = \frac{\|\hat{\mathbf{M}}_0\|}{\min_b \|\hat{\mathbf{M}}_1\|} \quad . \tag{11}$$

The process of minimizing the denominator is equivalent to maximizing the likelihood ratio. Using the following determinant identities,

$$(K+1)^N \|\hat{\mathbf{M}}_1\| = \|\mathbf{S}\|[1 + (\mathbf{z} - b\mathbf{p})^H \mathbf{S}^{-1}(\mathbf{z} - b\mathbf{p})] \tag{12}$$

and

$$(K+1)^N \|\hat{\mathbf{M}}_0\| = \|\mathbf{S}\|[1 + \mathbf{z}^H \mathbf{S}^{-1}\mathbf{z}] \quad , \tag{13}$$

and completing the square with respect to b, results in an expression where the unknown parameter b is isolated. The GLR can be written as

$$\Lambda(b) = \frac{1 + \mathbf{z}^H \mathbf{S}^{-1}\mathbf{z}}{\min_b \left\{ 1 + \mathbf{z}^H \mathbf{S}^{-1}\mathbf{z} + \mathbf{p}^H \mathbf{S}^{-1}\mathbf{p} \left| b - \frac{\mathbf{p}^H \mathbf{S}^{-1}\mathbf{z}}{\mathbf{p}^H \mathbf{S}^{-1}\mathbf{p}} \right|^2 - \frac{\left|\mathbf{p}^H \mathbf{S}^{-1}\mathbf{z}\right|^2}{\mathbf{p}^H \mathbf{S}^{-1}\mathbf{p}} \right\}} \quad . \tag{14}$$

The denominator in Equation (14) is minimized when $b = \hat{b}$, where $\hat{b}$ is

$$\hat{b} = \frac{\mathbf{p}^H \mathbf{S}^{-1}\mathbf{z}}{\mathbf{p}^H \mathbf{S}^{-1}\mathbf{p}} \quad . \tag{15}$$

The MLE $\hat{b}$ is identical in form to the minimum variance unbiased estimator of b given that $1/K$ $\mathbf{S}$ replaces $\mathbf{M}$. The estimate is essentially produced by whitening and cross-correlating the data vector with the steering vector. The substitution of $\hat{b}$ into the GLR yields

$$\Lambda = \frac{1 + \mathbf{z}^H \mathbf{S}^{-1}\mathbf{z}}{1 + \mathbf{z}^H \mathbf{S}^{-1}\mathbf{z} - \frac{\left|\mathbf{p}^H \mathbf{S}^{-1}\mathbf{z}\right|^2}{\mathbf{p}^H \mathbf{S}^{-1}\mathbf{p}}} \gtrless \alpha \quad , \tag{16}$$

where α is a threshold parameter.

Before deriving the statistics of the GLR test, an attempt was made to develop an intuitive feel for the behavior of the test. As found in Equation (16), the test is at a maximum when

the denominator term takes on its minimum value. The conditions under the signal-plus-noise hypothesis are examined. By setting the primary data vector z equal its expected value under H_1, $b\mathbf{p}$, and performing some algebraic manipulation, the denominator reduces to the scalar value 1, which is the minimum value of the denominator for any data vector $\mathbf{z}$. If the target is present and is collinear in space with the steering vector, the best chance of making the correct decision is at hand. The numerator term under this assumption takes on the form $1 + |b|^2\mathbf{p}^H\mathbf{S}^{-1}\mathbf{p}$. As the sample covariance approaches the true covariance, $|b|^2\mathbf{p}^H\mathbf{S}^{-1}\mathbf{p}$ becomes in the limit $|b|^2\mathbf{p}^H\mathbf{M}^{-1}\mathbf{p}$, which is the nonadaptive SNR of the array except for the factor $1/K$. Therefore, as the covariance estimate improves the SNR required to achieve a certain P_D decreases.

Alternatively, the presence of the sample covariance in the decision rule results in the actual SNR seen by the detector to be less than the optimal nonadaptive SNR. If the estimation loss factor is defined to be the random variable ρ, the performance of the test can be evaluated in terms of the product of the nonadaptive SNR and the complex β-distributed loss factor ρ. This loss factor was derived in Reed, Mallett, and Brennan [11], and its effect on the performance of the decision rule was first studied in Kelly [7]. The loss factor plays a critical role in the performance of the adaptive detector.

2.2 Statistical Representation of the Single-Pulse GLR Test

The analysis of the performance of the single-pulse GLR test is based upon several standard multivariate statistical techniques [16–18]. The primary and secondary data vectors are whitened so that each of the components of the N-element data vectors are statistically independent. The space spanned by the data vectors is then divided into two orthogonal subspaces of dimension 1 and $N-1$, respectively. A unitary transformation is then applied to the data vectors to rotate the mean value of the data completely into the first element of the data vectors. By fixing the components of the $N-1$ element subspace, the conditional distribution of the single-element subspace is obtained. After removing the conditioning effect through averaging, the unconditional distribution of the test provides the basis for the P_D and P_{FA} expressions.

For clarity, the GLR test is restated:

$$\Lambda = \frac{1 + \mathbf{z}^H\mathbf{S}^{-1}\mathbf{z}}{1 + \mathbf{z}^H\mathbf{S}^{-1}\mathbf{z} - \dfrac{\left|\mathbf{p}^H\mathbf{S}^{-1}\mathbf{z}\right|^2}{\mathbf{p}^H\mathbf{S}^{-1}\mathbf{p}}} \gtrless \alpha \quad , \tag{17}$$

and the analysis proceeds by whitening the data vectors. An appropriate choice for a whitening transformation is the inverse of the square root of the covariance matrix. Because the covariance matrix (and its inverse) is positive definite, a positive definite square root matrix $\mathbf{M}^{-1/2}$ can be introduced. By defining the whitened primary vector vector as $\tilde{\mathbf{z}} = \mathbf{M}^{-1/2}\,\mathbf{z}$, then $\mathrm{E}\tilde{\mathbf{z}}\,\tilde{\mathbf{z}}^H = \mathrm{E}\left\{\left(\mathbf{M}^{-1/2}\,\tilde{\mathbf{z}}\right)\left(\mathbf{M}^{-1/2}\,\tilde{\mathbf{z}}\right)^H\right\} = \mathbf{I}$. Thus the choice of $\mathbf{M}^{-1/2}$ as a transformation matrix results in

a statistically independent data set. Likewise, a similar definition of the whitened secondary vector $\bar{\mathbf{z}}(k)$ yields the same result. The steering vector is also whitened to obtain $\bar{\mathbf{p}} = \mathbf{M}^{-1/2}\mathbf{p}$.

A unitary transformation $\mathbf{U}_R$ is defined such that the whitened steering vector $\bar{\mathbf{p}} = \mathbf{M}^{-1/2}\mathbf{p}$ is transformed into a basis vector with 1 as the first coordinate and 0 as the remaining $N-1$ elements. This rotation can be accomplished by selecting the first row of $\mathbf{U}_R$ to be a normalized unit vector and the other rows of $\mathbf{U}_R$ to be orthogonal to the first row and to each other, i.e., $\mathbf{u}_i^H \mathbf{u}_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta function. The whitened, rotated steering vector is defined as

$$\mathbf{e} = \mathbf{U}_R \bar{\mathbf{p}} = [1 \;\; 0 \cdots 0]^H \quad . \tag{18}$$

The test statistic is unchanged by a unitary transformation because it does not change a vector norm. The premultiplication of $\bar{\mathbf{z}}$ and $\bar{\mathbf{z}}(k)$ by $\mathbf{U}_R$ affects only the mean of $\bar{\mathbf{z}}$. To simplify notation, the whitened, rotated primary and secondary data vectors are redefined in terms of the original notation as the vectors $\mathbf{z}$ and $\mathbf{z}(k)$, respectively. The GLR in terms of the whitened, rotated vectors is

$$\Lambda = \frac{1 + \mathbf{z}^H \mathbf{S}^{-1}\mathbf{z}}{1 + \mathbf{z}^H \mathbf{S}^{-1}\mathbf{z} - \dfrac{\left|\mathbf{e}^H \mathbf{S}^{-1}\mathbf{z}\right|^2}{\mathbf{e}^H \mathbf{S}^{-1}\mathbf{e}}} \quad , \tag{19}$$

where now

$$\mathbf{S} = \mathbf{U}_R \; \bar{\mathbf{S}} \; \mathbf{U}_R{}^H \quad ,$$

and

$$\mathrm{E}\{\mathbf{z}\} = b \, \mathbf{U}_R \; \mathbf{M}^{-1/2}\mathbf{p} \quad .$$

The data vectors are partitioned into an A component that corresponds to the first element and a B component that corresponds to the remaining $N-1$ elements. The partitioned primary and secondary vectors are

$$\mathbf{z} = [\mathbf{z}_A \mathbf{z}_B]^T = \mathbf{U}_R \, \bar{\mathbf{z}} \quad , \tag{20}$$

where $\mathbf{z}_A$ is a scalar, $\mathbf{z}_B$ is a $1 \times (N-1)$ vector, and

$$\mathbf{z}(k) = [\mathbf{z}_A(k) \;\; \mathbf{z}_B(k)]^T = \mathbf{U}_R \; \bar{\mathbf{z}}(k) \quad .$$

The choice of the transformed steering vector $\mathbf{p}$ results in this particular partitioning. The whitened, rotated $\mathbf{S}$ matrix, which is K times the sample covariance matrix based on the secondaries alone, is partitioned as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{AA} & \mathbf{S}_{AB} \\ \mathbf{S}_{BA} & \mathbf{S}_{BB} \end{bmatrix} \quad , \tag{21}$$

where

$$\mathbf{S}_{AA} = \sum_{i=1}^{K} \mathbf{z}_A(k)\mathbf{z}_A{}^H(k) \quad ,$$

$$\mathbf{S}_{AB} = \sum_{i=1}^{K} \mathbf{z}_A(k)\mathbf{z}_B{}^H(k) \quad ,$$

$$\mathbf{S}_{BA} = \sum_{i=1}^{K} \mathbf{z}_B(k)\mathbf{z}_A{}^H(k) \quad ,$$

$$\mathbf{S}_{BB} = \sum_{i=1}^{K} \mathbf{z}_B(k)\mathbf{z}_B{}^H(k) \quad .$$

The inverse of the sample covariance matrix is defined as $\mathbf{P}$ and is partitioned in a manner analogous to that used for the matrix $\mathbf{S}$, i.e.,

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{AA} & \mathbf{P}_{AB} \\ \mathbf{P}_{BA} & \mathbf{P}_{BB} \end{bmatrix} \quad . \tag{22}$$

The three inner product terms of the GLR test are now expressed in terms of the preceding definitions. First,

$$\begin{aligned} \mathbf{z}^H\mathbf{S}^{-1}\mathbf{z} &= \mathbf{z}^H\mathbf{P}\mathbf{z} = \\ &\left(\mathbf{z}_A + \mathbf{P}_{AA}^{-1}\mathbf{P}_{AB}\mathbf{z}_B\right)^H \mathbf{P}_{AA}\left(\mathbf{z}_A + \mathbf{P}_{AA}^{-1}\mathbf{P}_{AB}\mathbf{z}_B\right) \\ &+ \mathbf{z}_B^H\left(\mathbf{P}_{BB} - \mathbf{P}_{BA}\mathbf{P}_{AA}^{-1}\mathbf{P}_{AB}\right)\mathbf{z}_B \quad . \end{aligned} \tag{23}$$

Second,

$$\mathbf{e}^H\mathbf{S}^{-1}\mathbf{e} = \mathbf{e}^H\mathbf{P}\mathbf{e} = \mathbf{P}_{AA} \quad , \tag{24}$$

And third, the final inner product

$$\mathbf{e}^H\mathbf{S}^{-1}\mathbf{z} = \mathbf{e}^H\mathbf{P}\mathbf{z} = \mathbf{P}_{AA}\left(\mathbf{z}_A + \mathbf{P}_{AA}^{-1}\mathbf{P}_{AB}\mathbf{z}_B\right) \quad . \tag{25}$$

Using the Schur relations for partitioned matrices, the inner product $\mathbf{z}^H\mathbf{S}^{-1}\mathbf{z}$ is redefined:

$$\mathbf{z}^H\mathbf{S}^{-1}\mathbf{z} = \mathbf{z}_B^H\mathbf{S}_{BB}^{-1}\mathbf{z}_B + \frac{|y|^2}{T} \quad , \tag{26}$$

where

$$y = \mathbf{z}_A - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{z}_B \quad ,$$

and

$$T = \mathbf{S}_{AA} - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{S}_{BA} = \mathbf{P}_{AA}^{-1} \quad .$$

It is interesting to note that terms y and $\hat{b}$ can be shown to be identical using Equations (26) and (27).

To digress slightly, the discussion turns to the significance of the r.v.'s y and T. Suppose that a Gaussian-distributed N-vector $\mathbf{g}$ has a mean μ and covariance matrix $\mathbf{\Sigma}$. The vector $\mathbf{g}$ is partitioned into q and $N-q$ dimensional subvectors and the covariance matrix $\mathbf{\Sigma}$ is partitioned in the same way as $\mathbf{S}$. A linear transformation to the new components $\mathbf{h}_1$ and $\mathbf{h}_2$, which results in $\mathbf{h}_1$ being uncorrelated with $\mathbf{h}_2$ is

$$\mathbf{h}_1 = \mathbf{g}_1 - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{g}_2 \tag{27}$$

and

$$\mathbf{h}_2 = \mathbf{g}_2 \quad , \tag{28}$$

where $\mathbf{\Sigma}$ is partitioned as

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix} \quad .$$

The product $\boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1}$ is referred to as the "regression coefficient" in the statistical literature [16]. From a filtering or estimation theory viewpoint, $\mathbf{h}_1$ is the prediction error given $\mathbf{h}_2$.

The conditional distribution $\mathrm{p}(\mathbf{h}_1 \,|\, \mathbf{h}_2)$ is Gaussian-distributed with mean $\mu_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mu_2$ and covariance $\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. Observing that the form of y is identical to that of $\mathbf{h}_1$, the r.v. y is simply the prediction error of $\mathbf{z}_A$ given $\mathbf{z}_B$. The conditional covariance is not a function of $\mathbf{h}_2$; therefore, $\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1}\,\boldsymbol{\Sigma}_{21}$ is the unconditional covariance of $\mathbf{h}_1$. Because T is identical in structure to the unconditional covariance of y, T is clearly independent of y. The preceding discussion can be used to gain valuable insight into the final form of the decision rule.

The remaining inner products expressed in terms of the r.v.'s y and T are

$$\mathbf{e}^H\mathbf{S}^{-1}\mathbf{e} = \mathbf{e}^H\mathbf{P}\mathbf{e} = \frac{1}{T} \tag{29}$$

and

$$\mathbf{e}^H\mathbf{S}^{-1}\mathbf{z} = \mathbf{e}^H\mathbf{P}\mathbf{z} = \frac{y}{T} \quad . \tag{30}$$

The GLR expressed in terms of the Schur components is

$$\Lambda = \frac{1 + \mathbf{z}^H\mathbf{S}^{-1}\mathbf{z}}{1 + \mathbf{z}_B^H\mathbf{S}_{BB}^{-1}\mathbf{z}_B} \quad . \tag{31}$$

Equation (31) is now used along with the definitions of y and T to obtain the final form of the GLR. The r.v. v is defined as

$$v = \frac{y}{\left[1 + \mathbf{z}_B^H\mathbf{S}_{BB}^{-1}\mathbf{z}_B\right]^{1/2}} = \rho^{1/2}y \quad , \tag{32}$$

where the loss factor ρ is defined as

$$\rho = \frac{1}{\left[1 + \mathbf{z}_B^H\mathbf{S}_{BB}^{-1}\mathbf{z}_B\right]} \quad . \tag{33}$$

The GLR test for the single-pulse case expressed in terms of v and T is

$$\Lambda = 1 + \frac{|v|^2}{T} \gtrless \alpha \quad . \tag{34}$$

Equation (32) is the form of the decision rule used to obtain test performance criteria, i.e., P_D and P_{FA}.

The statistical characterization of the previous test is now examined. In general terms, given a vector $\mathbf{x}$ of length N distributed as $N(\mu, \Sigma)$ and a sample covariance matrix $\mathbf{S}$ composed of K zero-mean terms, the quadratic inner product $\mathbf{x}^H\mathbf{S}^{-1}\mathbf{x}$ is distributed as noncentral F with N and $K - N + 1$ complex degrees of freedom (dof) and $\mu^H\Sigma^{-1}\mu$ noncentrality parameter [16].[3] This definition of the complex Gaussian r.v. accounts for the correspondence of one complex to two real dof; therefore, the pdf for a real r.v. with an even number of dof is equivalent to the pdf of complex r.v. with one-half the number of dof. For distributions based upon the multivariate Gaussian distribution, the noncentrality parameter is the squared norm of the mean vector. If $\mathbf{x}$ is zero-mean, the resulting distribution is central F with the same dof. The noncentrality parameter is, of course, 0 for central distributions. Therefore, $\mathbf{z}_B^H\,\mathbf{S}_{BB}^{-1}\,\mathbf{z}_B$, the denominator of ρ, is distributed as central F with $N - 1$ and $K - N + 2$ complex dof. An important relationship between F and β r.v.'s is

$$\mathrm{x}_\beta(n; m) = \frac{1}{1 + \mathrm{x}_F(m; n)} \quad , \tag{35}$$

where m and n are the dof parameters; therefore, the loss factor ρ is distributed as central β with $K - N + 2$ and $N - 1$ dof.

Because v is a conditional Gaussian r.v., the likelihood ratio term $|v|^2$ is χ^2-distributed conditioned on the β-distributed loss factor ρ. Because v is composed of only a single term, it has just one complex dof. In terms of the original coordinates, the mean vector of $\mathbf{z}$ is $b\mathbf{M}^{-1/2}\mathbf{p}$. The squared norm of this vector is $(b\mathbf{M}^{-1/2}\mathbf{p})^H(b\mathbf{M}^{-1/2}\mathbf{p}) = |b|^2\mathbf{p}^H\mathbf{M}^{-1}\mathbf{p}$, which is simply the nonadaptive SNR. Under conditioning, the B components are held constant and the conditional mean of v is the product $b\mathbf{M}^{-1/2}\mathbf{p}\rho^{1/2}$. The resulting noncentrality parameter under conditioning is the product of the nonadaptive SNR and the loss factor ρ. The loss factor enters the test statistic only through the conditional mean.

The r.v. T, which equals $\mathbf{S}_{AA} - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{S}_{BA}$, is distributed as central χ^2 with $K - N + 1$ dof [5]. The distribution of T can be obtained using the properties of Wishart matrices and certain orthogonal transformations. T is not a function of the primary vector subcomponents, hence it is statistically independent of y and v.

The ratio $|v|^2/T$ is composed of two independent χ^2 r.v.'s. Because the ratio of independent χ^2 r.v.'s is defined to be F-distributed, the term $|v|^2/T$ is conditionally F-distributed with 1 and $K - N + 1$ dof. Using the relationship expressed in Equation (35), Λ is statistically equivalent

[3]Derivations are for real r.v.'s but can be extended to the complex r.v. case.

to the inverse of a conditional β-distributed random variable. The test statistic is now examined under hypotheses H_1 and H_0.

2.3 Single-Pulse P_{FA} and P_D

To obtain expressions for the P_D and P_{FA}, the conditional and unconditional distributions of the test statistic are examined first. As stated in Section 2.2, the inverse test statistic is complex β-distributed. The conditional distribution of the inverse test statistic Λ^{-1} is

$$p(x|\rho) = f_\beta(x; L, 1; c) \quad , \tag{36}$$

where

$$L = K - N + 1 \quad ,$$

$$c = a\rho = \left(|b|^2 \mathbf{p}^H \mathbf{M}_i \mathbf{p}\right) \rho \quad \text{under} \quad H_1 \quad ,$$

and

$$c = 0 \quad \text{under} \quad H_0 \quad .$$

The presence of the conditioning r.v. requires that the distribution of the test statistic, which is conditionally β-distributed, is averaged over ρ to obtain the unconditional distribution of the test. Under H_0 the noncentrality parameter is 0; therefore, the conditioning r.v. has no effect on the unconditional distribution because it appears only through the noncentrality parameter. As a result, the inverse test distribution is unconditionally central β-distributed under H_0.

Because its form is much simpler, the distribution of the test under H_0 is examined first. The analysis results in an exact expression for the P_{FA}. The general form of the central β distribution is [5]

$$f_\beta(x; m, n) = \frac{\Gamma(n+m)}{\Gamma(n)\Gamma(m)} x^{m-1}(1-x)^{n-1} \quad , \tag{37}$$

where n and m are the complex dof parameters.

In the current case where the first dof parameter is L and the second is 1, the central β distribution reduces to the simple form

$$f_\beta(x; L, 1) = L x^{L-1} \quad . \tag{38}$$

The P_{FA} is defined as:

$$P_{FA} = \int_{\alpha}^{\infty} f(\Lambda|H_0)d\Lambda \quad . \tag{39}$$

Using the relationship

$$Pr(\Lambda \geq \alpha) = Pr\left(\Lambda^{-1} \leq \frac{1}{\alpha}\right) \quad , \tag{40}$$

and the fact that Λ^{-1} is unconditionally distributed as $f_\beta(x; L, 1)$, P_{FA} is the cumulative distribution function of a central β r.v., i.e.,

$$P_{FA} = \int_0^{\alpha^{-1}} Lx^{L-1}dx = \frac{1}{\alpha^L} \quad . \tag{41}$$

Because P_{FA} is independent of the true covariance matrix and is only a function of the dimensional parameters of the problem, the test statistic possesses the desirable constant false alarm rate (CFAR) characteristic. The threshold value α is evaluated solving Equation (41) for a specific P_{FA} value.

Under the signal-plus-noise hypothesis, the presence of the conditioning r.v. results in a more complicated analysis. The unconditional distribution of the test statistic, which is obtained by averaging out over the distribution of the loss factor (contained in c), is

$$f(x|H_1) = \int_0^1 f_\beta(x; L, 1; c)\, f_\beta\,(\rho; L+1, N-1)\, d\rho \quad . \tag{42}$$

Per the P_{FA} calculation, the P_D can be expressed in terms of the unconditional test distribution as

$$P_D = Pr(\Lambda \geq \alpha) = \int_0^{\alpha^{-1}} f(x|H_1)dx \quad . \tag{43}$$

Using previously derived results [8], P_D is

$$P_D = 1 - \frac{1}{\alpha^L}\sum_{k=1}^{L}\binom{L}{k}(\alpha-1)^k H_k\left(\frac{a}{\alpha}\right) \quad , \tag{44}$$

where a is the nonadaptive SNR,

$$H_k(w) = \int_0^1 G_k(a\rho)\, f_\beta\,(\rho; L+1, N-1)\, d\rho \quad ,$$

and

$$G_k(x) = e^{-x} \sum_{m=0}^{k-1} \frac{x^m}{m!} \quad .$$

A complete discussion of the computational complexities involved in the evaluation of the probability of detection can be found elsewhere [19,20].

As discussed in greater detail in the remainder of the report, the formulation of the GLR test is equivalent to the Wilks' Λ test, which commonly occurs in multivariate hypothesis testing schemes [10]. Results obtained for the signal-plus-noise hypothesis differ from the standard results because of the presence of the conditioning r.v. Next is an examination of the performance of the decision rule as a function of the nonadaptive SNR and the excess number of secondary vectors L.

2.4 Performance of the Single-Pulse Detector

The statistical performance of the single-pulse decision rule is quantified through the use of receiver operating characteristic (ROC) plots. ROC is defined here as the plot of the P_D versus SNR for a fixed P_{FA}, which is set to a nominal value of 10^{-2}. The parameter L, which is related to the number of secondary vectors, is a controlling factor in the performance of the decision rule. As L tends to ∞, the loss factor tends to 1, hence, the covariance estimate approaches the true covariance. When there is no estimation loss, the Marcum-Q function can be used to evaluate P_D.

To illustrate, let the dimensional parameter $N = 4$. Figure 1 shows the performance of the single-pulse detector as a function of K. As K increases from 8 to 20, ROC approaches the known covariance case. Kelly has shown that ROC is only a function of L. For example, the case where $N = 16, K = 32$, and $L = 17$ yields the same results as the preceding case with $N = 4$ and $K = 20$.

2.5 Multiple-Pulse Test in a Stationary Noise Environment (Case 2)

The true covariance matrix is assumed to be constant over the N_p pulse observation intervals as does the mean value of the primary data vector. Analysis begins by forming an $N \times N_p$ data block $\mathbf{Z}_N$ where

$$\mathbf{Z}_N = [\mathbf{z}_1 \mathbf{z}_2 \ldots \mathbf{z}_{N_p}] \quad . \tag{45}$$

The columns of $\mathbf{Z}_N$ are the primary data vectors $\mathbf{z}_i$, $i = 1, \ldots, N_p$.

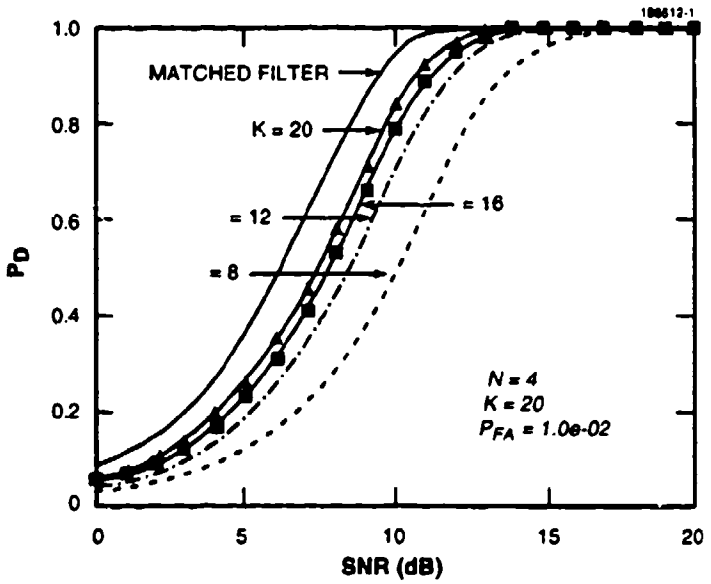

Figure 1. Single-pulse test ROC.

An $N_p \times N_p$ unitary matrix $\mathbf{U}_{N_p}$ can be appropriately selected to rotate the mean of the data solely into the first column of $\mathbf{Z}_N$. This operation is possible because all the mean vectors are the same. Performing the postmultiplication of $\mathbf{Z}_N$ by $\mathbf{U}_{N_p}^H$ results in

$$\mathbf{Z}_N \, \mathbf{U}_{N_p}^H = [\mathbf{Z}_a \;\; \mathbf{Z}_b] \quad , \tag{46}$$

where $\mathbf{Z}_a$ is a column vector with mean $b\,\mathbf{p}$ and $\mathbf{Z}_b$ is an $N \times (N_p - 1)$ zero-mean array.

The problem has now been reduced to the case 1 form with a single primary vector with a mean of $b\,\mathbf{p}$ and $K + N_p - 1$ secondary vectors. Test performance can be evaluated using the P_{FA} and P_D equations developed in Section 2.5.

2.6 Summary of Results

For the single-pulse case, the decision rule derived from a GLR test has been statistically characterized under both hypotheses. Under hypothesis H_0, the P_{FA} is solely a function of the excess number of secondaries; therefore, the test is a CFAR test. The expression for the P_D is complicated and requires considerable computation time. The ROC plots indicate that test performance approaches the matched filter levels as the the number of secondary data vectors used in the test increases.

3. NONCOHERENT ADAPTIVE DETECTION

This section examines the more difficult nonstationary multiple-observation case. Due to the nonstationarity assumption, the covariance matrix can vary over the complete observation interval. The true covariance matrix at each time instant is assumed unknown under both hypotheses and, for this case only, the model for the returned signal amplitude is time varying. A GLR test using data from N_p data observation periods is then formulated under these assumptions. The resulting decision rule is shown to be equivalent to a noncoherent detection process. The resulting test consists of the inverse product of N_p β-distributed r.v.'s and is independent of the unknown covariance parameters, implying that it possesses the desirable CFAR property [7]. An exact expression for the P_{FA} is obtained and an upper bound on the P_D is derived [13]. Theoretical and simulation results are compared and observations with respect to performance are made.

3.1 Multiple-Pulse GLR Test

This section formulates the GLR test for the multiple-observation case, where the covariance matrix and the returned signal amplitude are unknown parameters. To restate, the pdf of the primary vector for the ith pulse under H_1 is

$$p_1(\mathbf{z}_i|H_1) = \frac{1}{\pi^N \|\mathbf{M}_{1i}^{-1}\|} e^{-(\mathbf{z}_i - b_i\mathbf{p})^H \mathbf{M}_{1i}^{-1}(\mathbf{z}_i - b_i\mathbf{p})} \quad , \tag{47}$$

where

$$\mathbf{M}_{1i} = \mathrm{E}\left\{(\mathbf{z}_i - b_i\mathbf{p})(\mathbf{z}_i - b_i\mathbf{p})^H\right\} \quad ,$$

b_i is the complex returned signal amplitude, and $\mathbf{p}$ is the steering vector. The primary vector pdf under H_0 is

$$p_0(\mathbf{z}_i|H_0) = \frac{1}{\pi^N \|\mathbf{M}_{0i}^{-1}\|} e^{-\mathbf{z}_i^H \mathbf{M}_{0i}^{-1}\mathbf{z}_i} \quad , \tag{48}$$

where

$$\mathbf{M}_{0i} = \mathrm{E}\mathbf{z}_i\mathbf{z}_i^H \quad .$$

The secondary vectors are defined to be zero-mean and possess the same pdf as that of the primary vector under H_0.

As per the single-pulse case, the joint pdf of the primary and secondary vectors are formed at time instant i. Both likelihood functions are the product of $K+1$ multivariate Gaussian pdf's. Assuming the covariance matrices are totally unknown under both hypotheses, and the returned signal amplitudes are unknown parameters only under H_1, the likelihood ratio is

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i = \prod_{i=1}^{N_p} \frac{p_1(\mathbf{z}_i, \mathbf{z}_i(1), \ldots, \mathbf{z}_i(K); \mathbf{M}_{1i}, b_i)}{p_0(\mathbf{z}_i, \mathbf{z}_i(1), \ldots, \mathbf{z}_i(K)\mathbf{M}_{0i})} \tag{49}$$

$$= \prod_{i=1}^{N_p} \frac{\left(\frac{1}{\pi^N \|\mathbf{M}_{1i}\|}\right)^{K+1} e^{-\left(\mathbf{x}_i^H \mathbf{M}_{1i}^{-1} \mathbf{x}_i + Q_i\right)}}{\left(\frac{1}{\pi^N \|\mathbf{M}_{0i}\|}\right)^{K+1} e^{-\left(\mathbf{z}_i^H \mathbf{M}_{0i}^{-1} \mathbf{z}_i + Q_i\right)}} ,$$

where

$$\mathbf{x}_i = (\mathbf{z}_i - b_i \mathbf{p})$$

and

$$Q_i = \sum_{i=1}^{K} \mathbf{z}_i^H(k) \mathbf{M}_{0i}^{-1} \mathbf{z}_i(k) \quad .$$

According to the GLR test procedure, the numerator of Equation (49) must be maximized with respect to each $\mathbf{M}_{1i}$ and b_i. The denominator is maximized solely with respect to each $\mathbf{M}_{0i}$. Taking the $(K+1)$th root of the numerator and denominator pdf's for each Λ_i and applying the inner product identity $\mathbf{x}^H \mathbf{Y}^{-1} \mathbf{x} = \mathrm{Tr}\{\mathbf{Y}^{-1} \mathbf{x}\mathbf{x}^H\}$, the likelihood ratio at i is

$$\Lambda_i^{\frac{1}{K+1}} = \frac{\left(\frac{1}{\|\mathbf{M}_{1i}\|}\right) e^{-\mathrm{Tr}\left[\mathbf{M}_{1i}^{-1} \frac{1}{K+1} \{\mathbf{x}_i \mathbf{x}_i^H + \mathbf{S}_i\}\right]}}{\left(\frac{1}{\|\mathbf{M}_{0i}\|}\right) e^{-\mathrm{Tr}\left[\mathbf{M}_{0i}^{-1} \frac{1}{K+1} \{\mathbf{z}_i \mathbf{z}_i^H + \mathbf{S}_i\}\right]}} , \tag{50}$$

where

$$\mathbf{S}_i = \sum_{i=1}^{K} \mathbf{z}_i(k)\, \mathbf{z}_i^H(k) \quad .$$

Λ_i is redefined as the $(K+1)$th root of the test at i.

From Equation (50), while fixing b_i temporarily, the MLEs of the covariance under each hypothesis are

$$\hat{\mathbf{M}}_{1i} = \frac{1}{K+1}\left\{(\mathbf{z}_i - b_i\mathbf{p})\,(\mathbf{z}_i - b_i\mathbf{p})^H + \mathbf{S}_i\right\} \tag{51}$$

and

$$\hat{\mathbf{M}}_{0i} = \frac{1}{K+1}\left\{\mathbf{z}_i\mathbf{z}_i^H + \mathbf{S}_i\right\} \quad . \tag{52}$$

Again, the sample covariance matrix estimator form arises from the current problem formulation, not from an a priori assumption. After substituting the MLEs of the covariance matrix for each pulse interval into their respective likelihood functions, the following ratio of determinants, which still contains the unknown parameter b_i, is obtained:

$$\Lambda = \prod_{i=1}^{N_p} \frac{\|\hat{\mathbf{M}}_{0i}\|}{\|\hat{\mathbf{M}}_{1i}\|} \quad . \tag{53}$$

Equation (53) is maximized with respect to b_i by minimizing the denominator, i.e.,

$$\Lambda = \prod_{i=1}^{N_p} \frac{\|\hat{\mathbf{M}}_{0i}\|}{\min\limits_{b_i}\|\hat{\mathbf{M}}_{1i}\|} \quad . \tag{54}$$

Using the same determinant identities found in Section 2.1,

$$(K+1)^N\|\hat{\mathbf{M}}_{1i}\| = \|\mathbf{S}_i\|[1 + (\mathbf{z}_i - b_i\mathbf{p}^H)^H\mathbf{S}_i^{-1}(\mathbf{z}_i - b_i\mathbf{p})] \tag{55}$$

and

$$(K+1)^N\|\hat{\mathbf{M}}_{0i}\| = \|\mathbf{S}_i\|[1 + \mathbf{z}_i^H\mathbf{S}_i^{-1}\mathbf{z}_i] \quad , \tag{56}$$

and completing the square with respect to the b_i, an expression for the GLR, where the unknown parameters $b_i, i = 1, \ldots, N_p$ are isolated, is obtained:

$$\Lambda = \prod_{i=1}^{N_p} \frac{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i}{\min\limits_{b_i} \left\{ 1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i + \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p} \left| b_i - \frac{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}} \right|^2 - \frac{\left| \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i \right|^2}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}} \right\}} \quad . \tag{57}$$

Equation (57) is maximized when $b_i = \hat{b}_i$, $i = 1, \ldots, N_p$, as was the case in the single-pulse analysis. From the previous equation, the MLE for b_i is

$$\hat{b}_i = \frac{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}} \quad . \tag{58}$$

The estimator $\hat{b}_i$ possesses a form equivalent to the prediction error (see Section 2.2). Substituting $\hat{b}_i$ in place of b_i yields the final form of the LRT:

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i = \prod_{i=1}^{N_p} \frac{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i}{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i - \frac{\left| \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i \right|^2}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}}} \gtrless \alpha \quad , \tag{59}$$

where α is a threshold parameter. The above expression is just the product of N_p single-pulse tests.

The GLR test is solely a function of the primary vectors, the secondary vectors, and the known vector $\mathbf{p}$. Taking the logarithm of the GLR, a summation of N_p independent Λ_i terms is obtained. Because the data at each time instant are used to make decisions, the overall detection process is equivalent to noncoherent integration. An examination of LRT statistics given by Equation (59) follows.

3.2 Analysis of the Noncoherent GLR Test

This section examines the test statistic Λ under each hypothesis. The inverse test statistic under hypothesis H_0 possesses a central χ^2 distribution. The cumulative distribution function of the inverse test statistic, P_{FA}, is evaluated in terms of a closed-form expression. Under hypothesis H_1, an exact expression for the pdf of the inverse product of N_p independent noncentral β r.v.'s is unobtainable. Hence, P_D cannot be expressed in closed-form and an upper (Chernoff) bound is obtained.

The equivalence between the problem studied here and the well-known Wilks' Λ test [10] is discussed next. The general form of the Wilks' Λ test arises when the covariance matrix is unknown and the subsequent GLR procedure results in the ratio of determinants. Through the judicious use of several identities found in multivariate statistics and probability theory, the moments of the test

can be shown to be equal to the moment of the product of independent β-distributed r.v.'s (for real and complex r.v.'s, see Anderson [16] and Box [21], and Kelly and Forsythe [5], respectively). When the product consists of greater than three elements, however, an exact expression for the pdf is virtually impossible to obtain. Adding to the complexity of this problem is the presence of the loss factor ρ, which necessitates working with conditional distributions.

Many statistical researchers have examined the unconditional Wilks' Λ test. Several asymptotic approaches based on the χ^2 distribution have been developed, but these methods often require using a correction factor [15,22]. These approximations produce reliable results only when N_p is relatively small. Because this analysis may be extended to cases where the pulse train can be quite long, the χ^2 approaches do not appear to be useful in the current context. In terms of notation, the statistical community defines the likelihood ratio $\mathrm{p}|\mathrm{H}_0/\mathrm{p}|\mathrm{H}_1$, which is the inverse of the definition commonly used by communications engineers: $\mathrm{p}|\mathrm{H}_1/\mathrm{p}|\mathrm{H}_0$. The difference arises from what each group is testing. The goal of the communications engineer is to find differences between two hypotheses given the data, whereas the goal of the statistician is to determine whether the data fit a specified model. The test definition in this report results in the inverse product of independent β r.v.'s.

A single term of the GLR test, Λ_i, given in Equation (59), was shown in Section 2 to be statistically equivalent to 1 plus the ratio of two independent χ^2 r.v.'s, $|v_i|^2/T_i$. Under hypothesis H_1, $|v_i|^2$ possesses a complex noncentral χ^2 distribution with 1 complex dof and a noncentrality parameter $c_i = a_i\rho_i$, and is simply a complex central χ^2 r.v. under H_0. The noncentrality parameter c_i is the product of the nonadaptive SNR a_i and a β-distributed loss factor ρ_i. Hence, under H_1, $|v_i|^2$ is a complex noncentral χ^2 r.v. that is conditioned on ρ_i. The parameter T_i is an unconditional complex central χ^2 r.v. with L complex dof. In terms of these χ^2 parameters, the GLR test for the problem under study is

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i = \prod_{i=1}^{N_p} \left(1 + \frac{|v_i|^2}{T_i}\right) \gtrless \alpha \quad . \tag{60}$$

The complex central χ^2 pdf is defined:

$$f_{\chi^2}(x_i; m) = \frac{1}{\Gamma(m)} x_i^{m-1} \mathrm{e}^{-x_i} \quad , \tag{61}$$

where m is the dof parameter and the noncentral χ^2 pdf is

$$f_{\chi^2}(x_i; m; c_i) = \mathrm{e}^{-x_i - c_i} \left(\frac{x_i}{c_i}\right)^{\frac{m-1}{2}} I_{m-1}(2\sqrt{x_i c_i}) \quad , \tag{62}$$

where $c_i = a_i\rho_i$ is the noncentrality parameter and I_{m-1} is the modified Bessel function.

Because the ratio of the two independent χ^2 r.v.'s possesses an F distribution, the term $|v_i|^2/T_i$ is conditionally F-distributed. Using the previously defined relationship between β and F r.v.'s, the multiple-pulse GLR can be expressed as the inverse product of β r.v.'s that are conditioned on the loss factor ρ_i.

Under the null hypothesis, the general form of the complex Wilks' Λ test can be expressed as

$$\prod_{i=1}^{N_p} x_\beta(m+1-i,n) \gtrless \alpha \quad . \tag{63}$$

The GLR test in the noise-only case is a special form of the Wilks' Λ test in that the inverse of each Λ_i possesses the same β distribution, i.e., the current test is of the form

$$\prod_{i=1}^{N_p} \frac{1}{x_\beta(m,n)} \gtrless \alpha \quad . \tag{64}$$

The conditioning effect and the constant dof parameters are cases not found in the general statistical literature. The distribution of the GLR under each hypothesis is now examined.

3.2.1 The P_{FA} of the GLR Test

P_{FA} is defined $\int_\alpha^\infty p(\Lambda)|H_0\, d\Lambda$. The numerator term v_i is assumed to be zero-mean (therefore, the noncentrality parameter $c_i = 0$), which leads to the result that the distribution of the test is a function of central r.v.'s. For the dof specified in this problem, the GLR test statistic under hypothesis H_0 in terms of the product of central β r.v.'s is

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i = \prod_{i=1}^{N_p} \frac{1}{x_\beta(L,1)} \quad . \tag{65}$$

Using the definition of the complex β pdf, the pdf for each Λ_i reduces to

$$f_\beta(x;L,1) = Lx^{L-1} \quad . \tag{66}$$

By allowing the returned amplitude to be time varying, the single-pulse case results have been used to this point; however, the distribution of the inverse product of N_p r.v.'s must be determined.

In general, the distribution of the product of r.v.'s is extremely difficult to evaluate analytically. If a transformation can be performed so that the test consists of the sum of independent r.v.'s, several probability identities can then be applied to obtain the desired pdf [6]. Letting

$$\Lambda_i = \exp\left(\frac{y_i}{L}\right) \quad , \tag{67}$$

and performing the standard change of variable transformation procedure, the pdf of the newly defined r.v. y_i is

$$f(y_i) = e^{-y_i} \quad . \tag{68}$$

Hence, the term y_i is a central χ^2 r.v. with one complex dof. The test can now be expressed as

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i = \exp\left(\sum_{i=1}^{N_p} y_i/L\right) = e^{Y/L} \quad , \tag{69}$$

where

$$Y = \sum_{i=1}^{N_p} y_i \quad .$$

Taking the logarithm of Equation (69), the resulting test consists of the sum of independent χ^2 r.v.'s; therefore, Y is central χ^2-distributed with N_p dof. The pdf of Y is defined:

$$f(Y) = f_{\chi^2}(Y; N_p) = \frac{1}{\Gamma(N_p)} Y^{N_p-1} e^{-Y} \quad . \tag{70}$$

The cumulative distribution function (cdf) of a χ^2 r.v. with N_p dof is

$$F(Y) = 1 - G_{N_p}(Y) \quad , \tag{71}$$

where

$$G_{N_p}(Y) = \int_Y^\infty f_{\chi^2}(x; N_p)dx = e^{-Y} \sum_{k=0}^{N_p-1} \frac{Y^k}{k!} \quad .$$

This expression is the incomplete Gamma function, and the P_{FA} of the test can be expressed as

$$P_{FA} = G_{N_p}(Y_0) \quad , \tag{72}$$

where

$$Y_0 = L \log \alpha \quad .$$

The threshold that yields a particular P_{FA} value can be found by solving Equation (72) by means of a Newton-Raphson iteration procedure. The required derivative is given by

$$G'_{N_p} = -f_{\chi^2}(Y; N_p) \quad . \tag{73}$$

An exact expression for the P_{FA}, i.e., the error under hypothesis H_0, has been obtained. The P_{FA}, corresponding to a pulse train of arbitrary length is easily evaluated using the previous expressions. To complete the analysis, the test statistic under H_1 is now examined.

3.2.2 A Chernoff Bound for the P_D of the GLR Test

P_D is defined $\int_\alpha^\infty p(\Lambda)|H_1 \, d\Lambda$. Using the same notation as in the Section 3.2.1, the GLR test is composed of the inverse product of N_p noncentral β r.v.'s. The pdf of a single noncentral β r.v. with L and 1 complex dof and noncentrality parameter c_i is [5]

$$f_\beta(x_i; L, 1; c_i) = \sum_{k=0}^{L} \binom{L}{k} \frac{\Gamma(L+1)}{\Gamma(L+1+k)} c_i^k e^{-c_i x_i} f_\beta(x_i; L, 1+k) \quad . \tag{74}$$

Unfortunately, no simple transformation can be performed to enable the product of terms to be expressed in the preferable sum-of-terms form as was done in Section 3.2.1. As a result, the problem must be approached in a different way.

Statisticians have used asymptotic χ^2 approaches to obtain approximate solutions; however, many of these approaches have drawbacks, including the use of correction factors. This study chooses to obtain the Chernoff bound for the distribution of the LRT [23]. Although not exact, this upper bound will at least provide a measure of confidence in results obtained through simulations. The remainder of this section is devoted to obtaining an optimal (tightest) upper bound on the P_D for the inverse product of N_p independent, noncentral β r.v.'s. The derivation leading to the Chernoff bound involves several probability and integral identities [5,9].

Because the noncentrality parameter of each r.v. is conditioned on a β-distributed loss factor, the unconditional distribution of the test statistic must be obtained. The GLR under H_1, expressed as the inverse product of conditional noncentral β r.v.'s, is

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i = \prod_{i=1}^{N_p} \frac{1}{x_\beta(L, 1; c_i)} \quad . \tag{75}$$

The pdf and the moments of an r.v. that are composed of the sum of independent r.v.'s can be determined using moment generating function (MGF) properties. To take advantage of these properties, the following change of variable is made:

$$\Lambda_i = e^{-u_i} \quad . \tag{76}$$

The GLR test can now be expressed as

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i^{-1} = \prod_{i=1}^{N_p} e^{u_i} = \exp\left(\sum_{i=1}^{N_p} u_i\right) \gtrless \alpha \quad . \tag{77}$$

Taking the logarithm of Equation (77) and defining $U = \log\Lambda$, the log GLR test is

$$U = \sum_{i=1}^{N_p} u_i \gtrless \log\alpha \quad . \tag{78}$$

Equation (76) has the form that enables using MGF properties.

Basing the remaining statistical analysis on the r.v. U, an upper bound on $\Pr(U \geq \log\alpha$ is determined. Setting $\delta = \log\alpha$, the Chernoff bound in terms of U is defined:

$$\Pr(U \geq \delta) \leq e^{-\nu\delta} E[e^{\nu U}] \quad . \tag{79}$$

The optimization parameter $\hat{\nu}$ that yields the tightest possible upper bound is the solution of

$$\frac{dE[e^{\nu(U-\delta)}]}{d\nu} = 0 \quad . \tag{80}$$

To obtain the optimal upper bound for this problem, the term $E[e^{\nu U}]$, which is the MGF of the test statistic, and the parameter $\hat{\nu}$ need to be evaluated.

The Chernoff bound can obviously be expressed in terms of MGF, based on its definition. The r.v. U is the sum of independent r.v.'s; therefore, using the MGF summation-product relationship, the MGF of U is the product of the MGFs of each individual u_i:

$$M_U(\nu) = \prod_{i=1}^{N_p} M_{u_i}(\nu) = \prod_{i=1}^{N_p} E[e^{\nu u_i}] \quad . \tag{81}$$

Citing the definition of the r.v. u_i, $\mathrm{E}[e^{\nu u_i}]$ is equal to $\mathrm{E}[\Lambda_i^{-\nu}]$. The MGF of u_i is clearly equivalent to the $-\nu$th moment of a noncentral β r.v.

The $-\nu$th moment of the r.v. Λ is obtained by taking the expected value of $-\nu$th power of Λ_i. The pdf of a noncentral complex β r.v. with L and 1 complex dof is repeated for clarity:

$$f_\beta(x_i; L, 1; c_i) = \sum_{k=0}^{L} \binom{L}{k} \frac{\Gamma(L+1)}{\Gamma(L+1+k)} c_i{}^k e^{-c_i x_i} f_\beta(x_i; L, 1+k) \quad . \tag{82}$$

Given the pdf of Λ_i, the $-\nu$th conditional moment of this r.v. is

$$\begin{aligned} \mathrm{E}\left[\Lambda_i^{-\nu} | \rho_i\right] &= \sum_{k=0}^{L} \binom{L}{k} \frac{L\Gamma(L-\nu)}{\Gamma(L-\nu+1+k)} c_i{}^k \\ &\times \; e^{-c_i}{}_1F_1(1+k; L-\nu+1+k; c_i) \quad . \end{aligned} \tag{83}$$

The unconditional moment of Λ_i is now obtained. Conditioning is removed by averaging over the distribution of the loss factor. First set $c_i = a_i\rho_i$, then multiply Equation (83) by $f_\beta(\rho_i; L+1, N-1)$, and finally integrate over the range of ρ_i. The resulting expression for the $-\nu$th unconditional moment of Λ_i is

$$\begin{aligned} \mathrm{E}\left[\Lambda_i^{-\nu}\right] &= \sum_{k=0}^{L} \binom{L}{k} \frac{L\Gamma(L-\nu)}{\Gamma(L1k-\nu)} \frac{\Gamma(L+N)}{\Gamma(LNk)} \frac{\Gamma(L1k)}{\Gamma(L+1)} a_i{}^k \\ &\times \sum_{j=0}^{\infty} \frac{\Gamma(1+k+j)\Gamma(L1k+j)}{\Gamma(1+k)\Gamma(L1k)} \frac{\Gamma(LNk)\Gamma(L1k-\nu)}{\Gamma(LNk+j)\Gamma(L1k-\nu+j)} \\ &\times \; e^{-a_i}{}_1F_1(N-1; LNk+j; a_i) \frac{a_i^j}{j!} \quad , \end{aligned} \tag{84}$$

where $L1k = L+1+k$ and $LNk = L+N+k$.

The function ${}_1F_1(N-1; L+N+k+j; a_i)$ is the confluent hypergeometric function. It will converge to a finite value because the second parameter is greater than the first (the denominator gamma function values will grow faster than those of the numerator) [24]. The summation over the index j, which is of a form similar to a ${}_2F_2$ hypergeometric series, will also converge for the same reason given earlier.

A compact expression for the MGF of U is

$$M_U(\nu) = \prod_{i=1}^{N_p} \mathrm{E}\left[\Lambda_i^{-\nu}\right] \quad . \tag{85}$$

Assuming that the nonadaptive SNR is constant over the observation interval (remember that only the model for the returned amplitude is time varying, not the actual returned amplitude), the MGF of U reduces to

$$M_U(\nu) = \left(E\left[\Lambda_i^{-\nu}\right]\right)^{N_p} \quad . \tag{86}$$

Equations (84-86) are used for the computer-based evaluation of $M_U(\nu)$. The final form of the expression for the Chernoff bound requires that the optimization parameter $\hat{\nu}$ be found numerically. Computational difficulties encountered in obtaining the Chernoff bound in Section 3.3.

3.3 Simulation Results

This section discusses a comparative analysis that is performed using results obtained from Monte Carlo simulations, the Chernoff bound, and the benchmark matched filter case where the covariance matrix is assumed known. How the Chernoff bounds were obtained numerically is described first, followed by a discussion of the Monte Carlo simulation technique used for this investigation. Detector performance is then analyzed using ROC plots. For the purposes of this report, ROCs are defined as plots of P_D versus SNR for a fixed P_{FA}. The ROC for the known covariance case is produced using a modified Marcum-Q function [25].

The numerical evaluation of the Chernoff bound involves an expression that includes two infinite series. The ${}_1F_1(N-1; L+N+k+j; a_i)$ confluent hypergeometric series is well behaved and converges rapidly. The summation on i was truncated after the difference between two successive terms was $\leq 10^{-4}$. The summation on j , however, is not quite that simple. This summation is similar in form to a ${}_2F_2$ hypergeometric series [24] that monotonically increases to a peak value and then decreases until the summation reaches a steady-state; therefore, truncation bounds were not applied until after the maximum functional value had been reached. It was necessary to evaluate the logarithms of the terms in the Chernoff bound expression to avoid overflow. Also, recursive relationships were used to decrease the computation time. For each SNR value, the optimization parameter $\hat{\nu}$ was evaluated using a numerical search method with ν restricted to the range $0 \leq \nu \leq L$ due to the definition of the Γ function. The Chernoff bound was then computed for this particular value of ν and stored for plotting purposes.

A simple Monte Carlo approach was chosen to evaluate P_D. For the noncoherent decision rule, it was necessary to generate N_p noncentral β r.v.'s, form their inverse product, and compare it to a threshold. If the product was greater than the threshold (corresponding to a target present in the data), a counter was incremented by one.

A P_{FA} of 0.01 was selected and the corresponding threshold was easily computed using the incomplete gamma function. This relatively high (for radar applications) P_{FA} value was chosen to minimize the number of simulation runs. If the decision rule is used in conjunction with an M out of N detector, P_{FA} will be greatly reduced [26]. If one wanted to work with P_{FA} on the order of

10^{-6}, importance sampling or extreme value theory techniques could be applied to the problem under investigation [27–30]. Given the value of the current P_{FA}, at least 1,000 data sequences of N_p length were required for the simulation results to be statistically significant. The final counter value was divided by 1,000 with the resulting value defined as the simulation-generated P_D. This process was repeated for each SNR.

The generation of r.v.'s possessing the desired noncentral β distribution is now presented. A noncentral β r.v. is a function of the noncentrality parameter (c = SNR ρ) and the dimensional parameters of the problem, L and N. The nonadaptive SNR was chosen to range between 0 and 20 dB in 1-dB increments for simulation purposes. The loss factor ρ_i was set to its mean value, $E(\rho_i) = (L+1)/(L+N)$. Because ρ_i is a β-distributed r.v. and its distribution is sharply peaked about the mean for the given dof parameters. the use of the mean value in place of the r.v. is justified.

For each value of the noncentrality parameter, the random variates were generated using a stratified rejection method [14,31]. The rejection (also referred to as "acceptance-rejection") method was selected because the noncentral β probability distribution is not amenable to direct inversion. Direct inversion requires solving $X = F^{-1}(U)$, where $F(x)$ is the cumulative distribution function of the r.v. X, and U is uniformly distributed between 0 and 1. An examination of the noncentral β probability distribution indicates that $F^{-1}(U)$ would be difficult to obtain for this distribution; however, the rejection method can be used with any pdf. The stratification process significantly reduces the number of attempts required to generate the desired quantity of r.v.'s. The resulting computational savings is important considering the number of variates that were required for this study.

Analysis of detector performance capabilities proceeds by examining the ROCs for various parameter combinations. Figure 2 illustrates a typical relationship between the Chernoff bound, the Marcum-Q function, and the Monte Carlo simulation results.

Although it is somewhat loose (4 dB greater than the simulation results), the Chernoff bound provides some level of confidence in the Monte Carlo simulation results. In all cases examined, the bound is greater than the simulation result but tracks the shape of both the Marcum-Q and simulation curves. Notice that the simulation result is only 1 dB less than the optimal detector over the complete SNR range. The GLR-based detector with $N = 4$ and $K = 20$ performs well under these conditions.

The effect of the excess number of secondaries, $L = K - N + 1$, is now examined. Figure 3 portrays the asymptotic behavior of the detector as a function of K with number of array elements N fixed. As the number of secondary vectors K increases, detector performance clearly approaches the known covariance case, depicted as the Marcum-Q curve, with the loss compared to the optimal detector being significantly reduced. The detector provides excellent performance given a reasonable number of secondary data vectors. The improved performance as function of increasing K occurs for all pulse train lengths.

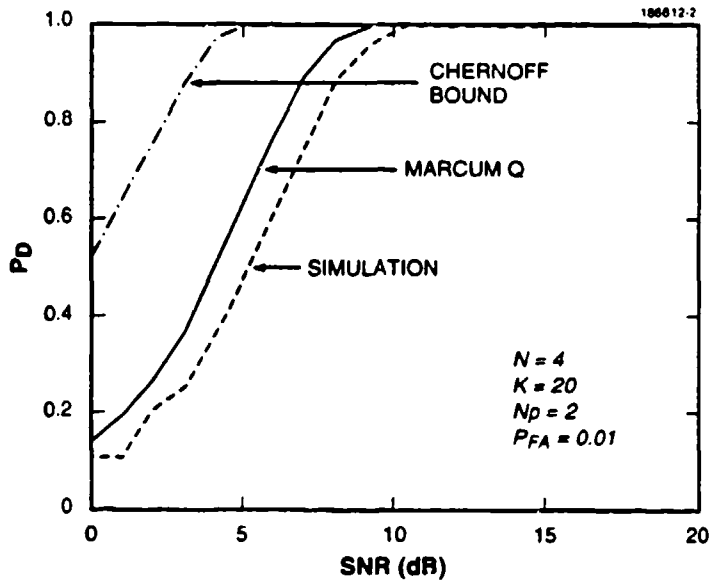

Figure 2. Noncoherent performance comparison for $N_p = 2$.

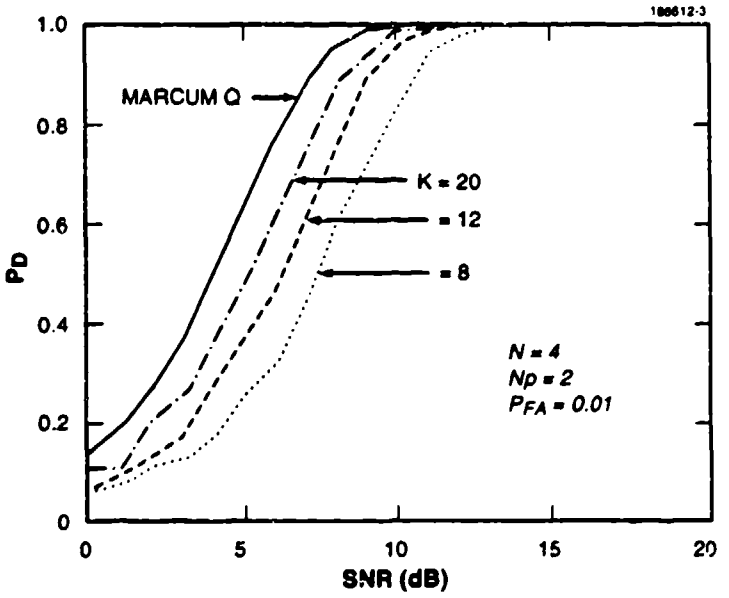

Figure 3. Noncoherent performance as a function of K for $N_p = 2$.

Figures 4 and 5 portray the same information as Figure 2 but with $N_p = 4$ and 10, respectively. The parameter L has the same value ($K = 20$, therefore, $L = 17$) as used for Figure 1 to make valid comparisons. The Chernoff bound tracks the Marcum-Q and simulation results in both cases. For both $N_p = 4$ and 10, the detection loss compared to the matched filter is the same 1-dB loss found in Figure 2.

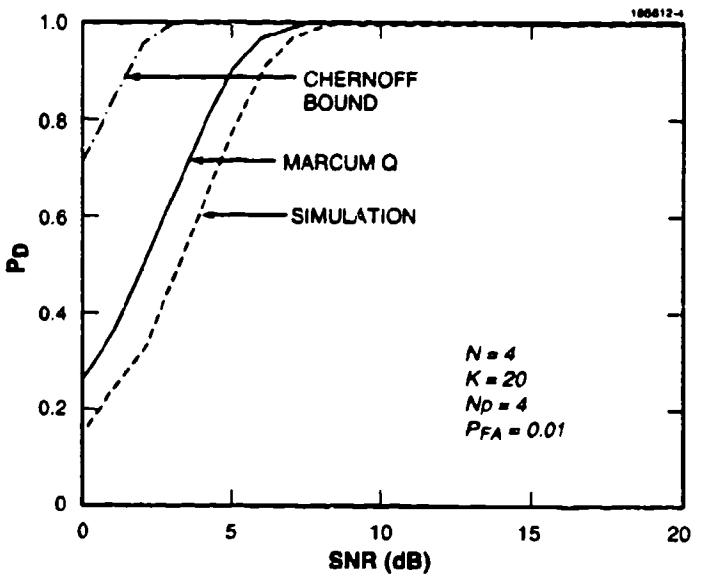

Figure 4. Noncoherent performance comparison for $N_p = 4$.

Figure 6 illustrates detector performance as a function of N_p. Notice the pronounced leftward shift of the ROC curves as a function of increasing N_p. The SNR required to achieve a certain P_D is significantly decreased as the pulse train is increased in length. These results are not unexpected because more information is being used to make the decision with a longer pulse train.

To further examine the effect of increasing N_p, the SNR required to achieve a $P_D = 0.9$ versus N_p for $N = 4$ and $K = 20$ is shown in Figure 7. The required SNR for a fixed P_D value decreases at a rate proportional to $1/\sqrt{N_p}$, which is consistent with the behavior of noncoherent detectors.

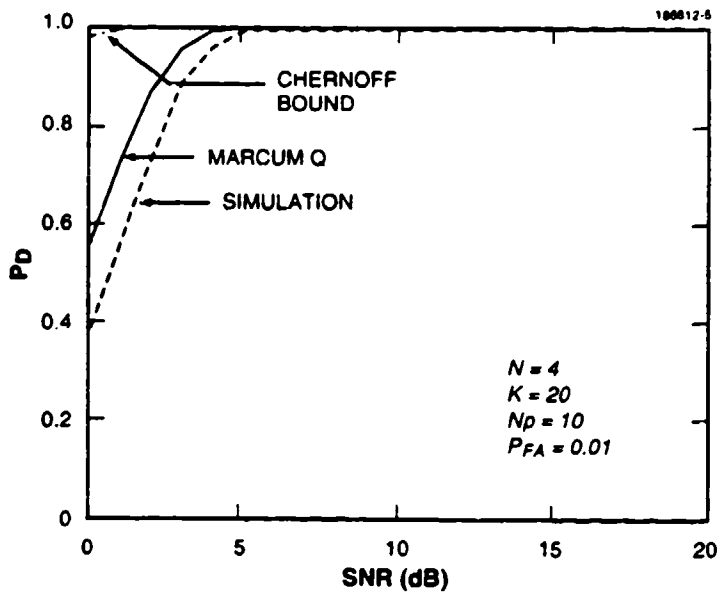

Figure 5. Noncoherent performance comparison for $N_p = 10$.

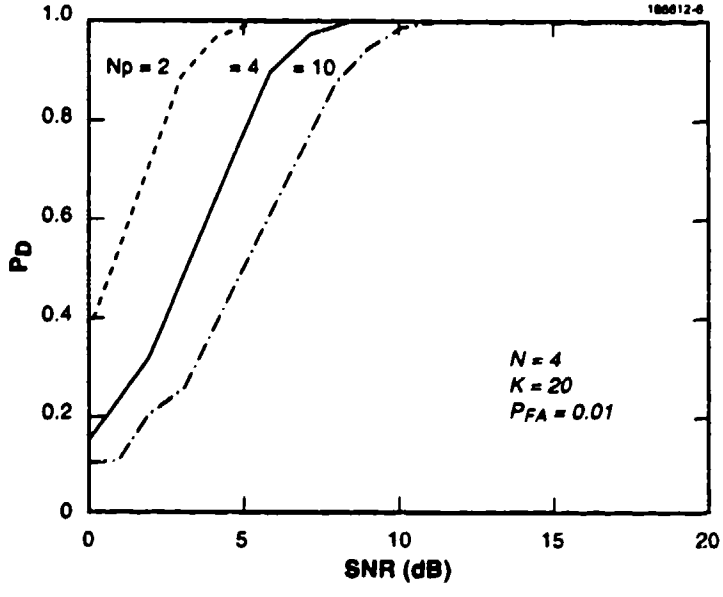

Figure 6. Noncoherent performance as a function of N_p.

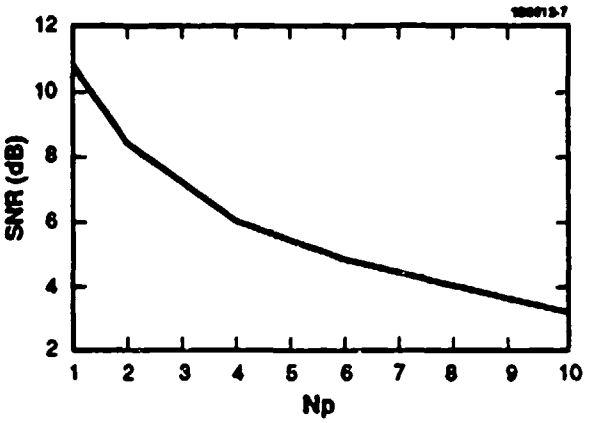

Figure 7. SNR required to obtain $P_D = 0.9$ for $N = 4$ and $K = 20$.

It is evident that the two parameters that control detector performance are the pulse train length N_p and the number of excess secondary vectors L. The array size N is now increased to evaluate its effect. The results illustrated in Figure 8 were produced under the conditions $N = 16, K = 32$, i.e., $L = 17$ and $N_p = 2$. The curves are essentially identical in form to the results portrayed in Figure 2 but with $L = 17$ and $N_p = 2$. Again, as L was increased, the performance approached that of the matched filter in the limit.

The results obtained clearly follow the patterns established for the single-pulse case examined in Kelly [7]. As the number of excess secondaries is increased, a marked improvement occurs in performance. In addition, the detector possesses the desirable CFAR characteristic in that it is statistically independent of the true covariance and is only a function of the dimensional parameters N and K. Given that the data can be processed from a large number of secondaries and a long pulse train, a detector that approaches the matched filter performance, even under low SNR conditions, has been developed.

3.4 Summary of Results

A noncoherent detector that is statistically identical to a special form of the Wilks' Λ test has been presented. The detector, the output of which is the inverse product of N_p complex β-distributed r.v.'s, has been analyzed under both the signal-plus-noise and noise-only hypotheses. Under hypothesis H_0, the detector is statistically equivalent to a central χ^2 r.v. with N_p dof. Under hypothesis H_1, an exact expression for the pdf of the LRT cannot be obtained, and a

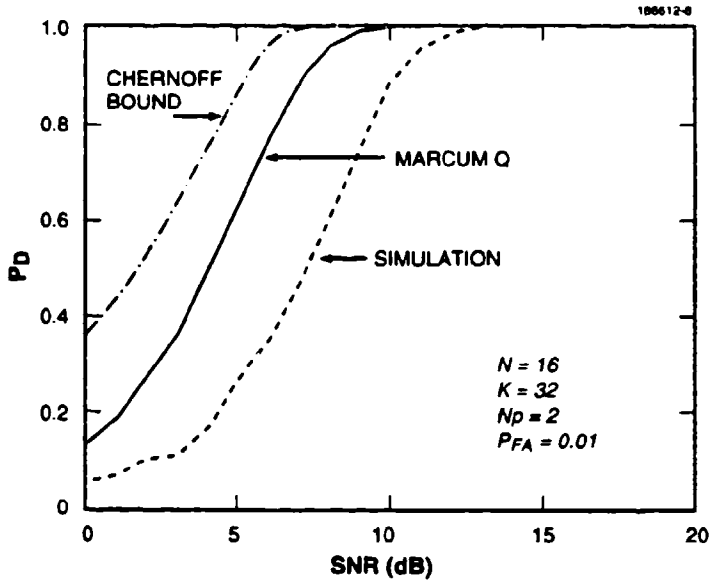

Figure 8. Noncoherent performance comparison for $N = 16$ and $K = 32$.

Chernoff bound on the P_D was derived. The Chernoff bound enabled some degree of confidence in the results obtained through Monte Carlo simulation runs. The detector was shown to provide excellent performance given that the number of excess secondaries is greater than at least $5N$. As the number of pulses is increased, detector performance improves as expected and less SNR is needed to achieve a certain P_D level. The use of additional signal-free data vectors was shown to provide a marked performance improvement in a detector derived from a GLR test formulation.

4. COHERENT ADAPTIVE DETECTION

An approximate solution for the coherent detection problem is developed in this section. Assuming a constant model for the returned amplitude, the maximization of the likelihood ratio is mathematically unmanageable. As a result, an approximation using a truncated series expansion of the GLR forms the basis for the coherent decision rule; however, the exact statistics of the coherent decision rule cannot be obtained. Both the threshold value corresponding to a specified P_{FA} and the ROC are evaluated using Monte Carlo simulations. Under most conditions, the coherent detector outperforms the noncoherent known covariance case as well as the noncoherent unknown covariance detector presented in Section 3.

4.1 An Approximate Multiple-Pulse Decision Rule

In Section 3, the model for the returned signal amplitude was allowed to vary as a function of time. Although it does not match actual conditions over short time intervals, this model enabled the maximization of the GLR with respect to the returned amplitude to be mathematically manageable. Noncoherent detector performance, however, cannot exceed the known covariance case, the performance of which is evaluated using the Marcum-Q function. The goal is to develop a detector that uses the complete N_p pulse observation interval to obtain the MLE (or a reasonable approximation) of the returned amplitude. An estimator of this form enables a coherent decision rule to be obtained as an end result.

Following the analysis used in the formation of the noncoherent GLR, the N_p -pulse likelihood ratio is formed:

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_i = \prod_{i=1}^{N_p} \frac{p_1(\mathbf{z}_i, \mathbf{z}_i(1), \ldots, \mathbf{z}_i(K); \mathbf{M}_{1i}, b)}{p_0(\mathbf{z}_i, \mathbf{z}_i(1), \ldots, \mathbf{z}_i(K); \mathbf{M}_{0i})} \quad , \tag{87}$$

where the returned amplitude b is a constant.

The likelihood functions in Equation (87) are first maximized with respect to the unknown covariances. After taking the $(K+1)$th root of the test statistic, substituting the MLEs of the covariance matrices into their respective likelihood functions and performing some algebraic manipulation, GLR is

$$\Lambda = \prod_{i=1}^{N_p} \frac{\|\hat{\mathbf{M}}_{0i}\|}{\min_b \|\hat{\mathbf{M}}_{1i}\|} \quad , \tag{88}$$

which is still a function of the unknown returned amplitude. Expanding the determinants in the preceding GLR and completing the square with respect to b, results in

$$\Lambda = \prod_{i=1}^{N_p} \frac{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i}{\min\limits_{b} \left\{ 1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i + \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p} \left| b - \frac{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}} \right|^2 - \frac{\left| \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i \right|^2}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}} \right\}} \quad , \tag{89}$$

where

$$\mathbf{S}_i = \sum_{k=1}^{K} \mathbf{z}_i(k) \mathbf{z}_i^H(k) \quad .$$

Observing Equation (89), the maximization of the GLR with respect to b requires the solution of a $2\ N_p$-degree polynomial.

A suboptimal solution technique suggested by Kelly [8,12] begins by defining a new parameter $\hat{b}_i$:

$$\hat{b}_i = \frac{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}} \quad . \tag{90}$$

In terms of $\hat{b}_i$, the GLR is

$$\Lambda = \prod_{i=1}^{N_p} \frac{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i}{\min\limits_{b} \left\{ 1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i + \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p} |b - \hat{b}_i|^2 - \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p} |\hat{b}_i| \right\}} \quad . \tag{91}$$

With the definitions

$$\Lambda_{0i} = \frac{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i}{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i - \frac{\left| \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i \right|^2}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}}} \tag{92}$$

and

$$V_i = \frac{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}}{1 + \mathbf{z}_i^H \mathbf{S}_i^{-1} \mathbf{z}_i - \frac{\left| \mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{z}_i \right|^2}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}}} \quad , \tag{93}$$

the likelihood ratio can be expressed as

$$\Lambda = \prod_{i=1}^{N_p} \Lambda_{0i} \left(1 + V_i \left|b - \hat{b}_i\right|^2\right)^{-1} = \prod_{i=1}^{N_p} \Lambda_{0i} \prod_{i=1}^{N_p} \left(1 + V_i \left|b - \hat{b}_i\right|^2\right)^{-1} \quad . \tag{94}$$

At each time instant i throughout the observation interval, the term $\hat{b}_i$ is equivalent in form to the minimum variance estimator of the true returned amplitude. Assuming reasonable SNRs (or alternatively, high quality covariance estimates), the difference $b - \hat{b}_i$ should be relatively small. Given this assumption, the GLR expressed in Equation (94) is expanded using the following identity. The series expansion for a fraction of the form $(1 + x)^{-1}$ is

$$\frac{1}{1+x} = 1 - x + x^2 - \cdots \quad . \tag{95}$$

If x is small, the series can be truncated after the first-order term without much loss in accuracy; therefore, equating x to $b - \hat{b}_i$, the GLR can be approximated by

$$\prod_{i=1}^{N_p} \Lambda_{0i} \prod_{i=1}^{N_p} \left(1 + V_i \left|b - \hat{b}_i\right|^2\right)^{-1} \approx \prod_{i=1}^{N_p} \Lambda_{0i} \left(1 - \sum_{i=1}^{N_p} V_i \left|b - \hat{b}_i\right|^2\right) \quad . \tag{96}$$

The approximate result in Equation (96) must still maximized with respect to the unknown parameter b. Only the summation is a function of b; therefore, concentration is solely on the term $\sum_{i=1}^{N_p} V_i |b - \hat{b}_i|^2$. Defining the term $S_B(b)$ as

$$S_B(b) = \sum_{i=1}^{N_p} V_i |b - \hat{b}_i|^2 \quad , \tag{97}$$

the magnitude squared term is expanded to obtain

$$S_B(b) = |b|^2 \sum_{i=1}^{N_p} V_i - 2\mathrm{Re}\left\{b^* \sum_{i=1}^{N_p} V_i \hat{b}_i\right\} + \sum_{i=1}^{N_p} V_i \left|\hat{b}_i\right|^2 \quad . \tag{98}$$

Completing the square with respect to b yields

$$S_B(b) = |b - \hat{b}|^2 \sum_{i=1}^{N_p} V_i + \sum_{i=1}^{N_p} V_i |\hat{b}_i|^2 - |\hat{b}|^2 \sum_{i=1}^{N_p} V_i \quad . \tag{99}$$

where

$$\hat{b} = \frac{\sum_{i=1}^{N_p} V_i \hat{b}_i}{\sum_{i=1}^{N_p} V_i} \quad .$$

The estimator $\hat{b}$ is a normalized, weighted sum that is an approximate MLE of the returned amplitude. In a stationary environment, the expected value of each V_i would be equal; therefore, under stationary conditions the expected value of the estimate would be the average of the N_p independent minimum variance unbiased estimates of the returned amplitude. The V_i terms may be viewed as weights that are proportional to the SNR. In a nonstationary environment the coherent estimator will place a greater emphasis on estimates obtained from high SNR conditions as opposed to conditions where there is much interference.

By substituting $\hat{b}$ for b in the expression for $S_B(b)$, the minimum value is obtained and is denoted as $\hat{S}_B$:

$$\hat{S}_B = \sum_{i=1}^{N_p} V_i \left|\hat{b}_i\right|^2 - \frac{\left|\sum_{i=1}^{N_p} V_i \hat{b}_i\right|^2}{\sum_{i=1}^{N_p} V_i} \quad . \tag{100}$$

As a result, the final form of the likelihood ratio is

$$\Lambda = \left(1 - \hat{S}_B\right) \prod_{i=1}^{N_p} \Lambda_{0i} \quad . \tag{101}$$

4.2 The Statistical Analysis of the Coherent GLR Test

The coherent decision rule bears some resemblance to the noncoherent decision rule. Each Λ_{0i} of Equation (101) is the likelihood ratio term found in Section 3. Hence, in statistical terms the coherent likelihood ratio is the product of $(1 - \hat{S}_B)$ and the product of N_p β-distributed r.v.'s. Because the coherent likelihood ratio contains another level of complexity beyond the noncoherent likelihood ratio, i.e., the presence of $(1 - \hat{S}_B)$, it is not expected that analysis will yield an exact statistical characterization of the test. For the sake of completeness, the statistics of $\hat{S}_B$ are now examined.

Following standard multivariate statistical analysis procedures and steps presented earlier, the primary and K secondary vectors are whitened and rotated. Coordinate rotation is accomplished via a unitary transformation that places the nonzero mean component of the primary data vector

into the vector first element. The whitened, rotated, Gaussian N-vectors defined in terms of the partitioned components are

$$\mathbf{p} = [1 \quad 0 \ldots 0] \quad , \tag{102}$$

$$\mathbf{z}_i = [\mathbf{z}_A \ \mathbf{z}_B]^{\mathrm{T}} \quad ,$$

$$\mathbf{S}_i = \begin{bmatrix} \mathbf{S}_{AA} & \mathbf{S}_{AB} \\ \mathbf{S}_{BA} & \mathbf{S}_{BB} \end{bmatrix} \quad ,$$

and

$$\mathbf{P}_i = \begin{bmatrix} \mathbf{P}_{AA} & \mathbf{P}_{AB} \\ \mathbf{P}_{BA} & \mathbf{P}_{BB} \end{bmatrix} \quad ,$$

where

$$\mathbf{P}_i = \mathbf{S}_i^{-1} \quad .$$

The matrix $\mathbf{P}_i$ is the inverse of K times the sample covariance formed from the secondary vectors at time instant i. The partitions of $\mathbf{z}_i, \mathbf{S}_i^{-1}$, and $\mathbf{P}_i$ are implicitly understood to be functions of time. The subscript i is suppressed to make subsequent formulas easier to read.

Next $\hat{\mathbf{S}}_B$ is expressed in terms of the partitions. (A more detailed description of the terms that follow can be found in Section 2.) For clarity the r.v. T_i is again defined:

$$T_i = \frac{1}{\mathbf{p}^H \mathbf{S}_i^{-1} \mathbf{p}} = \mathbf{P}_{AA}^{-1} = \mathbf{S}_{AA} - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{S}_{BA} \quad . \tag{103}$$

T_i is an unconditional central χ^2 r.v. with $L = K - N + 1$ dof. The complex amplitude estimator $\hat{b}_i$ possesses the form

$$\hat{b}_i = \mathbf{z}_A - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{z}_B \quad . \tag{104}$$

The parameter $\hat{b}_i$ is itself a conditional Gaussian r.v. and is the prediction error of $\mathbf{z}_A$ given $\mathbf{z}_B$. The β-distributed loss factor ρ_i is defined:

$$\rho_i = \frac{1}{1 + \mathbf{z}_B^H \mathbf{S}_{BB}^{-1} \mathbf{z}_B} \quad . \tag{105}$$

Using these definitions, V_i in terms of the partitions [see Equation (95)] is

$$V_i = \frac{\mathbf{P}_{AA}}{1 + \mathbf{z}_B^H \mathbf{S}_{BB}^{-1} \mathbf{z}_B} = \frac{\rho_i}{T_i} \quad , \tag{106}$$

which is conditionally distributed as the inverse of a χ^2 r.v. Expressing $\hat{\mathrm{S}}_B$ [see Equation (102)] in terms of the partitioned components results in

$$\hat{\mathrm{S}}_B = \sum_{i=1}^{N_p} \frac{\left| \frac{\mathbf{z}_A - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{z}_B}{\left(1+\mathbf{z}_B^H \mathbf{S}_{BB}^{-1}\mathbf{z}_B\right)^{\frac{1}{2}}} \right|^2}{T_i} - \frac{\left| \sum_{i=1}^{N_p} \frac{\left(\frac{\mathbf{z}_A - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{z}_B}{1+\mathbf{z}_B^H \mathbf{S}_{BB}^{-1}\mathbf{z}_B} \right)}{T_i} \right|^2}{\sum_{i=1}^{N_p} \frac{\rho_i}{T_i}} \quad , \tag{107}$$

or equivalently, as

$$\hat{\mathrm{S}}_B = \sum_{i=1}^{N_p} \frac{|\nu_i|^2}{T_i} - \frac{\left| \sum_{i=1}^{N_p} \frac{\rho_i^{1/2} \nu_i}{T_i} \right|^2}{\sum_{i=1}^{N_p} \frac{\rho_i}{T_i}} \quad , \tag{108}$$

where

$$\nu_i = \frac{\mathbf{z}_A - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{z}_B}{\left(1 + \mathbf{z}_B^H \mathbf{S}_{BB}^{-1}\mathbf{z}_B\right)^{\frac{1}{2}}} \quad . \tag{109}$$

Section 2 showed that each $|\nu_i|^2/T_i$ term is conditionally F-distributed; therefore, the first summation in $\hat{\mathrm{S}}_B$ in the sum of N_p conditional F r.v.'s. The second term in $\hat{\mathrm{S}}_B$ is examined in two parts. The denominator, where each V_i is equivalent to ρ_i/T_i, is conditionally distributed as the sum of the inverse of a χ^2 r.v. The individual terms in the numerator summation, $V_i\hat{b}_i$, are equivalent to

$$V_i\hat{b}_i = \frac{\frac{(\mathbf{z}_A - \mathbf{S}_{AB}\mathbf{S}_{BB}^{-1}\mathbf{z}_B)}{1+\mathbf{z}_B^H\mathbf{S}_{BB}^{-1}\mathbf{z}_B}}{T_i} \quad . \tag{110}$$

Each term in the summation $\sum_{i=1}^{N_p} V_i\hat{b}_i$ is the ratio of a conditional normal r.v. and a χ^2 r.v. The overall numerator term is the square of the sum of N_p such terms. Unfortunately, the resulting distribution is not of a common form. The distribution of the ratio of the square of this sum and the denominator term must also be determined. The second part of $\hat{\mathrm{S}}_B$ is practically impossible to analyze.

The exact statistical characterization of the coherent decision rule, let alone expressions for the $\mathrm{P_D}$ and $\mathrm{P_{FA}}$, is virtually impossible to obtain; therefore, the performance of the decision rule is evaluated experimentally, using Monte Carlo simulations.

4.3 Simulation Results

Because statistical analysis of the coherent decision rule is so cumbersome, Monte Carlo simulations were used to obtain performance measures. The threshold value corresponding to a specified $\mathrm{P_{FA}}$ was experimentally obtained and then used in the evaluation of the $\mathrm{P_D}$. The performance of the coherent decision rule is analyzed in both stationary and nonstationary environments. Coherent detector performance (assuming an unknown true covariance matrix) is shown to exceed the performance of a detector derived from the noncoherent, known covariance matrix.

The coherent decision rule upon which the subsequent results are based is

$$\Lambda = \left(1 - \hat{\mathrm{S}}_B\right) \prod_{i=1}^{N_p} \Lambda_{0i} \gtrless \alpha \quad , \tag{111}$$

where α is the threshold parameter. In Section 4.2, the exact statistics of $\hat{\mathrm{S}}_B$ proved to be unobtainable; therefore, the test statistic must be evaluated using the raw data $\mathbf{z}_i$ and $\mathbf{z}_i(k)$ which are multivariate Gaussian random vectors. Using Equations (4), (6), (7), and (15), the decision rule can be computed simply in terms of $\mathbf{z}_i$, $\mathbf{z}_i(k)$, which is used to evaluate $\mathbf{S}_i^{-1}$, and the steering vector $\mathbf{p}$.

Because the steering vector enters the decision rule, a particular array geometry must now be assigned. A uniformly (half-wavelength) spaced linear array was chosen for the sake of computational simplicity. The N-element steering vector for this array geometry is

$$\mathbf{p} = \left[1 \mathrm{e}^{-j\pi\sin\theta} \ldots \mathrm{e}^{-j(N-1)\pi\sin\theta}\right] \quad . \tag{112}$$

Again, for simplicity, the target is assumed to be located broadside to the array. The steering vector under this assumption has the following form:

$$\mathbf{p} = [1 \quad 1 \ldots 1] \quad . \tag{113}$$

Without changing the test, $\mathbf{p}$ is redefined as a unit vector.

The noise environment is composed of directional interference in addition to Gaussian background noise. Assuming that the jth directional interference source is narrowband, located in the far-field, and that its power is denoted as σ_j^2, the total noise covariance matrix can be represented as

$$\mathbf{M}_i = \sum_{j=1}^{N_j} \sigma_j^2 \mathbf{q}_j \mathbf{q}_j^H + \sigma_N^2 \mathbf{I} \quad , \tag{114}$$

where N_j is the number of interference sources, σ_j^2 is the interference power, and $\mathbf{q}_j$ is the steering vector of the interference.

The vectors $\mathbf{z}_i$ and $\mathbf{z}_i(k)$ both share the covariance matrix $\mathbf{M}_i$. To create multivariate Gaussian random vectors with covariance $\mathbf{M}_i$, the zero-mean N-vectors $\mathbf{z}_{wi}$ and $\mathbf{z}_{wi}(k)$ are formed first, using a random number generation program. Because the elements of these data vectors are statistically independent, the resulting covariance matrix is the identity matrix. The data vectors are then statistically transformed to possess the desired covariance using the Cholesky factorization of the true covariance matrix. The Cholesky factor $\mathbf{C}_i$, which is defined $\mathbf{C}_i^H \mathbf{C}_i = \mathbf{M}_i$, is applied to the data in the following manner:

$$\mathbf{z}_i = \mathbf{C}_i \mathbf{z}_{wi} \tag{115}$$

$$\mathbf{z}_i(k) = \mathbf{C}_i \mathbf{z}_{wi}(k) \quad . \tag{116}$$

The covariance matrix of the respective transformed data vectors is now the desired covariance $\mathbf{M}_i$. For Gaussian white noise processes, the transformation using the Cholesky factorization is not necessary.

Under H_1, the expected value of the primary vector $\mathbf{z}_i$ is $b\mathbf{p}$. To form a nonzero mean primary, $b\mathbf{p}$ is added to the random primary vector. The value of b, which corresponds to the target radar cross-sectional area, is computed assuming that the covariance matrix is the identity matrix. The known white-noise case is used as a reference for two reasons. First, assuming that the noise is white with unit variance and the steering vector is of unit length, the ideal nonadaptive SNR

reduces to $|b|^2$. Second, the reference SNR is now independent of jammer locations and powers. As a result, we can maintain constant SNR while varying the noise environment over an observation interval; therefore, all subsequent tests use the same target returned power for a particular SNR, which results in valid comparisons. The mean returned amplitude value is

$$b = 10^{\frac{SNR}{10}} \quad , \tag{117}$$

where the SNR is in decibels. At i, the matrix $\mathbf{S}_i^{-1}$ is obtained by forming the outer product of K secondary vectors, summing the outer products together and computing the inverse. The discussion of the creation of all the constituent components of the decision rule is now complete.

As previously stated, the threshold cannot be explicitly calculated and must be evaluated numerically. In the discussion of the simulations used to obtain the performance characteristics of the noncoherent detector, the reasons for the choice of $P_{FA} = 0.01$, although very high for radar systems, were explained. Assuming that the overall detection scheme incorporates M out of N decision rules, the resulting P_{FA} can be brought down into acceptable levels for radar. Also, the number of computations required to produce statistically significant results is drastically reduced when compared with the required number of tests conducted under extremely small P_{FA} conditions.

The P_{FA} is computed under the assumption that the primary data is zero-mean. For a given noise environment structure, 10,000 zero-mean primary data vectors were used as inputs to the coherent decision rule. At each i during the N_p pulse observation interval, a new covariance estimate was formed. The decision rule outputs were then used to create a sample pdf. The threshold value corresponding to $P_{FA} = 0.01$ was then calculated using the definition of the cumulative distribution function. This process was repeated several times, and the average of the threshold values was used to evaluate the ROC.

As per the analysis of the noncoherent detector, the performance of the coherent detector is analyzed as a function of K and N_p. In addition, the coherent detector is studied under both stationary and nonstationary noise conditions. For all cases examined in the sequel, the P_{FA} was set to 0.01. The P_D at a particular SNR was evaluated using 1,000 trials.

The performance of the coherent detector was evaluated with N, the number of array elements, set equal to 4. The particular value of N affects the performance of the detector only through the parameter L; therefore, the results for $N = 4$ can be extrapolated to larger-sized arrays by varying K. The added benefit of a small N value is that the number of computations is significantly reduced due to smaller vector sizes.

For the current analysis the known covariance case is used as the benchmark reference. The performance of the noncoherent detector under known covariance conditions is obtained using the Marcum-Q function. The Marcum-Q designation in the legends of Figures 9–11, 14, and 17 corresponds to this case. The performance of coherent detector under these conditions was obtained as follows. Because the covariance is assumed known, the secondaries do not provide any additional

information and are dropped from the likelihood functions (which are the product of N_p identical multivariate Gaussian distributions, where the true mean value is unknown under hypothesis H_1). After some algebraic manipulation, the final form of the GLR test takes the form

$$\Lambda = \frac{\left| \sum_{i=1}^{N_p} \mathbf{p}^H \mathbf{M}_i^{-1} \mathbf{z}_i \right|^2}{\sum_{i=1}^{N_p} \mathbf{p}^H \mathbf{M}_i^{-1} \mathbf{p}} \gtrless \alpha \quad . \tag{118}$$

After performing a whitening transformation and a coordinate rotation, the test can be shown under H_0 to be equivalent to a central χ^2 r.v. with one complex dof. The threshold value for the required P_{FA} is the logarithm of the P_{FA}. Under H_1, the test is distributed as a noncentral χ^2 r.v. with one complex dof and a noncentrality parameter $c = N_p$ SNR. The cdf of this Rician r.v. is evaluated using the Marcum-Q function. The ROC corresponding to the coherent detector under known covariance conditions is designated KNOWN COVARIANCE in the figures.

Detector performance is now examined in greater detail. It is assumed that the noise environment is known at each i whether it is stationary or nonstationary; therefore, as long as the type of noise environment that is present during the ith sample interval is known, the interference is always nulled (if present) due to the $\mathbf{S}_i^{-1}$ term found in the decision rule. Detector performance is essentially independent of the time-varying nature of the noise environment if the covariance matrix is continuously updated.

A typical case, $N_p = 4$ and $K = 20$, is shown in Figure 9. The coherent detector clearly outperforms the noncoherent detector examined in Section 3. In fact, the coherent detector requires about 1 dB less signal power in the low SNR region of the ROC than the known covariance matched filter represented by the Marcum-Q curve. For this value of K, coherent detector performance is down only 1 dB from optimal levels. The improved performance under the stated conditions provided by the coherent detector was found to occur for all values of N_p that were examined.

From previous analyses the number of secondary data vectors has been shown to play a major role in the performance of the GLR test-based detectors. Figures 1 and 3 illustrate the asymptotic nature of performance of these detectors as a function of K. For the coherent detector, Figure 10 shows that K is a controlling factor. Only when K exceeded 16 did the performance surpass that of the noncoherent, known covariance case. As K increases from 20 to 30, the ROC begins to approach the theoretical maximum. As the number of pulses in the observation interval is increased, the number of secondaries needed to surpass optimal noncoherent detector performance increases (see Figure 11).

Because of integration gain, as the number of pulses in the observation interval is increased the coherent detector should provide increased the performance (less SNR required to achieve a specified P_D). Figure 12 illustrates the ROC for the reported standard array size and $K = 30$

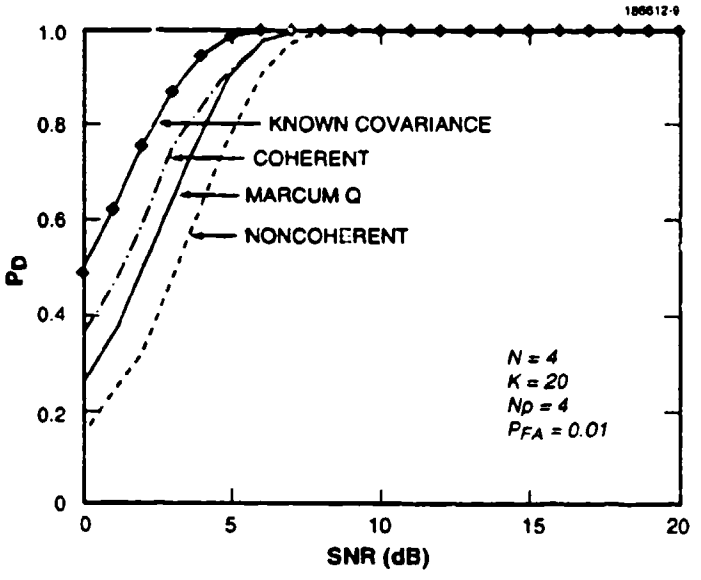

Figure 9. Performance comparison between detectors.

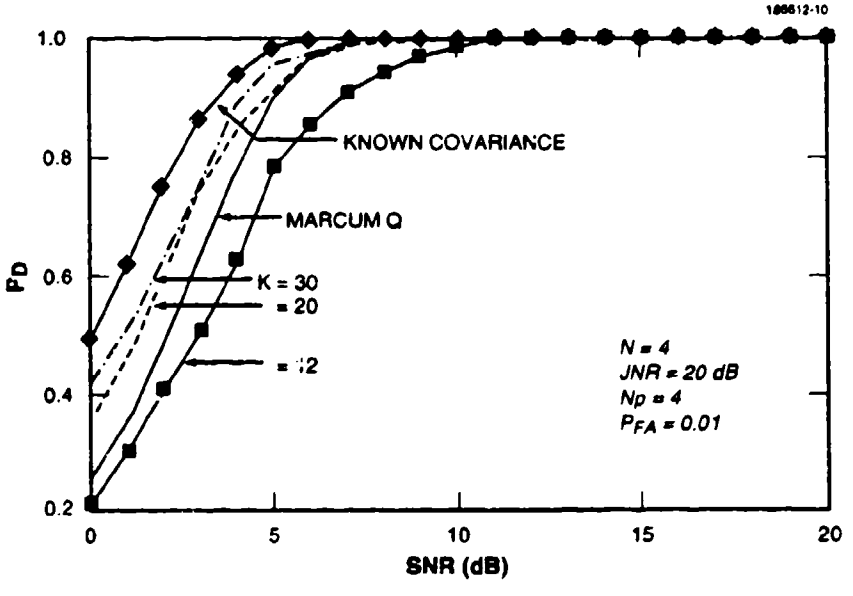

Figure 10. Coherent detector performance as a function of K for $N_p = 4$.

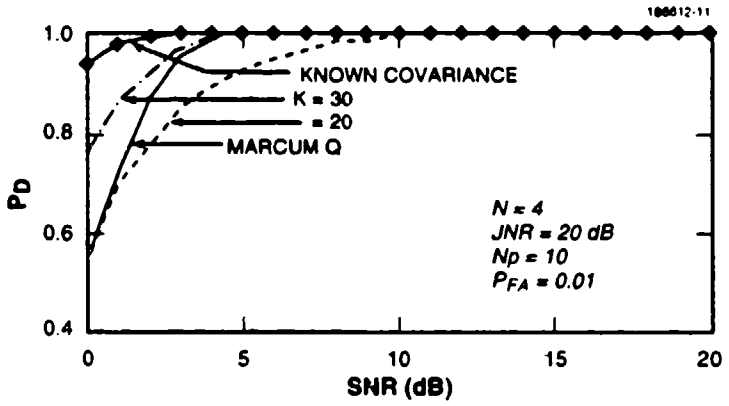

Figure 11. Coherent detector performance as a function of K for N_p= 10.

as a function of N_p. These simulation results follow the ROC curves of the matched filter as a function of the number of samples, i.e., the curves move upward and leftward as N_p increases. The observation that the integration gain effectively reduces the SNR required to achieve a specified P_D is verified in Figure 13 for $N = 4$ and $K = 30$.

The performance of the coherent detector, given a wide variety of noise and interference environments, is now examined. The presence of the inverse of K times the sample covariance in the decision rule suggests that directional interference should be effectively nulled if K is sufficiently large. The effect of different interference powers and the number of jammers on coherent detector performance are considered. To perform this analysis, a stationary environment is assumed by holding the jammer power and the number of jammers fixed over the observation interval. These constraints will be relaxed and their effects examined. Figure 14 shows detector performance for two different jammer-to-noise ratios (JNRs) as virtually identical. JNR is defined as σ_J^2/σ_N^2. Under both JNR scenarios and white noise conditions, coherent detector performance is superior to that of the noncoherent matched filter. Figure 15 portrays the effect of the number of interferers. The detector produces essentially identical performance levels regardless of interference structure.

At this point in the analysis it is clear that as long as K is sufficiently large and the statistics of the noise environment are known at each sample instant, the coherent detector outperforms the noncoherent detector and the noncoherent matched filter detector. It should be noted that the actual interference directions do not affect performance unless they lie inside the main lobe of the

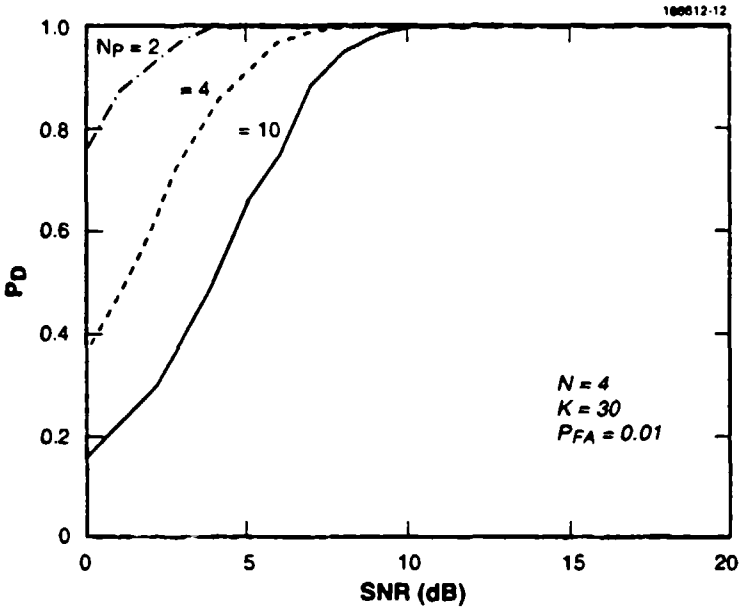

Figure 12. Coherent detector performance as a function of N_p.

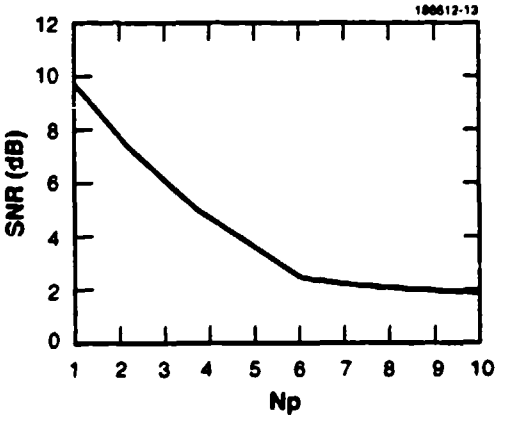

Figure 13. SNR Required for $P_D = 0.9$ as a function of N_p.

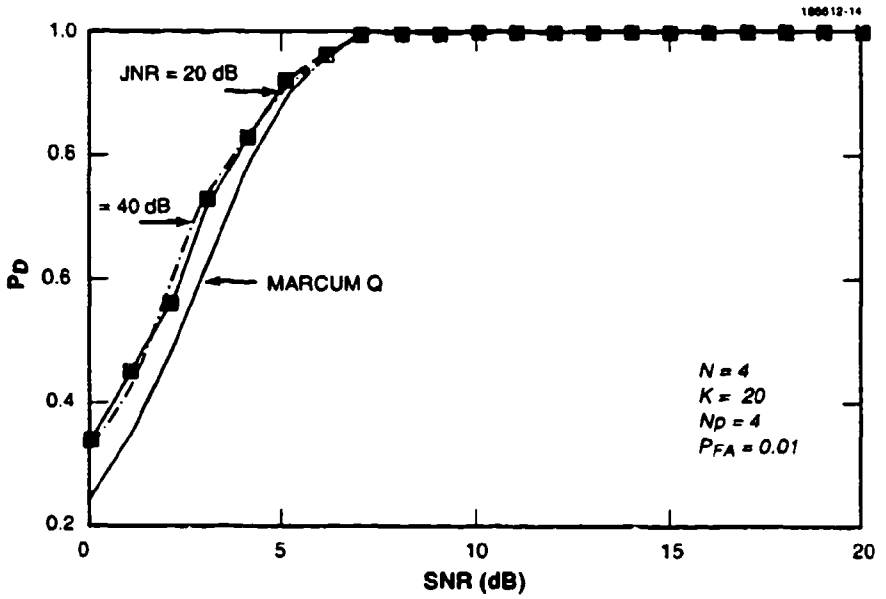

Figure 14. Coherent detector performance as a function of JNR.

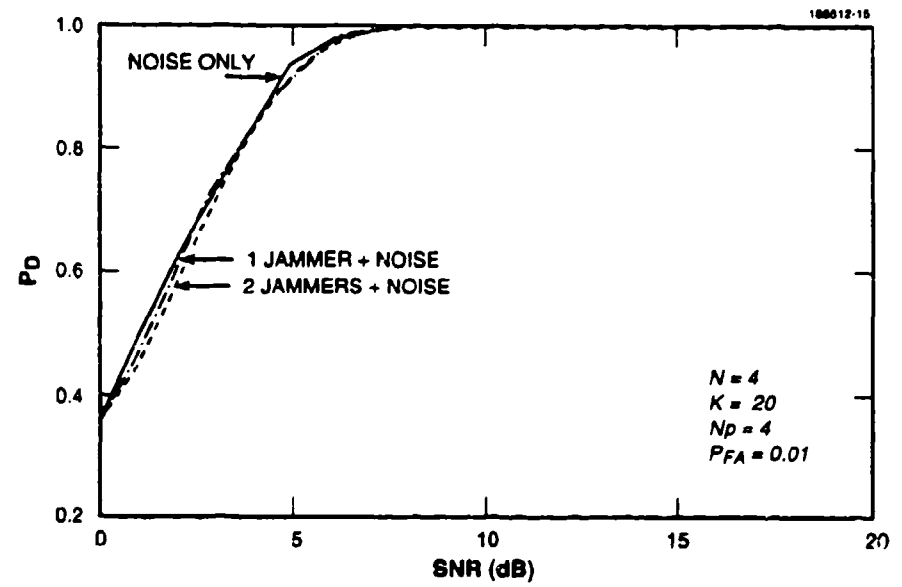

Figure 15. Coherent detector performance as a function of number of jammers.

array pattern. For $N = 4$, the main beam ranges between $\pm 30°$. If a jammer lays inside the main lobe, performance is seriously degraded.

So far, the ability to continuously update the covariance estimate has been a basic premise of this study. In actual practice this procedure involves collecting a large number of data vectors, forming outer products, and taking the inverse of an $N \times N$ matrix. Much processing is required and could be very expensive. The issue of a limited number of covariance estimates during an observation interval is now addressed.

The stationary noise environment is examined, and two extreme cases are considered. Figure 16 compares the performance of the coherent detector using a single and a continuously updated covariance estimate. In a stationary environment the additional covariance estimates provide little, if any, additional benefit. As in case 2, by rotating the data via a unitary transformation, the problem can be reduced to a single primary vector along with $K \times N_p$ secondary vectors. For a fixed number of array elements, however, a large increase in the number of secondaries (greater than 20 for $N = 4$) provides only small increases in performance; therefore, the results illustrated in Figure 16 are consistent with existing results.

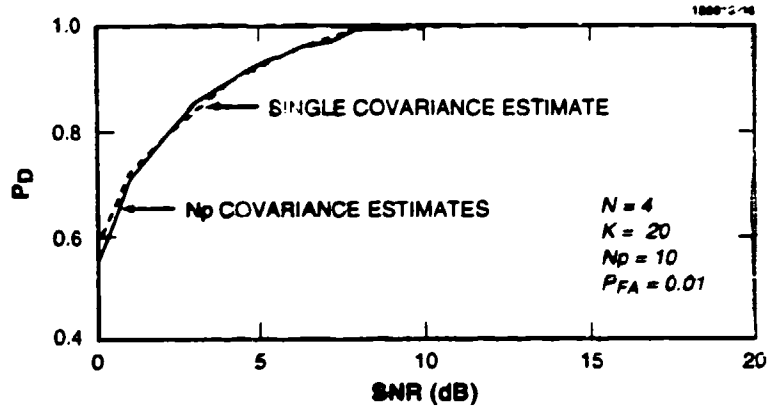

Figure 16. Coherent detector performance in a stationary environment.

The main goal of this study was to develop a detector that would exceed noncoherent matched filter performance levels in stationary and nonstationary noise environments. It remains to examine

the performance of the coherent detector in nonstationary environments with a limited number of covariance estimates. The blinking jammer is an extreme case of a nonstationary noise environment where the jammer randomly switches between ON and OFF. While OFF, the noise is simply white. If the covariance estimate can be continuously updated regardless of the jammer state, the coherent detector can achieve maximum performance because its structure enables it to effectively null out the jammer at each time instant. If the number of estimates is limited throughout the observation interval, the performance is expected to degrade for the following reasons. Suppose only a single covariance estimate is made at the beginning of the data collection period. If the jammer is ON, it could be nulled out at any time during this period; however, if the jammer is OFF during the single sample of the K secondaries, the sample covariance matrix under these conditions is not able to provide nulling capabilities. The result is performance loss.

Figure 17 provides a comprehensive review of these possible conditions in a nonstationary environment. (For reference purposes, the curve corresponding to continuous covariance updating is included.) Switching between ON and OFF was uniformly varied during the observation interval. The single covariance white noise curve corresponds to the case where the estimate contains only white noise. Performance under this condition is down 5 dB when compared with the reference at PD = 0.09; however, when the estimate includes the effect of the jammer, performance is equivalent to that of the reference case. Because the number of noise-only and jammer-plus-noise samples were equally likely, it is not surprising that the average performance curve lies roughly halfway between the two previously mentioned cases.

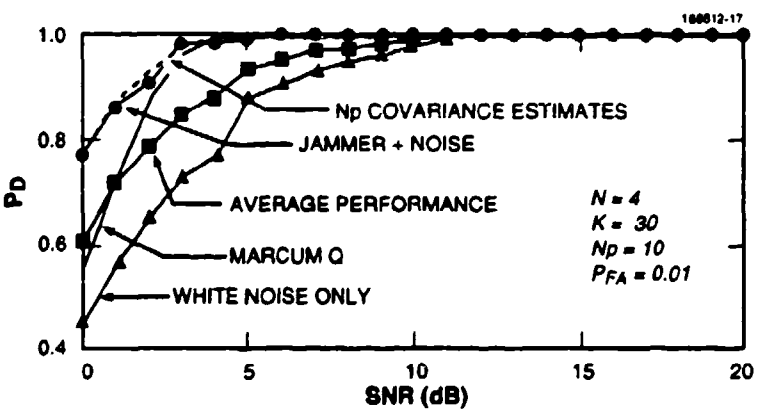

Figure 17. Coherent detector performance under blinking jammer conditions.

The conclusion reached from these results is that if the noise environment rapidly fluctuates, it pays to frequently update the covariance estimate; also if a sufficient number of secondaries is used, the coherent detector will exceed the noncoherent matched filter levels even though the true covariance matrix is both unknown and time varying.

4.4 Summary of Results

A coherent detector that is based upon an approximate GLR test has been presented. The approximation was required to enable the maximization of the test with respect to the constant unknown returned amplitude b. The detector was manipulated into a form such that the test is the product of the noncoherent decision rule and the term $(1 - \hat{S}_B)$. Beyond a certain number of secondaries, the coherent detector performed very well, always surpassing noncoherent matched filter and noncoherent detector performance levels. Even under extreme conditions of nonstationarity, the coherent detector produced very good results, especially when continuous covariance estimate updating was allowed. In conclusion, the coherent decision rule yields excellent adaptive detection performance in both stationary and nonstationary noise environments.

5. CONCLUSIONS

Two new multiple-pulse GLR test-based detectors that provide excellent performance in nonstationary noise environments have been presented. Due to the complexity of the analysis, little work has been done in the analysis of nonstationary systems; however, because nonstationary noise environments are quite typical in the radar context, the results obtained in this study are important. One of the strengths of this analysis is that only precise analytical techniques were considered as opposed to commonly used ad hoc methods.

The analysis of the coherent and noncoherent detectors are extensions of the single-pulse detector that was completely analyzed by Kelly. The single-pulse decision rule was derived under the assumptions that the true covariance matrix and the complex returned amplitude were unknown parameters. To overcome covariance estimation losses that degraded detection performance, additional signal-free data vectors were incorporated into the decision rule that resulted in markedly improved performance. The decision rule was shown to be statistically equivalent to the inverse of a conditional β-distributed r.v.

The multiple-pulse GLR test was also formulated using additional secondary vectors. Also, the true covariance matrix was allowed to vary over the N_p pulse sampling intervals. Unfortunately, the maximization of the likelihood ratio with respect to the returned amplitude proved to be mathematically cumbersome, and to cope with this problem two different solution techniques were considered.

The first solution was to allow the model for the signal amplitude to be time varying with the resulting test statistically equivalent to the product of N_p β-distributed r.v.'s. By taking the logarithm of the product, the test involved the summation of N_p independent r.v.'s being compared to a threshold, hence, the noncoherent detector nomenclature. Through the judicious use of a transformation, the threshold value was readily obtained. The exact analysis of the test under H_1 was shown to be virtually impossible, but a Chernoff bound was derived. As the number of secondary vectors used for covariance estimation increased, simulations showed that noncoherent detector performance approached matched filter levels.

The second solution technique involved the use of an algebraic identity that enabled the test to be expressed as a series expansion. Truncating the series after the first term allowed the test to be maximized with respect to the unknown returned amplitude. Because the complete observation interval (corresponding to coherent processing) was used to evaluate the maximum likelihood estimate of b, the detector was designated as coherent. The exact statistics of the coherent decision rule were not obtained and Monte Carlo simulation was used to generate ROCs. As expected, the coherent detector surpassed the performance of the noncoherent detector for the same number of secondary vectors and the coherent detector asymptotically approached the coherent, known covariance detector performance levels. The coherent detector even proved capable of handling the highly nonstationary blinking jammer.

In conclusion, both the coherent and noncoherent structures provide a framework to analyze detection performance in nonstationary noise environments. The results show that given sufficient processing power and a reasonable number of terms available for covariance estimation, detectors can be designed to operate in rapidly fluctuating environments. In particular, the coherent detector can provide almost optimal performance levels despite the necessity of estimating the covariance matrix and that the true noise statistics are time varying.

REFERENCES

1. R.A. Monzingo and T.W. Miller, *Introduction to Adaptive Arrays*, New York: John Wiley (1980).
2. H.L. Van Trees, *Detection, Estimation, and Modulation Theory, Part 1*, New York: John Wiley (1968).
3. M.I. Skolnick, *Radar Handbook*, 2nd ed, New York: McGraw-Hill (1990).
4. P. Swerling, "Probability of detection for fluctuating targets," *IRE Trans. Inf. Theory* **IT-6**, 269–308 (1960).
5. E.J. Kelly and K.M. Forsythe, "Adaptive Detection and Parameter Estimation for Multidimensional Signal Models," Technical Rep. 848, Lexington, Mass.: MIT Lincoln Laboratory (19 April 1988), DTIC AD-A208971.
6. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill (1965).
7. E.J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron Syst.*, **AES-22** (1986).
8. E.J. Kelly, "Adaptive Detection in Non-Stationary Interference, Part I and Part II," Technical Rep. 724, Lexington, Mass.: MIT Lincoln Laboratory, (25 June 1985), DTIC AD-A158810.
9. E.T. Copson, *Theory of Functions of a Complex Variable*, Oxford: Oxford University Press (1935).
10. S.S Wilks, "Certain Generalizations in the Analysis of Variance," *Biometrika*, **24**, 471–494 (1932).
11. I.S. Reed, J.D. Mallett, and L.E. Brennan, "Rapid convergence rates in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, **AES-10** (1974).
12. E.J. Kelly, Private communication (Technical Memorandum, 7 November 1988).
13. P. Monticciolo, E.J. Kelly, and J.G. Proakis, "A noncoherent adaptive detection scheme," *IEEE Trans Aerosp. Electron. Syst.* (1992).
14. M. Abramovitz and I.A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards Applied Mathematics Series 55, Washington, D.C.: U.S. Government Printing Office (1972).
15. M. Schatzoff, "Exact distributions of Wilks' likelihood ratio criterion," *Biometrika*, **53**, 347–358 (1966).
16. T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley (1958).

REFERENCES
(Continued)

17. J.H. Maindonald, *Statistical Computation*, New York: John Wiley (1984).

18. S.F. Arnold, *The Theory of Linear Models and Multivariate Analysis*, New York: John Wiley (1981).

19. E.J. Kelly, "Adaptive Detection in Non-Stationary Interference, Part III," Technical Rep. 761, Lexington, Mass.: MIT Lincoln Laboratory, (24 August 1987), DTIC AD-A185622.

20. E.J. Kelly, "Performance of an adaptive detection algorithm: rejection of unwanted signals," *IEEE Trans. Aerosp. Electron. Syst.*, **AES-25** (1989).

21. G.E.P. Box, "A general distribution theory for a class of likelihood criteria," *Biometrika*, **36**, 317–341 (1949).

22. R. Muirhead, "Asymptotic distributions of some multivariate tests," *Annals of Math. Stat.*, **4**, 1002–1010 (1970).

23. J.G. Proakis, *Digital Communications*, New York: McGraw-Hill (1983).

24. A. Erdyli, *Higher Transcendental Functions*, New York: McGraw-Hill (1953).

25. D.A. Shnidman, "Efficient evaluation of the probabilities of detection and the generalized Q-function," *IEEE Trans. Inf. Theory*, **22**, 746–751 (1976).

26. J. Margolin, "Detection Algorithms for Partially Correlated, Fluctuating Targets," Paris: Int. Conf. on Radar (1989).

27. R.L. Mitchell, "Importance sampling applied to simulation of false alarm rates," *IEEE Trans. Aerosp. Electron. Syst.*, **AES**, 15–24 (1981).

28. G.W. Lank, "Theoretical aspects of importance sampling applied to false alarm rates," *IEEE Trans. Inf. Theory*, **IT-29**, 73–83 (1983).

29. D. Lu and K. Yao, "Improved importance sampling techniques for efficient simulation of digital communications systems," *IEEE J. Sel. Areas Commun.*, **6**, 67–75 (1988).

30. M. Guida, D. Iovino, and M. Longo, "Comparative performance analysis of some probability tails," *IEEE J. Sel. Areas Commun.*, **6**, 76–84 (1988).

31. B.J.T. Morgan, *Elements of Simulation*, London: Chapman and Hall (1984).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	24 February 1994	Technical Report

4. TITLE AND SUBTITLE

Adaptive Detection in Stationary and Nonstationary Noise Environments

5. FUNDING NUMBERS

C — F19628-90-C-0002
PR — 373

6. AUTHOR(S)

Paul Monticciolo

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Lincoln Laboratory, MIT
P.O. Box 73
Lexington, MA 02173-9108

8. PERFORMING ORGANIZATION REPORT NUMBER

TR-937

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Air Force ESC
Hanscom AFB
Bedford, MA 01730

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ESC-TR-93-238

11. SUPPLEMENTARY NOTES

None

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

This report is a study of the statistical performance of several radar-based adaptive detection schemes in both stationary and nonstationary noise and interference environments. The detectors under study must be able to correctly determine the presence of a target in a range gate with a high degree of probability given that the probability of misclassification is a fixed small value. The hostile noise environment is assumed to consist of possibly time varying, spatially correlated interference along with Gaussian background noise. In a typical radar environment, the mean value of the returned radar signal and the noise covariance matrix are unknown parameters. Therefore, generalized likelihood ratio procedures were used to develop decision rules that meet the Neyman-Pearson criterion. Three major cases of interest were examined. First, the single-pulse test developed by Kelly[1] is reviewed. The multiple-pulse return test case is extremely complicated and was divided into distinct detector forms: noncoherent and coherent. The performance of each detector is a function of the signal-to-noise ratio, the number of radar pulse returns used in the decision rule, and the quality of the covariance estimate.

[1]E.J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron Syst.*, **AES-22** (1986).

14. SUBJECT TERMS		15. NUMBER OF PAGES
adaptive detection	Chernoff bound	76
generalized likelihood ratio	nonstationary noise environments	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Same as Report