ESD-TR-85-197

Technical Report
724

AD-A158 810

Adaptive Detection in Non-Stationary Interference, Part I and Part II

E.J. Kelly

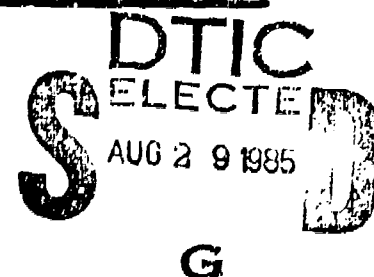25 June 1985

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS

DTIC
ELECTE
AUG 2 9 1985

G

DTIC FILE COPY

85 8 20 125

The views and conclusions contained in this document are those of the contractor and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States Government.

The ESD Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

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

FOR THE COMMANDER

Thomas J. Alpert, Major, USAF
Chief, ESD Lincoln Laboratory Project Office

Accession For

NTIS GRA&I

DTIC TAB

Unannounced

Justification

Distribution/

Availability Codes

Dist | Avail and/or Special

A

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

ADAPTIVE DETECTION IN NON-STATIONARY INTERFERENCE, PART I AND PART II

E.J. KELLY

Group 96

TECHNICAL REPORT 724

25 JUNE 1985

LEXINGTON   MASSACHUSETTS

ABSTRACT

In this report, which consists of two parts, the problem of radar target detection in a background of non-stationary external interference is considered. The object of the analysis is to treat this problem from the point of view of statistical decision theory, and to derive a signal processing algorithm which accepts the totality of inputs on which final decision is to be based, and performs both interference suppression and target detection. It is assumed that the radar is provided with multiple RF input channels and that target-free samples, from range gates other than the one in which a target is being sought, can be used for the estimation of the interference statistics.

In Part I a general formulation is given and a likelihood ratio detection rule is derived. The probabilities of detection and false alarm are evaluated exactly and the performance of the test is illustrated numerically. In Part II, a more specific interference model is introduced, which gives a more realistic representation of the situations likely to be encountered in practice. A decision rule is derived which approximates the exact likelihood ratio test for this case, and approximations for the detection and false alarm probabilities are found.

TABLE OF CONTENTS
PART I

PART II

LIST OF ILLUSTRATIONS

PART I

PART II

PART I

(1)   INTRODUCTION

This two-part study is devoted to certain aspects of the problem of radar target detection in a background of external interference which is non-stationary in character. In many discussions of this problem, interference rejection is treated as a distinct stage of the processing, target detection taking place as a subsequent operation. It is the object of the present analysis to deal with this problem in a unified way, leading to a single algorithm which accepts the totality of inputs on which final decision is to be based, and performs both suppression of interference and target detection.

The physical means with which a radar copes with interference, which usually take the form of multiple RF input channels and target-free data which can be used for the estimation of the interference statistics, are taken as givens. The target-free data is assumed to be provided by the samples from range gates other than the one in which a target is being sought. It is convenient to refer to the former as secondary data, while the samples from the range gate being processed for targets are called the primary data.

The analysis of Part I of this study is essentially a generalization of that of the well-known paper[1] of Reed, Mallett and Brennan (RMB). These authors discuss an adaptive

I-1

procedure for the detection of a signal of known form in the presence of noise (interference) which is assumed to be Gaussian, but whose covariance matrix is totally unknown. The possibility of signal presence is accepted for the primary data, while the secondary inputs are assumed to contain only noise, independent of and statistically identical to the noise components of the primary data. In the RMB procedure, the secondary inputs are used to form an estimate of the noise covariance, from which a weight vector is determined. This weight vector is then applied to the primary data in the form of a standard colored noise matched filter.

If the output of this filter were compared to a threshold, a complete detection procedure would be obtained. However, no predetermined threshold can be assigned to achieve a given PFA, since the detector is supposed to operate in an interference environment of unknown form and intensity. Instead, in the RMB paper an analysis is given of the signal to noise ratio (SNR) of this filter output, for given values of the secondary data. This SNR is a function of the secondary data and is therefore a random variable. The probability density function (PDF) of this SNR is deduced, which has the remarkable property of being independent of the actual noise covariance matrix; it is a function only of the dimensional parameters of the problem.

In the present study, the problem is reconsidered as an exercise in hypothesis testing, and the ad hoc RMB procedure is

replaced by a likelihood ratio test. No optimality properties are claimed for this test, involving as it does the maximization of two likelihood functions over a set of unknown parameters. The form of the test is, however, reasonable, and the RMB matched filter output appears as a portion of the likelihood ratio detection statistic. This test exhibits the desirable property that its PFA is independent of the covariance matrix (level and structure) of the actual noise encountered. This is a generalization of the familiar constant false alarm rate (CFAR) behavior of detectors using scalar input data, in which only the level of the noise is unknown. In addition, it is shown that the effect of signal presence depends only on the dimensional parameters of the problem and a parameter which is the same as the SNR of a conventional colored noise matched filter.

A brief outline of Part I of this study follows. The detection problem is formulated in a particular radar context in Section 2, where it is also pointed out that the actual mathematical analysis is considerably more general in nature. The likelihood ratio test is derived in Section 3, where it is compared to the RMB procedure mentioned above. In Section 4, the general form of the test is discussed, and its basic properties, in particular its CFAR property, are exhibited. The performance of the test is treated in Section 5, where the probabilities of detection and false alarm are obtained. This performance is

illustrated numerically in Section 6, which also contains a discussion of the results. Supporting material, including an alternate derivation of the RMB result, is provided in four appendices.

The analysis in Part I is self-contained, dealing with a single well-defined problem. A more specific (and more difficult) version of the detection problem is addressed in Part II, in which the results of Part I are freely used. The analytical methods used here lean heavily on the techniques of the RMB paper (which in turn is largely based on the analysis of Capon and Goodman[2]) with the difference that the matrix transformations required are carried out here directly on the variables of the problem, so that much less reliance is placed on known properties of the Wishart distribution.

(2)   FORMULATION OF ThE PROBLEM

The mathematical setting for the formulation of this detection problem will actually be quite general, but it is introduced here first in a relatively specific way, in order to lend concreteness by way of example. Suppose that the antenna system of a radar provides a number, say M, of RF signals. These may be the outputs of array elements, subarrays, beamformers or any mix of the above. The radar waveform is supposed to be a simple burst of identical pulses, say J in number, and target detection is to be based upon the returns from this burst. Further processing is possible, of course, and the 'detections' of the burst processor may be taken as inputs for subsequent binary integration over a string of bursts. In any case, the burst processor makes a decision, comparing some function of its input data (called the detection statistic) to a threshold, and the design of this processor is the same as if its decision were to be the final one.

In effect, the radar front end carries out amplification, filtering and reduction to base band, at which point the quadrature signals are subjected to pulse compression, the final stage of filtering. The order in which these things are carried out is immaterial to the present model, since it is not addressed to the problems of realization and channel matching, although these are of great importance in practice. The in-phase and

quadrature output pairs are next sampled to form range gate samples for each pulse, say, G range gates per pulse . This results in a total of MJG complex samples for the burst. Signal presence is sought in one range gate at a time, hence the primary data consists of the MJ samples from a single, unnamed range gate. These samples are arranged in a column vector, z, of dimension $N = MJ$. The secondary data consist of the outputs of K range gates, forming a subset of the G-1 remaining ones, and these are described by the set of vectors, $z(k)$, $(k = 1...K)$. The decision rule will be formulated in terms of the totality of input data, without the a priori assignment of different functions to the primary and secondary inputs.

The secondary data are assumed to be free of signal components, at least in the design of the algorithm, and any selection rules applied to make this assumption more plausible are ignored. The primary data may contain a signal vector, written in the form bs, where b is an unknown complex scalar amplitude, and s is a column vector of N components describing the signal which is sought. The modeled variation of signal amplitude and phase among the array inputs is included in s, as well as pulse to pulse variations, such as those relating to a particular target doppler velocity. The problem of unknown doppler, or other unknown signal parameters, is mentioned briefly below. It should be noted that the signal vector, s, can be

I-6

normalized in any convenient way, since an unknown amplitude factor is already included, and we retain the freedom to assign a norm to s at a later point, where it will be most advantageous to make a specific choice.

The total noise components of the data vectors, representing all sources of internal and external noise and interference, are modeled as zero-mean complex Gaussian random vectors. The noise component of the primary vector, z, is characterized by the unknown covariance matrix, M. Each of the z(k) is assumed to share this NxN covariance matrix, and the vectors z and the z(k) are all mutually independent. All Gaussian vectors are assumed to have the 'circular' property usually associated with I and Q pairs.

The key features of this model are the Gaussian assumption, the independence of the primary and secondary inputs, and the assumption that these share a common covariance matrix. The structure of the N-vectors, in particular the doubly indexed model used to describe multiple pulses and multiple array outputs, is not used in the following. One may equally well think of J as being to unity, in which case the analysis corresponds to a situation in which detection is based on the returns from a single pulse.

(3)   THE LIKELIHOOD RATIO TEST

Consider a single input vector from the secondary data set, say $z(k)$. If the covariance matrix of this vector is M:

$$M = E\{ z(k) z(k)^\dagger \},$$

then the N-dimensional Gaussian PDF of this complex random vector will be

$$f[z(k)] = \frac{1}{\pi^N \| M \|} e^{- z(k)^\dagger M^{-1} z(k)}.$$

In the notation used here the double bars signify the determinant of a matrix, and the superscript dagger symbolizes the conjugate transpose of a vector. Each of the secondary data vectors has this same PDF, and under the 'noise-alone' hypothesis, the primary vector does so as well, hence the joint PDF of all the input data is the product:

$$f_0[z, z(1), ..., z(K)] = f[z] \prod_{k=1}^{K} f[z(k)].$$

If v is any N-vector, we can write the following inner product in the form of a matrix trace (Tr):

$$v^\dagger M^{-1} v = Tr(M^{-1} V)$$

where V is the open product matrix

$$V = v v^\dagger$$

When this equivalence is applied to all the factors of the joint PDF, it will be seen that the latter may be written in the convenient form

$$f_0[z,z(1),...,z(K)] = \left\{ \frac{1}{\pi^N \|M\|} e^{-\operatorname{Tr}(M^{-1} T_0)} \right\}^{K+1}$$

where

$$T_0 = \frac{1}{K+1} \left(z\, z^\dagger + \sum_{k=1}^{K} z(k)\, z(k)^\dagger \right).$$

Under the 'signal-plus-noise' hypothesis, the $z(k)$ have the same PDF as before, and the PDF of the primary vector is obtained by replacing z by

$$z - E\{z\} = z - bs.$$

The resulting joint PDF of the inputs is then

$$f_1[z,z(1),...,z(K)] = \left\{ \frac{1}{\pi^N \|M\|} e^{-\operatorname{Tr}(M^{-1} T_1)} \right\}^{K+1}$$

where now

$$T_1 = \frac{1}{K+1} \left((z-bs)(z-bs)^\dagger + \sum_{k=1}^{K} z(k)\, z(k)^\dagger \right).$$

In the likelihood ratio testing procedure, the PDF of the inputs is maximized over all unknown parameters, separately for each of the two hypotheses. The ratio of these maxima is the detection statistic, and the hypothesis whose PDF is in the numerator is accepted as true if it exceeds some preassigned threshold. The maximizing parameter values are, by definition, the maximum likelihood (ML) estimators of these parameters, hence the maximized PDF's are obtained by replacing the unknown parameters by their ML estimators.

We begin with the noise-alone hypothesis, maximizing over the unknown covariance matrix, M. Of all positive definite M matrices, the one which maximizes the expression inside the curly brackets of this PDF is simply T_0. This is equivalent to the statement that the ML estimator of a covariance matrix is equal to the sample covariance matrix, which is well known[3]. When this estimator is substituted in the PDF, the trace which appears there becomes the trace of the NxN unit matrix, which is just N, and we find

$$\underset{M}{\text{Max}} \ f_0 = \left(\frac{1}{(e\pi)^N \| T_0 \|} \right)^{K+1} .$$

The same procedure, applied to the signal-plus-noise hypothesis yields the formula

$$\underset{M}{\text{Max}} \ f_1 = \left(\frac{1}{(e\pi)^N \| T_1 \|} \right)^{K+1} ,$$

and it remains to maximize this expression over the complex unknown signal amplitude, b. Since b appears only in this PDF, we can form a likelihood ratio, $L(b)$, at this point and subsequently maximize it over b. It is more convenient to work with the $(K+1)^{st}$ root of this ratio, and we put

$$L(b) = \{\ell(b)\}^{K+1} .$$

Obviously,

$$\ell(b) = \frac{\| T_0 \|}{\| T_1 \|} .$$

and the final likelihood ratio test takes the form

$$\underset{b}{\text{Max}} \; \ell(b) = \frac{\| T_0 \|}{\underset{b}{\text{Min}} \| T_1 \|} > \ell_0 \; .$$

The threshold parameter on the right will evidently be greater than unity, since the denominator on the left equals the numerator for the choice $b = 0$, and we are maximizing over b.

To proceed, we define the matrix

$$S = \sum_{k=1}^{K} z(k) \, z(k)^\dagger$$

which involves only the secondary data. This matrix is K times the sample covariance matrix of these data, and it satisfies the well-known Wishart distribution. The only property of this distribution that we need here is the fact that for $K > N$, a condition we now impose, the matrix S is non-singular with probability one. S is, of course, positive definite, and hence Hermitian. We use a lemma proved in Appendix B to evaluate the determinants of both sides of the equation

$$(K+1) \, T_0 = S + z \, z^\dagger$$

with the result

$$(K+1)^N \| T_0 \| = \| S \| \left(1 + z^\dagger S^{-1} z \right) .$$

Similarly, we have

$$(K+1)^N \| T_1 \| = \| S \| \left(1 + (z - bs)^\dagger S^{-1} (z - bs) \right) .$$

Now is a good time to minimize this quantity over b, and we do this by completing the square:

$$(z-bs)^\dagger S^{-1}(z-bs) = (z^\dagger S^{-1}z) + |b|^2(s^\dagger S^{-1}s) - 2\,\text{Re}\{b\,(z^\dagger S^{-1}s)\}$$

$$= (z^\dagger S^{-1}z) + (s^\dagger S^{-1}s)\left| b - \frac{(s^\dagger S^{-1}z)}{(s^\dagger S^{-1}s)} \right|^2 - \frac{|(s^\dagger S^{-1}z)|^2}{(s^\dagger S^{-1}s)}.$$

The minimum is clearly attained when the positive factor
containing b is made to vanish, and the resulting likelihood
ratio is given by

$$\ell = \underset{b}{\text{Max}}\ \ell(b) = \frac{1 + (z^\dagger S^{-1}z)}{1 + (z^\dagger S^{-1}z) - \frac{|(s^\dagger S^{-1}z)|^2}{(s^\dagger S^{-1}s)}}\ .$$

It is convenient to introduce the quantity η, defined by

$$\eta = \frac{|(s^\dagger S^{-1}z)|^2}{(s^\dagger S^{-1}s)\,[1 + (z^\dagger S^{-1}z)]}$$

so that

$$\ell = \frac{1}{1-\eta}\ .$$

Then the test

$$\ell > \ell_0$$

is equivalent to the test

$$\eta > \eta_0 = \frac{\ell_0 - 1}{\ell_0}\ .$$

We note that η_0 lies between the values zero and one.

If the target model is generalized, so that the signal
vector still contains one or more unknown parameters (such as
target doppler), the likelihood ratio obtained above must next be
maximized over these parameters. It is clear that this is

1-12

equivalent to maximizing η itself over the remaining target
parameters. This maximization generally cannot be carried out
explicitly, and the standard technique is to approximate it by
evaluating the test statistic, in this case η, for a discrete set
of target parameters, forming a 'filter bank', and declaring
target presence if any filter output exceeds the threshold. Our
purpose in discussing this here is only to show how our test can
be generalized in this straightforward way, but from now on we
ignore any additional target parameters, which is equivalent to
concentrating on the performance of a single member of the filter
bank.

For comparison with the RMB procedure, we introduce $\hat{M}$, the
ML estimator of the noise covariance, based on the secondary data
alone. We have already noted that this estimator is equal to

$$\hat{M} = \frac{1}{K} S .$$

The likelihood ratio test can then be written in the form

$$\frac{|(s^{\dagger}\hat{M}^{-1}z)|^2}{(s^{\dagger}\hat{M}^{-1}s)[1 + \frac{1}{K}(z^{\dagger}\hat{M}^{-1}z)]} > K\eta_0 .$$

We note that the secondary inputs enter this test only through
the sample covariance matrix, $\hat{M}$, and also that

$$(s^{\dagger}\hat{M}^{-1}z) = (\hat{w}^{\dagger}z)$$

where $\hat{w}$ is the RMB weight vector

$$\hat{w} = \hat{M}^{-1} s .$$

The RMB test itself is just

$$|(\hat{w} z)|^2 > \text{threshold},$$

which has the form of the colored noise matched filter test, with $\hat{M}$ replacing the usual known covariance matrix of the noise.

The presence of the signal-dependent factor in the denominator of the expression for η causes this detection statistic to be unchanged if the signal vector is altered by a scalar factor. Since the normalization of this vector has been left arbitrary, this invariance is highly desirable. In effect, this factor in the denominator is normalizing s for us, in terms of the estimated noise covariance. The entire detection statistic is also invariant to a common change of scale of all the input data vectors, a minimal CFAR requirement. Further properties of η will be developed in the following section.

In the limit of very large K, one expects the estimator, $\hat{M}$, to converge to the true covariance matrix, M, at least in probability. Moreover, it can be shown that the quantity

$$(z^\dagger M^{-1} z)$$

an inner product utilizing the actual covariance matrix instead of its estimator, obeys the chi-squared distribution, with 2N degrees of freedom, and hence this term, when divided by K, converges to zero in probability, as K grows without bound. In this sense the likelihood ratio test passes over into the

I-14

conventional colored noise matched filter test, as the number of sample vectors in the secondary data set becomes very large.

(4)   PROPERTIES OF THE LIKELIHOOD RATIO TEST

The likelihood ratio test will be discussed in terms of th random variable η, the decision statistic eventually obtained in the preceding section. The definition of η, as well as that of the matrix S on which it depends, are reproduced here for convenience:

$$\eta = \frac{|(s^\dagger S^{-1} z)|^2}{(s^\dagger S^{-1} s)[1 + (z^\dagger S^{-1} z)]}$$

$$S = \sum_{k=1}^{K} z(k)\, z(k)^\dagger .$$

The random variable η is, of course, a function of both the primary and secondary data, and as a preliminary to discussing its actual PDF, some useful properties are first derived. We begin with the noise-alone case, and assume that the actual noise covariance matrix is M.

The matrix M is positive definite, and hence a positive definite square root matrix can be defined. Since M can be diagonalized by a unitary transformation, it can be represented in the form

$$M = U \Lambda U^\dagger$$

where the columns of the unitary matrix, U, are eigenvectors of M, and Λ is diagonal. The diagonal elements of Λ, say $\lambda(n)$,

(n=1...N), are the real, positive eigenvalues of M. In case of degeneracy of an eigenvalue, the corresponding eigenvectors are assumed to have been orthogonalized. The square root may be defined by the representation

$$M^{\frac{1}{2}} = U \Lambda^{\frac{1}{2}} U^{\dagger}$$

where $\Lambda^{1/2}$ is diagonal, with diagonal elements $[\lambda(n)]^{1/2}$. $M^{-1/2}$ is similarly defined in terms of $\Lambda^{-1/2}$, and it is easily seen to be the inverse of $M^{1/2}$. Uniqueness of the square roots is not necessary for our purpose, only their existence and positive definite (hence also Hermitian) character.

Now consider the vector

$$\gamma \equiv M^{-\frac{1}{2}} z ,$$

and the similarly transformed secondaries

$$\gamma(k) \equiv M^{-\frac{1}{2}} z(k) .$$

The new vectors are zero-mean Gaussian variables, but with covariance matrix equal to I_N, the NxN identity matrix. This follows directly from the definitions:

$$E\{\gamma \gamma^{\dagger}\} = M^{-\frac{1}{2}} E\{z z^{\dagger}\} M^{-\frac{1}{2}} = M^{-\frac{1}{2}} M M^{-\frac{1}{2}} = I_N ,$$

with identical reasoning for the transformed secondaries. The linear transformation introduced here is, of course, a whitening transformation.

We note that the scalar, η, depends on the data and signal vectors only through inner products. By inverting the whitening transformation we may evaluate, for example, the product

I-17

$$(z^\dagger S^{-1} z) = (\gamma^\dagger M^{\frac{1}{2}} S^{-1} M^{\frac{1}{2}} \gamma) = \left(\gamma^\dagger (M^{-\frac{1}{2}} S M^{-\frac{1}{2}})^{-1} \gamma \right) .$$

We define the new matrix

$$\mathscr{S} \equiv M^{-\frac{1}{2}} S M^{-\frac{1}{2}}$$

and substitute for S, finding

$$\mathscr{S} = \sum_{k=1}^{K} M^{-\frac{1}{2}} z(k) z(k)^\dagger M^{-\frac{1}{2}} = \sum_{k=1}^{K} \gamma(k) \gamma(k)^\dagger .$$

Therefore, the new S-matrix is K times the sample covariance matrix of the whitened secondaries, and the random variable

$$\Sigma \equiv (z^\dagger S^{-1} z) = (\gamma^\dagger \mathscr{S}^{-1} \gamma)$$

is seen to be independent of M, being expressible as a function of K+1 independent Gaussian vectors, each of dimension N, and each sharing the covariance matrix, I_N. The PDF of Σ, like that of the RMB signal-to-noise ratio, is therefore a universal function of the dimensional parameters, N and K, alone.

The other inner products in the decision statistic are handled in an analogous manner; thus

$$(s^\dagger S^{-1} z) = (s^\dagger M^{-\frac{1}{2}} \mathscr{S}^{-1} \gamma) = (t^\dagger \mathscr{S}^{-1} \gamma)$$

where t stands for the whitened signal vector

$$t \equiv M^{-\frac{1}{2}} s .$$

At this point we make the deferred definition of signal normalization, by taking t to be a unit vector:

$$(t^\dagger t) = (s^\dagger M^{-1} s) \equiv 1 .$$

This choice gives specific meaning to the signal amplitude

parameter, b, whose square is now a proper signal-to-noise ratio, for which we introduce the symbol a:

$$a \equiv |b|^2 = \left(E\{z\}^\dagger M^{-1} E\{z\} \right).$$

When the obvious substitutions are made in the final inner product, we obtain

$$\eta = \frac{|(\ell^\dagger \mathcal{S}^{-1} \gamma)|^2}{(\ell^\dagger \mathcal{S}^{-1} \ell) \, [1 + (\gamma^\dagger \mathcal{S}^{-1} \gamma)]}.$$

The dependence on M is now confined to t, and it will be shown below that even this dependence on the true covariance matrix is illusory. When a signal is present, z is replaced by

$$z = bs + n$$

where n has all the properties attributed to z in the noise-alone case. In this situation, the whitened data vector is

$$\gamma = M^{-\frac{1}{2}} z = b\ell + \nu$$

where

$$\nu \equiv M^{-\frac{1}{2}} n,$$

which is statistically identical to the whitened data vector in the noise-alone case. We have therefore found that when signal is present, the PDF of n depends on M only through b and t, and the dependence on the unit vector, t, is again only apparent, as we now proceed to show.

Suppose the whitening transformation is followed by a unitary one, in which the whitened vectors are expressed as the

products of a unitary matrix and a new set of random vectors. These new random vectors are statistically indistinguishable from their predecessors, and it would only be confusing to introduce a new notation for them. Tracing this transformation through the inner products, we find that only the normalized signal vector is changed: t is replaced by

$$t_1 \equiv U_1 t$$

where U_1 is the unitary matrix characterizing this last transformation. Any unit vector in the complex N-space can be realized, as t_1, by such a transformation. In particular, we can cause t_1 to be a 'coordinate vector', for which a single element is unity, the remaining (N-1) elements vanishing. It is for this reason that the PDF of η depends on M only through the meaning of the signal amplitude parameter, b. In fact, this PDF can depend only on b, N and K, and hence the false alarm probability of the likelihood ratio detector, namely

$$PFA = Prob\{\eta > \eta_0\},$$

is independent of M, and this is the generalized CFAR property claimed in section 1.

(5)   PROBABILITY DISTRIBUTION OF THE TEST STATISTIC

We now take advantage of our freedom to make a unitary transformation, and choose for t_1 a vector whose first element is unity, all others being zero. This can be accomplished by choosing for U_1 a matrix whose first row is the conjugate transpose of t, and whose other rows are the conjugates of unit vectors orthogonal to t. Understanding that this choice has been made, we drop the subscript on t, so that n is still given by the formula of the preceding section.

This form for t makes it expedient to decompose all vectors into two components, an A-component consisting of the first element only, and a B-component consisting of the rest of the vector. Thus we write

$$\gamma = \begin{bmatrix} \gamma_A \\ \gamma_B \end{bmatrix}$$

where the A-component is a scalar and the B-component is an (N-1)-vector. In this notation, the signal vector is just

$$t = \begin{bmatrix} 1 \\ 0 \end{bmatrix} ,$$

the zero being (N-1) dimensional. Matrices are decomposed in analogous fashion, and we write

$$\mathscr{S} = \begin{bmatrix} \mathscr{S}_{AA} & \mathscr{S}_{AB} \\ \mathscr{S}_{BA} & \mathscr{S}_{BB} \end{bmatrix} .$$

Note that the AA-element is a scalar, the BA-element is an (N-1)

dimensional column vector, and so on. We also give a name to the inverse of this matrix, decomposing it as well:

$$\mathcal{S}^{-1} \equiv \mathcal{P} = \begin{bmatrix} \mathcal{P}_{AA} & \mathcal{P}_{AB} \\ \mathcal{P}_{BA} & \mathcal{P}_{BB} \end{bmatrix} .$$

With this notation we have, simply,

$$(\iota^{\dagger} \mathcal{S}^{-1} \iota) = (\iota^{\dagger} \mathcal{P} \iota) = \mathcal{P}_{AA} ,$$

while

$$(\iota^{\dagger} \mathcal{S}^{-1} \gamma) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \mathcal{P}_{AA} & \mathcal{P}_{AB} \\ \mathcal{P}_{BA} & \mathcal{P}_{BB} \end{bmatrix} \begin{bmatrix} \gamma_A \\ \gamma_B \end{bmatrix}$$

$$= \mathcal{P}_{AA} \gamma_A + \mathcal{P}_{AB} \gamma_B .$$

It is important to keep in mind that we now have a four-fold decomposition of the total input data set into primary and secondary vectors, each of which is divided into A- and B-components.

According to the Frobenius relations for partitioned matrices,

$$\mathcal{P}_{AA} = (\mathcal{S}_{AA} - \mathcal{S}_{AB} \mathcal{S}_{BB}^{-1} \mathcal{S}_{BA})^{-1} ,$$

which is a scalar, and also

$$\mathcal{P}_{BA} = -\mathcal{S}_{BB}^{-1} \mathcal{S}_{BA} \mathcal{P}_{AA} .$$

Since the NxN sample covariance matrix and its inverse are Hermitian, we obtain

$$\mathcal{P}_{AB} = \mathcal{P}_{BA}^\dagger = -\mathcal{P}_{AA}\mathcal{Y}_{AB}\mathcal{Y}_{BB}^{-1}$$

and therefore

$$(t^\dagger \mathcal{Y}^{-1}\gamma) = \mathcal{P}_{AA}(\gamma_A - \mathcal{Y}_{AB}\mathcal{Y}_{BB}^{-1}\gamma_B) \ .$$

The final inner product is expanded as follows:

$$(\gamma^\dagger \mathcal{Y}^{-1}\gamma) = \mathcal{P}_{AA}|\gamma_A|^2 + 2\,\mathrm{Re}\{\gamma_A^*\mathcal{P}_{AB}\gamma_B\} + (\gamma_B^\dagger \mathcal{P}_{BB}\gamma_B) \ ,$$

where we have applied the identity

$$(u^\dagger v) = (v^\dagger u)^*$$

to the (N-1) dimensional inner product

$$\gamma_B^\dagger \mathcal{P}_{BA} \ .$$

Next, we complete the square in this last expression, writing

$$(\gamma^\dagger \mathcal{Y}^{-1}\gamma) = \mathcal{P}_{AA}|\gamma_A + \mathcal{P}_{AA}^{-1}\mathcal{P}_{AB}\gamma_B|^2$$

$$+ \gamma_B^\dagger(\mathcal{P}_{BB} - \mathcal{P}_{AA}^{-1}\mathcal{P}_{BA}\mathcal{P}_{AB})\gamma_B \ ,$$

and by using a Frobenius relation in reverse we see that

$$\mathcal{P}_{BB} - \mathcal{P}_{AA}^{-1}\mathcal{P}_{BA}\mathcal{P}_{AB} = \mathcal{Y}_{BB}^{-1} \ .$$

Finally, combining these results, we find

$$(\gamma^\dagger \mathcal{Y}^{-1}\gamma) = \mathcal{P}_{AA}|\gamma_A - \mathcal{Y}_{AB}\mathcal{Y}_{BB}^{-1}\gamma_B|^2 + (\gamma_B^\dagger \mathcal{Y}_{BB}^{-1}\gamma_B) \ .$$

When these evaluations are substituted into our expression for η, the result can be expressed in the apparently simple form:

$$\eta = \frac{X}{1 + X + \Sigma_B} \ .$$

We have introduced here the notation

$$\Sigma_B = (\gamma_B^\dagger \mathcal{Y}_{BB}^{-1}\gamma_B) \ ,$$

which will be retained, and the temporary notation

$$X = \mathscr{P}_{AA} |\gamma_A - \mathscr{S}_{AB} \mathscr{S}_{BB}^{-1} \gamma_B|^2 \ .$$

Note that Σ_B is just like the quantity Σ defined earlier, except that the dimensionality of the vectors involved is now N-1. The last form of the decision test, namely

$$\eta > \eta_0 \ .$$

is evidently equivalent to

$$X > \frac{\eta_0}{1 - \eta_0} (1 + \Sigma_B) \ .$$

We shall leave the test in this form for a time, while we examine the statistical properties of the quantities which enter into it.

A previous evaluation for the leading factor in X can be used to obtain the following form:

$$X = \frac{|\gamma_A - \mathscr{S}_{AB} \mathscr{S}_{BB}^{-1} \gamma_B|^2}{\mathscr{S}_{AA} - \mathscr{S}_{AB} \mathscr{S}_{BB}^{-1} \mathscr{S}_{BA}} \ .$$

We make use of the definitions to express the denominator as a sum:

$$\mathscr{S}_{AA} - \mathscr{S}_{AB} \mathscr{S}_{BB}^{-1} \mathscr{S}_{BA} = \sum_{k=1}^{K} \left(\gamma_A(k) - \mathscr{S}_{AB} \mathscr{S}_{BB}^{-1} \gamma_B(k) \right) \gamma_A(k)^* \ .$$

This is the same as the sum of squares

$$\sum_{k=1}^{K} |\gamma_A(k) - \mathscr{S}_{AB} \mathscr{S}_{BB}^{-1} \gamma_B(k)|^2 \ ,$$

because the terms supplied to complete this square add up to zero:

$$\sum_{k=1}^{K} \left(\gamma_A(k) - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\gamma_B(k) \right) \left(\mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\gamma_B(k) \right)^*$$

$$= \sum_{k=1}^{K} \gamma_A(k)\gamma_B(k)^{\dagger}\mathcal{S}_{BB}^{-1}\mathcal{S}_{BA} - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1} \sum_{k=1}^{K} \gamma_B(k)\gamma_B(k)^{\dagger}\mathcal{S}_{BB}^{-1}\mathcal{S}_{BA}$$

$$= \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathcal{S}_{BA} - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\mathcal{S}_{BB}\mathcal{S}_{BB}^{-1}\mathcal{S}_{BA} = 0 \ .$$

The evaluation of the sums here follows from the definitions of the partitioned matrix elements. We introduce the notation

$$y(k) \equiv \gamma_A(k) - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\gamma_B(k)$$

for the terms of the sum, and the analogous notation

$$y \equiv \gamma_A - \mathcal{S}_{AB}\mathcal{S}_{BB}^{-1}\gamma_B$$

for the quantity appearing in the numerator of X, so that the likelihood ratio test can be written in the more explicit form

$$X = \frac{|y|^2}{\sum_{k=1}^{K} |y(k)|^2} > \frac{\eta_0}{1 - \eta_0} \ (1 + \Sigma_{\bullet}) \ .$$

We proceed by fixing the B-vectors temporarily, and consider the probability densities of all quantities entering into the decision statistic to be conditioned on these values. The conditional probabilities of detection and false alarm will be evaluated first, and the condition will then be removed by taking expectation values over the joint PDF of the B-vectors. With the B-vectors fixed, only the K+1 scalar A-components are random, and we show now, under this condition, that y and the

$y(k)$ are Gaussian variables, that y is uncorrelated with the $y(k)$, and that the latter have a covariance matrix with simple properties.

Using the definitions of the y's and of the AB matrix element which enters there, we can express these quantities in the form

$$y = \gamma_A - \sum_{k=1}^{K} \gamma_A(k) \gamma_B(k)^\dagger \mathscr{S}_{BB}^{-1} \gamma_B$$

and

$$y(k) = \gamma_A(k) - \sum_{l=1}^{K} \gamma_A(l) \gamma_B(l)^\dagger \mathscr{S}_{BB}^{-1} \gamma_B(k) \ .$$

This represents the y's as linear combinations of the A-components, and hence proves their conditional Gaussian character. Moreover, the $y(k)$ have zero mean in all cases, while the conditional mean, written E_B, of y in the general case is

$$E_B y = E \gamma_A = b \ ,$$

as a result of our choice of signal vector.

The linear dependence of the y's on the A-components is best expressed in terms of the quantities

$$q(k) \equiv \gamma_B(k)^\dagger \mathscr{S}_{BB}^{-1} \gamma_B$$

and

$$Q(l,k) \equiv \gamma_B(l)^\dagger \mathscr{S}_{BB}^{-1} \gamma_B(k) \ ,$$

which are constants under the conditioning. Obviously,

$$y = \gamma_A - \sum_{k=1}^{K} \gamma_A(k) \, q(k)$$

1-26

and

$$y(k) = \gamma_A(k) - \sum_{k=1}^{K} \gamma_A(l)\, Q(l,k) \ ,$$

The q(k) may be considered as the components of a K-vector, q, and the Q(i,k) as the elements of a KxK matrix, Q. The desired properties of the y's flow from the following facts about this new vector and matrix:

$$Q q = q$$

and

$$Q^2 = Q \ .$$

To prove the first of these, we write it out in component form:

$$\sum_{l=1}^{K} Q(k,l)q(l) = \sum_{l=1}^{K} \gamma_B(k)^\dagger \mathscr{S}_{BB}^{-1} \gamma_B(l) \gamma_B(l)^\dagger \mathscr{S}_{BB}^{-1} \gamma_B \ .$$

The sum over i regenerates the BB matrix element:

$$\sum_{l=1}^{K} \gamma_B(l)\gamma_B(l)^\dagger = \mathscr{S}_{BB} \ .$$

(as happened when the denominator of X was expresed as a sum of squares), and the result follows immediately. The idempotent character of Q is proved in the same way. We also note that Q is Hermitian, and that its trace is N-1:

$$\begin{aligned}
\mathrm{Tr}(Q) &= \sum_{k=1}^{K} \gamma_B(k)^\dagger \mathscr{S}_{BB}^{-1} \gamma_B(k) \\
&= \mathrm{Tr}\!\left(\mathscr{S}_{BB}^{-1} \sum_{k=1}^{K} \gamma_B(k)\gamma_B(k)^\dagger \right) \\
&= \mathrm{Tr}(I_{N-1}) = N - 1 \ .
\end{aligned}$$

Note that we are dealing with the trace of a K$\times$K matrix on the left side here, and of (N-1)x(N-1) matrices on the right. All of these results will be required in the following.

The fact that y and the y(k) are conditionally uncorrelated now follows easily from the independence of the A-components themselves:

$$E_{\mathtt{a}}y\,y(k)^* = -E_{\mathtt{a}}\left(\sum_{i=1}^{K} q(i)\gamma_A(i)\gamma_A(k)^*\right)$$

$$+ E_{\mathtt{a}}\left(\sum_{i=1}^{K} q(i)\gamma_A(i) \sum_{n=1}^{K} \gamma_A(n)^* Q(n,k)^*\right)$$

$$= - q(k) + \sum_{i=1}^{K} q(i)Q(k,i) = 0 .$$

Next, consider the conditional variance of y:

$$E_{\mathtt{a}}|y-b|^2 = 1 + \sum_{k=1}^{K} |q(k)|^2 .$$

Substituting for the q(k), we have

$$\sum_{k=1}^{K} |q(k)|^2 = \sum_{k=1}^{K} \gamma_B{}^\dagger \mathcal{Y}_{BB}{}^{-1}\gamma_B(k)\gamma_B(k)^\dagger \mathcal{Y}_{BB}{}^{-1}\gamma_B$$

$$= \gamma_B{}^\dagger \mathcal{Y}_{BB}{}^{-1}\gamma_B = \Sigma_{\mathtt{a}} ,$$

and hence

$$E_{\mathtt{a}}|y-b|^2 = 1 + \Sigma_{\mathtt{a}} .$$

This last result is responsible for a significant simplification of the statistics of the likelihood ratio test.

Finally, we compute the conditional covariance of the $y(k)$. We use the notation $\delta(i,k)$ for the elements of the unit matrix, so that $y(k)$ can be written

$$y(k) = \sum_{l=1}^{K} \hat{r}_A(l)[\delta(l,k) - Q(l,k)] \ .$$

Using the independence of the A-variables again, we obtain

$$E\,y(k)\,y(n)^* = \sum_{l=1}^{K} [\delta(l,k) - Q(l,k)][\delta(l,n) - Q(l,n)]^*$$

$$= \delta(k,n) - Q(n,k) - Q(k,n)^* + \sum_{l=1}^{K} Q(l,k)Q(l,n)^* \ .$$

Since Q is Hermitian and idempotent, we find the simple result

$$E\,y(k)\,y(n)^* = \delta(n,k) - Q(n,k) \ .$$

The likelihood ratio test is now rearranged slightly to read

$$\frac{|y|^2}{1 + \Sigma_a} > \frac{\eta_0}{1 - \eta_0} \sum_{k=1}^{K} |y(k)|^2 \ .$$

In view of fact that the conditional variance of y equals the denominator on the left, it makes sense to define a normalized variable

$$w \equiv \frac{y}{(1 + \Sigma_a)^{1/2}} \ .$$

Conditioned on the B-vectors, w is Gaussian and independent of the $y(k)$. It has a conditional variance of unity, and a conditional mean value:

$$E_a w = \frac{b}{(1 + \Sigma_a)^{1/2}} .$$

In the noise-alone case, the conditioning has no effect on the PDF of w. The sum over k is also given a name:

$$T \equiv \sum_{k=1}^{K} |y(k)|^2 .$$

and the test is now written

$$|w|^2 > (\ell_0 - 1) T .$$

where the original threshold constant has been reintroduced. In fact, it is easily verified that our original likelihood ratio is given by

$$\ell = 1 + \frac{|w|^2}{T} .$$

We now turn to the properties of T. Given the B-vectors, the joint PDF of the y(k) is zero-mean Gaussian with covariance matrix J:

$$J(l,k) \equiv \delta(l,k) - Q(k,l) .$$

The conditional characteristic function of T is therefore

$$\Phi_a(\lambda) \equiv E_a\{e^{i\lambda T}\} = \|I - i\lambda J\|^{-1} .$$

Since Q is idempotent, its eigenvalues are either zero or one, and from the value of its trace we see that Q must have exactly N-1 unit eigenvectors. It follows that J has K+1-N unit eigenvectors, the others being zero, and thus

$$\Phi_a(\lambda) = (1 - i\lambda)^{-(K+1-N)}$$

This is the characteristic function of a chi-squared random

I-30

variable, and the PDF of T is simply

$$f_s(T) = \frac{T^{K-N}}{(K-N)!} e^{-T}$$

It is remarkable that the statistical properties of T are independent of the actual values of the conditioning B-vectors, and we can consequently drop the subscript on its PDF. Moreover, T is statistically equivalent to the sum of the squares of K+1-N independent, complex Gaussian variables, each of which has zero mean and unit variance. If we let w(k), (k=1...K+1-N), be such a set, then T is statistically indistinguishable from the sum

$$\sum_{k=1}^{K+1-N} |w(k)|^2 .$$

The properties of the likelihood ratio test are therefore identical to the properties of the simple test

$$|w|^2 > (\ell_0 - 1) \sum_{k=1}^{K+1-N} |w(k)|^2 .$$

where the w(k) are now also taken to be independent of w. The probability of the truth of this inequality is still conditioned on the B-vectors, but this conditioning appears only through the quantity Σ_B, which is contained in the conditional mean of w.

This equivalent decision rule represents the behavior of a simple scalar CFAR test, in which the power in one complex sample (a single radar hit), being tested for signal presence, is compared to a threshold proportional to the sum of the powers of

1-31

K+1-N samples of noise. This problem is quite familiar, and the test just described is also a likelihood ratio test in the corresponding situation. The performance of the scalar CFAR detector is very simple, and in particular, its PFA is just

$$\left(\frac{1}{\ell_0}\right)^{K+1-N} .$$

In this case, when the signal amplitude is zero, the conditioning B-vectors do not appear at all, and hence this simple formula gives the PFA for our original likelihood ratio test.

The probability of detection (PD) of the scalar CFAR test is also well known, and in our case it depends on the conditional SNR, which is the squared magnitude of the conditional mean of w. In terms of the colored noise matched filter SNR, a, defined earlier, and the quantity

$$r = \frac{1}{1 + \Sigma_\bullet} ,$$

this conditional SNR is just ra. The factor r represents a loss factor, applied to the SNR, and caused by the necessity of estimating the noise covariance matrix. The PD of the CFAR detector can be expressed in a particularly convenient way as a finite sum[4]:

$$P_D = 1 - \frac{1}{\ell_0{}^L} \sum_{k=1}^{L} \binom{L}{k} (\ell_0 - 1)^k G_k\left(\frac{ra}{\ell_0}\right) ,$$

where L = K+1-N. The function G which enters here is itself a finite sum:

$$G_k(y) = e^{-y} \sum_{n=0}^{k-1} \frac{y^n}{n!}$$

In order to complete our computation of the PD of the likelihood ratio test, we must take the expectation value of this conditional PD over the joint PDF of the B-vectors. These, however, enter the final result only through the loss factor, r, which acts as a fluctuation model for the signal. Unlike more familiar fluctuation models, this one is characterized by a factor lying in the range zero to one. The present situation is similar to that discussed in the RMB paper, except that besides a SNR loss, our test will suffer a CFAR loss as well, when compared to a colored noise matched filter test in which everything is known concerning the noise or interference.

Although the loss factor found here depends on the primary data, through its B-component, while the RMB loss factor is a function only of the secondary data, it turns out that the two factors have exactly the same PDF. The proof of this interesting result is deferred to Appendix A, in which the RMB loss factor is also expressed in our notation, and the evaluation of the PDF's of both these quantities is carried out in parallel.

The PDF shared by these loss functions is the beta distribution:

$$f(r) = \frac{(N+L-1)!}{L!(N-2)!} (1-r)^{N-2} r^L \, .$$

I-33

and the final expression for the PD of our test can be written

$$P_D = 1 - \frac{1}{\ell_0^L} \sum_{k=1}^{L} \binom{L}{k} (\ell_0 - 1)^k H_k\left(\frac{a}{\ell_0}\right).$$

In this formula, the H-functions are the expected values of the G's:

$$H_k(y) = \int_0^1 G_k(ry) f(r) dr.$$

These integrals are elementary, although not simple, and their detailed evaluation is presented in Appendix C, where the result is also recast in a form more suitable for computation.

(6)   NUMERICAL RESULTS AND DISCUSSION

The performance of the likelihood ratio test depends only on the dimensional integers, N and K, and the SNR parameter, a. The latter is a function of the true signal strength and the intensity and character of the actual noise and interference. Our analysis deals with a very general problem, and nothing can be said about the anticipated values of a. The ability of a system to function effectively in interference depends principally on the arrangements which have been made in its design to achieve a good colored noise matched filter SNR in its intended environment. These arrangements will usually take the form of diversity of RF inputs in one form or another. An additional requirement is the need to have inputs available from which the actual noise characteristics can be estimated, and this is the aspect of the problem which has been addressed here. In particular, for given values of PD and PFA, we can determine what SNR is actually required to achieve those values using the likelihood ratio detector, and compare that number to the SNR which would be adequate to achieve identical performance if the noise covariance matrix were known in advance. The difference is the penalty for having to estimate the noise covariance, and we expect that penalty to vary sharply with the number, K, of available secondary input vectors.

This penalty has two components: one due to the CFAR character of the decision rule and another due to the effective

SNR loss factor. The latter is expected to behave much as the results of the RMB analysis would predict, based on the statistical properties of the loss factor alone. The CFAR loss will decrease as the value of K increases, and it may be expected to depend largely on that parameter, while the SNR loss effect depends roughly on the ratio of K to N.

These expectations are borne out by the numerical consequences of our analysis, as shown in the accompanying figures. In Figs. 1 through 4, probability of detection is shown as a function of a, the SNR, for three detectors (PFA is fixed at 10^{-6} for these curves). The detector performing best is a matched filter with known noise covariance, and the worst is the likelihood ratio detector which, of course, is estimating the noise covariance. The middle curve (dashed) in these plots shows the performance of a simple, scalar CFAR detector using L=K+1-N noise samples, and it differs from the behavior of the likelihood ratio detector only in that the SNR loss factor has been ignored. This detector is included in the comparison in order to show how much of the degradation imposed by noise estimation is due to each of the two contributing effects.

We note that doubling both K and N has little effect on the portion of the degradation due to SNR loss, while the CFAR part is reduced, simply because K is being increased. The curves also show the significant improvement which results from

PROBABILITY of DETECTION vs SNR
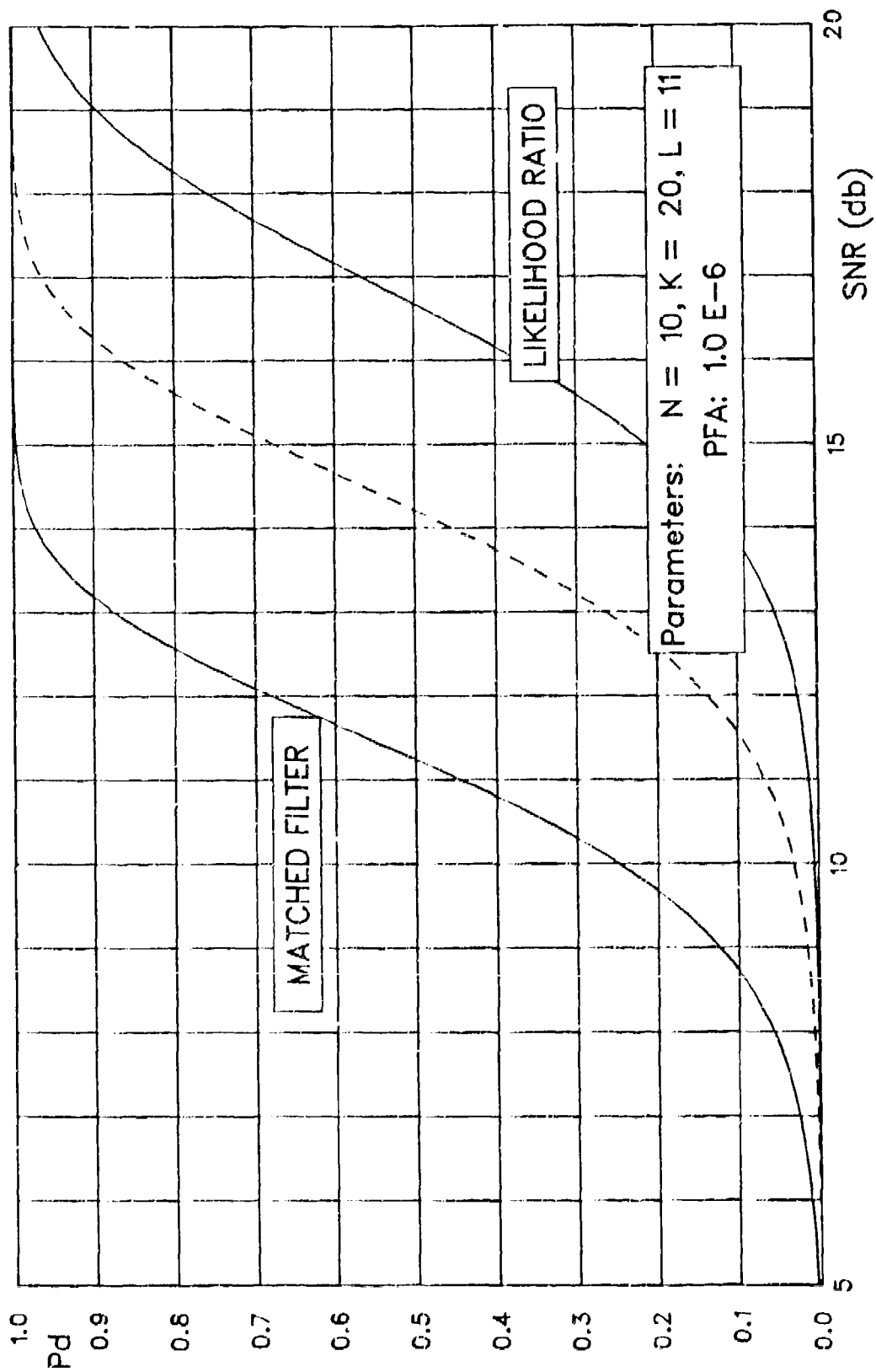
Fig. I-1 Probability of Detection vs SNR
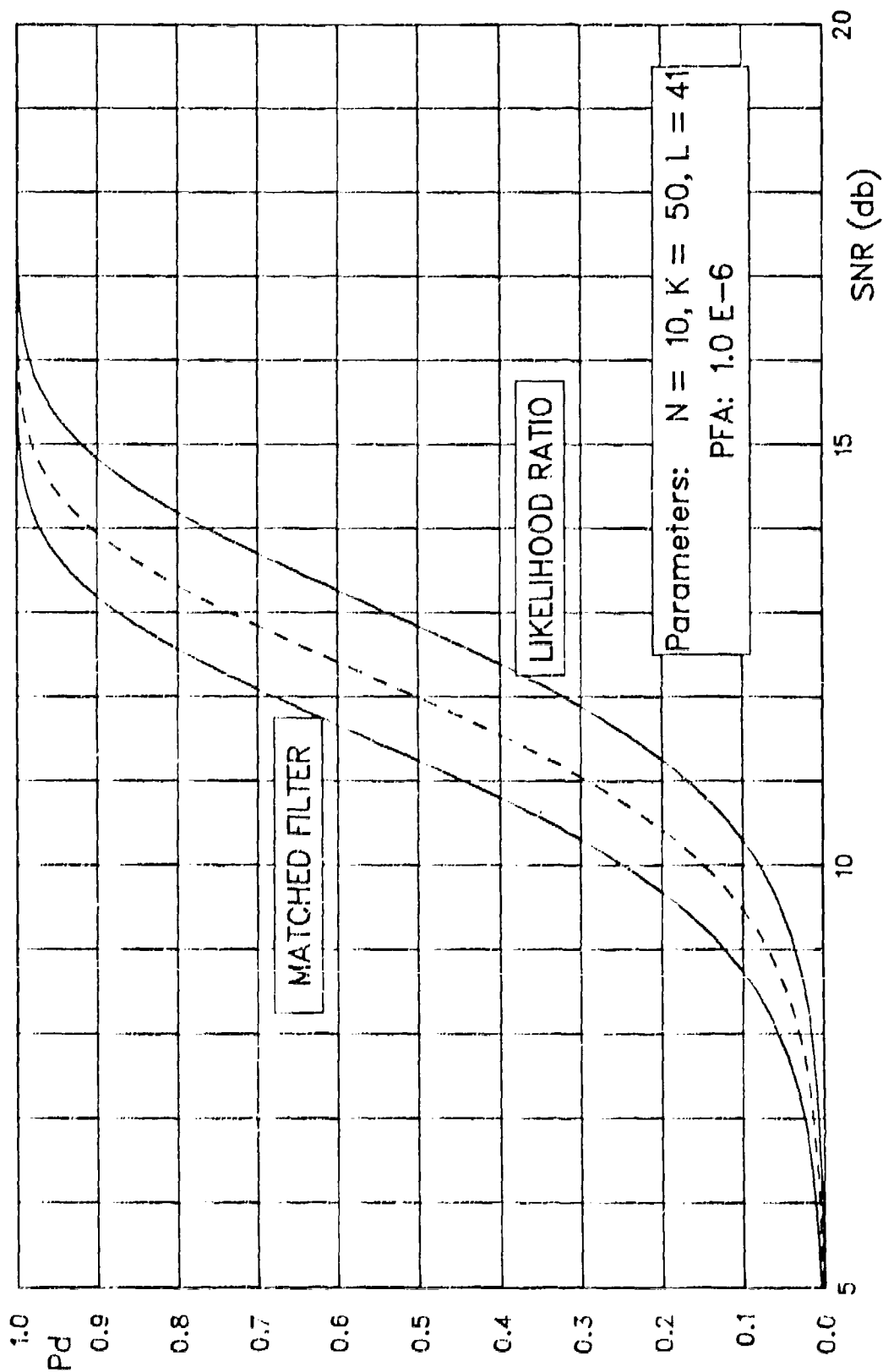
PROBABILITY of DETECTION vs SNR

MATCHED FILTER

LIKELIHOOD RATIO

Parameters: N = 10, K = 50, L = 41
PFA: 1.0 E−6

Fig. I−2 Probability of Detection vs SNR

1-38

PROBABILITY of DETECTION vs SNR

LIKELIHOOD RATIO

MATCHED FILTER

Parameters: N = 20, K = 40, L = 21
PFA: 1.0 E-6

SNR (db)

Pd

Fig. 1-3 Probability of Detection vs SNR

PROBABILITY of DETECTION vs SNR

MATCHED FILTER

LIKELIHOOD RATIO

Parameters: N = 20, K = 100, L = 81
PFA: 1.0 E−6

SNR (db)

Pd
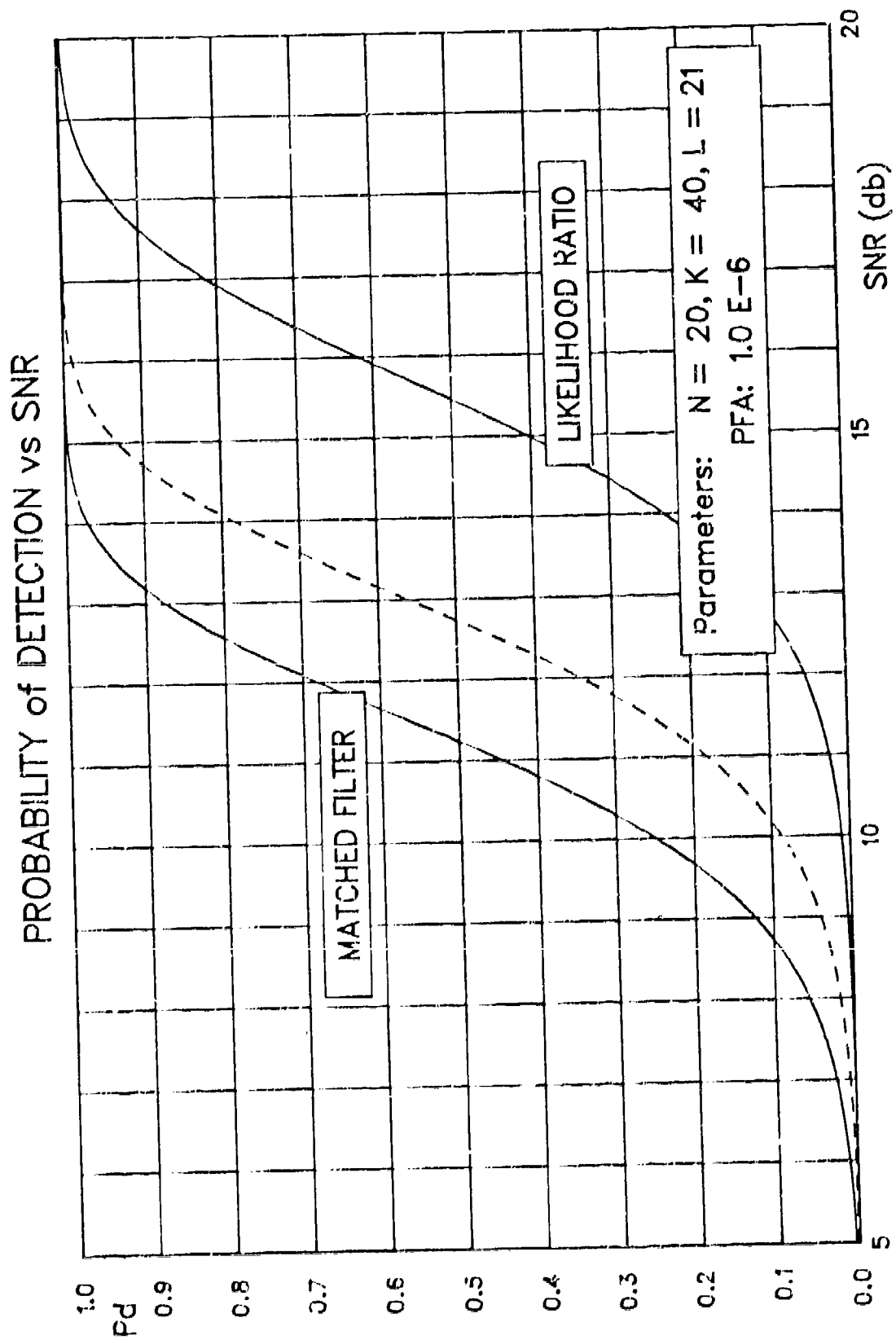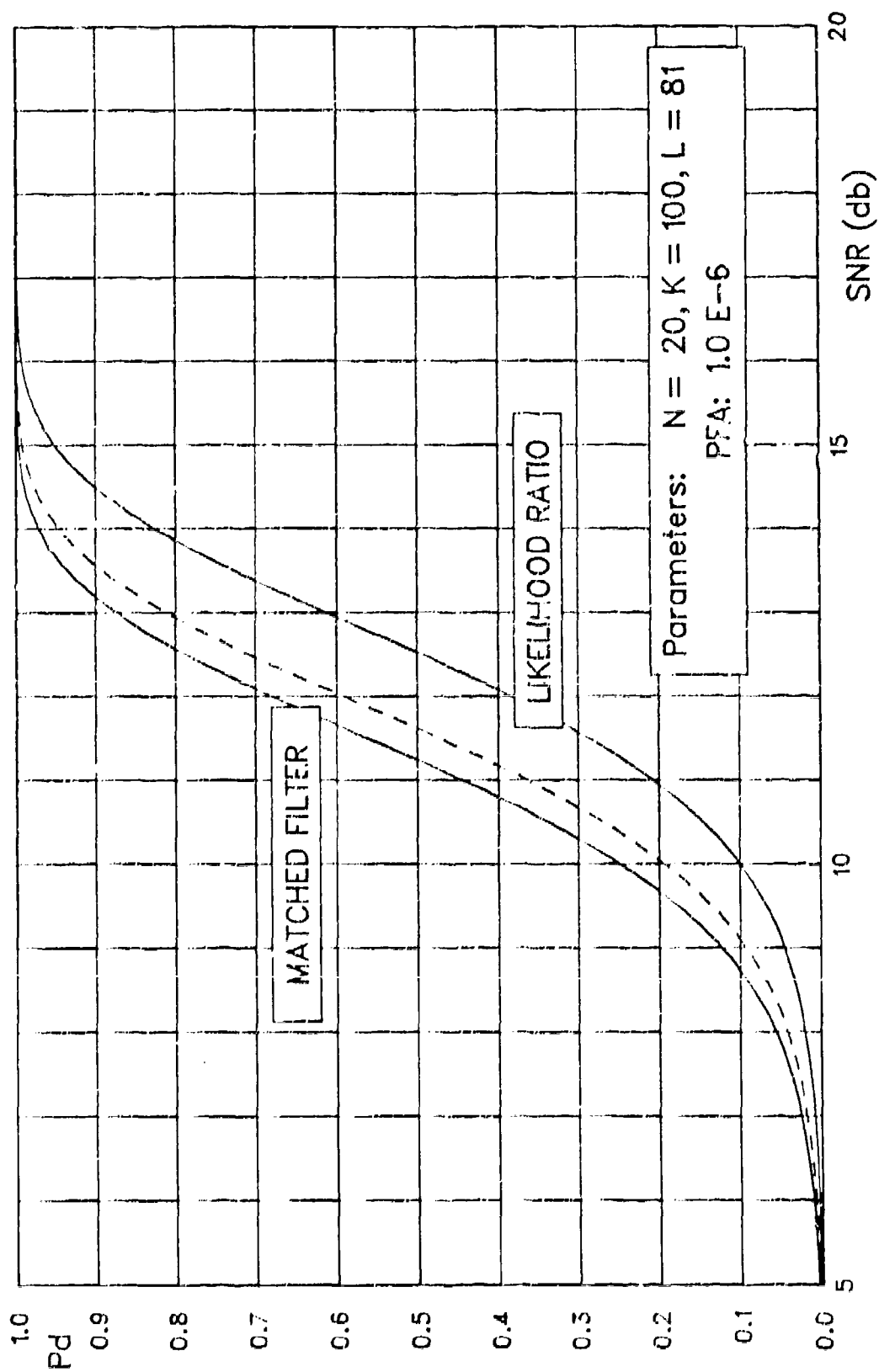
Fig. 1-4 Probability of Detection vs SNR

increasing the ratio of K to N. When this ratio is equal to five, the SNR loss contribution to the performance degradation is about 0.9 dB, in agreement with the mean value of the SNR loss as obtained from the beta distribution. Likewise, the CFAR contributions are directly comparable with the ordinary CFAR loss for a detector of nonfluctuating targets with no noncoherent integration (i.e. a single radar hit).

The detector performance is characterized in a different way in Figs. 5 through 8, which show the additional SNR required, when estimating the noise covariance, to achieve the same PD and PFA as a matched filter for known noise. In all these figures, the PD is specified at 0.9, but the results will not depend strongly on the chosen PD level, since the curves of PD vs SNR are nearly parallel for the two detectors. Three PFA values are represented on each plot. The independent variable for these curves is the number of secondary vectors, and this variable always covers the range 2N through 5N, for four different N values. It can be seen that SNR loss is not strictly a function of the ratio K/N, but generally decreases with increasing N, with this ratio held constant. The loss shown is the total loss, due to the CFAR effect and the SNR loss factor itself.

It was noted earlier that K, the number of secondary vectors, must exceed N, the dimension of each of the data vectors, in order to have a non-singular sample covariance

LOSS IN SNR vs NUMBER OF SECONDARY VECTORS

Vector dimension: 5
Detection Probability: 0.9
PFA = 10^{-C}

C values:

Fig. I-5 Loss in SNR vs Number of Secondary Vectors

Fig. 1-6 Loss in SNR vs Number of Secondary Vectors
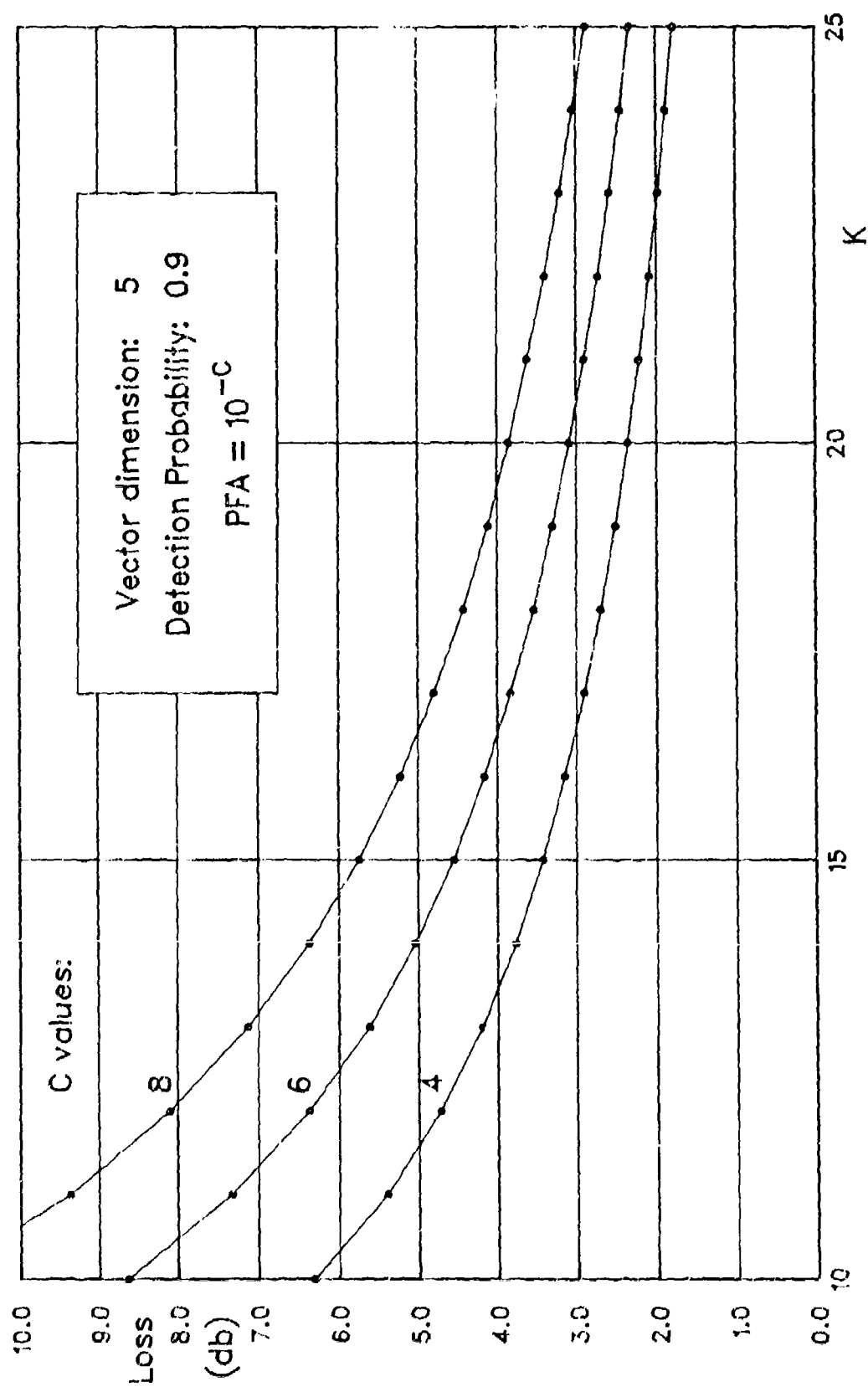
LOSS IN SNR vs NUMBER OF SECONDARY VECTORS

Vector dimension: 20
Detection Probability: 0.9
PFA = 10^{-C}

C values:

Fig. I-7, Loss in SNR vs Number of Secondary Vectors
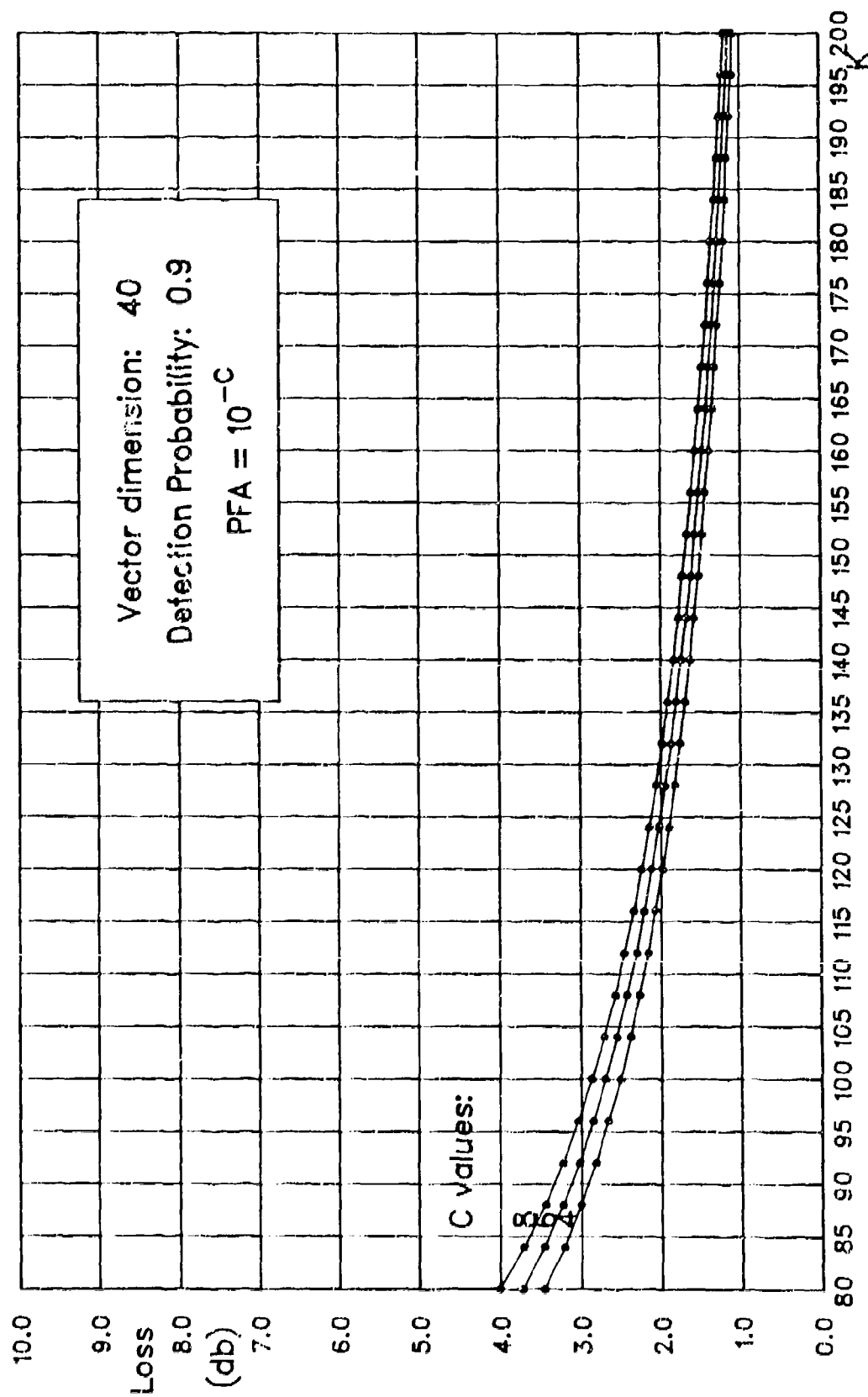
Fig. I-8 Loss in SNR vs Number of Secondary Vectors

matrix. It is clear from the results just discussed that K must exceed N by a significant factor if noise estimation is not to cause a serious loss in performance. Since N is the dimension of the total vector of data used for detection, the requirements on the number of secondaries can become very large. In the radar example mentioned in Section 2, N was the product of the number of RF channels (M) and the number of pulses (J) in a coherent processing interval, usually a large number. On the other hand, our results are equally valid for the case J=1, which represents a situation in which detection is based on the inputs from a single pulse. In the latter case, the RF inputs could be the elements of an adaptive array, and N might then be a much smaller quantity.

In the original application, the reason that so many secondaries are required is the generality of our formulation, in which any interference covariance matrix is allowed. In the radar example, this includes the possibility of arbitrary correlation between interference inputs from pulse returns widely spaced in time, although it is more realistic to assume independence (but not statistical identity) of the interference inputs accompanying distinct pulse returns. The fact that we have allowed correlation between separate pulse returns, but still assumed that the secondary data is independent of the primaries is somewhat inconsistent, since the secondaries are taken from adjacent range

gate samples. Instead, the present analysis should be viewed as a formulation of the problem which does not preclude pulse to pulse independence, but does not incorporate it as a feature of the model.

If the problem is reformulated with pulse to pulse independence as a specific assumption, there will be only JM^2 real unknown parameters for the noise matrix instead of $(MJ)^2$ parameters, which could be an enormous difference. One expects the possibility of improved performance when fewer unknown, or nuisance, parameters are being estimated, and one would also expect that fewer secondaries would be required. This case is the subject of Part II of this study, where it is shown that these expectations are indeed borne out.

Returning to the general problem addressed in the present analysis, it should be mentioned that the likelihood ratio decision statistic can be reformulated in a way which significantly reduces the number of matrix inversions (or Cholesky factorizations) required. Suppose that a set of K+1 data vectors is specified, and one of these is singled out as a primary vector for the likelihood ratio test. Since targets can usually appear in any range gate output, one would next return the selected vector to the data pool and choose another as primary, and so on. Each stage requires the inversion of a different NxN matrix, or K+1 inversions to test for signals in

all the vectors of the original set. It is shown in Appendix D that one can form a single sample covariance matrix, using all K+1 vectors, then invert it (or factor it), and use this matrix in all the K+1 tests for target presence in the individual range gates.

Although the reformulated test is precisely equivalent t) the original, the question naturally arises as to the effect of signal presence in more than one of the range gate outputs. This is a problem in any CFAR, and it is usually minimized by the application of some screening procedure to keep signal-bearing vectors out of the set used for noise estimation. Such a procedure would limit the applicability of the reformulated test, except among the vectors which passed the screening test themselves. Without screening, the presence of unwanted signals in the secondaries will degrade performance. This degradaton could be evaluated, but the PDF of the likelihood ratio decision statistic would become much more complicated than it is without unwanted signals, and this topic will not be discussed further here.

APPENDIX A   THE PROBABILITY DENSITY FUNCTION OF THE LOSS FACTOR

The SNR loss factor derived in the text was expressed in the form

$$r = \frac{1}{1 + \Sigma_B}$$

where

$$\Sigma_B = \gamma_B^{\dagger} \mathcal{Y}_{BB}^{-1} \gamma_B .$$

Before discussing the PDF of Σ_B, from which the PDF of r follows easily, we express the RMB loss factor, ρ, in our notation. From the RMB paper,

$$\rho = \frac{|(\widehat{w}^{\dagger} s)|^2}{(s^{\dagger} M^{-1} s)(\widehat{w}^{\dagger} M^{-1} \widehat{w})}$$

where M is the actual noise covariance matrix, s is the signal vector, and $\widehat{w}$ is a weight vector:

$$\widehat{w} = k \widehat{M}^{-1} s .$$

In this last formula, $\widehat{M}$ is the sample covariance matrix of the secondary data and k is an arbitrary constant. The loss factor itself is the ratio of the conditional SNR of the output of a filter which uses $\widehat{w}$ as a weight, relative to the SNR of the colored noise matched filter for known M. The conditioning in this case corresponds to given values of the secondary data.

Choosing k=1/K, we obtain

$$\widehat{w} = S^{-1} s ,$$

in terms of the S matrix used in the text, and then

$$\rho = \frac{(s^\dagger S^{-1}s)^2}{(s^\dagger M^{-1}s)(s^\dagger S^{-1}MS^{-1}s)} \; .$$

Note that ρ is unaffected if s is changed by a constant factor.

We now carry out the whitening transformation, as in the text, and normalize the signal vector as before. The result is

$$\rho = \frac{(\ell^\dagger \mathscr{Y}^{-1}\ell)^2}{(\ell^\dagger \mathscr{Y}^{-2}\ell)}$$

$$= \frac{(\ell^\dagger \mathscr{P}\ell)^2}{(\ell^\dagger \mathscr{P}^2\ell)} \; .$$

A unitary transformation is now applied to convert the signal vector to the final one used in the text, and the matrices are decomposed in the same way. This gives the simple expression

$$\rho = \frac{(\mathscr{P}_{AA})^2}{(\mathscr{P}^2)_{AA}} \; .$$

It is clear at this point that the PDF of ρ will be independent of the actual covariance matrix M.

Using the Frobenius relations, we obtain

$$(\mathscr{P}^2)_{AA} = (\mathscr{P}_{AA})^2 + \mathscr{P}_{AB}\mathscr{P}_{BA}$$

$$= (\mathscr{P}_{AA})^2 \left(1 + \mathscr{Y}_{AB}\mathscr{Y}_{BB}^{-2}\mathscr{Y}_{BA}\right) \; ,$$

and therefore

$$\rho = \frac{1}{1 + \Sigma_\rho}$$

where

$$\Sigma_\rho = \mathscr{Y}_{AB}\mathscr{Y}_{BB}^{-2}\mathscr{Y}_{BA} \; ,$$

I-50

Note that the RMB loss factor depends on the secondary data only, both A- and B-components, while r depends on the B-components of both primary and secondary data.

We proceed to analyze the two loss factors together, and begin by conditioning on the B-components of the secondary data vectors, on which both loss factors depend. Then

$$\mathcal{Y}_{BB} = \sum_{k=1}^{K} \gamma_B(k)\gamma_B(k)^{\dagger}$$

is a constant matrix, positive definite and non-singular for all sets of conditioning vectors (except for a set of probability zero). We can therefore introduce the square root of this matrix and define the vectors

$$\xi_B \equiv \mathcal{Y}_{BB}^{-1/2}\gamma_B$$

and

$$\xi_\rho \equiv \mathcal{Y}_{BB}^{-1}\mathcal{Y}_{BA} .$$

With the conditioning, these quantities are zero-mean Gaussian vectors; the former is a linear function of the elements of the B-component of the whitened primary vector, and the latter is expressible in terms of the secondary A-components:

$$\xi_\rho = \mathcal{Y}_{BB}^{-1}\sum_{k=1}^{K} \gamma_B(k)\gamma_A(k)^{\dagger} .$$

We use the subscript C to denote the present conditioning, and compute the conditional covariance matrices of these vectors:

$$E_c \xi_B \xi_B^\dagger = \mathscr{S}_{BB}^{-1/2} E_c \gamma_B \gamma_B^\dagger \mathscr{S}_{BB}^{-1/2} = \mathscr{S}_{BB}^{-1}$$

and

$$E_c \xi_\rho \xi_\rho^\dagger = \mathscr{S}_{BB}^{-1} W \mathscr{S}_{BB}^{-1} \; .$$

where

$$W = E_c \sum_{k=1}^{K} \gamma_B(k)\gamma_A(k)^\dagger \sum_{l=1}^{K} \gamma_A(l)\gamma_B(l)^\dagger$$

$$= \sum_{k=1}^{K} \gamma_B(k)\gamma_B(k)^\dagger = \mathscr{S}_{BB} \; .$$

Therefore

$$E_c \xi_\rho \xi_\rho^\dagger = \mathscr{S}_{BB}^{-1} \; .$$

and the two ζ-vectors are statistically equivalent under the conditioning. They therefore share the same final PDF when the conditioning is removed by averaging over the secondary B-vectors. Since

$$\Sigma_B = (\xi_B^\dagger \xi_B)$$

and

$$\Sigma_\rho = (\xi_\rho^\dagger \xi_\rho) \; .$$

this proves that the loss factors themselves are statistically identical, and we continue with the loss factor, r.

Since ξ_B is a Gaussian $(N-1)$-vector, the conditional joint PDF of its components is

$$f_c(\xi_B) = \frac{\|\mathscr{S}_{BB}\|}{\pi^{N-1}} e^{-(\xi_B^\dagger \mathscr{S}_{BB} \xi_B)} \; .$$

The S matrix which enters here is itself subject to the Wishart

I-52

PDF, which in the present case (in which the sample vectors are of dimension N-1 and the covariance matrix is equal to the identity) takes the form

$$f_W(A) = \frac{\|A\|^{K+1-N}}{C(N-1,K)} e^{-Tr(A)} \quad .$$

In this formula

$$C(N,K) = \pi^{\frac{N(N-1)}{2}} \prod_{n=1}^{N} (K-n)!$$

is the Wishart normalization factor. The volume element for this PDF will be written $d(A)$. It is $(N-1)^2$ dimensional, ranging over the diagonal elements and the real and imaginary parts of the upper off-diagonals, of all positive definite matrices, A. For our purpose, only the normalization integral of the Wishart PDF is required, hence we need not dwell on the detailed properties of this fascinating distribution.

Since ξ_B depends on the conditioning data only through the S matrix, its unconditioned PDF can be written

$$f(\xi_B) = \frac{1}{\pi^{N-1}} \int \cdots \int \|A\| e^{-Tr(A\mathcal{B})} f_W(A) d(A)$$

As in the text, we have replaced the exponential part of the Gaussian PDF by a trace, this time involving the open product matrix

$$\mathcal{B} = \xi_B \xi_B^\dagger \quad .$$

When we substitute for the Wishart PDF in the expresson above, we encounter the integral

I-53

$$\int \cdots \int \|A\|^{K+2-N} e^{-Tr[A(I_{N-1} + \mathscr{B})]} \, d(A) \, .$$

This is the same as the normalization integral for another
Wishart PDF, of dimensions N-1 and K+1, and for which the
underlying sample vectors share the covariance matrix

$$\mathscr{M} \equiv (I_{N-1} + \mathscr{B})^{-1} \, .$$

The normalization factor for this slightly more general Wishart
PDF is just

$$C(N-1,K+1) \, \|\mathscr{M}\|^{K+1}$$

$$= \frac{C(N-1,K+1)}{\|I_{N-1} + \mathscr{B}\|^{K+1}} = \frac{C(N-1,K+1)}{[1 + (\xi_{\mathbf{B}}^{\dagger} \xi_{\mathbf{B}})]^{K+1}}$$

(the evaluation of the determinant uses the same lemma utilized in
Section 3). Combining these facts, we obtain the simple result

$$f(\xi_{\mathbf{B}}) = \frac{1}{\pi^{N-1}} \frac{C(N-1,K+1)}{C(N-1,K)} [1 + (\xi_{\mathbf{B}}^{\dagger} \xi_{\mathbf{B}})]^{-(K+1)}$$

$$= \frac{K!}{\pi^{N-1}(K+1-N)!} (1 + \Sigma_{\mathbf{B}})^{-(K+1)} \, .$$

The remainder of the derivation is identical to the final
few steps given in the Appendix of the RMB paper. The norm of the
vector ξ_{B} is interpreted as the square of the radial coordinate
in a (2N-2)-dimensional Cartesian space, a change to polar
coordinates is made, and the angular coordinates integrated out.
This process yields the PDF of Σ_{B}, and then a simple change of
variable provides the desired PDF of r:

$$f(r) = \frac{K!}{(K+1-N)!(N-2)!} (1 - r)^{N-2} r^{K+1-N} \, .$$

APPENDIX B. EVALUATION OF A DETERMINANT

In the main derivation in the text, and again in Appendix A, a lemma was used which may be stated as follows:

$$\|A + ab^\dagger\| = \|A\|(1 + b^\dagger A^{-1} a)$$

where A is assumed to be nonsingular. Because of this assumption, we can write

$$a = A c$$

and factor out the matrix A, and hence its determinant. It remains to be shown that

$$\|I_N + cb^\dagger\| = 1 + b^\dagger c$$

where I_N is the NxN identity matrix. The desired evaluation then follows by elimination of c.

The above result is proved by induction, and it is obvious for N=2. In general, the matrix

$$M \equiv I_N + cb^\dagger$$

is decomposed as follows

$$M = \begin{bmatrix} M_{11} & v^\dagger \\ w & L \end{bmatrix}$$

where the first row and column have been singled out. In terms of the (N-1)-vectors

$$\bar{c} \equiv \begin{bmatrix} c_2 \\ \vdots \\ c_N \end{bmatrix}$$

and

$$\bar{b} \equiv \begin{bmatrix} b_R \\ \vdots \\ b_N \end{bmatrix}$$

we can write

$$w = b_1^* \bar{c}$$

$$v = c_1^* \bar{b}$$

and

$$L = I_{N-1} + \bar{c}\,\bar{b}^\dagger$$

Using the Frobenius relations again, we have

$$\|M\| = M_{11} \, \|L - w M_{11}^{-1} v^\dagger\|$$

Of course,

$$M_{11} = 1 + c_1 b_1^*$$

and we also have

$$w v^\dagger = c_1 b_1^* \bar{c}\,\bar{b}^\dagger$$

Therefore

$$L - w M_{11}^{-1} v^\dagger = I_{N-1} + \left(1 - \frac{c_1 b_1^*}{1 + c_1 b_1^*} \right) \bar{c}\,\bar{b}^\dagger$$

$$= I_{N-1} + \frac{\bar{c}\,\bar{b}^\dagger}{1 + c_1 b_1^*}$$

Assuming the validity of the evaluation for the N-1 dimensional case, we get

$$\|M\| = M_{11} \left(1 + \frac{\bar{b}^\dagger \bar{c}}{1 + c_1 b_1^*} \right)$$

$$= 1 + b_1^* c_1 + \bar{b}^\dagger \bar{c} = 1 + b^\dagger c$$

and the lemma is proved.

APPENDIX C. COMPUTATION OF THE PROBABILITY OF DETECTION

The probability of detection obtained in the text has the
form

$$P_D = 1 - \frac{1}{\ell_0^L} \sum_{k=1}^{L} \binom{L}{k} (\ell_0 - 1)^k H_k\left(\frac{a}{\ell_0}\right)$$

where L=K+1-N, a is the colored noise matched filter SNR and ℓ_0
is the likelihood ratio threshold parameter. The functions H_k
are the expected values of the G_k, averaged over the PDF of r,
the likelihood ratio SNR loss factor:

$$H_k(y) = E\, G_k(ry) = \int_0^1 G_k(ry) f(r)\, dr \ .$$

The G_k functions are the sums

$$G_k(y) = e^{-y} \sum_{n=0}^{k-1} \frac{y^n}{n!}$$

(which are directly related to the incomplete gamma function),
and the PDF of the loss factor is the beta distribution

$$f(r) = F(N,L)(1-r)^{N-2} r^L \ .$$

The normalizing factor here is

$$F(N,L) = \frac{(N+L-1)!}{L!(N-2)!} \ .$$

The false alarm probability is given by the very simple
formula

$$PFA = \left(\frac{1}{\ell_0}\right)^L .$$

and it can be seen that PD reduces to PFA when a vanishes, since $G_k(0)=1$. This implies that $H_k(0)=1$, and the result then follows from the fact that PD becomes an incomplete binomial sum.

To obtain an explicit solution for PD, we substitute for the $G_k(y)$ and write

$$H_k(y) = \sum_{n=0}^{k-1} I_n(y)$$

where

$$I_n(y) \equiv \frac{y^n}{n!} F(N,L) \int_0^1 e^{-ry}(1-r)^{N-2} r^{L+n} dr .$$

Next, a binomial expansion provides the formula

$$I_n(y) = \frac{y^n}{n!} F(N,L) \sum_{p=0}^{N-2} \binom{N-2}{p}(-1)^p \int_0^1 e^{-ry} r^{L+n+p} dr .$$

The integrals which enter here are elementary, and it proves useful to define the functions

$$T_m(y) \equiv e^y \int_0^1 e^{-ry} r^m dr .$$

In terms of the T functions, we have

$$I_n(y) = e^{-y} \frac{y^n}{n!} F(N,L) \sum_{p=0}^{N-2} \binom{N-2}{p}(-1)^p T_{L+n+p}(y) ,$$

and it is easily verfied that

I-58

$$T_m(y) = \frac{m!}{y^{m+1}}\left(e^y - \sum_{s=0}^{m} \frac{y^s}{s!}\right)$$

$$= \frac{m!}{y^{m+1}} e^y\left(1 - G_{m+1}(y)\right) .$$

The combination of these steps provides a formal solution to our problem, but the result obtained is not useful as a basis for numerical evaluation. Especially when y is small, the direct evaluation of the T functions for large and growing m-values involves the products of factors, one of which is increasingly large while the other is inceasingly small. The alternating sum into which the T evaluations must be substitued further stresses the numerical precision of the computation.

To avoid the alternating sum and its attendant numerical problems, a different approach has been taken to the evaluation of the I functions. In the integral which appears in the definition of these functions, the variable of integration is changed from r to (1-r), and we write

$$I_n(y) = \frac{y^n}{n!} e^{-y} J_n(y)$$

where

$$J_n(y) \equiv F(N,L) \int_0^1 e^{ry}(1-r)^{L+n} r^{N-2} dr .$$

The exponential is now expanded in a power series and the integral evaluated term by term. Each of the integrals

encountered is the normalizing integral of another beta
distribution, and is easily evaluated from the definition. The J
functions then appear as infinite series, which can be
expressed in the form

$$J_n(y) = D_n \sum_{k=0}^{\infty} T_n(y,k)$$

where

$$D_n \equiv \frac{K!(L+n)!}{(K+n)!L!}$$

and

$$T_n(y,k) \equiv \frac{(N-2+k)!(K+n)!}{(N-2)!(K+n+k)!} \frac{y^k}{k!} .$$

These quantities are easily generated recursively, and the
series is terminated after a total of, say, k terms. The
truncation error is then

$$\varepsilon_n(y,k) = D_n \sum_{s=k}^{\infty} T_n(y,s)$$

$$= D_n \sum_{s=0}^{\infty} T_n(y,k+s) .$$

It is not difficult to show that

$$\varepsilon_n(y,k) = D_n T_n(y,k) \sum_{s=0}^{\infty} Q_n(k,s) \frac{y^s}{s!}$$

where

$$Q_n(k,s) \equiv \frac{(N-2+k+s)!}{(N-2+k)!} \frac{(K+n+k)!}{(K+n+k+s)!} \frac{k!s!}{(k+r)!} .$$

In our application K exceeds N, and this ensures that the Q sequence decreases with increasing values of 1. Since $Q_n(k,0)=1$, we see that

$$\varepsilon_n(y,k) < D_n \, T_n(y,k) e^y \ .$$

If a truncation error bound, ε, is prescribed, then the series for $J_n(y)$ can be terminated when

$$D_n \, T_n(y,k) < \varepsilon e^{-y} \ .$$

and this provides a simple algorithm for the computation of the I functions. The sum sequence of the I's is the sequence of H functions, required in the final finite summation for P_D. The other coefficients needed in this sum (the first formula of this Appendix) are easily generated by recursion. This procedure has been used to obtain the numerical results illustrated in the text.

APPENDIX D. ALTERNATIVE FORMULATION OF THE DECISION RULE

At first it seems unlikely that a test will perform as well if the primary vector, thought to contain a signal component, is included with the secondaries in the estimation of the noise covariance matrix. A simple CFAR example will, however, show that this can be an entirely reasonable procedure. Suppose that K+1 complex samples are available, denoted by z(k), k=0,1,..K. A CFAR test for signal presence in z(0), using the other samples for the estimation of noise level, would have the form

$$|z(0)|^2 > \alpha\left(|z(1)|^2 + \cdots + |z(K)|^2\right)$$

where the constant α is to be determined by the assigned value of PFA. However, by simply adding

$$\alpha|z(0)|^2$$

to both sides of this inequality, and then dividing through by $1+\alpha$, we find the equivalent form

$$|z(0)|^2 > \frac{\alpha}{1+\alpha}\left(|z(0)|^2 + |z(1)|^2 + \cdots + |z(K)|^2\right)$$

All the samples now enter into the noise level estimate, which can be used unchanged for tests of signal presence in each of the other samples in turn. The performance attained by this detector, and in particular the losses caused by the presence of unwanted targets in the samples used for noise estimation, is

identical to that of the original form of the test. Formulas for the performance of such a CFAR detector, when unwanted targets of various kinds are present in the noise estimate, may be found in Ref. 4.

We expect that a similar reformulation is possible in the present problem, and begin by returning to the analysis of Section 3. The likelihood ratio test was expressed there in the form

$$\frac{\| T_0 \|}{\underset{b}{Min} \| T_1 \|} > \ell_0$$

where

$$(K+1) T_0 = S + z z^\dagger$$

and

$$(K+1) T_1 = S + (z-bs)(z-bs)^\dagger .$$

We now define

$$\tilde{S} \equiv S + z z^\dagger$$

and note that this matrix is K+1 times the sample covariance matrix formed from all the input data vectors.

Clearly, we can express T_1 in the form

$$(K+1) T_1 = \tilde{S} + (z-bs)(z-bs)^\dagger - z z^\dagger ,$$

and, of course,

$$(K+1) T_0 = \tilde{S} .$$

Next, we factor out the new S-matrix, and write

$$(K+1)\, T_1 = \widetilde{S}\left(I_N + a\, b^\dagger + c\, d^\dagger\right),$$

where

$$a \equiv \widetilde{S}^{-1}(z - bs)$$

$$b \equiv z - bs$$

$$c \equiv \widetilde{S}^{-1}z$$

and

$$d \equiv -z \ .$$

The matrix ratio is now simply

$$\frac{\|T_1\|}{\|T_0\|} = \|I_N + a\, b^\dagger + c\, d^\dagger\| \ ,$$

and we can again use the lemma of Appendix B in the following way. Let P be the matrix

$$P \equiv I_N + a\, b^\dagger \ .$$

then

$$\|I_N + a\, b^\dagger + c\, d^\dagger\| = \|P + c\, d^\dagger\|$$

$$= \|P\|\left(1 + \left(d^\dagger P^{-1} c\right)\right) \ .$$

By the same lemma,

$$\|P\| = 1 + \left(b^\dagger a\right),$$

and it is easily verified that

$$P^{-1} = I_N - \frac{ab^\dagger}{1 + (b^\dagger a)} \ .$$

When these evaluations are used, we obtain

$$\frac{\|T_1\|}{\|T_0\|} = \left(1+(b^\dagger a)\right)\left(1+(d^\dagger c)\right) - (d^\dagger a)(b^\dagger c)$$

$$= \left(1 + [(z-bs)^\dagger \widetilde{S}^{-1}(z-bs)]\right)\left(1 - (z^\dagger \widetilde{S}^{-1} z)\right)$$

$$+ [z^\dagger \widetilde{S}^{-1}(z-bs)][(z-bs)^\dagger \widetilde{S}^{-1} z]$$

After this expression is developed, it is a simple matter to
complete the square, much as was done before, with the result:

$$\frac{\|T_1\|}{\|T_0\|} = 1 + Q \left| b - \frac{(s^\dagger \widetilde{S}^{-1} z)}{Q} \right|^2 - \frac{|(s^\dagger \widetilde{S}^{-1} z)|^2}{Q} \ ,$$

where

$$Q \equiv (s^\dagger \widetilde{S}^{-1} s)\left(1 - (z^\dagger \widetilde{S}^{-1} z)\right) + |(s^\dagger \widetilde{S}^{-1} z)|^2 \ .$$

The minimization over b is now trivial, and we see that the test
assumes the form

$$\frac{Q}{Q - |(s^\dagger \widetilde{S}^{-1} z)|^2} > \ell_0 \ ,$$

or

$$\frac{|(s^\dagger \widetilde{S}^{-1} z)|^2}{(s^\dagger \widetilde{S}^{-1} s)[1 - (z^\dagger \widetilde{S}^{-1} z)]} > \ell_0 - 1 \ .$$

This is the desired expression for the decision rule,
which is very much like its predecessor, but now involves the

sample covariance formed from all the data vectors. By using the evaluation

$$\tilde{S}^{-1} = [S + z\,z^\dagger]^{-1}$$

$$= \left(S^{\frac{1}{2}} (I_N + S^{-\frac{1}{2}} z\,z^\dagger S^{-\frac{1}{2}}) S^{\frac{1}{2}} \right)^{-1}$$

$$= S^{-\frac{1}{2}} \left(I_N - \frac{S^{-\frac{1}{2}} z\,z^\dagger S^{-\frac{1}{2}}}{1 + (z^\dagger S^{-1} z)} \right) S^{-\frac{1}{2}} \ ,$$

it is not difficult to recover the likelihood ratio test in its original form.

REFERENCES

1. I.S. Reed, J.D. Mallett and L.E. Brennan, 'Rapid Convergence Rate in Adaptive Arrays,' IEEE Trans Aerospace Electron. Systems, AES-10, 853, (1974).

2. J. Capon and N.R. Goodman, 'Probability Distributions for Estimators of the Frequency-Wavenumber Spectrum,' Proc. IEEE, Vol 58, pp. 1785-1786, (Oct. 1970).

3. N.R. Goodman, 'Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution,' Ann. Math. Stat., Vol. 34, pp. 152-177, (Mar. 1963).

4. E.J. Kelly, "Finite-Sum Expressions for Signal Detection Probabilities," Technical Report 566, Lincoln Laboratory, M.I.T. (20 May 1981) DTIC-AD-A102143

(1) INTRODUCTION

In Part I of this study, a general problem of radar detection was discussed which was characterized by the presence of unknown non-stationary interference. The radar is assumed to have a number of RF channels, and target detection is based on their outputs for a train of pulses which form a coherent processing interval. The methods of likelihood ratio decision theory were applied to the derivation of a detection algorithm for this problem. Arbitrary correlation was permitted for the interference between different pulses, as well as between different RF channels. A decision rule was derived in closed form and exact expressions for the system performance were obtained.

In Part II, the problem is reformulated with the single additional assumption that the interference is statistically independent from pulse to pulse. The correlation properties of the interference are still unknown, and are allowed to vary arbitrarily from pulse to pulse. The present discussion is largely based on the methods and results of Part I, but the more specific problem is more difficult to solve, and we have not obtained a closed form expression for the exact likelihood ratio decision rule.

An approximation to the likelihood ratio test statistic has been obtained however, and this will be derived and discussed below. The approximation appears to be a reasonable one, and the resulting test reduces to the exact decision rule for a single pulse, and also to the likelihood ratio test for the analogous multiple-pulse problem in which the noise covariance matrices are presumed known. The form of this test itself provides considerable insight into the detection problem.

The exact probability density function (PDF) of the approximate test statistic appears to be very difficult to obtain. An approximation is developed, however, which contains several of the features of the exact PDF, and expressions for the probability of false alarm (PFA) and probability of detection (PD) are derived for this approximation in the present report.

An essential feature of the analyses of both Part I and the present Part II lies in the assumed existence of so-called secondary inputs which can be used to estimate the covariance matrices of the interference. The samples from range gates adjacent to the one being examined for targets are used for this purpose, and it is worth specifying in detail the assumptions made in our model regarding them. It is assumed that all external sources of interference are sufficiently wide-band in character so that, like internal noise, samples separated in time by the duration of a (compressed) radar pulse are independent. The key

assumption lies in the degree of non-stationarity which is allcwed.

In the present model, we assume that the covariance matrices of the total interference can be arbitrarily different for samples separated in time by an interval as large as the pulse repetition interval (PRI) of the radar. On the other hand, the correlation properties between the RF channels are not supposed to change so rapidly that those of successive samples (separated by one pulse length) are widely different. The non-stationarity of the interference is actually assumed to be slow compared to the pulse length, but can be fast relative to the PRI. Samples from range gates adjacent to the one being tested for target presence can then be assumed to be statistically identical to the latter, as long as they are not too large in number, while for successive pulses, the covariance matrices shared by this set of samples can be entirely different.

This model is admittedly somewhat contrived, but if the interference is allowed to change arbitrarily from sample to sample, then noise covariance estimation is not possible, and adaptive detection, in the sense studied here, will not be feasible. It should be noted that the interference provided by a group of sources with time-varying output levels does not result in a completely arbitrary time variation of the covariance

matrices, and represents an intermediate case not addressed here.

The model used in Part I assumed statistically identical secondaries, but allowed correlation from pulse to pulse in both primary and secondary inputs. This is not a good model of an actual situation which might be encountered by a radar, but should be viewed as a way of solving the problem in which the pulse to pulse independence property is not explicitly employed in the derivation of a decision rule. The results of Part I are, however, perfectly applicable to the problem of detection using a single pulse.

A brief summary of the following sections of this report is given here. In Section 2, a likelihood ratio test is derived for the simpler problem in which the interference covariance matrices are assumed to be known, although they may vary from pulse to pulse. This represents the limiting case of perfect noise estimation, and the results provide useful insight into the problem of interest. The likelihood ratio decision rule for the original problem is derived in Section 3, where an approximation is introduced which is necesary to obtain a closed form for the test. A number of properties of the approximate test are derived in Section 4, and these tend to show the reasonableness of the approximation made. The performance of this test is analyzed in Section 5, where a further approximation is needed to obtain

numerical results. The limitations of the analysis resulting from
this approximation are also discussed there. Numerical results
are included in Section 5, and a brief general summary and
discussion is presented in Section 6. Supporting analysis of the
final formula for the probability of detection and its numerical
evaluation are the subject of the Appendix.

(2)   TARGET DETECTION IN KNOWN NON-STATIONARY INTERFERENCE

In the detection problem addressed in this study, a target is sought in the samples corresponding to a single range gate, while K other signal-free range gate outputs are available as well. If K is extremely large and if all these signal-free samples share the covariance matrix of the primary sample, we can use these secondary data vectors to determine an accurate estimate of the noise covariance matrices for each pulse. In the limit, these estimates become the actual covariances, and the problem reduces to the detection of a target in the primary data, in a background of interference whose statistical properties are completely known. The present section is devoted to the study of this simpler problem, which is useful because it provides context and a performance comparison for the more general problem, and also makes a convenient vehicle for the introduction of some of the notation required.

The inputs to the processor are all data vectors, of dimension M, corresponding to the M RF input channels of the radar. The sample vectors for a particular range gate (the one being tested for signal presence) form a sequence, denoted by z_j, where j runs from 1 through J, and J is the number of pulses in the pulse train. The sum of internal noise and external interference is assumed to be a zero-mean, circular Gaussian process, and its sample vectors for the jth pulse have the

covariance matrix

$$E_0 \, z_j \, z_j^\dagger = M_j \ .$$

The symbol E_0 stands for expectation value on the 'noise-alone' hypothesis and the superscript dagger represents the Hermitian conjugate of a vector or matrix. For different pulses, the data vectors are assumed to be independent, each pulse being characterized by the corresponding covariance matrix, M_j. In the present secton, all these covariance matrices are considered to be known.

We use the subscript '1' to denote the 'signal-plus-noise' hypothesis, and we characterize the signal component in the following way:

$$E_1 \, z_j = b \, \sigma_j \, s$$

where b is an unknown, complex scalar amplitude parameter, s is an M-vector which represents the relative signal amplitudes among the RF inputs, and the σ_j form a sequence of complex scalars which describes the pulse-to-pulse variation in signal amplitude and phase.

The signal direction vector, s, is normalized as follows

$$(s^\dagger s) = 1 \ ,$$

i.e. the sum of the absolute squares of its elements is unity. No specific normalization convention is assumed for the sequence σ_j. An example of the latter would be a Doppler progression:

$$\sigma_j = e^{ij\theta}$$

where θ is the phase change per pulse caused by target motion. It is physically reasonable to assume that s is unchanged from pulse to pulse, but the analysis is easily modified to allow this vector to be a function of j.

The joint probability density function (PDF) of the data vectors under the noise-alone hypothesis is

$$f_0(z_1,...,z_J) = \prod_{j=1}^{J} \left(\frac{1}{\pi^M \|M_j\|} e^{-(z_j^\dagger M_j^{-1} z_j)} \right)$$

where the double bar represents a determinant. There are no unknown parameters in this PDF. Under hypothesis one, the joint PDF is the same, except that z_j is replaced by the quantity $z_j - b\sigma_j s$ in the exponent:

$$f_1(z_1,...,z_J|b) = \prod_{j=1}^{J} \left(\frac{1}{\pi^M \|M_j\|} e^{-[z_j-b\sigma_j s]^\dagger M_j^{-1} [z_j-b\sigma_j s]} \right) .$$

Before maximizing over the unknown b, we form a likelihood ratio and take its logarithm:

$$\ell(b) \equiv \log \frac{f_1(z_1,...,z_J|b)}{f_0(z_1,...,z_J)}$$

$$= \sum_{j=1}^{J} \left((z_j^\dagger M_j^{-1} z_j) - [z_j-b\sigma_j s]^\dagger M_j^{-1} [z_j-b\sigma_j s] \right)$$

$$= -|b|^2 \sum_{j=1}^{J} |\sigma_j|^2 (s^\dagger M_j^{-1} s) + 2 \,\mathrm{Re}\left(b^* \sum_{j=1}^{J} \sigma_j^* (s^\dagger M_j^{-1} z_j) \right) .$$

We introduce the notation

$$A_j^2 \equiv (s^\dagger M_j^{-1} s)$$

and

$$u_j \equiv \frac{(s^\dagger M_j^{-1} z_j)}{(s^\dagger M_j^{-1} s)} = \frac{(s^\dagger M_j^{-1} z_j)}{A_j^2} ,$$

and substitute in the last formula. The result is

$$\ell(b) = -|b|^2 \sum_{j=1}^{J} A_j^2 |\sigma_j|^2 + 2 \, \text{Re}\left(b^* \sum_{j=1}^{J} A_j^2 \sigma_j^* u_j \right)$$

$$= - \sum_{j=1}^{J} A_j^2 |\sigma_j|^2 \left| b - \frac{\sum_{j=1}^{J} A_j^2 \sigma_j^* u_j}{\sum_{j=1}^{J} A_j^2 |\sigma_j|^2} \right|^2$$

$$+ \frac{\left| \sum_{j=1}^{J} A_j^2 \sigma_j^* u_j \right|^2}{\sum_{j=1}^{J} A_j^2 |\sigma_j|^2} .$$

The unknown parameter, b, is now chosen to make the quadratic term in which it appears vanish, and we have, simply

$$\ell = \underset{b}{\text{Max}} \, \ell(b) = \frac{\left| \sum_{j=1}^{J} A_j^2 \sigma_j^* u_j \right|^2}{\sum_{j=1}^{J} A_j^2 |\sigma_j|^2} .$$

Since the quantities in the denominator of this expression are all known, the likelihood ratio test reduces to

$$\left| \sum_{j=1}^{J} A_j^2 \sigma_j^* u_j \right|^2 > \text{constant.}$$

In this simple problem the likelihood ratio test is a 'uniformly most powerful test', and hence it yields an optimum detector.

The quantities, u_j, are recognized as colored noise matched filter outputs for the individual pulses, normalized in a particular way. In the absence of signal they have zero mean values:

$$E_0 \, u_j = 0 \; ,$$

but when a signal is present, we have

$$E_1 \, u_j = b \, \sigma_j \; .$$

The effect of the denominator in the definition of u_j is to cause these matched filters to pass the signal components with unaltered amplitudes.

The variance of the matched filter output is easily computed:

$$E_0 |u_j|^2 = \frac{1}{A_j^4} \, (s^\dagger M_j^{-1} E_0 z_j z_j^\dagger M_j^{-1} s) = \frac{1}{A_j^2} \; ,$$

and this expression gives meaning to the parameter, A_j, as the inverse of the residual noise after the matched filter processing (i.e. nulling) on each pulse. Another way of viewing the colored noise matched filter processing is based on a decomposition of each data vector into a scalar component in the signal direction and an orthogonal component, of dimension M-1. The orthogonal component is used to predict the noise in the signal component, and this predictor is subtracted out. The variance of the residue after subtraction is exactly the inverse of A_j.

Combining the means and variances just obtained, we see that the signal-to-noise ratio (SNR) of each matched filter output is just

$$(SNR)_j = |b|^2 A_j^2 |\sigma_j|^2 ,$$

and this is also the SNR of the jth term in the sum on which detection is based. If we give a name to this sum:

$$w \equiv \sum_{j=1}^{J} A_j^2 \sigma_j^* u_j ,$$

then we have

$$E_0 w = 0$$

$$E_1 w = b \sum_{j=1}^{J} A_j^2 |\sigma_j|^2$$

and

$$E_0|w|^2 = \sum_{j=1}^{J} A_j^4 |\sigma_j|^2 E_0 |u_j|^2 = \sum_{j=1}^{J} A_j^2 |\sigma_j|^2 .$$

The actual SNR of the likelihood ratio test is therefore the sum of the individual SNR's:

$$SNR = |b|^2 \sum_{j=1}^{J} A_j^2 |\sigma_j|^2 = \sum_{j=1}^{J} (SNR)_j .$$

This processor is simply a coherent integrator of matched filter outputs, in which account is taken of the varying signal amplitudes and residual noise variances of the individual pulses. If the signal amplitudes are constant and the total noise

is stationary, the processor provides a gain equal to J, the number of pulses, relative to the SNR achieved on each pulse after matched filtering (nulling). We will find analogues of all these properties and parameters in the more general problem, whose analysis begins in the next section.

(3)   DERIVATION OF THE LIKELIHOOD RATIO TEST

        We now return to the problem described in Section 1 and

assume that the covariance matrices, M_j, are all unknown. We

also introduce the secondary data, consisting of K vectors for

each pulse, denoted $z_j(k)$, $(k=1,..,K)$. These vectors have zero

mean under both hypotheses, and share the covariance matrix of

the primary vector for the corresponding pulse:

$$E \, z_j(k) \, z_j(k)^\dagger = M_j \ .$$

Under the noise-alone hypothesis, the joint PDF of the K+1

vectors associated with the jth pulse is

$$f_{0j}[z_j, z_j(1),...,z_j(K)] = \left\{ \frac{1}{\pi^M \|M_j\|} e^{-Tr(M_j^{-1} T_{0j})} \right\}^{K+1}$$

where

$$T_{0j} \equiv \frac{1}{K+1} \left(z_j z_j^\dagger + \sum_{k=1}^{K} z_j(k) z_j(k)^\dagger \right) \ .$$

This representation of the PDF exactly parallels that given in

Part I, and the corresponding expression in the signal-plus-noise

case is

$$f_{1j}[z_j, z_j(1),...,z_j(K)|b] = \left\{ \frac{1}{\pi^M \|M_j\|} e^{-Tr(M_j^{-1} T_{1j}(b))} \right\}^{K+1}$$

where

$$T_{1j}(b) \equiv \frac{1}{K+1} \left((z_j - b\sigma_j s)(z_j - b\sigma_j s)^\dagger + \sum_{k=1}^{K} z_j(k) z_j(k)^\dagger \right) \ .$$

The signal parameters which enter here have the same significance

as in Section 2. The joint PDF's of all the data under the two

hypotheses have the form of products:

$$f_0 = \prod_{j=1}^{J} f_{0j}$$

and

$$f_1(b) = \prod_{j=1}^{J} f_{1j}(b) .$$

To form the likelihood ratio test, each PDF is maximized over the unknown parameters separately, and then a ratio is computed. The unknown signal parameter, b, appears only in the 'number one' hypothesis, and we defer maximization over this parameter until last, forming a ratio first. As in the similar derivation in Part I, the maximization over each unknown covariance matrix is simple; we obtain

$$\underset{M_j}{\text{Max}} \, f_{0j} = \left\{ \frac{1}{(e\pi)^M \|T_{0j}\|} \right\}^{K+1}$$

and

$$\underset{M_j}{\text{Max}} \, f_{1j} = \left\{ \frac{1}{(e\pi)^M \|T_{1j}(b)\|} \right\}^{K+1} .$$

With b still to be varied, the likelihood ratio, L(b), is a product of factors, each raised to the power (K+1). It is convenient to work with the (K+1)st root of the likelihood ratio instead, and we write

$$L(b) \equiv \{\ell(b)\}^{K+1}$$

and

$$\ell(b) = \prod_{j=1}^{J} \ell_j(b) \ ,$$

where

$$\ell_j(b) = \frac{\|T_{0j}\|}{\|T_{1j}(b)\|} \ .$$

The sample covariance matrix of the secondary vectors for pulse j is given by the sum

$$\widehat{M}_j \equiv \frac{1}{K} \sum_{k=1}^{K} z_j(k) z_j(k)^\dagger \ ,$$

and this quantity is also the maximum likelihood estimator of the unknown covariance matrix for this pulse, based on the secondary data alone. In terms of this matrix, we have

$$(K+1) T_{0j} = K \widehat{M}_j + z_j z_j^\dagger \ .$$

and by the determinant lemma

$$\|A + a b^\dagger\| = \|A\| \left(1 + b^\dagger A^{-1} a \right) \ ,$$

for which a proof is given in Part I, we obtain

$$(K+1)^M \|T_{0j}\| = K^M \|\widehat{M}_j\| \left(1 + \frac{1}{K}(z_j^\dagger \widehat{M}_j^{-1} z_j) \right) \ .$$

By the same reasoning we find

$$(K+1)^M \|T_{1j}(b)\| = K^M \|\widehat{M}_j\| \left(1 + \frac{1}{K}[z_j - b\sigma_j s]^\dagger \widehat{M}_j^{-1}[z_j - b\sigma_j s] \right)$$

for the signal-plus-noise case, and thus we have the ratio

$$\ell_j(b) = \frac{1 + \frac{1}{K}(z_j^\dagger \widehat{M}_j^{-1} z_j)}{1 + \frac{1}{K}[z_j - b\sigma_j s]^\dagger \widehat{M}_j^{-1}[z_j - b\sigma_j s]} \ .$$

In analogy to the quantities introduced in Section 2, we make the definitions

$$\widehat{A}_j^{\,2} \equiv (s^\dagger \widehat{M}_j^{-1} s)$$

and

$$\widehat{u}_j \equiv \frac{(s^\dagger \widehat{M}_j^{-1} z_j)}{(s^\dagger \widehat{M}_j^{-1} s)} = \frac{(s^\dagger \widehat{M}_j^{-1} z_j)}{\widehat{A}_j^{\,2}} \ .$$

It should be noted that these new variables involve the estimator of the covariance matrix, and not the unknown M_j itself. It is also convenient to use the temporary notation

$$N_j \equiv (z_j^{\,\dagger} \widehat{M}_j^{-1} z_j) \ .$$

With the help of these definitions we can make the evaluation

$$[z_j - b\sigma_j s]^\dagger \widehat{M}_j^{-1} [z_j - b\sigma_j s] = N_j - 2\,\mathrm{Re}\{b^* \sigma_j^* \widehat{A}_j^{\,2} \widehat{u}_j\} + |b|^2 |\sigma_j|^2 \widehat{A}_j^{\,2}$$

$$= \widehat{A}_j^{\,2} |b\sigma_j - \widehat{u}_j|^2 + N_j - \widehat{A}_j^{\,2} |\widehat{u}_j|^2 \ .$$

The likelihood ratio for the jth pulse is then given by

$$\ell_j(b) = \frac{1 + \dfrac{1}{K} N_j}{1 + \dfrac{1}{K} N_j - \dfrac{1}{K} \widehat{A}_j^{\,2} |\widehat{u}_j|^2 + \dfrac{1}{K} \widehat{A}_j^{\,2} |b\sigma_j - \widehat{u}_j|^2} \ .$$

Next, we introduce the quantity

$$\Sigma_j \equiv \frac{1}{K} (N_j - \ddot{A}_j^{\,2} |\widehat{u}_j|^2)$$

and eliminate N_j, so that

$$\ell_j(b) = \frac{1 + \Sigma_j + \dfrac{1}{K} \widehat{A}_j^{\,2} |\widehat{u}_j|^2}{1 + \Sigma_j + \dfrac{1}{K} \widehat{A}_j^{\,2} |b\sigma_j - \widehat{u}_j|^2} \ .$$

The signif.cance of Σ_j can be seen by defining an inner product as follows:

$$(a,b) \equiv (a^\dagger \widehat{M}_j^{-1} b) \ .$$

Here, a and b are arbitrary complex M-vectors, and the corresponding norm is

$$\|a\|^2 \equiv (a^\dagger \widehat{M}_j^{-1} a) \ .$$

This definition is possible because the sample covariance matrix and its inverse are positive definite with probability one (so long as K exceeds M, a condition we now impose). We can now write

$$N_j = \|z_j\|^2$$

$$\widehat{A}_j^2 = \|s\|^2$$

and

$$\widehat{A}_j \widehat{u}_j = \frac{(s,z_j)}{\|s\|} \ .$$

This last expression is the component of z_j in the direction of the 'unit signal vector'

$$\frac{s}{\|s\|} \ .$$

The portion of z_j which is orthogonal to the signal direction (in the sense of this new inner product) is

$$z_{\perp j} \equiv z_j - \frac{(s,z_j)}{\|s\|^2} s$$

and in terms of this projection we obtain the desired expression

$$K\Sigma_j = N_j - \widehat{A}_j{}^2 |\widehat{u}_j|^2$$

$$= \|z_j\|^2 - \frac{|(s,z_j)|^2}{\|s\|^2} = \|z_{\perp j}\|^2 \ .$$

Thus $K\Sigma_j$ is the norm of that portion of the data vector which is orthogonal to the unit signal vector. This separation of a data vector into a 'signal' component and an orthogonal component, relative to some noise-dependent inner product, will appear again in connection with the coherent integration of pulses.

It will be shown in a later section that Σ_j is exactly analogous to the quantity Σ_B of Part I, and also that the variables

$$r_j \equiv \frac{1}{1 + \Sigma_j}$$

represent SNR loss factors, described statistically by the Beta distribution.

Returning to our likelihood ratio, we divide numerator and denominator by $(1 + \Sigma_j)$ and make use of the above definition to obtain the formula

$$\ell_j(b) = \frac{1 + \frac{1}{K} r_j \widehat{A}_j{}^2 |\widehat{u}_j|^2}{1 + \frac{1}{K} r_j \widehat{A}_j{}^2 |b\sigma_j - \widehat{u}_j|^2} \ .$$

The final likelihood ratio is the product of J of these factors, maximized over b, and we note that r_j, $\widehat{A}_j$ and $\widehat{u}_j$ are

sufficient statistics for the decision process. These scalar
parameters are, in turn, simple functions of the norms of s and
z_j, and the inner product (s, z_j), as defined above.

We cannot derive an exact likelihood ratio decision rule,
because the required maximization of the product of factors (over
b) cannot be carried out in closed form. However, we note that
the parameter K, the number of secondary data vectors, will
control the accuracy of the noise covariance estimate, and we
know that K will have to be large compared to M in order to avoid
significant SNR loss. For large K, the random variables $\hat{A}_j$ and
$\hat{u}_j$ will tend toward the constants A_j and u_j, as the sample
covariance matrix tends to the true noise covariance matrix. In
the same limit, the loss factors, r_j, will approach unity
because of the properties of the Beta distribution to which they
are subject. Thus all the quantities in the likelihood ratio,
other than K itself, remain bounded as K increases and we will
assume that for practical values of K, both the numerator and
denominator of the likelihood ratio (for each pulse) take the
form of unity plus a small quantity.

Motivated by this reasoning, we make the approximation

$$\ell(b) = \prod_{j=1}^{J} \ell_j(b)$$

$$\approx \frac{1 + \frac{1}{K} \sum_{j=1}^{J} r_j \hat{A}_j{}^2 |\hat{u}_j|^2}{1 + \frac{1}{K} \sum_{j=1}^{J} r_j \hat{A}_j{}^2 |b\sigma_j - \hat{u}_j|^2} .$$

products of a unitary matrix and a new set of random vectors.
These new random vectors are statistically indistinguishable from
their predecessors, and it would only be confusing to introduce a
new notation for them. Tracing this transformation through the
inner products, we find that only the normalized signal vector is
changed: t is replaced by

$$t_1 \equiv U_1 t$$

where U_1 is the unitary matrix characterizing this last
transformation. Any unit vector in the complex N-space can be
realized, as t_1, by such a transformation. In particular, we
can cause t_1 to be a 'coordinate vector', for which a single
element is unity, the remaining (N-1) elements vanishing. It is
for this reason that the PDF of η depends on M only through the
meaning of the signal amplitude parameter, b. In fact, this PDF
can depend only on b, N and K, and hence the false alarm
probability of the likelihood ratio detector, namely

$$PFA = Prob\{\eta > \eta_0\},$$

is independent of M, and this is the generalized CFAR property
claimed in section 1.

$$X = 1 + \frac{1}{K} \sum_{j=1}^{J} r_j \widehat{A}_j{}^2 |\widehat{u}_j|^2$$

and

$$Y = \frac{|\sum_{j=1}^{J} r_j \widehat{A}_j{}^2 \sigma_j{}^* \widehat{u}_j|^2}{\sum_{j=1}^{J} r_j \widehat{A}_j{}^2 |\sigma_j|^2} \quad .$$

The likelihood ratio test

$$\ell > \ell_0$$

becomes

$$X > \ell_0 X - \frac{\ell_0}{K} Y \quad ,$$

which is equivalent to

$$Y > K \frac{\ell_0 - 1}{\ell_0} X \quad ,$$

or, in terms of the original variables:

$$\frac{|\sum_{j=1}^{J} r_j \widehat{A}_j{}^2 \sigma_j{}^* \widehat{u}_j|^2}{\sum_{j=1}^{J} r_j \widehat{A}_j{}^2 |\sigma_j|^2} > K \frac{\ell_0 - 1}{\ell_0} \left(1 + \frac{1}{K} \sum_{j=1}^{J} r_j \widehat{A}_j{}^2 |\widehat{u}_j|^2 \right) \quad ,$$

which is the desired form. Some of the properties of this test are derived in the next section.

(4)   PROPERTIES OF THE APPROXIMATE LIKELIHOOD RATIO TEST

The approximate likelihood ratio (ALR) test, derived in the previous section, has a number of properties which suggest that it is a reasonable detection procedure, in spite of the approximation made in its derivation. Before discussing these, it is perhaps worthwhile to point out just what is involved in its actual implementation. The procedure requires the estimation of a covariance matrix for each pulse, carried out by averaging the open (dyadic) products of the secondary data vectors. Two sets of linear equations must then be solved, which we may write in the form

$$\widehat{M}_j w_j = s$$

and

$$\widehat{M}_j q_j = z_j \ .$$

The three inner products are then given by

$$(s^\dagger \widehat{M}_j^{-1} s) = (w_j^\dagger s)$$

$$(s^\dagger \widehat{M}_j^{-1} z_j) = (w_j^\dagger z_j)$$

and

$$(z_j^\dagger \widehat{M}_j^{-1} z_j) = (q_j^\dagger z_j) \ .$$

The variables r_j, $\widehat{A}_j$ and $\widehat{u}_j$ are, as we have seen, defined directly in terms of these inner products, and only these variables are needed for the construction of the ALR test.

By far the greatest task is the solution of the linear
equations, normally carried out by means of a Cholesky
factorization technique. A nulling scheme which employs the
colored-noise matched filter method directly, and which is
designed to compute new weights for each pulse, will have to
solve the first of the above equations to find the appropriate
weight vector. After factorization of the sample covariance
matrix, it does not require much extra computation to solve the
second equation as well. The inner products are, of course,
simple complex sums, as are the sums over j which enter into the
test itself. Therefore, the implementation of the ALR decision
rule would require only a moderate increase in computational
complexity over such a matched-filter nulling processor.

From the form of the ALR test, and the definitions of
$\hat{r}_j$, $\hat{A}_j$ and u_j, it can be seen that the test is unaltered if
the signal vector, s, is changed by a scale factor. Thus the
detection rule is not affected by our choice of signal vector
normalization. More obviously, the test is unchanged if the
sequence σ_j is multiplied by a common scale factor as well.
Finally, the test is invariant to a scaling of all the data
vectors, primary and secondary, by a common scalar factor. This
last represents a weak form of CFAR behavior. It will be shown
later that the test is not a true CFAR, in the sense that its
probability of false alarm (PFA) is completely insensitive to the

actual covariance matrices of the noise on all of the pulses. However, it will turn out that the test is approximately a true CFAR, and we will be able to get some idea of the nature of its departure from the desirable true CFAR performance.

If K, the number of secondaries, becomes very large, the ALR test passes over into the likelihood ratio test derived in Section 2, for the case of known noise. This follows from the convergence of the noise covariance estimators to the true covariance matrix in this limit, and, as discussed before, the convergence of the Beta-distributed loss factors, r_j, to unity. In the final form of the ALR test, given as the last equation of Section 3, the summation over j on the right side will remain finite as K increases, and hence this sum, divided by K, will tend to zero. If the quantity

$$K \frac{\ell_0 - 1}{\ell_0}$$

is then defined as a new threshold constant, the limiting form is seen to be identical with the known-noise test of Section 2.

In the case of a single pulse the ALR test is exact, since the approximation we have made is irrelevant in this situation. This case is actually the same as the general problem treated in Part I, since we have made use of no special structure for our M-vectors. If we put $J = 1$, the ALR test becomes

$$r \widehat{A}^2 |\widehat{u}|^2 > K \frac{\ell_0 - 1}{\ell_0} \left(1 + \frac{1}{K} r \widehat{A}^2 |\widehat{u}|^2 \right)$$

II-24

or

$$\frac{1}{K} r \widehat{A}^2 |\widehat{u}|^2 > \ell_0 - 1 .$$

We have dropped the subscript j in these expressions. Substituting for r, $\widehat{A}$ and $\widehat{u}$, we obtain

$$\frac{|(s^\dagger \widehat{M}^{-1} z)|^2}{K (s^\dagger \widehat{M}^{-1} s)} > (\ell_0 - 1)(1 + \Sigma) ,$$

again without the subscripts. Finally, substituting for Σ, we find

$$\frac{|(s^\dagger \widehat{M}^{-1} z)|^2}{K (s^\dagger \widehat{M}^{-1} s)} > \frac{\ell_0 - 1}{\ell_0} \left(1 + \frac{1}{K} (z^\dagger \widehat{M}^{-1} z)\right) ,$$

which is identical to the decision rule derived in Part I.

This test can be put in another form which will provide an interesting analogy to the multiple-pulse ALR decision rule. Using the inner product notation introduced in Section 3 and the definition of the orthogonal component of the data vector with respect to that inner product, we can write the next to last form for the single-pulse test as

$$\frac{|(s,z)|^2}{(s,s)} > (\ell_0 - 1)\left(K + (z_\perp, z_\perp)\right) .$$

Returning to the multiple-pulse test, we define a new inner product, using square brackets to distinguish it, as follows:

$$[a,b] \equiv \sum_{j=1}^{J} r_j \widehat{A}_j^2 a_j^* b_j .$$

Here a and b are arbitrary sequences of complex numbers, or

J-vectors. Using this inner product, the ALR test can be written

$$\frac{|[\sigma,\hat{u}]|^2}{[\sigma,\sigma]} > \frac{\ell_0 - 1}{\ell_0}\left(K + [\hat{u},\hat{u}]\right) .$$

Again we can separate the data vector, this time the J-vector $\hat{u}$, into two components relative to the signal vector, σ, defining an orthogonal component as follows:

$$\hat{u}_\perp \equiv \hat{u} - \frac{[\sigma,\hat{u}]}{[\sigma,\sigma]}\sigma .$$

The norm of $\hat{u}$ is then the sum

$$[\hat{u},\hat{u}] = \frac{|[\sigma,\hat{u}]|^2}{[\sigma,\sigma]} + [\hat{u}_\perp,\hat{u}_\perp]$$

and the ALR test reduces to

$$\frac{|[\sigma,\hat{u}]|^2}{[\sigma,\sigma]} > (\ell_0 - 1)\left(K + [\hat{u}_\perp,\hat{u}_\perp]\right) ,$$

which stands in remarkable analogy to the single-pulse form of the test, just derived above.

To procede with the analysis of the ALR test we follow the procedure of Part I and introduce a change of variables, to 'whitened' coordinates. Suppose the actual noise covariance matrix for the data vectors of the jth pulse is M_j. New vectors are defined by the equations

$$\gamma_j \equiv M_j^{-\frac{1}{2}} z_j$$

and

$$\gamma_j(k) \equiv M_j^{-\frac{1}{2}} z_j(k) .$$

The covariance matrix shared by all K+1 of these vectors is the M×M identity matrix, and the same is true for the data vectors for all J pulses. The sample covariance matrix of the whitened secondaries is given by

$$\mathcal{M}_j \equiv \frac{1}{K} \sum_{k=1}^{K} \gamma_j(k)\gamma_j(k)^\dagger = M_j^{-\frac{1}{2}} \widehat{M}_j M_j^{-\frac{1}{2}} \ .$$

The signal vector is also changed by the whitening transformation into a different vector for each pulse:

$$t_j \equiv M_j^{-\frac{1}{2}} s \ ,$$

and here we see it would be very simple to allow the original s vector to be a function of j. The ordinary norm of the transformed signal vector is

$$(t_j^\dagger t_j) = (s^\dagger M_j^{-1} s) = A_j^{\ 2} \ ,$$

the quantity encountered in Section 2, which is the inverse of the residual noise that would remain after nulling with a filter matched to the actual noise.

In Part I we were able to choose a convenient signal norm which made this quantity unity, since there was, in effect, only one pulse. That choice cannot be made here, and hence a simple norm has been chosen for s itself. The presence of the A_j in our expressions will be the chief difference between the present analysis and that of Part I. The evaluations that are made next follow very closely those of Part I, and only enough detail will be given to make clear the small differences. Besides the

appearance of the A_j, we are now working with the sample covariance matrices themselves, instead of the S-matrices of Part I, which differ by a factor of K.

We introduce the inverses of the whitened sample covariance matrices:

$$\mathcal{P}_j \equiv \mathcal{M}_j^{-1} = M_j^{\frac{1}{2}} \, \widehat{M}_j^{-1} M_j^{\frac{1}{2}}$$

and substitute the definitions in our three inner products. We obtain

$$\widehat{A}_j^{\,2} = (s^\dagger \widehat{M}_j^{-1} s) = (t_j^{\,\dagger} \mathcal{P}_j t_j)$$

$$(s^\dagger \widehat{M}_j^{-1} z_j) = (t_j^{\,\dagger} \mathcal{P}_j \gamma_j)$$

and

$$(z_j^{\,\dagger} \widehat{M}_j^{-1} z_j) = (\gamma_j^{\,\dagger} \mathcal{P}_j \gamma_j) \; .$$

The new random vectors are, of course, still Gaussian, and the secondaries still have zero means under either hypothesis. The whitened primary vector has zero mean in the absence of signal, and mean value

$$E_1 \gamma_j = M_j^{-\frac{1}{2}} E_1 z_j = b \, \sigma_j t_j$$

under the 'signal-plus-noise' hypothesis. These variables are thus completely characterized statistically, and we note that the actual noise covariance matrices, M_j, appear only in the new signal vectors, the t_j.

Following the method of Part I, we next make unitary transformations, one for each pulse, which will leave all the

Gaussian random vectors statistically unaltered except for the mean values of the primary vectors. These mean values are proportional to the transformed signal vectors, and we choose the unitary transformations to make each signal vector proportional to the 'coordinate unit vector' whose transpose is $[1,0,\ldots,0]$. Since the norms of the signal vectors are unchanged in this process, the transformed t_j will be given by

$$A_j \begin{bmatrix} 1 \\ 0 \end{bmatrix} .$$

In this vector, the 'one' is a scalar and the '0' is an $(M-1)$ vector. We can think of these unitary tranformations as being combined with the preceding whitening transformations (which were in any case not unique), and hence we make no change in the notation to reflect the unitary transformations.

We see now that the matrices, M_j, affect the statistical character of the inner products only through the numbers, A_j, a single scalar for each pulse. Moreover, these inner products are sufficient statistics for decision, since the ALR test itself depends on the data vectors only through them. The only other parameters which enter the test are the signal amplitude constants, the σ_j, and we shall find that the test statistics depend only on the products, $A_j \sigma_j$.

The data vectors and sample covariance matrices are now decomposed into blocks, separating out the 'one' component,

called the A portion, and the (M-1) dimensional remainder, the B portion. Thus we write

$$\gamma_j = \begin{bmatrix} \gamma_{jA} \\ \gamma_{jB} \end{bmatrix}$$

$$\gamma_j(k) = \begin{bmatrix} \gamma_{jA}(k) \\ \gamma_{jB}(k) \end{bmatrix}$$

and

$$\mathcal{M}_j = \begin{bmatrix} \mathcal{M}_{jAA} & \mathcal{M}_{jAB} \\ \mathcal{M}_{jBA} & \mathcal{M}_{jBB} \end{bmatrix} \,,$$

with a similar decomposition for the inverse of the latter matrix. These forms are substituted in the inner products, and use is made of the Frobenius relations for partitioned matrices, exactly as in Part I.

When this is done, it turns out that

$$\Sigma_j = \frac{1}{K} \left(\gamma_{jB}{}^{\dagger} \mathcal{M}_{jBB}{}^{-1} \gamma_{jB} \right) \,,$$

which is the exact counterpart of Σ_B of Part I. It follows that the loss factors, i.e. the r_j, all obey the Beta distribution

$$f(r) = \frac{(N+L-1)!}{(N-2)!\,L!} \, (1-r)^{N-2} r^L$$

where

$$L \equiv K + 1 - M \,.$$

This last parameter plays the same basic role here as did its counterpart before, but in addition to the number of secondaries,

L now depends on the dimension, M, of the data vectors for each pulse, instead of the total number, MJ, of data vectors which enter in the detection process. This difference is due, of course, to the more specific assumption made here concerning the covariance structure of the interference, namely pulse-to-pulse independence.

When all these substitutions are carried out, new combinations of variables appear, which motivate the definitions

$$y_j \equiv \gamma_{jA} - \mathcal{M}_{jAB}\,\mathcal{M}_{jBB}^{-1}\,\gamma_{jB}$$

and

$$y_j(k) \equiv \gamma_{jA}(k) - \mathcal{M}_{jAB}\,\mathcal{M}_{jBB}^{-1}\,\gamma_{jB}(k) \ .$$

These variables are the precise analogues of y and y(k), encountered in Part I, and we also obtain the relation

$$\mathcal{M}_{jAA} - \mathcal{M}_{jAB}\,\mathcal{M}_{jBB}^{-1}\,\mathcal{M}_{jBA} = \frac{1}{K}\sum_{k=1}^{K} |y_j(k)|^2 \ .$$

Note the appearance of the factor 1/K in this formula, which results because the left side is expressed in terms of the sample covariance matrices themselves.

When conditioned on the B components of the primary and secondary vectors of pulse j, the y_j and $y_j(k)$ are Gaussian, and y_j is independent of the $y_j(k)$. The K×K conditional covariance matrix of the $y_j(k)$ has the same structure found earlier; it is an idempotent matrix with L unit eigenvalues, the remaining M-1 eigenvalues being zero. Conditionally, the means of

the $y_j(k)$ are always zero, and that of y_j is given in general by

$$E_B y_j = E_B r_{jA} = b A_j \sigma_j \; .$$

The subscript B denotes the conditioning, and obviously this last mean is zero in the absence of signal. Finally, the conditional variance of y_j is related to the loss factor, as follows:

$$E_B |y_j - b A_j \sigma_j|^2 = 1 + \Sigma_j = \frac{1}{r_j} \; .$$

In terms of these y-variables, it is easily shown that

$$\widehat{u}_j = \frac{y_j}{A_j}$$

and

$$\widehat{A}_j^{\,2} = \frac{K A_j^{\,2}}{\displaystyle\sum_{k=1}^{K} |y_j(k)|^2} \; .$$

The variables which enter into the three basic inner products have now been written in terms of the y's and Σ_j, whose statistical properties have been completely described. The 'square-bracket' inner products themselves are expressible in terms of these quantities, and we find

$$[\widehat{u}, \widehat{u}] = \sum_{j=1}^{J} r_j \widehat{A}_j^{\,2} |\widehat{u}_j|^2 = K \sum_{j=1}^{J} \frac{r_j |y_j|^2}{\displaystyle\sum_{k=1}^{K} |y_j(k)|^2}$$

$$[\sigma, \sigma] = \sum_{j=1}^{J} r_j \widehat{A}_j^{\,2} |\sigma_j|^2 = K \sum_{j=1}^{J} \frac{r_j A_j^{\,2} |\sigma_j|^2}{\displaystyle\sum_{k=1}^{K} |y_j(k)|^2}$$

and

$$[\sigma,\hat{u}] = \sum_{j=1}^{J} r_j \hat{A}_j^2 \sigma_j^* \hat{u}_j = K \sum_{j=1}^{J} \frac{r_j A_j \sigma_j^* y_j}{\sum_{k=1}^{K} |y_j(k)|^2} \ .$$

We have already seen that the ALR test takes a very simple form when written in terms of these inner products.

Following the analysis of Part I, we find that the sums

$$T_j \equiv \sum_{k=1}^{K} |y_j(k)|^2$$

are Chi-squared random variables, independent of conditioning on the B-components, and each T_j is subject to the PDF

$$f(T) = \frac{T^{L-1}}{(L-1)!} e^{-T} \ .$$

Note that the number of degrees of freedom is related to L and not K, because of the mutual correlation of the $y_j(k)$. We now introduce new, normalized random variables instead of the T_j, namely

$$\tau_j \equiv \frac{T_j}{L}$$

These new variables have unit mean values, and each has variance equal to $1/L$. These properties are central to an approximation we shall make in the next section, in connection with the performance of the ALR test. Like all variables relating to different pulses, the τ_j are mutually independent.

Again following Part I, we replace the y_j by new variables, $'_j$, which are defined so that they have unit variance, conditioned on the B-components:

$$w_j \equiv \frac{y_j}{(1 + \Sigma_j)^{\frac{1}{2}}} = r_j^{\frac{1}{2}} y_j \ .$$

The loss factors now appear in the conditional means:

$$E_B w_j = b \, r_j^{\frac{1}{2}} A_j \sigma_j \ .$$

To express the 'square' inner products in compact form, it is natural to use the definition

$$\mu_j \equiv r_j^{\frac{1}{2}} A_j \sigma_j \ .$$

These quantities are random variables, but they are dependent only on the B-components of the data vectors. This last definition allows us to write

$$E_B \bar{w}_j \approx b \, \mu_j$$

and the inner products are then given by the equations

$$\frac{1}{K}\left[\hat{u},\hat{u}\right] = \frac{1}{L} \sum_{j=1}^{J} \frac{|w_j|^2}{\tau_j}$$

$$\frac{1}{K}\left[\sigma,\sigma\right] = \frac{1}{L} \sum_{j=1}^{J} \frac{|\mu_j|^2}{\tau_j}$$

and

$$\frac{1}{K}\left[\sigma,\hat{u}\right] = \frac{1}{L} \sum_{j=1}^{J} \frac{\mu_j^* w_j}{\tau_j} \ .$$

The ALR test itself is therefore statistically equivalent to the test

II-34

$$\frac{|\sum_{j=1}^{J} \frac{\mu_j^* w_j}{\tau_j}|^2}{\sum_{j=1}^{J} \frac{|\mu_j|^2}{\tau_j}} > L \frac{\ell_0 - 1}{\ell_0} \left(1 + \frac{1}{L} \sum_{j=1}^{\cdot} \frac{|w_j|^2}{\tau_j}\right) .$$

Of course, the transformations we have made were based on the actual noise covariance matrix, and the form of the test just given will be used only to study its performance; it is not an alternative representation to the original form, which was written in terms of the observables themselves.

Conditioned on the B-components of the whitened vectors, the μ_j are constants and the w_j are simple Gaussian variables. We note that the conditioning is represented now only by the presence of the loss factors, hence we can interpret the conditional probabilities as being conditioned on a set of values of these J independent random variables. Besides these, only the products $A_j \sigma_j$ appear in the test, and also in the statistics, since the conditional means of the w_j are equal to the numbers μ_j, which are proportional to these same products.

Because of the presence of the products $A_j \sigma_j$ in the decision rule, it is not a true CFAR, in the sense that the probability of false alarm is not totally insensitive to the actual noise covariance matrices. However, the PFA is invariant to any permutation of these numbers, since the random variables which enter the test are independent from pulse to pulse and statistically indentical for different pulses. The PFA is also

unchanged if all the A_j are changed by a common factor. We conclude that the PFA depends on the variability of the sequence of products, $A_j \sigma_j$, in some normalized way, such as the ratio of standard deviation to mean of these numbers. It seems likely that the dependence of the PFA on this variability is not strong, and in the next section an approximation will be introduced which causes this dependence to disappear entirely. This approximation can be used to set a unique threshold for the test, for a given assigned PFA, and the actual PFA will then vary somewhat from this assigned value, according to the actual pulse to pulse variations in the level of the residual noise after nulling.

(5)   PERFORMANCE OF THE APPROXIMATE LIKELIHOOD RATIO TEST

To evaluate the performance of the ALR test exactly, it would be necessary to obtain the PDF of the quantity

$$\xi \equiv \frac{\left|\sum_{j=1}^{J} \frac{\mu_j^* w_j}{\tau_j}\right|^2}{\sum_{j=1}^{J} \frac{|\mu_j|^2}{\tau_j}} - \frac{\ell_0 - 1}{\ell_0}\left(L + \sum_{j=1}^{J} \frac{|w_j|^2}{\tau_j}\right) .$$

The probability that the detection threshold is exceeded is then equal to

$$\text{Prob}(\xi > 0) .$$

Evaluation of this probability under the noise-alone hypothesis yields the PFA of the decision rule, and under the signal-plus-noise hypothesis we obtain the PD. We have shown that the τ_j are independent, normalized Chi-squared variables, independent of any conditioning, and that the w_j are independent. The w_j are also Gaussian, when conditioned on the loss factors contained in the μ_j. Finally, the loss factors themselves are independent and satisfy the Beta distribution given in Section 4. We would like to be able to evaluate the conditional PDF of ξ, and then remove the conditioning by taking the expectation value of this probability with respect to the loss factors. This procedure was feasible for the general problem analyzed in Part I, but appears to be intractable here.

It was pointed out in Section 4 that the PDF of the τ_j has mean value unity and variance equal to $1/L$, where

L=K+1-M. Since L will have to be large compared to unity in order to control SNR losses (i.e. to keep the loss factors close to unity), it may be expected that the τ_j will not differ greatly from unity themselves. A family of PDF curves for this normalized Chi-squared distribution is presented in Fig. 1, for various values of L. It can be seen that the PDF becomes relatively narrow and quite symmetrical about the mean, for values of L in excess of about 50. This suggests making the simplifying approximation

$$\tau_j \approx 1$$

in the expression for ξ, which then becomes

$$\xi \approx \frac{\left|\sum\limits_{j=1}^{J} \mu_j^* w_j\right|^2}{\sum\limits_{j=1}^{J} |\mu_j|^2} - \frac{\ell_0 - 1}{\ell_0}\left(L + \sum\limits_{j=1}^{J} |w_j|^2\right) .$$

This approximation greatly simplifies the PDF of ξ, and it is only this simplified form that will be discussed here. The penalty associated with the simplification is, of course, a failure of our results to describe all aspects of the original problem. In Part I it was shown that the likelihood ratio test derived there was equivalent to a simple CFAR detection process, in which the threshold was estimated from secondary data, and the target exhibited a Beta-distributed fluctuation. The threshold estimation in that problem was represented by a single variable exactly like one of our τ_j. The resulting performance

NORMALIZED CHI—SQUARED PROBABILITY DENSITY FUNCTIONS
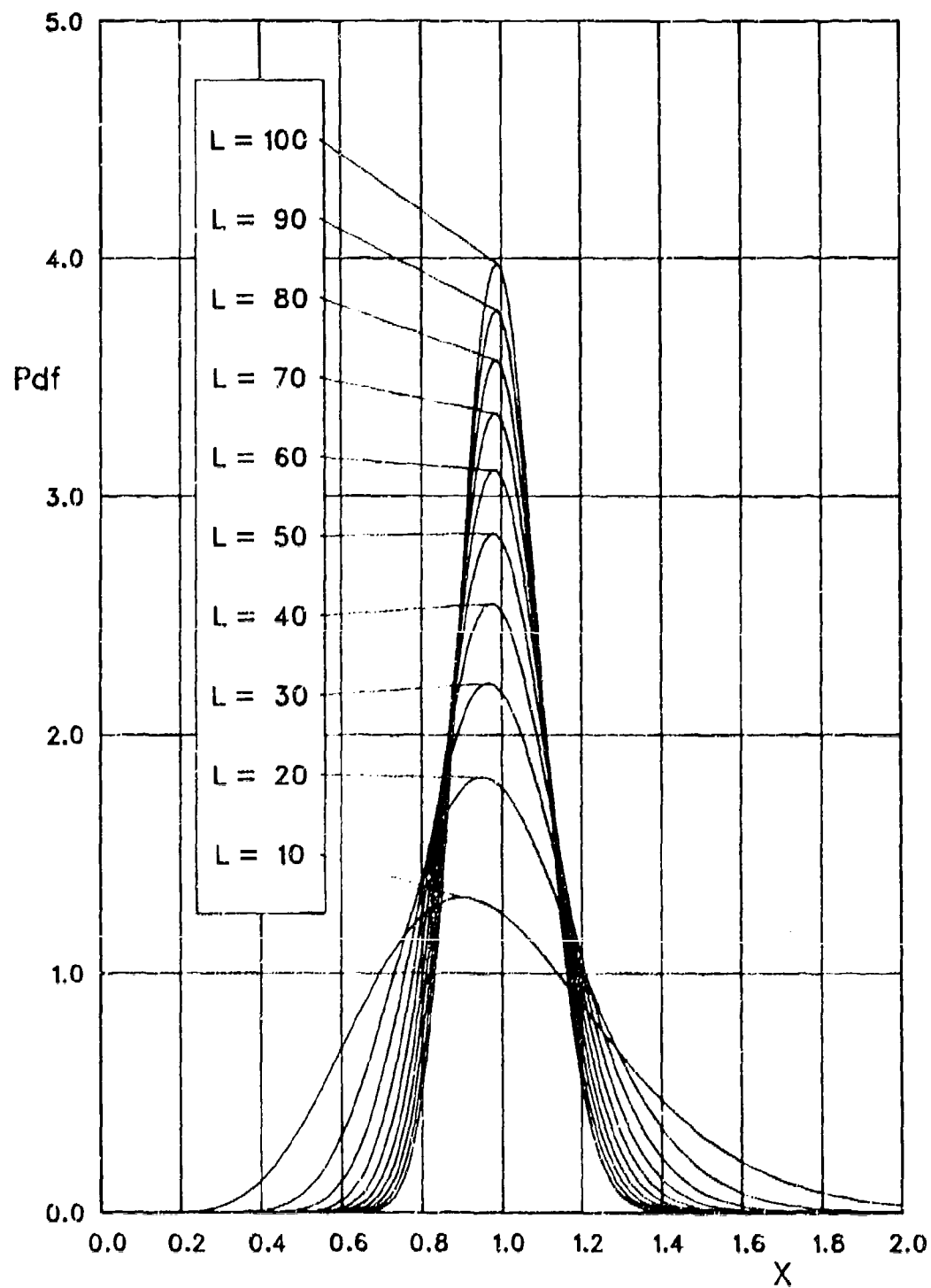

Fig. 11-1   Normalized Chi-Squared Probability Density
Functions

analysis (in Part I) also showed that the total SNR loss could be recognized as the sum of two contributions: one due to the target fluctuation and another due to the threshold estimation. The latter effect, which provided the CFAR behavior, accounted for a SNR loss very close to the standard CFAR loss for a non-fluctuating target and threshold estimation by summation of noise sample powers. The number of terms entering in this summation is exactly L. In the present problem there is a τ_j factor for each term in the sums which enter in the decision statistic, since noise estimation is carried out for each pulse. Each pulse will then show a CFAR loss, and we expect that the overall performance of the actual ALR test will have a CFAR loss of the same order of magnitude (in db) as that for a typical pulse. It is this aspect of performance that will not be described by our results when the simplifying approximation is made.

In the approximating expression for ξ we can interpret the w_j and μ_j as components of complex J-vectors and, as before, make use of the inner product notation. Then we have

$$\sum_{j=1}^{J} \mu_j^* w_j = [\mu, w]$$

$$\sum_{j=1}^{J} |\mu_j|^2 = [\mu, \mu]$$

and

$$\sum_{j=1}^{J} |w_j|^2 = [w, w] \ .$$

In this J-dimensional complex vector space, we now introduce new vectors, obtained from w and μ by means of a unitary matrix U:

$$w' \equiv U\,w$$

and

$$\mu' \equiv U\,\mu \ .$$

Inner products and norms are unchanged by this transformation, and we can choose U so that μ' is aligned with a coordinate vector, just as was done in a different space in Section 3. In particular, we can choose U so that

$$\mu' = [\mu,\mu]^{\frac{1}{2}} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ .$$

The norm of μ cancels out in the expression for ξ, which is now simply

$$\xi \approx |w'_1|^2 - \frac{\ell_0-1}{\ell_0}\left(L + \sum_{j=1}^{J} |w'_j|^2 \right) \ .$$

$$= \frac{1}{\ell_0}|w'_1|^2 - \frac{\ell_0-1}{\ell_0}\left(L + \sum_{j=2}^{J} |w'_j|^2 \right) \ .$$

We define

$$\xi' \equiv \ell_0\,\xi$$

so that

$$\xi' = |w'_1|^2 - (\ell_0-1)\sum_{j=2}^{J} |w'_j|^2 - (\ell_0-1)L$$

and note that the probability of exceeding the detection threshold is equal to

$$\mathrm{Prob}(\xi' > 0) \ .$$

The components of w' are independent and Gaussian, given the loss-factor conditioning, and each has unit variance. The conditional mean of the vector w' is

$$E_B w' = b \mu' ,$$

and therefore all components of w' except the first have zero mean. The latter component has mean value

$$E_B w'_1 = b[\mu, \mu]^{\frac{1}{2}} = a^{\frac{1}{2}}$$

where

$$a \equiv |b|^2 [\mu, \mu] = |b|^2 \sum_{j=1}^{J} r_j A_j^2 |\sigma_j|^2 .$$

This last parameter is the basic signal parameter of the simplified problem, and we see that it is just like the SNR of the known-noise test (see Section 2) except for the presence of the loss factors, the r_j.

In the noise-alone situation, $a = 0$, and the simplified test performance depends only on L and the threshold parameter; it is in this approximation a true CFAR. When a signal is added, its presence is felt only through the parameter, a, and we shall calculate the conditional PD for this case, as a function of a. It will not be possible to average this PD over the PDF of a, in order to remove the conditioning, because of the difficulty of dealing with a weighted sum of Beta-distributed variables. Again we appeal to the results of Part I, where it was shown that the

exact SNR loss caused by the loss factor was very similar to the expected value of the loss factor itself. We therefore note that

$$E a = E r \, |b|^2 \sum_{j=1}^{J} A_j^2 |\sigma_j|^2$$

where

$$E r = \int_0^\infty r \, f(r) \, dr$$

$$= \frac{L+1}{N+L} = \frac{K+2-N}{K+1}$$

is the expected loss. The expected value of a is thus equal to the known-noise SNR value times the expected value of a typical loss factor, as found from the Beta distribution.

It remains to derive the probability of threshold crossing for fixed a, from the simplified version of the test statistic. We define the characteristic function of the random variable ξ':

$$\Phi(\lambda) \equiv E \, e^{i\lambda\xi'} \, .$$

and note that the desired probability can be expressed as an integral in the complex λ-plane, as follows

$$\text{Prob}(\xi' > 0) = \frac{1}{2\pi i} \int_{-\infty-i\epsilon}^{\infty-i\epsilon} \Phi(\lambda) \, \frac{d\lambda}{\lambda} .$$

The interchange of order of integration required to obtain this result is made valid by the displacement of the contour below the real axis in the complex plane. The characteristic function itself can be expressed in the form

$$\Phi(\lambda) = e^{-i(\ell_0-1)L\lambda} \, \Phi_1(\lambda) \, \Phi_2[-(\ell_0-1)\lambda] \, .$$

II-43

where

$$\Phi_1(\lambda) \equiv E\,e^{\,i\lambda|w'_1|^2}$$

and

$$\Phi_2(\lambda) \equiv E\,e^{\,i\lambda\sum_{j=2}^{J}|w'_j|^2}$$

is the characteristic function of the sum

$$S \equiv \sum_{j=2}^{J} |w'_j|^2 \ .$$

Since S is the sum of the squares of (J-1) complex Gaussian variables, each with mean zero and variance unity, it is a Chi-squared variable of 2J-2 degrees of freedom, and its characteristic function is therefore

$$\Phi_2(\lambda) = (1 - i\lambda)^{-(J-1)} \ .$$

The random variable w'_1 is conditionally Gaussian, with unit variance and a mean squared value of a, hence its characteristic function is

$$\Phi_1(\lambda) = e^{-a}\,\frac{e^{\frac{a}{1-i\lambda}}}{1 - i\lambda} \ .$$

Substituting, we obtain

$$\Phi(\lambda) = e^{-a - i(\ell_0-1)L\lambda}\,\frac{e^{\frac{a}{1-i\lambda}}}{(1 - i\lambda)\{1 + i(\ell_0-1)\lambda\}^{J-1}} \ .$$

The final complex integral is evaluated in the Appendix, where the desired probability is derived as a series of Marcum Q-functions. A numerical analysis of this series is also given

there, on the basis of which the figures of the present section have been produced. When there is no signal present, the characteristic function simplifies greatly, and the contour (closed in the lower half plane) includes only the simple pole at $\lambda = -i$. The residue at this pole is easily evaluated, with the resulting formula for the probability of false alarm

$$PFA = (1/\ell_0)^{J-1} e^{-(\ell_0-1)L} .$$

For a single pulse, this formula and the corresponding formula for P_D and identical to the expressions for PFA and PD for conventional detection with a single radar hit. This is to be expected, since they are conditioned on given loss factors, and the CFAR effect associated with threshold estimation has been eliminated by our simplying approximation.

The general character of the formulas is illustrated in Figs. 2 through 7, in which $P_D(a)$ is plotted vs a, for several sets of values for the parameters L and J. In all cases, the PFA has been set to $10^{-6}.$ The dashed curves on these plots represent the PD vs SNR for a conventional detector using a single radar hit, or Marcum's Q-function for N = 1. The two curves are extremely close when J and L are large, and when the number of pulses is large, they are relatively insensitive to L. We conclude that the performance of the ALR test can be expected to be very similar to that of the known-noise test of Section 2, with two additional losses: a CFAR loss typical of the parameter
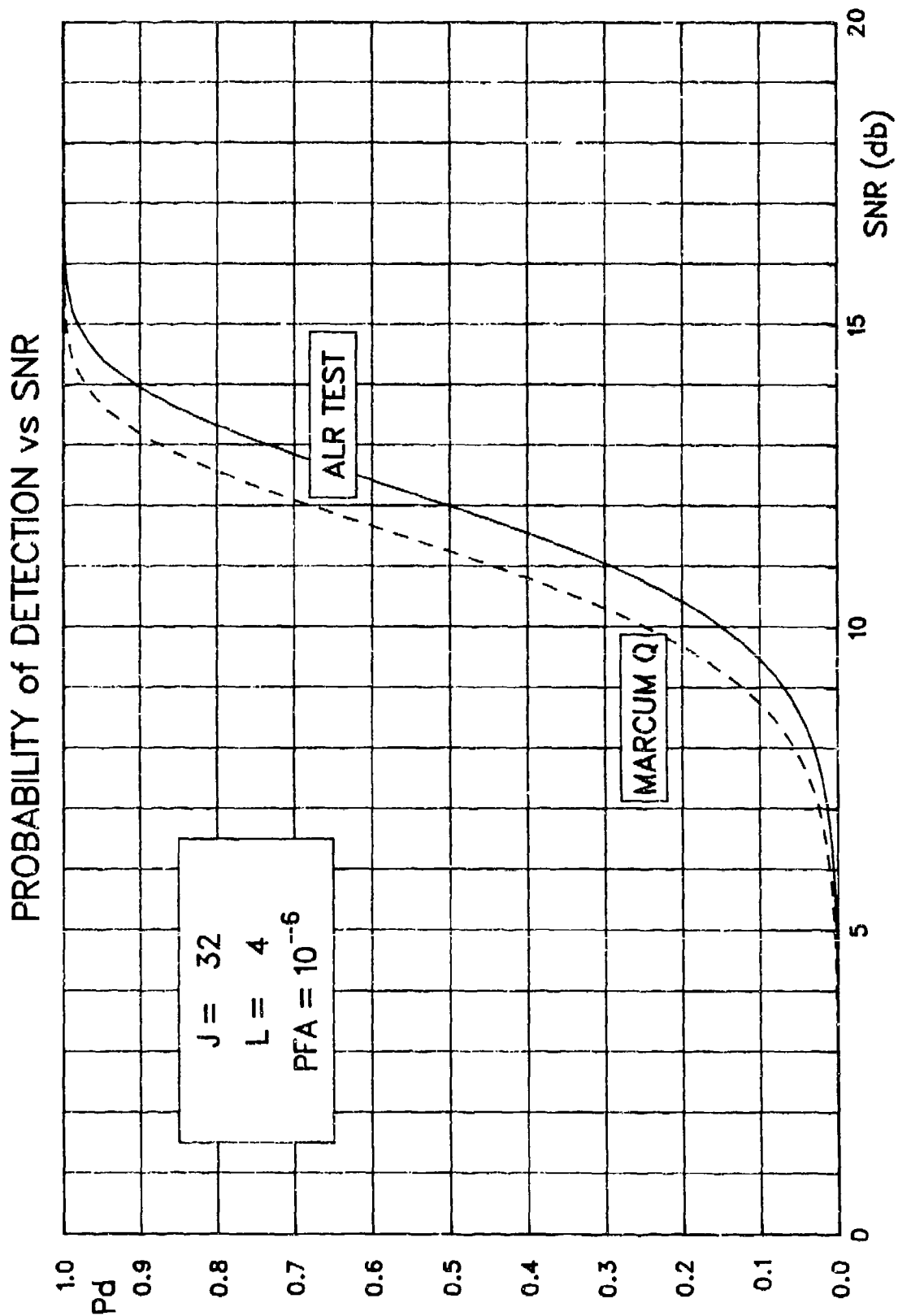
PROBABILITY of DETECTION vs SNR

J = 32
L = 4
PFA = 10^{-6}

ALR TEST

MARCUM Q

Pd

SNR (db)

Fig. II-2 Probability of Detection vs SNR

II-46

PROBABILITY of DETECTION vs SNR

J = 128
L = 4
PFA = 10^{-6}

ALR TEST

MARCUM Q

Pd: 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.0

SNR (db): 0, 5, 10, 15, 20

Fig. II-3 Probability of Detection vs SNR

PROBABILITY of DETECTION vs SNR

J = 32
L = 10
$PFA = 10^{-6}$

ALR TEST

MARCUM Q

Pd

Fig. II-4 Probability of Detection vs SNR

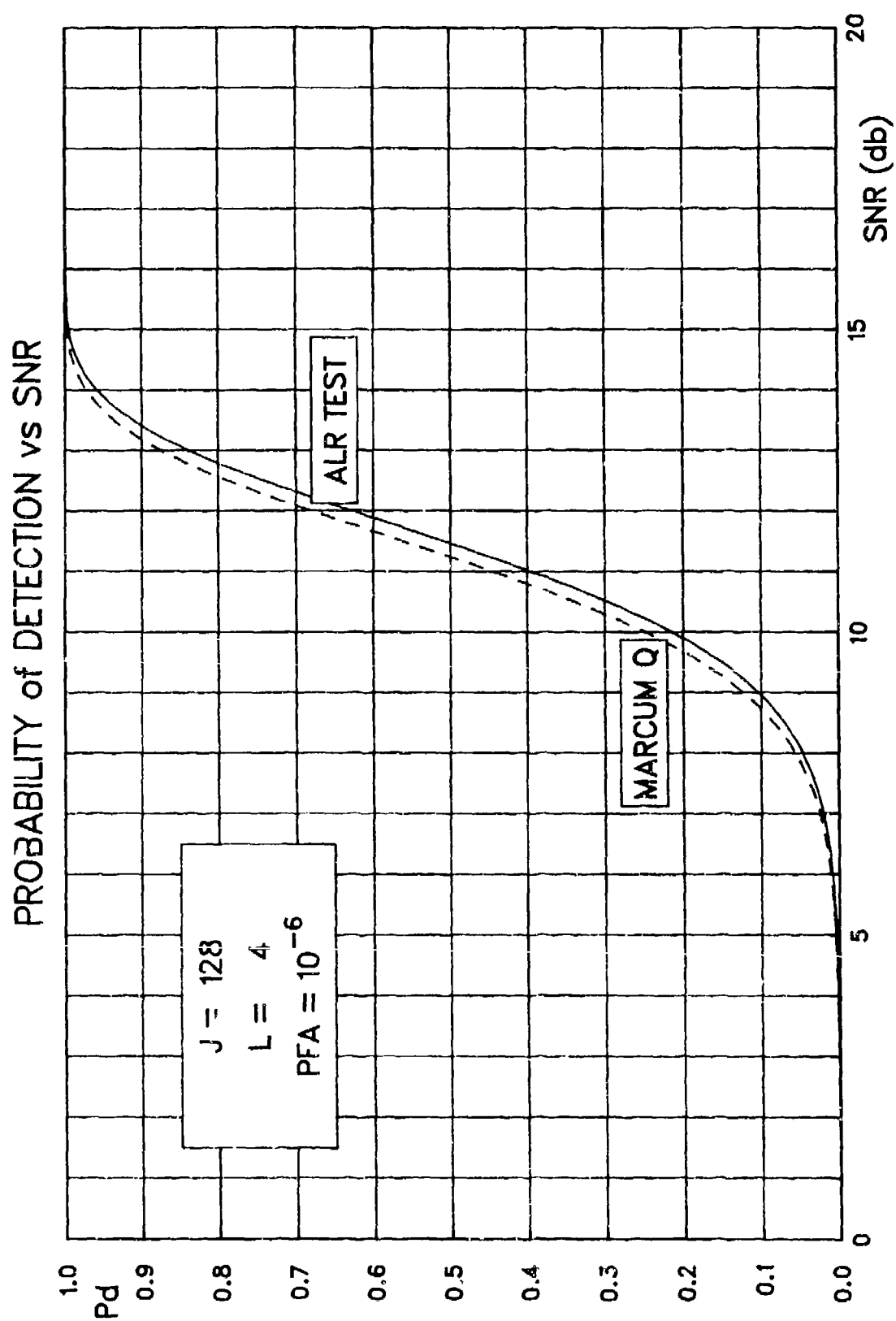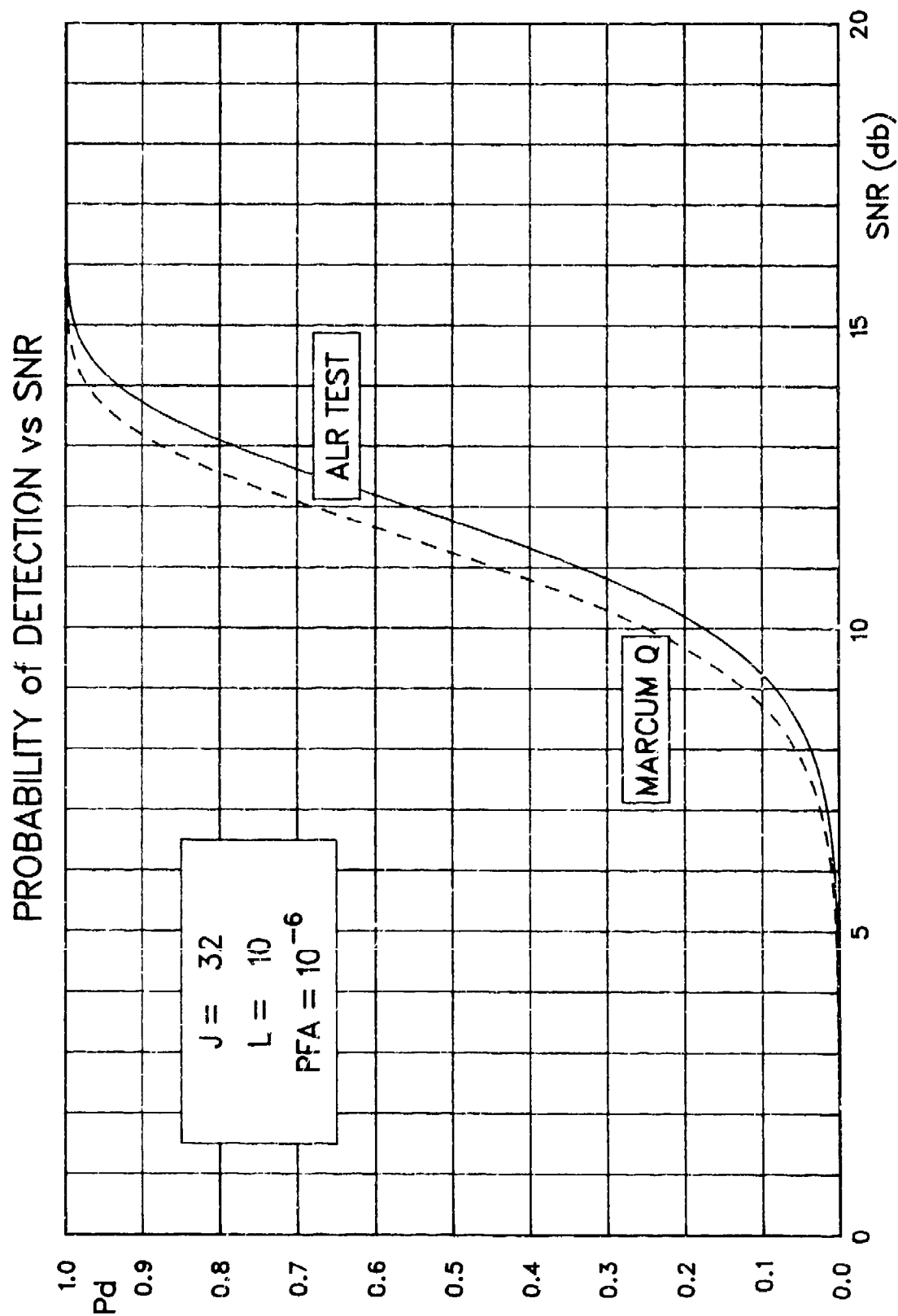## PROBABILITY of DETECTION vs SNR

J = 128
L = 10
PFA = 10^{-6}

ALR TEST

MARCUM Q

Fig. II-5 Probability of Detection vs SNR

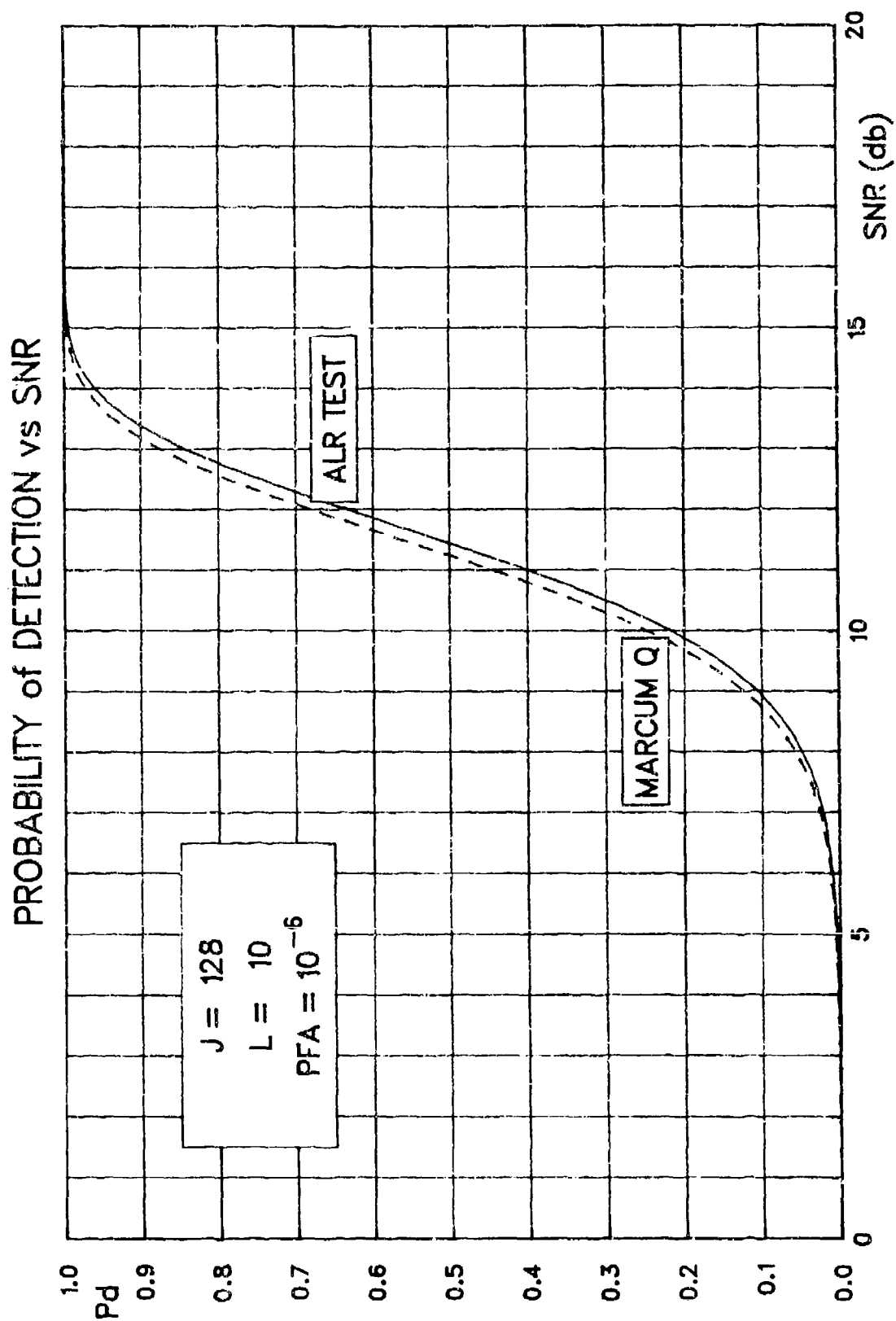Fig. II-6 Probability of Detection vs SNR

PROBABILITY of DETECTION vs SNR

J = 128
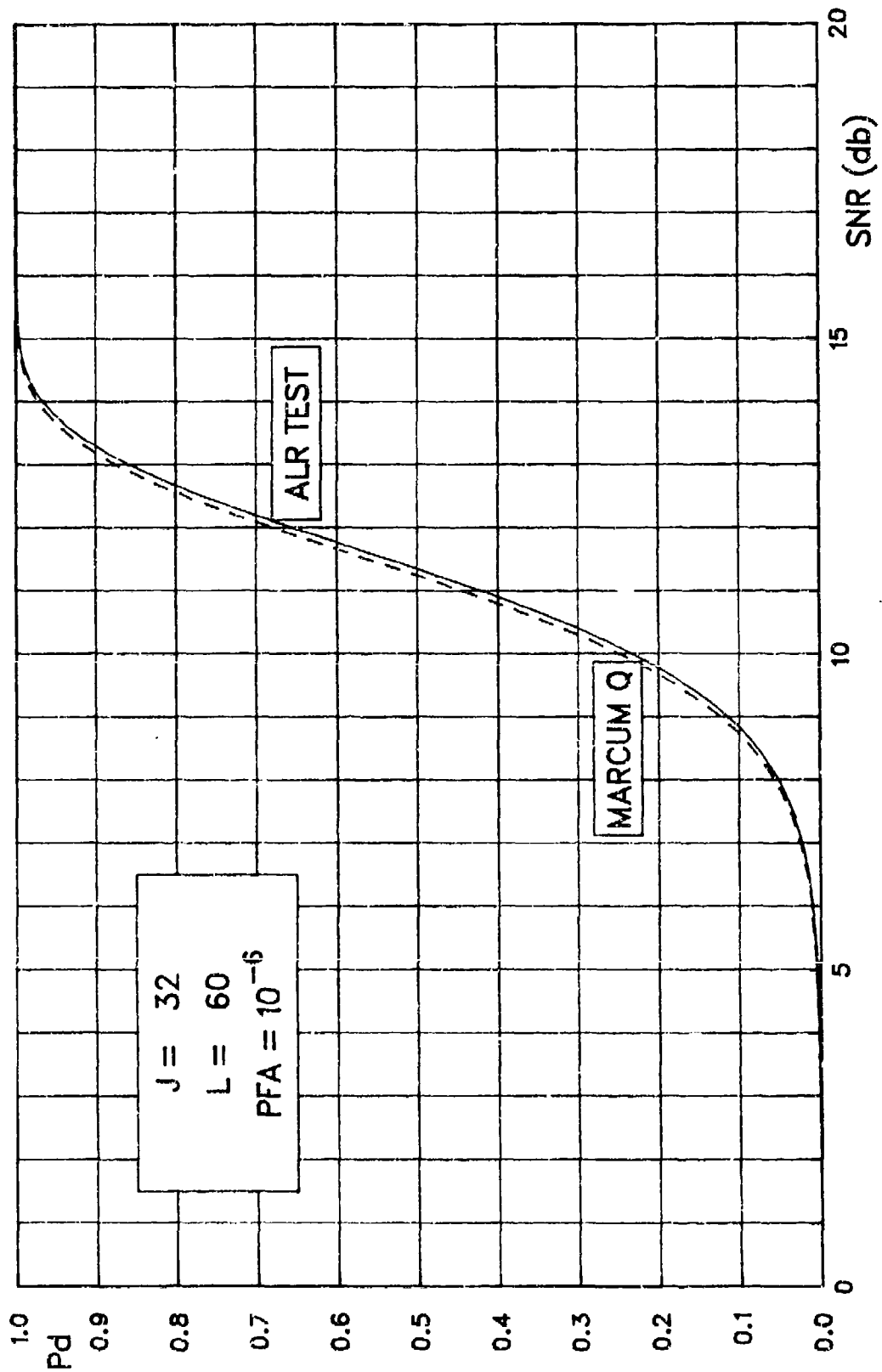L = 60
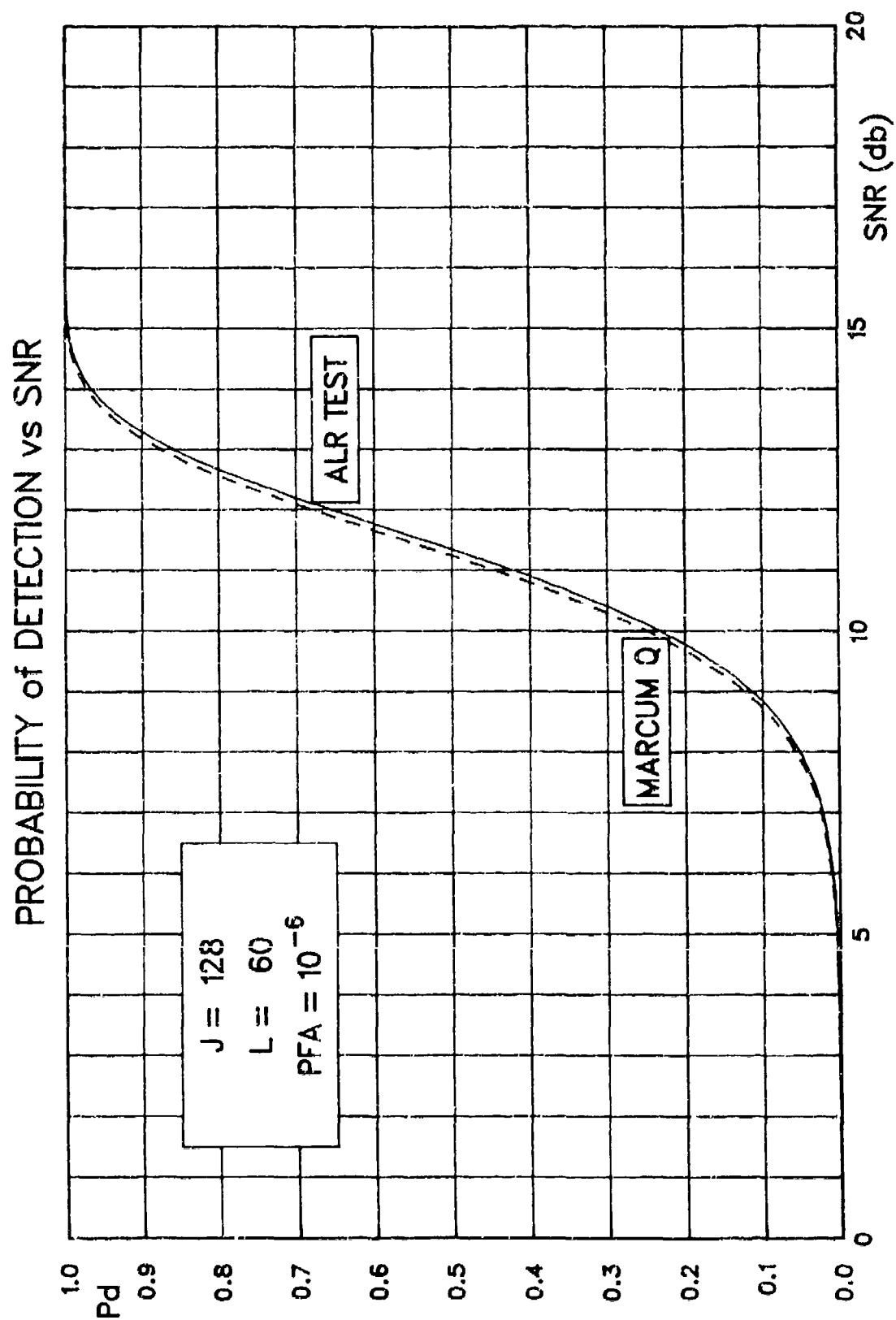PFA = 10^{-6}

ALR TEST

MARCUM Q

Fig. II-7 Probability of Detection vs SNR

II-51

L, and a SNR loss which is essentially that predicted by the Beta
distribution for a single pulse.

(6)   SUMMARY

It has been the intention of this two-part study to discuss the problem of radar operation in non-stationary interference, from the point of view of final target detection. Those aspects of the radar design which make it possible to achieve a satisfactory degree of interference rejection are not analyzed here, but are taken as given. We refer here to the provision of an adequate number of auxiliary RF channels and the means to produce precisely controlled weighted sums of their outputs. The choice of these weights is often discussed in terms of interference rejection, or nulling, on a pulse by pulse basis. By viewing the problem as one of target detection, utilizing the returns from a sequence of pulses which form a coherent processing interval (CPI), we have obtained a more complete decision algorithm. This algorithm contains both the rule for choosing the auxiliary weights and the procedure for combining the weighted outputs for each pulse to form an integrated resultant for target declaration.

Although the total interference is modeled as Gaussian noise, the correlation properties of this interference are presumed to be unknown. Detection is only possible if these correlation properties can be estimated, and for this purpose we have assumed the availability of other data, namely the outputs of adjacent range gates, which are taken to be signal free. The

important additional assumption is made that the interference
present in these other data is independent of, but statistically
identical to that of the main range gate in which target
detection is attempted. They can therefore provide a valid data
base from which the covariance matrices of the interference can
be estimated.

In this analysis, the available input data has been
characterized statistically and the methods of statistical
decision theory have been applied to derive a detection
procedure, based on the totality of original inputs. The
resulting procedure can be considered to have three components,
namely covariance estimation, interference rejection and coherent
integration. The first two components of this algorithm are in
agreement with standard procedures. Covariance estimation is
accomplished by means of the sample covariance matrix of the
secondary (target-free) data, and interference rejection is
performed by sample matrix inversion and the application of the
corresponding colored noise matched filter weights. The third
component of the decision procedure is a form of weighted
coherent integration of all the pulses of the CPI, in which the
weights are dependent on estimates of the residual noise level
after interference rejection on each individual pulse. It was
this aspect of the problem that was of most interest at the
outset, although it is reassuring to have the eminently

reasonable and conventional noise estimation and interference rejection procedures appear as derived results from the theory.

In Part I the interference was allowed to have arbitrary correlation, not only among the multiple RF input channels, but also among all the pulses of the CPI. Because of the great number of unknowns this entails, the number of secondary data inputs required for noise estimation must be very great. On the other hand, due to the generality of the formulation, an exact decision rule was obtained, together with an exact evaluation of its performance. The latter was given in terms of probability of detection and false alarm, as functions of the system parameters. Since no special structure was assumed by which the data from separate pulses could be distinguished, all these data were, in effect, lumped into one input vector in which a target is sought, and a set of secondary target-free vectors for noise estimation. Thus, by a suitable re-interpretation, Part I can also be said to deal with the problem of target detection based on a single pulse.

In Part II, the multiple pulse problem is specifically treated, with the additional and reasonable assumption that the interference is independent from pulse to pulse. The correlation properties of this interference are otherwise totally unknown, and must be estimated from the target-free inputs. The number of these secondary inputs needed to assure adequate performance is

now much smaller, being related to the number of RF channels

themselves, and not to the product of this number and the number

of pulses, as in the former case. On the other hand, the analysis

is inherently more difficult, and two distinct approximations had

to be made to obtain useful results. The first was an

approximation to the decision rule itself, which seems to be

justified not only by the quantitative arguments made in its

selection, but also by the reasonable character of the resulting

decision rule, the approximate likelihood ratio (ALR) test. The

other approximation was made in the performance analysis, as a

result of which these results are somewhat incomplete. The

limitations of this analysis were discussed in section 5, where

the insight gained from Part I was used to assess their probable

impact.

The general conclusions of the study are best expressed in

terms of the performance of an idealized radar which uses the

same primary inputs, but has the advantage of knowing the

correlation properties of the total noise. This radar can

therefore dispense with the secondary inputs of the real system.

The decision procedure for the idealized case, derived in Section

2, consists of interference rejection on each pulse, followed by

a weighted coherent integration of the pulses of the CPI.

Interference rejection is based on the known noise covariance

matrices, and the integrator weights are inversely proportional

to the residual noise levels after this portion of the
processing. This detector could be described as a conventional
nulling processor applied to each pulse, followed by a colored
noise matched filter which performs integration over the CPI. If
the external interference is nulled well below the internal
noise, the integrator weights will be equal, and the second
portion of the processor becomes a simple integrator. The
advantage of using weighted integration over the CPI, compared to
the use of constant weights, will depend on the residual noise
levels and their variation from pulse to pulse. No general rule
can be given for the performance gain in this case, but it will
generally represent an improvement, relative to the use of an
incorrect set of weights.

In the actual radar processor, noise estimates are used,
both in the nulling of individual pulses and in the establishment
of the final detection threshold. The use of estimated covariance
matrices in the nulling process results in a loss of signal to
noise ratio on each pulse. This loss has been shown to be
statistically identical to that derived by Reed, Mallett and
Brennan, in a well-known paper[3]. The relationship between
this SNR loss and the number of secondary inputs (range gates in
our case) is well understood, on the basis of the Beta
distribution to which these losses are subject. The second use of
noise estimation relates to the normalization of weights for the

coherent integrator, and the final detection threshold itself. The threshold estimation provides an approximately CFAR detection test, and also leads to a CFAR loss in performance. Because of the approximations referred to already, the CFAR loss is not accurately assesed by our results, but arguments have been given that it should not be large. It has also been shown that this loss should be approximated by the loss of a simple linear CFAR detector, whose threshold is based on K-M+1 noise samples, where K is the number of secondary inputs and M the number of RF channels of the actual system.

APPENDIX   EVALUATION OF THE APPROXIMATE DETECTION PROBABILITY

In Section 5 an expression was derived for the approximate probability of detection, conditioned on given values of the loss factors. This probability was expressed as the contour integral

$$\text{Prob}(\xi' > 0) \equiv P_D(a) = \frac{1}{2\pi i} \int_{-\infty - i\varepsilon}^{\infty - i\varepsilon} \Phi(\lambda) \frac{d\lambda}{\lambda}$$

where

$$\Phi(\lambda) = e^{-a} \frac{e^{-i(\ell_0 - 1)L\lambda + \frac{a}{1 - i\lambda}}}{(1 - i\lambda)[1 + i(\ell_0 - 1)\lambda]^{J-1}} \quad .$$

The detection probability depends on the conditioning only through the signal parameter, a, defined in Section 5.

For $a > 0$, the integrand has an essential singularity when λ equals $-i$, a simple pole at the origin and a pole of order $J-1$ at $i/(\ell_0 - 1)$. The path of integration is completed in the lower half plane and then shrunk to a small circle about the essential singularity. We can therefore write

$$P_D(a) = -\frac{1}{2\pi i} \int \Phi(\lambda) \frac{d\lambda}{\lambda}$$

where the contour is a small circle, now enclosing the point $-i$ in a positive sense.

We make the definition

$$y \equiv (\ell_0 - 1)L$$

and the change of variable

$$\lambda \equiv -i(1-t) \ ,$$

which brings the singularity to the origin. The result is

$$P_D(a) = \frac{e^{-a-y}}{2\pi i} \int_{|t|=\epsilon} \frac{e^{yt+\frac{a}{t}}}{[\ell_0 - (\ell_0-1)t]^{J-1}} \frac{dt}{t(1-t)} \quad .$$

For a = 0, the singularity at the origin becomes a simple pole, and the residue easily yields the formula for PFA given in Section 5. In general, the integral will not yield a closed form answer, and a series expansion is required. The difficulty is caused by the factor exp(yt), which causes another essential singularity at infinity.

We make the following expansion

$$\frac{1}{[\ell_0 - (\ell_0-1)t]^{J-1}} = \left(\frac{1}{\ell_0}\right)^{J-1} \sum_{m=0}^{\infty} \binom{J+m-2}{m}\left(\frac{\ell_0-1}{\ell_0}\right)^m t^m$$

whose convergence is assured by the fact that the magnitude of t is constant and arbitrarily small on the contour. Then we find

$$P_D(a) = \left(\frac{1}{\ell_0}\right)^{J-1} \sum_{m=0}^{\infty} \binom{J+m-2}{m}\left(\frac{\ell_0-1}{\ell_0}\right)^m \times$$

$$\times \frac{e^{-a-y}}{2\pi i} \int_{|t|=\epsilon} e^{yt+\frac{a}{t}} \frac{t^{m-1}dt}{1-t} \quad .$$

The contour integrals now remaining have been discussed in detail in Ref. 1, in relation to the Marcum Q-function. Let

$$f_N(u,a) \equiv e^{-u-a}\left(\frac{u}{a}\right)^{\frac{N-1}{2}} I_{N-1}(2\sqrt{au})$$

and

$$F_N(y,a) \equiv \int_y^\infty f_N(u,a)\,du \ .$$

Then F_N is equivalent to Marcum's Q-function, the probability of detection after non-coherent integration of N hits. The parameter y is the normalized threshold and a is the normalized total (i.e. integrated) signal to noise ratio.

In Ref. 1 it is shown that

$$\frac{e^{-a-y}}{2\pi i} \int_{|t|=\varepsilon} e^{yt+\frac{a}{t}} \frac{t^m dt}{1-t} = \begin{cases} 1 - F_{m+1}(a,y) & ; m \geq 0 \\ F_{-m}(y,a) & ; m < 0 \end{cases}$$

where the evaluation for non-negative m has been made by replacing the variable of integration, t, by $1/t$. Using this result, we obtain

$$P_D(a) = \left(\frac{1}{\ell_0}\right)^{J-1} F_1(y,a)$$

$$+ \left(\frac{1}{\ell_0}\right)^{J-1} \sum_{m=1}^{\infty} \binom{J+m-2}{m}\left(\frac{\ell_0-1}{\ell_0}\right)^m [1 - F_m(a,y)] \ .$$

When a equals zero, this formula reduces properly, since

$$F_N(0,y) = 1 \ ,$$

and

$$F_N(y,0) = e^{-y} \ .$$

The resulting expression for the probability of false alarm,

$$PFA = \left(\frac{1}{\ell_0}\right)^{J-1} e^{-(\ell_0-1)L}$$

is easliy solved for ℓ_0, given PFA, by the Newton-Raphson method.

For J = 1, the detection probability is simply

$$P_D(a) = F_1(y,a)$$

since all the binomial coefficients vanish in this case (no expansion was necessary for J = 1). This formula describes the detection performance of a simple coherent integrator, with total SNR equal to a and fixed threshold, and we see that the CFAR loss is. not accounted for, an expected consequence of the use of our simplifying approximation.

For other values of J, and in particular for large J, this formula looks quite different, since $(\ell_0)^{-(J-1)}$ will be a small factor, and many terms of the series will contribute. As the results presented in Section 5 show, however, our formula produces a rather conventional PD vs SNR curve for a wide range of values of the parameters, L and J.

Numerical results are obtained by truncation of the series and the use of a truncation bound, which will now be derived. We write the series in the form

$$P_D(a) = \left(\frac{1}{\ell_0}\right)^{J-1}\left\{F_1(y,a) + \sum_{m=1}^{M-1} T_m\right\} + \varepsilon_M$$

where

$$T_m \equiv \binom{J+m-2}{m}\left(\frac{\ell_0-1}{\ell_0}\right)^m[1 - F_m(a,y)]$$

is the general term of the series and ε_M is the error of
truncation at the Mth term:

$$\varepsilon_M = \left(\frac{1}{\ell_0}\right)^{J-1} \sum_{m=M}^{\infty} T_m \; .$$

We now make use of the identity

$$F_N(x,y) = F_{N-1}(x,y) + f_N(x,y) \; ,$$

which follows easily from the integral representation given
above, and which implies that the sequence $F_N(a,y)$ increases
monotonically with N, since the PDF f_N is necessarily
non-negative. The factors

$$[1 - F_m(a,y)] \; .$$

then decrease monotonically, and we shall use this property to
write

$$\varepsilon_M < \left(\frac{1}{\ell_0}\right)^{J-1} [1 - F_M(a,y)] \sum_{m=M}^{\infty} \binom{J+m-2}{m}\left(\frac{\ell_0-1}{\ell_0}\right)^m \; .$$

After replacing the index m by M+m, we find that the
inequality becomes

$$\varepsilon_M < \left(\frac{1}{\ell_0}\right)^{J-1} [1 - F_M(a,y)]\left(\frac{\ell_0-1}{\ell_0}\right)^M \binom{J+M-2}{M} S \; ,$$

where S stands for the series

$$S \equiv \sum_{m=0}^{\infty} \frac{(J+M+m-2)! M!}{(J+M-2)!(M+m)!}\left(\frac{\ell_0-1}{\ell_0}\right)^m \; .$$

But

$$S = \sum_{m=0}^{\infty} \binom{J+M+m-2}{m} \frac{M!\,m!}{(M+m)!} \left(\frac{\ell_0-1}{\ell_0} \right)^m$$

$$< \sum_{m=0}^{\infty} \binom{J+M+m-2}{m} \left(\frac{\ell_0-1}{\ell_0} \right)^m$$

$$= \left(1 - \frac{\ell_0-1}{\ell_0} \right)^{-(J+M-1)} = \ell_0^{J+M-1}$$

and hence

$$\varepsilon_M < [1 - F_M(a,y)] \binom{J+M-2}{M} (\ell_0-1)^M = \ell_0^M T_M \ .$$

The truncation bound is therefore easily computed along with the general term of the series, using recursion for the binomials and simple powers, and Shnidman's algorithm[2] for the Marcum functions. This procedure has been used to produce the figures of Section 5.

REFERENCES

1. E.J. Kelly, "Finite-Sum Expressions for Signal Detection Probabilities," Technical Report 566, Lincoln Laboratory, M.I.T., (20 May 1981) DTIC-AD-A102143.

2. D.A. Shnidman, 'Efficient Evaluation of Probabilities of Detection and the Generalized Q-Function,' IEEE Trans. Inf. Theory IT-22, 746, (1976).

3. I.S. Reed, J.D. Mallett and L.E. Brennan, 'Rapid Convergence Rate in Adaptive Arrays,' IEEE Trans. Aerospace Electron. Systems, AES-10, 853 (1974).

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-85-197	**2. GOVT ACCESSION NO.** AD-A158810	**3. RECIPIENT'S CATALOG NUMBER**
4. TITLE *(and Subtitle)* Adaptive Detection in Non-Stationary Interference, Part I and Part II		**5. TYPE OF REPORT & PERIOD COVERED** Technical Report
		6. PERFORMING ORG. REPORT NUMBER Technical Report 724
7. AUTHOR(s) Edward J. Kelly		**8. CONTRACT OR GRANT NUMBER(s)** F19628-85-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, M.I.T. P.O. Box 73 Lexington, MA 02173-0073		**10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS** Program Element No.63250F
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Systems Command, USAF Andrews AFB Washington, DC 20331		**12. REPORT DATE** 25 June 1985
		13. NUMBER OF PAGES 148
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)* Electronic Systems Division Hanscom AFB, MA 01731		**15. SECURITY CLASS.** *(of this Report)* Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

adaptive detection	adaptive nulling
non-stationary interference	maximum likelihood

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The problem of radar target detection in a background of non-stationary external interference is considered from the point of view of statistical decision thory. A signal processing algorithm is derived which accepts the totality of inputs on which final decision is to be based, and performs both interference suppression and target detection. In Part I a general formulation is given, the probabilies of detection and false alarm are evaluated exactly and the performance of the test is illustrated numerically. In Part II, a more specific interference model is introduced and a decision rule is derived which approximates the exact likelihood ratio test for this case. Approximations for the detection and false alarm probabilities are also found for this case.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 Jan 73