MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

ADAPTIVE DETECTION AND PARAMETER ESTIMATION FOR MULTIDIMENSIONAL SIGNAL MODELS

E.J. KELLY
Group 61

K.M. FORSYTHE
Group 44

DTIC
COPY
INSPECTED
6

Accession For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist Avail and/or Special
A-1

TECHNICAL REPORT 848

19 APRIL 1989

Approved for public release; distribution is unlimited.

LEXINGTON MASSACHUSETTS

ABSTRACT

The problem of target detection and signal parameter estimation in a background of unknown interference is studied, using a multidimensional generalization of the signal models usually employed for radar, sonar, and similar applications. The required techniques of multivariate statistical analysis are developed and extensively used throughout the study, and the necessary mathematical background is provided in Appendices. Target detection performance is shown to be governed by a form of the Wilks' Lambda statistic, and a new method for its numerical evaluation is given which applies to the probability of false alarm of the detector. Signal parameter estimation is shown to be directly related to known techniques of adaptive nulling, and several new results relevant to adaptive nulling performance are obtained.

TABLE OF CONTENTS

1. INTRODUCTION AND PROBLEM FORMULATION

The basic physical model which motivates this study corresponds to an array of sensors of some kind, positioned in an arbitrary way in space, and providing inputs to a processor whose nature is the subject of the analysis. One "sample" is a set of outputs from this array, arranged as a vector. These samples may come directly from the elements of the array, or they may be the outputs from a beamforming network of some kind. We use complex variables to represent the data since we are concerned with signals which modulate a carrier. Then the real and imaginary parts of a complex quantity represent the in-phase and quadrature components of such a signal. The modifications required to deal with real data are generally straightforward.

The basic data set upon which a processor will operate is a collection of sample vectors, arranged as the columns of a rectangular data array. We do not wish to specify the physical arrangements in greater detail because the mathematical model itself is applicable to many diverse systems which may use adaptive processing of array outputs in the radar, optical, and acoustical fields, and so on. Indeed, the elements of the sample vectors could easily have a significance other than the direct outputs of some set of sensors. However, we wish to draw attention to certain basic assumptions made in our model which, in certain cases, will limit its relevance.

We model the data array as a set of Gaussian random variables, and the covariance structure of the model is used to characterize the "noise" component of the data, including both system noise and any random external interference. On the other hand, "signals" are considered to be more structured contributions to the input, and these are modeled by making appropriate assumptions about the mean values of the elements of the data array. The emphasis here is on the detection of these signals and the estimation of their parameters, and the most natural applications are to radar or active sonar, where coherent processing is possible due to the known form of the signals. In this study, a general linear model is used to represent signals.

Our strongest assumption concerning the covariance structure is a postulate of stationarity: the sample vectors are assumed to be statistically independent and to share a common covariance matrix. If the samples correspond to successive times, then this is stationarity in the usual sense. However, the concept can be applied in other ways. For example, in the radar case the samples may correspond to successive range bins; but the data may already have been subjected to some form of processing embracing a larger interval of time, such as Fourier transformation (Doppler processing) of the array outputs before the adaptive phase of the process in which we are interested.

Another strong assumption is that the covariance matrix of the sample vectors is completely unknown. The advantage of this assumption is that it makes the mathematics more tractable, and also leads to a decision rule for which the probability of false alarm is independent of the actual covariance structure of the interference. This is a highly desirable feature, much stronger than the usual constant-false-alarm-rate (CFAR) property in which the false alarm rate is independent of the *level* of the noise. The disadvantage of our model in this respect is that it includes no constraint on the structure of the covariance matrix, other than the obvious one of positivity. This generality results in a restriction on the signal parametrization to assure a meaningful decision rule, a point discussed more fully in Appendix 5. We now proceed to a detailed description of the model.

Let Z be a complex $N \times L$ data array whose elements are modeled as circular complex Gaussian random variables. The columns of Z (i.e., the sample vectors) are assumed to be independent and to share the covariance matrix Σ. This is expressed by the formula

$$\mathrm{Cov}(Z) = \Sigma \otimes I_L . \tag{1-1}$$

where $\otimes$ stands for the Kronecker product, and I_L is the $L \times L$ identity matrix. This notation is defined in Appendix 1, where several basic properties of random arrays needed in this analysis are derived. The more general problem, in which the matrix I_L is replaced by a given positive definite matrix in Equation (1-1), is easily transformed into the model used here by post-multiplication of the data array by a suitable "whitening" matrix.

The mean of Z is assumed to have the form

$$EZ = \sigma B \tau , \tag{1-2}$$

where σ $(N \times J)$ is a given array, B $(J \times M)$ is an array of signal amplitude parameters, and τ $(M \times L)$ is also a given array. The fixed arrays σ and τ describe the assumed signal structure, as will be illustrated by examples. It is further postulated that the rank of σ is $J \leq N$, while that of τ is $M \leq L$. The mathematical setting we have just described is a generalization to complex random variables of a formulation often used in multivariate statistics to model quite different kinds of problems.

The basic task is to decide between two hypotheses concerning this statistical model: H_0, in which $B = 0$ and Σ is unknown; and H_1, in which both B and Σ are unknown. An unknown B matrix is completely arbitrary, but the covariance matrix must be positive definite, a property we denote by $\Sigma > 0$. The decision will be based on the Generalized Likelihood Ratio (GLR) principle,[1] and a GLR test is derived below. An estimate of the signal parameter array B is also of considerable interest, and the Maximum Likelihood (ML) estimator of B is automatically obtained in the derivation of the test statistic.

As noted earlier, the GLR test has the CFAR property in that its probability of false alarm (PFA) is completely independent of the actual covariance matrix of the data. Under the null hypothesis, the GLR test turns out to be a complex version of Wilks' Λ-statistic,[2] which is well known in the literature of multivariate statistical analysis. The PFA for this test will be evaluated by a technique of numerical integration in the complex plane. Complete results for the probability of detection (PD) are obtained only in special cases, but certain general properties of the PD will be established in Section 6.

The signal model introduced above allows considerable flexibility. The simplest case corresponds to $J = 1$ and $M = 1$, in which the signal array is represented as a single dyadic product. The σ array becomes a column vector of N elements, and τ is then a row vector of L elements. A specific example of this case, in which σ is a general vector and

$$\tau = [1,0,\ldots,0] \, ,$$

is discussed in References 3, 4, and 5. In this specialization, σ may represent a steering vector, as that concept is usually applied for adaptive arrays, and the model allows signal contributions in only one sample vector. In this special case, it is often convenient to normalize the σ and τ vectors to unity, which amounts to a simple redefinition of the parameter B.

A dual version, featuring a general τ vector and a σ vector of the form

$$\sigma = [1,0,\ldots,0]^T \, ,$$

is treated in Reference 6, on the basis of a totally different physical model. Although these special cases are really different versions of the same problem, and can be transformed into one another by a coordinate change of the kind discussed below, their analyses take rather different forms when they are carried out in the original coordinates.

In the general model, the σ array controls the distribution of signal contributions among the rows of the data array, while τ controls their appearance among the columns. If the components of the sample vectors represent the outputs from the sensors of an array, then σ will relate to the spatial character of the signals. Similarly, if the sample vectors themselves correspond to successive instants of time (snapshots), then τ will describe the temporal aspects of the signals.

Two other cases, which are natural duals of one another, are direct generalizations of the examples given above. In the first, σ is an arbitrary fixed array which satisfies the rank constraint mentioned earlier and τ is taken to be

$$\tau = [\, I_M \quad 0 \,] , \tag{1-3}$$

where I_M is the $M \times M$ identity, and the zero array here is $M \times (L - M)$ in dimension. With this model signals appear in the first M columns only, and each of these is represented as a different linear combination of the columns of σ. These latter columns determine a J-dimensional subspace of the N-dimensional complex vector space $\mathbb{C}^N$. This represents a generalization of the ordinary notion of an array steering vector. An example of such a model for signals is provided by multipath, which commonly occurs in seismic, acoustic, and "over the horizon" radar applications. For our model to be to be directly applicable, however, the multipath characteristics associated with a given principal signal component must be predictable, except for a set of complex amplitude factors. Another example is one in which the signal spatial structure is totally unknown, which corresponds to the special case $J = N$.

In the dual version, σ is taken to have the form

$$\sigma = \begin{bmatrix} I_J \\ 0 \end{bmatrix} \tag{1-4}$$

and τ is arbitrary (but full-rank), so that signals are described as row vectors, confined to the first J rows of Z. These row signals are independent linear combinations of the rows of τ which determine an M-dimensional subspace of the L-dimensional complex vector space $\mathbb{C}^L$. The characteristic feature of the general problem is the restriction of signals to subspaces in both the row and column directions, and the key to its analysis is the use of mathematical techniques which are adapted to this geometrical structure.

By changing coordinates, the general problem can be put in a "canonical form," which provides further insight into the postulated signal structure. We note first that the data array Z can be simultaneously pre- and post-multiplied by unitary matrices without changing the form of the problem. We write

$$Z_1 \equiv W_N Z W_L , \tag{1-5}$$

where W_N and W_L are unitary matrices whose dimensions are indicated by their subscripts. The new array is characterized by the properties

$$\mathrm{Cov}(Z_1) = W_N \Sigma W_N^H \otimes I_L \tag{1-6}$$

[see Appendix 1, Equation (A1-44)] and

$$E Z_1 = W_N \sigma B \tau W_L . \tag{1-7}$$

Since the matrix Σ is unknown and the unitary transformations are reversible, the new matrix

$$\Sigma_1 \equiv W_N \Sigma W_N^H$$

can be taken as the unknown covariance matrix of the columns of the new data array Z_1, instead of Σ; hence, the only real effect of this change of coordinates is on the signal components, as expressed by the mean of Z_1.

Now we introduce the singular value decompositions[7] of σ and τ:

$$\sigma = X_1 \begin{bmatrix} D_\sigma \\ 0 \end{bmatrix} X_2$$

$$\tau = Y_1 \left[D_\tau \;\; 0 \right] Y_2 ,$$

where D_σ and D_τ are diagonal matrices of dimension $J \times J$ and $M \times M$, respectively, and the arrays X_1, X_2, Y_1, and Y_2 are unitary. If we choose $W_N = X_1^H$, $W_L = Y_2^H$, and then set

$$B_1 = D_\sigma X_2 B Y_1 D_\tau .$$

we obtain the desired canonical form for the signal matrix:

$$EZ_1 = \begin{bmatrix} I_J \\ 0 \end{bmatrix} B_1 [\, I_M \;\; 0 \,] = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix} . \tag{1-8}$$

The new signal parameters now appear only in the upper left-hand corner of the data array, uniting the dual forms of the problem into one. In this formulation, the logic of our restrictions on the ranks of the original σ and τ arrays can be seen, since rank deficiencies in these arrays would lead to zero singular values in D_σ or D_τ. As a consequence, some of the signal parameters in the original B array would be redundant. The canonical form of the problem will not be used as a basis for analysis. It seems preferable to derive the decision rule in the original coordinates, since they will retain some physical meaning from the initial formulation of the problem. The canonical form then appears as a special case. In some situations, of course, a change of coordinates may be quite useful, and examples of this will be provided in Section 2.

We mentioned above that a certain limitation must be applied to the signal model in order to derive a GLR test. This takes the form of an inequality relating three of the dimensional parameters of the problem, namely:

$$L \geq M + N . \tag{1-9}$$

If this inequality is not satisfied, then the GLR procedure does not lead to a meaningful test statistic. In effect, there are too many free parameters in the model, and the likelihood function under the H_1 hypothesis can be made infinite. The point at which this occurs will be noted in passing, where the sufficiency of our condition will be evident. A proof of its necessity is given in Appendix 5.

In the decision problem formulated above, the null hypothesis (H_0) represents the complete absence of signal components in the data array. Following the example of multivariate statistics, a more general null hypothesis can be introduced in which a homogeneous linear constraint on the signal parameter array B replaces the original H_0. This constraint takes the form

$$\alpha B \gamma = 0 , \tag{1-10}$$

where α and γ are fixed arrays of dimension $r \times J$ and $M \times t$, respectively. The more general decision problem will be treated in Section 7, where the physical significance of this model will be discussed. Here, we mention only that it represents the presence of "nuisance signals," in addition to the desired signals in the data. These nuisance signals may be present under either hypothesis, but the desired signals are either present or totally absent. A decision rule will be found in this case whose PFA retains the CFAR property and is also completely insensitive to the presence or absence of these nuisance signals.

We have seen that the σ array determines a J-dimensional subspace of $\mathbb{C}^N$ which contains all permissible signal vectors. If $\mathbb{C}^N$ is decomposed into a subspace $\mathbf{A}$ and its orthogonal complement, where $\mathbf{A}$ contains this J-dimensional "signal subspace," then the covariance matrix Σ will automatically be partitioned into four components. Partitionings of this kind play a prominent role in our analysis. Suppose it is now assumed that the off-diagonal blocks of the partitioned covariance matrix vanish, thus adding some structure to the original interference model. This means that the interference in the subspace $\mathbf{A}$ is independent of that in its orthogonal complement, while the diagonal blocks of the covariance matrix are still considered to be unknown. In this model, the components of the data vectors which lie outside the subspace $\mathbf{A}$ play no role in signal detection or signal parameter estimation, and a GLR test for this problem disregards them completely. It is usually advantageous to reduce the dimensionality of the data model, if possible, and this kind of supplementary knowledge of the covariance structure will facilitate such a reduction. This is one way in which our model can be extended to allow some structure in the covariance of the interference.

The model can be generalized in other ways as well. For example, the arrays σ and τ may contain "internal" parameters which are also free under the H_1 hypothesis. To deal with these, we first obtain the GLR test statistic for fixed σ and τ, and then proceed to maximize it over the internal parameters. If an internal parameter takes on only discrete values, then estimation of this parameter is equivalent to carrying out a multiple-hypothesis test. Some examples of these generalizations will be mentioned briefly later, but discussion of them will be limited to the character of the GLR test itself.

The special case $J = N$, with $\sigma = I_N$ and $\alpha = I_N$, represents a complex version of the classical multivariate linear regression problem, which is thoroughly treated in several textbooks.[8,9,10] (The same name is often given to the special case in which $\gamma = I_M$.) In the literature, the regression problem is frequently discussed in terms of a data array which is the transpose of ours, so that its rows are independent instead of its columns. The analog of our general problem in terms of real variables also appears in the statistical literature[11,12] under other names, such as the generalized multivariate analysis of variance (GMANOVA). In statistics, the interest is usually centered on the null hypothesis, which corresponds to the PFA in our context. The detection problem, described in terms of complex variables, has recently been studied by Khatri and Rao.[13,14] The explicit results we have obtained concerning detection probability and the statistical character of the signal parameter estimates are specific to the class of problems we are modeling here, and many of these are new.

Our study is organized as follows. In Section 2, the GLR test itself is obtained, and the test statistic is expressed in several different forms. The basic statistical character of the test statistic is derived in Section 3, and the probability of false alarm is discussed in Section 4. In Section 5, the probability density function of the estimator of the signal amplitude parameter array is treated, and the probability of detection of the GLR test is discussed in Section 6. In these two sections, complete results are obtained only in the special cases $J=1$, any M, and $M=1$, any J. Certain properties of the solution of the general problem are also obtained. In Section 7, the generalization mentioned above is analyzed, with the result that this problem is reduced to the original one by means of straightforward transformations which eliminate the redundant data.

The Appendices are of two kinds: the first three contain mathematical results of a background nature, all used freely in the main portion of the text. The other Appendices contain special topics, separated out for readability. In Appendix 1, a collection of known results concerning matrices and random arrays is assembled. Perusal of this Appendix is recommended, since it contains a number of identities and lemmas indispensable to an understanding of the analysis. Appendix 2 is a collection of formulas for distributions related to the Gaussian in complex form. The corresponding real distributions are well known; some of the formulas derived here are less frequently seen. More background material is included in Appendix 3. The latter results relate mainly to integral properties of multivariate complex distributions, and they are less essential to the main development than are those of Appendix 1.

In Appendix 4, an alternate derivation of the GLR test is presented. The resulting test statistic is of a different form than those obtained in Section 2, but it is shown in this Appendix that it is statistically completely equivalent to the others. In Appendix 5, a proof of the necessity of the condition expressed in Equation (1-9) is provided. The probability of false alarm for the GLR test is evaluated explicitly in Section 4 only for certain special cases. In Appendix 6, a procedure is described by which numerical evaluation of this probability for arbitrary values of the parameters can be carried out. Finally, in Appendix 7, computational algorithms applicable to the GLR test in either of two forms are presented.

2. THE GENERALIZED LIKELIHOOD RATIO (GLR) TEST

This section contains a derivation of the GLR test for the original problem described in Section 1, in which the null hypothesis corresponds to a mean of zero for the data array. Background material on the complex multivariate Gaussian probability density function will be found in Appendix 1.

Under the null hypothesis, the joint probability density function (pdf) of the elements of the data array is given by

$$f_0(Z;\Sigma) = \frac{1}{\pi^{NL}|\Sigma|^L} e^{-\mathrm{Tr}(\Sigma^{-1}ZZ^H)} , \tag{2-1}$$

where Tr stands for trace, the superscript H represents Hermitian transpose, and the bars surrounding Σ denote its determinant. According to the model described in Section 1, the pdf under hypothesis H_1 is

$$f_1(Z;\Sigma,B) = \frac{1}{\pi^{NL}|\Sigma|^L} e^{-\mathrm{Tr}[\Sigma^{-1}(Z-\sigma B\tau)(Z-\sigma B\tau)^H]} . \tag{2-2}$$

Each of these density functions must be maximized over the unknown covariance matrix Σ, and, for the H_0 hypothesis, we obtain the ML estimator[15]

$$\hat{\Sigma}_0 = \frac{1}{L} ZZ^H . \tag{2-3}$$

The square array ZZ^H is subject to the complex Wishart distribution, with dimension N and L "complex" degrees of freedom. A discussion of complex Wishart matrices and some of their properties is given in Appendix 1. A derivation of the complex Wishart distribution itself will be found in Appendix 3. By the assumption expressed in Equation (1-9), we are assured that this matrix is positive definite with probability one. Substituting in Equation (2-1), we obtain

$$f_0(Z;\hat{\Sigma}_0) = [(e\pi)^N|\hat{\Sigma}_0|]^{-L} . \tag{2-4}$$

The analogous ML estimator of Σ under H_1 is, of course,

$$\hat{\Sigma}_1(B) = \frac{1}{L}(Z - \sigma B \tau)(Z - \sigma B \tau)^H , \tag{2-5}$$

which is a function of B. The final estimator of the covariance matrix under the H_1 hypothesis will be obtained when B is replaced by its estimator, which must still be derived. The formula analogous to Equation (2-4) is, of course,

$$f_1[Z;\hat{\Sigma}_1(B),B] = [(e\pi)^N |\hat{\Sigma}_1(B)|]^{-L} . \tag{2-6}$$

The GLR test statistic is, by definition,

$$\frac{\underset{\Sigma,B}{\text{Max}}\ f_1(Z;\Sigma,B)}{\underset{\Sigma}{\text{Max}}\ f_0(Z;\Sigma)} = \frac{\underset{B}{\text{Max}}\ f_1[Z;\hat{\Sigma}_1(B),B]}{f_0(Z;\hat{\Sigma}_0)} . \tag{2-7}$$

A test using this statistic is evidently equivalent to a test based on

$$\ell \equiv \frac{|\hat{\Sigma}_0|}{\underset{B}{\text{Min}}|\hat{\Sigma}_1|} ,$$

which is the L^{th} root of the GLR statistic, after substitution from Equations (2-4) and (2-6). Combining results, we obtain

$$\ell = \frac{|ZZ^H|}{\underset{B}{\text{Min}}\ |(Z - \sigma B \tau)(Z - \sigma B \tau)^H|} , \tag{2-8}$$

and H_1 is accepted if $\ell \geq \ell_0$.

We now introduce some tools which will allow us to manipulate the various arrays in a manner directly related to certain subspace projections associated with the given signal arrays, σ and τ. Beginning with τ, we note that the $M \times M$ array $\tau\tau^H$ is positive definite, since τ itself has rank M. Therefore, we can introduce a square-root array

$$(\tau\tau^H)^{1/2} > 0 ,$$

the notation indicating that a positive-definite square root has been chosen. Square roots of positive-definite matrices are used frequently in the ensuing work. An equivalent procedure would be to represent such matrices in terms of Cholesky factors. It should be emphasized that these factorizations always occur in intermediate stages of the analysis, and that none of the results will depend on which choice which has been made.

Using the above definition, we introduce the array

$$p \equiv (\tau\tau^H)^{-1/2}\,\tau \,. \tag{2-9}$$

If $M=1$, p reduces to a unit vector in the direction of the row vector τ. In general, the following properties follow directly from the definition:

$$\begin{aligned} p p^H &= 1_M \\ p^H p &= \tau^H (\tau\tau^H)^{-1}\,\tau \\ \tau &= (\tau\tau^H)^{1/2}\,p \,. \end{aligned} \tag{2-10}$$

The first of these equations shows that the rows of p are orthonormal, and the right side of the second equation (which is idempotent and Hermitian) is a standard form for a projection matrix[16] onto the subspace of $\mathbb{C}^L$ which is spanned by the rows of τ. This is the M-dimensional row space of τ, and the rows of p form a basis in it. The last equation is the analog of the representation of a vector as the product of its norm and an appropriate unit vector. When $M=L$, τ is invertible, p is unitary, and the last of Equations (2-10) is a polar decomposition of τ. It is characteristic of our approach that basis arrays for subspaces are used directly, rather than the projection operators themselves, to carry out the analysis.

The subspace of $\mathbb{C}^L$ which is orthogonal to the space spanned by p is of dimension $L-M$, and we can introduce an orthonormal set of $L-M$ row vectors to serve as a basis for it in many ways. Let q be an $(L-M)\times L$ array whose rows form such a basis. The relations

$$\begin{aligned} q q^H &= 1_{L-M} \\ q\,p^H &= 0 \end{aligned} \tag{2-11}$$

express these properties, and p and q together will form a unitary matrix of dimension $L \times L$:

$$\begin{bmatrix} p \\ q \end{bmatrix} = U_L . \tag{2-12}$$

The unitary property of U_L contains the orthonormality rules already given, and also the relation

$$p^H p + q^H q = I_L , \tag{2-13}$$

which expresses the fact that the rows of p and q together span $\mathbb{C}^L$.

If we multiply Z by I_L on the right and make use of Equation (2-13), we obtain the decomposition

$$Z = Z_p p + Z_q q = [Z_p \; Z_q] \begin{bmatrix} p \\ q \end{bmatrix} , \tag{2-14}$$

where the "components" of Z are defined by the equations

$$Z_p \equiv Z p^H$$

$$Z_q \equiv Z q^H . \tag{2-15}$$

Note that Z_p has dimension $N \times M$, while Z_q is an $N \times (L-M)$ array. This decomposition may be introduced in an equivalent way by writing

$$Z\, U_L^H = Z\, [p^H \; q^H] = [Z_p \; Z_q] . \tag{2-16}$$

which shows that the components of Z are formed by first rotating the coordinates in $\mathbb{C}^L$ (by means of the unitary transformation) and then partitioning it into two subspaces.

The complex vector space $\mathbb{C}^N$ is also decomposed, based on the structure of the σ array. Since σ has rank J, we can introduce the positive-definite square-root matrix

$$(\sigma^H \sigma)^{1/2} > 0 ,$$

and the corresponding array

$$e \equiv \sigma\,(\sigma^H \sigma)^{-1/2} \; . \tag{2-17}$$

The properties

$$e^H e = I_J$$

$$e\,e^H = \sigma\,(\sigma^H \sigma)^{-1}\,\sigma^H$$

$$\sigma = e\,(\sigma^H \sigma)^{1/2} \tag{2-18}$$

then follow directly from the definitions. The e array forms a basis for the J-dimensional subspace of $\mathbb{C}^N$ spanned by the columns of σ (the column space of σ). The second of Equations (2-18) contains a projection matrix which projects onto this column space.

Next, we introduce a basis in the $(N - J)$-dimensional subspace orthogonal to the span of e. These new vectors will form the columns of an array of dimension $N \times (N - J)$ which will be called f, and which satisfies the orthonormality relations

$$f^H f = I_{N-J}$$

$$f^H e = 0 \; . \tag{2-19}$$

The unit arrays e and f together form another unitary matrix, this time of dimension $N \times N$, as follows

$$[\,e \quad f\,] = U_N \; , \tag{2-20}$$

and the analog of Equation (2-13) is then

$$e\,e^H + f\,f^H = I_N \; . \tag{2-21}$$

Using this apparatus, we can express the signal model in terms of e and p, writing

$$E Z = \sigma B \tau = e b p \; . \tag{2-22}$$

where b is defined by

$$b \equiv (\sigma^H \sigma)^{1/2} B (\tau \tau^H)^{1/2} . \tag{2-23}$$

We now work with b as the array of unknown signal amplitude parameters, returning to B only at the end of the derivation. In terms of the new quantities, Equation (2-5) can be written

$$\hat{\Sigma}_1(b) = \frac{1}{L}(Z - ebp)(Z - ebp)^H , \tag{2-24}$$

and Equation (2-8) is the same as

$$\ell = \frac{|ZZ^H|}{\underset{b}{\text{Min}} |(Z - ebp)(Z - ebp)^H|} . \tag{2-25}$$

The denominator of this equation is now written

$$\underset{b}{\text{Min}} |F(b)| ,$$

where F(b) is given by

$$F(b) \equiv (Z - ebp)(Z - ebp)^H$$

$$= ebb^H e^H - ebZ_p^H - Z_p b^H e^H + ZZ^H . \tag{2-26}$$

In the second line we have used the new definitions and also the first of Equations (2-10). It follows directly from Equation (2-14) that

$$ZZ^H = Z_p Z_p^H + Z_q Z_q^H . \tag{2-27}$$

and, therefore, we can write

$$F(b) = (eb - Z_p)(eb - Z_p)^H + S , \tag{2-28}$$

in which we have introduced the new quantity

$$S \equiv Z_q Z_q^H . \tag{2-29}$$

Like ZZ^H, the S array is subject to a complex Wishart distribution of dimension N, but this time with $L-M$ complex degrees of freedom, in accordance with the dimensionality of Z_q. S is positive definite (with probability one) as a consequence of Equation (1-9), and is therefore an invertible matrix.

Returning to the minimization problem, we note the following fact:

$$\operatorname*{Min}_{u} |A_1 + u^H A_2 u| = |A_1| , \tag{2-30}$$

which is valid when A_1 and A_2 are positive-definite matrices (not necessarily of the same dimension) and u (in general rectangular) is an arbitrary array. To prove this result, we introduce positive-definite square roots of A_1 and A_2 and define

$$w \equiv A_2^{1/2} u A_1^{-1/2} .$$

Then

$$|A_1 + u^H A_2 u| = |A_1| \, |I + w^H w| ,$$

and the minimization can be carried out over w instead of u. But

$$\operatorname*{Min}_{w} |I + w^H w| = 1 , \tag{2-31}$$

because $w^H w$, being positive semidefinite, has non-negative eigenvalues. It follows that the determinant in Equation (2-31) is a product of eigenvalues, all of which are greater than or equal to unity. A unique minimum is therefore achieved for $w = 0$, which corresponds to $u = 0$ in the original notation.

In order to apply this result, we make use of an elementary determinant identity [Equation (A1-2) of Appendix 1] to write

$$|F(b)| = |S| \, |J(b)| ,$$

where J(b) is given by

$$J(b) \equiv I_M + (eb - Z_p)^H S^{-1}(eb - Z_p) . \quad (2\text{-}32)$$

It is clear that the second term on the right side of this expression for J(b) is positive semi-definite, hence J(b) itself is positive definite for any array b. Multiplying out the terms of Equation (2-32), we obtain

$$J(b) = I_M + b^H e^H S^{-1} eb - b^H e^H S^{-1} Z_p - Z_p^H S^{-1} eb + Z_p^H S^{-1} Z_p . \quad (2\text{-}33)$$

Since $S > 0$ and e has full rank, it follows that

$$e^H S^{-1} e > 0 .$$

This allows us to define the array

$$\hat{b} \equiv (e^H S^{-1} e)^{-1} e^H S^{-1} Z_p , \quad (2\text{-}34)$$

and, using this definition, we can "complete the square" with respect to b in Equation (2-33). The result is the formula

$$J(b) = I_M + Z_p^H S^{-1} Z_p - \hat{b}^H (e^H S^{-1} e) \hat{b} + (b-\hat{b})^H (e^H S^{-1} e)(b-\hat{b}) .$$

We have noted that J(b) is always positive definite; hence, in particular,

$$J(\hat{b}) > 0 .$$

Thus, we can apply Equation (2-30) to the determinant of J(b), since the conditions for its validity are satisfied. The result is

$$\underset{b}{\text{Min}} \, |J(b)| = |I_M + Z_p^H S^{-1} Z_p - \hat{b}^H (e^H S^{-1} e) \hat{b}| . \quad (2\text{-}35)$$

The ML estimator of b is therefore given by Equation (2-34), and the final estimator of covariance under hypothesis H_1 is given by

$$\hat{\Sigma}_1 = \frac{1}{L} \left[S + (Z_p - e\hat{b})(Z_p - e\hat{b})^H \right] . \quad (2\text{-}36)$$

It is an interesting fact that this estimator can be substituted for S in Equation (2-34), and the result is still a valid representation of the amplitude parameter array estimator. To see this, we first observe that

$$Z_p - e\hat{b} = Z_p - e(e^H S^{-1} e)^{-1} e^H S^{-1} Z_p ,$$

from which it follows that

$$e^H S^{-1}(Z_p - e\hat{b}) = 0 . \tag{2-37}$$

Next, we use the generalized Woodbury identity,[7] which is derived as Equation (A1-5) in Appendix 1, to write

$$(L\hat{\Sigma}_1)^{-1} = S^{-1} - S^{-1}(Z_p - e\hat{b})\left[I_M + (Z_p - e\hat{b})^H S^{-1}(Z_p - e\hat{b})\right]^{-1}(Z_p - e\hat{b})^H S^{-1} . \tag{2-38}$$

Using the Hermitian transpose of Equation (2-37), we see from Equation (2-38) that

$$(L\hat{\Sigma}_1)^{-1} e = S^{-1} e .$$

When this equivalence is used in Equation (2-34), the result is

$$\hat{b} = \left(e^H (\hat{\Sigma}_1)^{-1} e\right)^{-1} e^H (\hat{\Sigma}_1)^{-1} Z_p , \tag{2-39}$$

which is the desired form.

Returning to the derivation of the test statistic, we substitute from Equation (2-34) to obtain

$$Z_p^H S^{-1} Z_p - \hat{b}^H (e^H S^{-1} e)\hat{b} = Z_p^H P Z_p ,$$

where

$$P \equiv S^{-1} - S^{-1} e (e^H S^{-1} e)^{-1} e^H S^{-1} . \tag{2-40}$$

Combining these results and substituting in Equation (2-35), we obtain the desired minimization

$$\underset{b}{\mathrm{Min}}\, |F(b)| = |S|\, \underset{b}{\mathrm{Min}}\, |J(b)| = |S|\,|I_M + Z_p^H P Z_p| \,. \tag{2-41}$$

The numerator of Equation (2-25) can be developed in the form

$$|ZZ^H| = |Z_p Z_p^H + S| = |S|\,|I_M + Z_p^H S^{-1} Z_p| \,,$$

and then, finally, the GLR test statistic is obtained as a ratio of determinants:

$$\ell = \frac{|I_M + Z_p^H S^{-1} Z_p|}{|I_M + Z_p^H P Z_p|} \,. \tag{2-42}$$

In the special case described by Equation (1-3), where the signal contributions are confined to the first M columns of the data array, the decomposition of Z into the components Z_p and Z_q is simply a separation of columns into two groups, and formula (2-42) has a natural interpretation in this case. In Appendix 4 a derivation of the GLR test is carried out, by a variation of the technique used here, which leads to a result of quite different form than Equation (2-42), although completely equivalent to it. This other form is naturally suited to the dual special case, described by Equation (1-4), in which signals are confined to the first J rows of the data array.

Working back through the definitions, we obtain the relations

$$Z_p = Z\tau^H(\tau\tau^H)^{-1/2}$$

$$S = Zq^H q Z^H = Z[I_L - \tau^H(\tau\tau^H)^{-1}\tau]Z^H \tag{2-43}$$

and

$$P = S^{-1} - S^{-1}\sigma(\sigma^H S^{-1}\sigma)^{-1}\sigma^H S^{-1} \,. \tag{2-44}$$

With their help, the test statistic can be expressed directly in terms of quantities which appear in the original formulation of the problem. In particular, none of the arrays introduced as bases in the various subspaces appears in the final result. Formula (2-42) is a direct generalization of the GLR test obtained for the special case treated in References 3 and 4.

To facilitate comparison with these previously obtained results, the GLR test can be recast in a different form. If we make the definitions

$$D \equiv I_M + Z_p^H S^{-1} Z_p$$

$$G \equiv \sigma^H S^{-1} \sigma$$

$$A \equiv \sigma^H S^{-1} Z_p \,, \tag{2-45}$$

and also make use of Equation (2-44), we can write Equation (2-42) as

$$\ell = \frac{|D|}{|D - A^H G^{-1} A|} \,. \tag{2-46}$$

Since D is positive definite, we can multiply both numerator and denominator by $D^{-1/2}$, both on the right and on the left, and thus convert the test statistic to the form

$$\ell = \frac{1}{|I_M - \eta|} \,,$$

where η is given by

$$\eta \equiv D^{-1/2} A^H G^{-1} A \, D^{-1/2} \,.$$

If $M = 1$, η is a scalar, and the test statistic is simply

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

Moreover,

$$\eta = \frac{A^H G^{-1} A}{D} = \frac{Z_p^H S^{-1} \sigma \, (\sigma^H S^{-1} \sigma)^{-1} \, \sigma^H S^{-1} Z_p}{1 + Z_p^H S^{-1} Z_p} \,,$$

in this case.

On the other hand, if $J = 1$ (and M is unrestricted), then G is a scalar and we can apply identity (A1-3) of Appendix 1 to obtain

$$|I_M - \eta| = 1 - \eta' ,$$

where

$$\eta' \equiv \frac{A D^{-1} A^H}{G} = \frac{\sigma^H S^{-1} Z_p (I_M + Z_p^H S^{-1} Z_p)^{-1} Z_p^H S^{-1} \sigma}{\sigma^H S^{-1} \sigma} .$$

If $J = 1$ *and* $M = 1$, then η and η' coincide and the test becomes

$$\frac{|\sigma^H S^{-1} Z_p|^2}{(\sigma^H S^{-1} \sigma)\,(1 + Z_p^H S^{-1} Z_p)} \geq \frac{\ell_0 - 1}{\ell_0} , \tag{2-47}$$

which is the form obtained in Reference 3.

For general values of J and M, the A array introduced above can be expressed as

$$A = \sigma^H S^{-1} Z p^H = w^H Z p^H ,$$

where

$$w \equiv S^{-1} \sigma .$$

Post-multiplication of the data array by p^H corresponds to ordinary coherent integration of the elements of Z, in the row direction, using a set of matched filters determined by the τ array. Similarly, pre-multiplication by w^H corresponds to *adaptive* whitening and coherent integration in the column direction, by means of a "weight array" w, formed from the signal "steering array" σ and the S matrix. Except for a constant factor, the matrix S is a sample covariance matrix based on the signal-free vectors which comprise the array Z_q. We introduce the notation

$$(L - M)^{-1} S \equiv \hat{\Sigma}_q \tag{2-48}$$

for this estimator, indicating that it is formed from the Z_q component alone.

The ML estimator of the signal amplitude array B is recovered by the use of Equations (2-23), (2-34), and (2-43). The result is

$$\hat{B} = (\sigma^H S^{-1} \sigma)^{-1} \sigma^H S^{-1} Z \tau^H (\tau \tau^H)^{-1} . \tag{2-49}$$

This expression represents a direct generalization of a standard algorithm used for adaptive nulling. To illustrate this more explicitly, consider the case in which the τ array has the simple form expressed by Equation (1-3). This models a situation in which Z_p consists of the first M columns of the data array, representing the data vectors which may contain signals, while the others constitute the Z_q array. We can then write Equation (2-49) in the form

$$\hat{B} = (\sigma^H \hat{\Sigma}_q^{-1} \sigma)^{-1} \sigma^H \hat{\Sigma}_q^{-1} Z_p ,$$

which expresses the columns of the B estimator array as matrix products involving a "weight array" and the columns of Z_p. In this interpretation, the columns of the B estimator array represent the outputs of a generalized adaptive nulling processor whose inputs are the sample vectors which form the columns of Z_p. If $J = 1$, the weight array reduces to a weight vector, and the correspondence with the standard adaptive nulling technique, based on sample matrix inversion, is complete. In Section 5, the joint probability density of the elements of the B estimator array (which is a row vector in this case) will be obtained, and the relation to adaptive nulling will be pursued further.

For the special case $J = N$, the matrix σ is square and, by hypothesis, it has full rank. From Equations (2-18) we see that the array e is unitary under this assumption, and our formulas will simplify accordingly. In particular, the matrix P will vanish in this case, leaving only the numerator in Equation (2-42) for the GLR test. In addition, the estimator of the amplitude array, given by Equation (2-49), will assume the simple form

$$\hat{B} = \sigma^{-1} Z \tau^H (\tau \tau^H)^{-1} ,$$

when $J = N$. As noted in Section 1, the complex version of the multivariate linear regression problem (without the generalized null hypothesis) is characterized by $\sigma = I_N$; hence, our results are easily specialized to this problem.

In an extension of our model, of the type mentioned in Section 1, τ is allowed to contain a discrete internal parameter. In other words, τ is actually one of several given τ arrays, and the problem is to decide which of these arrays best describes the signal, if signal is actually deemed to be present. One can evaluate the GLR test statistic for each τ, and if the largest of these exceeds a threshold for signal detection, then use it to decide which signal was received.

A simple example, in which $M=1$, would arise if the sample vectors corresponded to regular instants of time and the parametrized τ arrays, each a row vector, described different possible temporal sequences, such as those corresponding to the Doppler phase variations of a moving radar target. One could test for one value of the Doppler parameter at a time, using the remaining part of the data array, described by Z_q, for noise estimation via the matrix S. As noted earlier, the GLR test involves post-multiplication of the data array by p^H, and p is just a normalized version of the τ vector in this case; hence, this represents coherent integration in the ordinary sense.

The formation of a conventional "Doppler filter bank," based on L time samples, is equivalent to post-multiplication of the original data array by a suitable unitary matrix. The new τ vectors will then be unit vectors, each containing a single component equal to unity, and the rest all zero. Each of the multiple hypotheses in this case amounts to placing the signal in a different column of Z. This is an example of a situation in which a change of coordinates, mentioned in Section 1, is a natural thing to do.

Added insight into the significance of the GLR test statistic and the associated ML signal parameter estimator is gained by considering the simpler version of our problem in which the covariance matrix Σ is known. The hypotheses concerning the signal components remain the same. From Equations (2-1) and (2-2), together with Equations (2-22) and (2-26), it follows that the logarithm of the likelihood ratio for this problem is given by

$$\begin{aligned}\lambda(b) &= -\mathrm{Tr}\{\Sigma^{-1}[F(b) - F(0)]\} \\ &= -\mathrm{Tr}[\Sigma^{-1}(eb - Z_p)(eb - Z_p)^H - \Sigma^{-1}Z_pZ_p^H] \\ &= -\mathrm{Tr}[(eb - Z_p)^H\Sigma^{-1}(eb - Z_p) - Z_p^H\Sigma^{-1}Z_p] . \qquad (2\text{-}50)\end{aligned}$$

We define

$$\hat{b}_\Sigma \equiv (e^H\Sigma^{-1}e)^{-1} e^H\Sigma^{-1}Z_p . \qquad (2\text{-}51)$$

using the subscript to indicate that Σ is known, and complete the square in Equation (2-50). The result is

$$\lambda(b) = -\mathrm{Tr}\left[(b-\hat{b}_\Sigma)^H (e^H \Sigma^{-1} e)(b-\hat{b}_\Sigma) - \hat{b}_\Sigma^H (e^H \Sigma^{-1} e)\hat{b}_\Sigma\right] ,$$

which is clearly maximized by the choice

$$b = \hat{b}_\Sigma ,$$

thus establishing the ML estimator of b. This is, of course, the classical solution, expressed here in terms of the component Z_p. Formula (2-34) is a direct generalization of this result.

For the non-adaptive test statistic itself, we have

$$\lambda \equiv \underset{b}{\mathrm{Max}}\ \lambda(b) = \mathrm{Tr}\left[\hat{b}_\Sigma^H (e^H \Sigma^{-1} e)\hat{b}_\Sigma\right] , \tag{2-52}$$

or

$$\lambda = \mathrm{Tr}\left\{Z_p^H \Sigma^{-1} \sigma (\sigma^H \Sigma^{-1} \sigma)^{-1} \sigma^H \Sigma^{-1} Z_p\right\} . \tag{2-53}$$

These formulas will be developed further in Sections 3 and 5, and the relationship to the GLR test statistic for the general problem will be elucidated.

We close this section with the derivation of some alternative expressions for the GLR test statistic which exhibit the roles of the subspace projections in a rather nice way. To obtain the first of these forms, we apply identity (A1-2) of Appendix 1,

$$|G||D - A^H G^{-1} A| = |D||G - A D^{-1} A^H| , \tag{2-54}$$

to Equation (2-46), with the result

$$\ell = \frac{|G|}{|G - A D^{-1} A^H|} .$$

Eliminating the new definitions, we have

$$\ell = \frac{|\sigma^H S^{-1} \sigma|}{\left|\sigma^H [S^{-1} - S^{-1} Z_p (I_M + Z_p^H S^{-1} Z_p)^{-1} Z_p^H S^{-1}] \sigma \right|} . \tag{2-55}$$

Applying the generalized Woodbury identity [Equation (A1-5)] to the denominator of Equation (2-55), we obtain the desired result:

$$\ell = \frac{|\sigma^H S^{-1} \sigma|}{|\sigma^H (S + Z_p Z_p^H)^{-1} \sigma|} = \frac{|e^H S^{-1} e|}{|e^H (S + Z_p Z_p^H)^{-1} e|} . \tag{2-56}$$

Equivalent versions of this test statistic are:

$$\ell = \frac{|\sigma^H (Z_q Z_q^H)^{-1} \sigma|}{|\sigma^H (Z Z^H)^{-1} \sigma|} = \frac{|\sigma^H (Z Z^H - Z_p Z_p^H)^{-1} \sigma|}{|\sigma^H (Z Z^H)^{-1} \sigma|} . \tag{2-57}$$

Note that the second form above makes use of a sample covariance matrix based on the full data array Z.

Equation (2-57) is a generalization of a formula stated by Brillinger.[17] For the case $J = 1$, in which σ is a column vector, Equation (2-57) may be interpreted as the ratio of maximum-likelihood (Capon) spectral estimates,[18] in the direction of σ, using either all the data in the Z array or only its projection onto the orthogonal complement of the row space of τ.

The simple form which the GLR test assumes when $J = N$ is easily reproduced from Equation (2-56). Since σ is then square and non-singular, its determinant may be factored out of the numerator and the denominator of this ratio, with the result

$$\ell = \frac{|S + Z_p Z_p^H|}{|S|} = |I_M + Z_p^H S^{-1} Z_p| . \tag{2-58}$$

Equation (2-54) has been applied to obtain the final form, which is the same as that to which Equation (2-42) reduces when $J = N$. If the eigenvalues of the matrix $Z_p^H S^{-1} Z_p$ are called λ_m, then, obviously,

$$|I_M + Z_p^H S^{-1} Z_p| = \prod_{m=1}^{M} (1 + \lambda_m) . \tag{2-59}$$

If $M > N$, some of these eigenvalues will vanish since the corresponding matrix will not have full rank, but Equation (2-59) will remain valid.

A generalization of our basic problem will be mentioned briefly here, since Equation (2-57) is especially suitable to its analysis and a result very similar to Equation (2-59) can be obtained. In this model, everything is the same as already postulated, but the σ array is now allowed to be an arbitrary full- rank array of dimension $N \times J$. In the original model, the signals are drawn from the *given* J-dimensional subspace of $\mathbb{C}^N$ which is determined by the σ array. In the generalization, the signals are drawn from *any* subspace of dimension J. The structure imposed by τ, which controls the distribution of signals among the columns of the data array, is not changed.

A likelihood-ratio test for the new problem is evidently obtained by maximizing the statistic expressed by Equation (2-57) over the σ array, since the likelihood ratio itself is directly related to ℓ. Suppose that A_1 and A_2 are positive-definite matrices of order N. Then, it can be shown that

$$\operatorname*{Max}_{\sigma} \frac{|\sigma^H A_1 \sigma|}{|\sigma^H A_2 \sigma|} = \prod_{j=1}^{J} \mu_j . \tag{2-60}$$

where the maximization is carried out over all full-rank $N \times J$ arrays σ, and the μ_j are the eigenvalues of the matrix $A_1(A_2)^{-1}$, ordered from largest to least:

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_N .$$

For application to Equation (2-57), this matrix product is

$$(Z_q Z_q^H)^{-1} Z Z^H = S^{-1}(S + Z_p Z_p^H) = I_N + S^{-1} Z_p Z_p^H$$

But the matrices $S^{-1}Z_pZ_p^H$ and $Z_p^HS^{-1}Z_p$ share the same non-zero eigenvalues and, consequently, the product of the eigenvalues of $I_N + S^{-1}Z_pZ_p^H$ is the same as the product of the eigenvalues of $I_M + Z_p^HS^{-1}Z_p$. A proof of this result, and also of Equation (2-60), will be found in Appendix 1. The CLR test statistic for the generalized problem therefore takes the form

$$\operatorname*{Max}_{\sigma} \frac{|\sigma^H (Z_q Z_q^H)^{-1} \sigma|}{|\sigma^H (Z Z^H)^{-1} \sigma|} = \prod_{j=1}^{J} (1 + \lambda_j) . \tag{2-61}$$

If $J \geq M$ in the generalized model, the test co incides with that obtained for the special case of the original problem in which [illegible] understand this feature, it is useful to imagine the special form of the τ a[illegible] bed by formula (1-3), in which signals are confined to the first M columns o[illegible] a array. For $M=1$, the two models are equivalent ways of allowing the signal in [illegible] first column to be arbitrary, and the equality of the tests is obvious. For $M>1$, the models coincide only if the freedom conferred by the dimensionality of the subspace $\mathbb{C}^J$ (in the generalized problem) is sufficient to overcome the fact that the signals from the first M columns must lie in the same J-dimensional subspace.

Equation (2-57) was a convenient starting point for the problem generalization just discussed, because al' the dependence of the test statistic on the σ array appears in a simple and explicit way in this formula. An analogous expression, in which all the τ-dependence is exhibited in the same simple way, also can be obtained. This form will not contain the matrix S explicitly, since the formation of that matrix carries with it an implicit dependence on τ through the p and q arrays.

We begin with Equation (2-42), and rewrite it in the form

$$l = \frac{|(I_M + Z_p^H P Z_p)^{-1}|}{|(I_M + Z_p^H S^{-1} Z_p)^{-1}|} \, , \tag{2-62}$$

where P is the matrix defined in Equation (2-40). We use the generalized Woodbury identity, Equation (A1-5), to evaluate the matrix in the denominator:

$$(I_M + Z_p^H S^{-1} Z_p)^{-1} = I_M - Z_p^H (S + Z_p Z_p^H)^{-1} Z_p \, .$$

We make the definition

$$S_+ \equiv S + Z_p Z_p^H = Z Z^H \, , \tag{2-63}$$

thereby giving a name to a matrix which has already entered our previous form for the test statistic. S_+ is proportional to the sample covariance matrix based on all the data vectors which comprise the Z array. We make use of the first of Equations (2-10), together with the new definition, and write

$$(I_M + Z_p^H S^{-1} Z_p)^{-1} = I_M - Z_p^H S_+^{-1} Z_p = p(I_L - Z^H S_+^{-1} Z) p^H \, .$$

We can therefore write the test statistic in the form

$$\ell = \frac{|(I_M + Z_p^H P Z_p)^{-1}|}{|p(I_L - Z^H S_+^{-1} Z) p^H|} \qquad (2\text{-}64)$$

The denominator now has the desired structure, with all the p-dependence in the outer factors of a matrix product. The numerator, however, requires a little coercion. We introduce some temporary notation to simplify the writing, as follows:

$$G \equiv e^H S_+^{-1} e$$

$$F \equiv e^H S_+^{-1} Z_p$$

$$W \equiv I_M - Z_p^H S_+^{-1} Z_p \,. \qquad (2\text{-}65)$$

Next, we use the Woodbury formula again, this time to express the inverse of S in terms of S_+:

$$S^{-1} = (S_+ - Z_p Z_p^H)^{-1} = S_+^{-1} + S_+^{-1} Z_p W^{-1} Z_p^H S_+^{-1} \,. \qquad (2\text{-}66)$$

To evaluate the numerator of Equation (2-64), we require the following results, which are direct consequences of Equation (2-66) and the new definitions:

$$e^H S^{-1} e = G + F W^{-1} F^H$$

$$e^H S^{-1} Z_p = F + F W^{-1} (I_M - W) = F W^{-1} \,. \qquad (2\text{-}67)$$

We have already seen that

$$I_M + Z_p^H S^{-1} Z_p = (I_M - Z_p^H S_+^{-1} Z_p^H)^{-1} = W^{-1} \,.$$

Combining all these results, and recalling definition (2-40), we obtain

$$I_M + Z_p^H P Z_p = W^{-1} - W^{-1} F^H (G + F W^{-1} F^H)^{-1} F W^{-1}$$

$$= (W + F^H G^{-1} F)^{-1} \,,$$

again with the help of the indispensable Woodbury identity. We now substitute from definitions (2-65) and write

$$W + F^H G^{-1} F = I_M - Z_p^H Q Z_p = p(I_L - Z^H Q Z) p^H ,$$

where

$$Q \equiv S_+^{-1} - S_+^{-1} e (e^H S_+^{-1} e)^{-1} e^H S_+^{-1} . \tag{2-68}$$

The new matrix Q is closely analogous to P, but Q involves S_+ where P has S itself. Finally, we obtain the desired form

$$\ell = \frac{|p(I_L - Z^H Q Z) p^H|}{|p(I_L - Z^H S_+^{-1} Z) p^H|} = \frac{|\tau(I_L - Z^H Q Z) \tau^H|}{|\tau(I_L - Z^H S_+^{-1} Z) \tau^H|} . \tag{2-69}$$

From definition (2-63), we see that

$$I_L - Z^H S_+^{-1} Z = I_L - Z^H (Z Z^H)^{-1} Z$$

is a projection matrix. In fact, it projects onto the orthogonal complement of the row space of the data array Z. For fixed p, the denominator of Equation (2-69) is positive with probability one, since its inverse is the numerator of Equation (2-42). The latter is finite (with probability one), so long as our basic constraint $L \geq N + M$ is satisfied. For fixed data, however, we cannot generalize our GLR test by letting τ be arbitrary (as we were able to generalize it earlier by letting σ be an arbitrary array), since the rows of τ could always be chosen from the row space of Z, thus making the denominator of Equation (2-69) vanish. This is another example of a statistical model which provides too much freedom in the parameters to sustain a meaningful decision rule.

With suitable constraints on τ, Equation (2-69) could be made the basis of a generalization of our basic GLR test, but this topic will not be pursued further here. This equation does, however, provide us with a useful property of the basic GLR test, which may be mentioned at this point. Suppose that τ can be expressed in the form $\tau = \tau_1 W_L^H$, where τ_1 is another $M \times L$ array of rank M, and where W_L is a unitary matrix of order L. If this representation for τ is substituted in Equation (2-69), the equation will have the same form as before, but with τ replaced by τ_1, and with Z replaced by $Z_1 = Z U_L$. This replacement for Z also may be made in the formula for S_+

without changing that matrix. Since Q depends on Z only through S_+, we see that the simultaneous replacement of τ by τ_1 and Z by Z_1 leaves the test statistic unaltered. When Z is considered as a random array, the post-multiplication by W_L does not change its covariance, as we can see from Equations (1-5) and (1-6) of Section 1. The mean value of Z is altered, of course, as shown by Equation (1-7). The effect is simply to replace τ by $\tau W_L = \tau_1$ in the formula for the mean. We conclude that the performance of the GLR test, as a detection criterion, is unchanged if the τ array is post-multiplied by any unitary matrix. In particular, τ can be converted to a form in which all but the first M columns are identically zero, by means of a suitable unitary transformation. We will encounter this invariance property again in Section 6.

In Section 3, the performance of the GLR test will be studied starting from Equation (2-42). An algorithm for the efficient computation of this expression is presented in Appendix 7. This is a "square-root" algorithm which uses standard signal processing techniques applied to the data arrays themselves, and it avoids the computation and inversion of the sample covariance matrices. In Appendix 1, we show that the same performance results can be derived directly from Equation (2-56), and a square-root algorithm for the computation of the GLR test statistic in this form also can be devised. This algorithm is also discussed in Appendix 7.

3. STATISTICAL PROPERTIES OF THE GLR TEST STATISTIC

We turn now to the statistical properties of the test statistic, given by Equation (2-42). Recall the arrays e and f, defined in Equations (2-17) and (2-19), with their properties as derived in Section 2. Together they form a unitary matrix U_N [Equation (2-20)], which we now use to decompose both Z_p and Z_q into further components. We define

$$U_N^H Z_p = \begin{bmatrix} e^H Z_p \\ f^H Z_p \end{bmatrix} \equiv \begin{bmatrix} Z_A \\ Z_B \end{bmatrix} , \tag{3-1}$$

and

$$U_N^H Z_q = \begin{bmatrix} e^H Z_q \\ f^H Z_q \end{bmatrix} \equiv \begin{bmatrix} W_A \\ W_B \end{bmatrix} , \tag{3-2}$$

in analogy to Equation (2-16), so that

$$\begin{aligned} Z_p &= e Z_A + f Z_B \\ Z_q &= e W_A + f W_B . \end{aligned} \tag{3-3}$$

We have now resolved the data array Z into four components:

$$U_N^H Z U_L^H = \begin{bmatrix} Z_A & W_A \\ Z_B & W_B \end{bmatrix} = \begin{bmatrix} e^H Z p^H & e^H Z q^H \\ f^H Z p^H & f^H Z q^H \end{bmatrix} , \tag{3-4}$$

where U_L is the unitary matrix defined in Equation (2-12). The A-components of the new arrays have J rows, and the B-components consist of the remaining $(N - J)$ rows.

We also define

$$U_N^H S U_N \equiv \begin{bmatrix} S_{AA} & S_{AB} \\ S_{BA} & S_{BB} \end{bmatrix} = \begin{bmatrix} e^H S e & e^H S f \\ f^H S e & f^H S f \end{bmatrix} , \tag{3-5}$$

and its inverse

$$U_N^H S^{-1} U_N \equiv \begin{bmatrix} S^{AA} & S^{AB} \\ S^{BA} & S^{BB} \end{bmatrix} = \begin{bmatrix} e^H S^{-1} e & e^H S^{-1} f \\ f^H S^{-1} e & f^H S^{-1} f \end{bmatrix} . \qquad (3\text{-}6)$$

The AA-portions of these arrays are (J × J) in dimension, and the BB-parts are also square, of dimension (N − J). The transformed S array may also be expressed in terms of the W-components, as follows:

$$U_N^H S U_N = \begin{bmatrix} W_A W_A^H & W_A W_B^H \\ W_B W_A^H & W_B W_B^H \end{bmatrix} . \qquad (3\text{-}7)$$

We introduce a s: n .ar notation for the components of the actual covariance matrix after transformation by U_N:

$$U_N^H \Sigma U_N \equiv \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{bmatrix} , \qquad (3\text{-}8)$$

together with an analogous terminology for its inverse:

$$U_N^H \Sigma^{-1} U_N \equiv \begin{bmatrix} \Sigma^{AA} & \Sigma^{AB} \\ \Sigma^{BA} & \Sigma^{BB} \end{bmatrix} . \qquad (3\text{-}9)$$

In terms of the new components, we have

$$Z_p^H S^{-1} Z_p = [Z_A^H \; Z_B^H] \begin{bmatrix} S^{AA} & S^{AB} \\ S^{BA} & S^{BB} \end{bmatrix} \begin{bmatrix} Z_A \\ Z_B \end{bmatrix} ,$$

and, using Equation (A1-9) of Appendix 1, we obtain

$$Z_p^H S^{-1} Z_p = Y^H S^{AA} Y + Z_B^H S_{BB}^{-1} Z_B , \qquad (3\text{-}10)$$

where

$$Y \equiv Z_A - S_{AB} S_{BB}^{-1} Z_B \,. \tag{3-11}$$

Using these results, the numerator of the test statistic [Equation (2-42)] becomes

$$|I_M + Z_B^H S_{BB}^{-1} Z_B + Y^H S^{AA} Y| \,.$$

From the definition of P [Equation (2-40)] we see that $Pe = 0$ and $e^H P = 0$. Then, using the first of Equations (3-3), it follows that

$$Z_p^H P Z_p = Z_B^H f^H P f Z_B \,. \tag{3-12}$$

Moreover, with the help of Equations (A1-8) of Appendix 1 and Equation (3-6), we find that

$$f^H P f = S^{BB} - S^{BA} (S^{AA})^{-1} S^{AB} = S_{BB}^{-1} \,,$$

and, consequently, the GLR test statistic can be written

$$\ell = \frac{|I_M + Z_B^H S_{BB}^{-1} Z_B + Y^H S^{AA} Y|}{|I_M + Z_B^H S_{BB}^{-1} Z_B|} \,.$$

This expression obviously depends on the subspace decompositions which have been introduced, but it is invariant to any changes in the actual bases defined in them.

According to Equation (3-6), we have

$$e^H S^{-1} e = S^{AA} \,,$$

and we also find that

$$e^H S^{-1} Z_p = S^{AA} Z_A + S^{AB} Z_B \,.$$

These results allow us to evaluate the ML estimator of the signal amplitude array [Equation (2-34)] in terms of quantities introduced in this section. Using Equation (A1-8) again, we obtain

$$\hat{b} = Z_A + (S^{AA})^{-1} S^{AB} Z_B$$

$$= Z_A - S_{AB} S_{BB}^{-1} Z_B = Y . \quad (3\text{-}13)$$

We now define the quantities

$$C_M \equiv I_M + Z_B^H S_{BB}^{-1} Z_B$$

$$V \equiv Y C_M^{-1/2}$$

$$T \equiv (S^{AA})^{-1} , \quad (3\text{-}14)$$

which allow us to express the test in the desired form:

$$\ell = \frac{|C_M + Y^H T^{-1} Y|}{|C_M|} = |I_M + V^H T^{-1} V| . \quad (3\text{-}15)$$

This quantity is a complex analog of the so-called Wilks' Lambda statistic, which arises in many applications of the multivariate analysis of variance. For the case of real variables, a test statistic analogous to Equation (3-15) is known.[2,9] It should be noted that the definition of V depends on the particular way in which C_M was factored to form a square-root matrix. The matrix C_M could also have been represented in terms of Cholesky factors, and an equation identical to Equation (3-15) obtained, with an appropriate V array. This freedom of choice cannot affect the statistical character of the GLR test statistic, and it is actually a useful feature in some cases. The point is taken up again in Section 6.

It is interesting to compare the form of this GLR test with the simpler result found in Section 2 for the non-adaptive problem (i.e., the case of known Σ). With the notation introduced here, we can express the non-adaptive ML estimator of b [Equation (2-51)] in the form

$$\hat{b}_\Sigma = (\Sigma^{AA})^{-1} (\Sigma^{AA} Z_A + \Sigma^{AB} Z_B)$$

$$= Z_A - \Sigma_{AB} \Sigma_{BB}^{-1} Z_B . \quad (3\text{-}16)$$

The second line of this equation expresses the estimator as the difference between Z_A and its conditional expectation given Z_B. The latter term is the predictable portion of the random noise part of Z_A, and the estimator can be viewed as the prediction error. This makes sense as an estimator, since the expected value of Z_A is the true value of b (see below). Conditional expectations and linear prediction are discussed in Appendix 1. Formula (3-13) shows that the estimator in the adaptive case (unknown Σ) has the same form as Equation (3-16), but with Σ replaced by an estimator of covariance, namely the one defined in Equation (2-48).

In the non-adaptive problem, the GLR test statistic is given by Equation (2-52), which may be restated as

$$\lambda = \mathrm{Tr}\left[\, \hat{b}_\Sigma^H \Sigma^{AA} \hat{b}_\Sigma \,\right] . \tag{3-17}$$

The trace operation describes non-coherent integration over the columns of $\hat{b}_\Sigma$, and these, in turn, depend only on Z_A and Z_B, the components of Z_p. The Z_q component of the data array is not used at all in the test, since, in the non-adaptive case, it contains no information of use for the detection problem.

As noted in Appendix 1, the matrix Σ^{AA} is the inverse of the covariance matrix shared by the independent columns of $\hat{b}_\Sigma$, inasmuch as they may be interpreted as prediction errors. Thus, each term of the trace on the right side of Equation (3-17) itself represents a form of non-coherent integration (following a suitable whitening operation) over the J components of each column of the estimator. This is a logical way of detecting the presence of a signal specified only as a vector in a subspace of dimension greater than unity. The formation of $\hat{b}_\Sigma$ itself is an application of *coherent* integration, which takes account of the structure of the actual signals that determine the subspace. This may be seen by referring to the original definition [Equation (2-51)] of this estimator, which depends on the data array through the term

$$e^H \Sigma^{-1} Z_p = (\sigma^H \sigma)^{-1/2} \sigma^H \Sigma^{-1} Z_p .$$

The array $\sigma^H \Sigma^{-1} Z_p$ which appears on the right side of this formula may be interpreted as comprising the outputs of a set of colored-noise matched filters, which are matched to the columns of the signal array σ and applied to the columns of Z_p. These, in turn, are formed by coherent integration along the rows of Z.

In the adaptive problem, the columns of the ML signal parameter estimator $\hat{b}$ are correlated, because they all use the same estimator of covariance. It will be shown below that this correlation is described by the matrix C_M, and it is removed in the formation of the GLR test statistic by the transition from the Y array to V. Except for a constant factor, the matrix T of this statistic is just like the inverse of Σ^{AA}, but using the estimated covariance matrix instead of the known one. Thus, the general GLR test is built with structures quite similar to those which appear in its non-adaptive analog. The final form, however, appears to be quite different, since it involves a determinant instead of a trace. This distinction disappears when we consider the limiting process by which the adaptive problem tends toward the non-adaptive one, namely the unbounded growth of $L-M$. This is the number of data array columns in excess of M, the dimensionality of the signal-defining τ array.

Without attempting to be precise, we can say that the covariance estimator given in Equation (2-48) will tend to the true covariance in this limit, and write

$$S \to (L-M)\Sigma .$$

The inverse of S therefore becomes smaller as L increases. In the limit, C_M becomes the identity matrix, as the second term in its definition [see Equation (3-14)] becomes vanishingly small. Hence, in this limit, the correlation between the columns of $\hat{b}$ disappears. Then

$$\hat{b} \to \hat{b}_\Sigma$$

and also

$$T^{-1} = S^{AA} \to \frac{1}{L-M}\Sigma^{AA} ,$$

so that

$$V^H T^{-1} V \to \frac{1}{L-M} \hat{b}_\Sigma^H \Sigma^{AA} \hat{b}_\Sigma .$$

In this form, the GLR test statistic is the determinant of the sum of the identity matrix and a "small" term, so that we obtain

$$\ell \to |I_M + \frac{1}{L-M} \hat{b}_\Sigma^H \Sigma^{AA} \hat{b}_\Sigma|$$

$$\to 1 + \frac{1}{L-M} \mathrm{Tr}\left[\hat{b}_\Sigma^H \Sigma^{AA} \hat{b}_\Sigma \right] = 1 + \frac{\lambda}{L-M} .$$

Thus, heuristically at least, the GLR test for the adaptive problem goes over into that for the non-adaptive case in the appropriate limiting situation.

From this discussion, it follows that the simpler decision rule

$$\mathrm{Tr}(V^H T^{-1} V) \geq \text{Constant} \ ,$$

should perform well for large values of L. The analog of this detector with real variables is known as the Lawley-Hotelling test.[19]

Until now, the data array has been considered as a given set of complex numbers, while the parameters characterizing the statistical model, namely B and Σ, have been treated as variables for the derivation of the GLR test. To evaluate the performance of the test, these parameters must be considered fixed and given, while the elements of the data array are considered to be random variables. The remainder of this section is devoted to establishing the statistical properties of the test statistic.

Suppose that the true signal parameter array is B and that the actual covariance matrix of the columns of Z is Σ. Then,

$$E\,Z = \sigma B \tau = ebp \ , \tag{3-18}$$

and

$$\mathrm{Cov}(Z) = \Sigma \otimes I_L \ . \tag{3-19}$$

The mean value of the transformed data array will be

$$E\,U_N^H Z\, U_L^H = \begin{bmatrix} e^H \\ f^H \end{bmatrix} ebp \, [p^H \ q^H] \quad \begin{bmatrix} b & 0 \\ & 0 \end{bmatrix} , \tag{3-20}$$

and its covariance, using formula (A1-44) of Appendix 1, will be

$$\mathrm{Cov}(U_N^H Z\, U_L^H) = (U_N^H \Sigma\, U_N) \otimes I_L \ .$$

Comparing Equation (3-20) with Equation (3-4), we see that the expected value of Z_A is just b, while the other three components of the transformed array have zero mean. The columns of the transformed array are still independent, and they now share the covariance matrix which has been expressed in component form in Equation (3-8).

Note that the situation corresponding to the "true" parameters, as described by Equations (3-18) and (3-19) above, coincides exactly with the model postulated in Equations (1-1) and (1-2) of Section 1. We refer to this as the "matched" situation. It is interesting to consider the effect of various departures from this matched condition on the performance of the GLR test. At the end of this section we introduce a particular form of "mismatch" which proves to be amenable to analysis, and take up its implications in Sections 5 and 6.

To proceed, we first fix the arrays Z_B and W_B, and we refer to this conditioning by using the subscript B. Referring again to Appendix 1 for details, we have the following conditional expectations:

$$E_B Z_A = b + \Sigma_{AB} \Sigma_{BB}^{-1} Z_B , \tag{3-21}$$

and

$$E_B W_A = \Sigma_{AB} \Sigma_{BB}^{-1} W_B . \tag{3-22}$$

From Equation (3-7), we see that

$$Y = Z_A - W_A W_B^H S_{BB}^{-1} Z_B ,$$

hence, Y is a Gaussian array under the conditioning, with conditional expectation

$$E_B Y = b + \Sigma_{AB} \Sigma_{BB}^{-1} Z_B - \Sigma_{AB} \Sigma_{BB}^{-1} W_B W_B^H S_{BB}^{-1} Z_B .$$

But $W_B W_B^H = S_{BB}$; therefore,

$$E_B Y = b . \tag{3-23}$$

Finally, using Equations (3-14), we obtain

$$E_B V = b C_M^{-1/2} . \tag{3-24}$$

We note that the matrix C_M depends only on quantities fixed under the conditioning, and it may therefore be treated as a constant as long as the conditioning holds. Therefore, V itself is conditionally a Gaussian random array.

Since the columns of the transformed data array are all independent, the conditioning variables only affect their own columns. It follows that the conditional covariance of any of the columns of Z_A or W_A is given by

$$(\Sigma^{AA})^{-1} = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA} ,$$

using a standard property of Gaussian random variables, reviewed in Appendix 1. We can therefore describe the conditional covariance properties of Z_A and W_A together by the statement

$$\mathrm{Cov}_B\big([Z_A \; W_A]\big) = (\Sigma^{AA})^{-1} \otimes I_L . \tag{3-25}$$

To evaluate the covariance properties of Y and V, it is convenient to write

$$Y = Z_A - W_A Q = [Z_A \; W_A] \begin{bmatrix} I_M \\ -Q \end{bmatrix} ,$$

where Q is defined by

$$Q \triangleq W_B^H S_{BB}^{-1} Z_B . \tag{3-26}$$

Then, using Equation (A1-44) of Appendix 1, we have

$$\mathrm{Cov}_B(Y) = (\Sigma^{AA})^{-1} \otimes (I_M + Q^H Q)^* .$$

But it is easily verified that

$$Q^H Q = Z_B^H S_{BB}^{-1} Z_B ,$$

and, thus, Y has the covariance matrix

$$\mathrm{Cov}_B(Y) = (\Sigma^{AA})^{-1} \otimes C_M^* . \tag{3-27}$$

Since Y is actually the ML estimator of the signal parameter array b, Equation (3-27) expresses the conditional correlation between the columns of this estimator which was mentioned earlier. A fuller discussion, including the effect of removing the conditioning for this estimator, is given in Section 5.

Recalling the definition of V, and using the same covariance identity, we find that the columns of V are independent under the conditioning:

$$\mathrm{Cov}_B(V) = (\Sigma^{AA})^{-1} \otimes I_M . \tag{3-28}$$

It is useful to write V as the sum of two parts:

$$V = V_s + V_n , \tag{3-29}$$

where

$$V_s \equiv b\, C_M^{-1/2} . \tag{3-30}$$

Then V_n, which we may call the "noise component," is a complex Gaussian array, with zero mean and covariance given by Equation (3-28), independent of the conditioning variables which appear only in the "signal component" V_s.

Turning now to the T array, we recall that T is the inverse of S^{AA}, and that the matrix

$$\begin{bmatrix} S_{AA} & S_{AB} \\ S_{BA} & S_{BB} \end{bmatrix} = \begin{bmatrix} W_A \\ W_B \end{bmatrix} \begin{bmatrix} W_A^H & W_B^H \end{bmatrix}$$

is a complex Wishart matrix, of order N, with L − M complex degrees of freedom. The unconditioned means of W_A and W_B are zero, and from Equations (3-2) and (A1-44) we obtain the unconditioned covariance

$$\mathrm{Cov}\left(\begin{bmatrix} W_A \\ W_B \end{bmatrix} \right) = (U_N^H \Sigma U_N) \otimes I_{L-M} .$$

The partitioned form of the transformed Σ matrix is given by Equation (3-8).

Some of the properties of Wishart matrices are discussed in Appendix 1, where it is proved that a matrix such as T, which is the inverse of a diagonal block of a partitioned complex Wishart matrix, is also a complex Wishart matrix of an appropriate order and with a reduced number of complex degrees of freedom. In the case of T, the dimension is J and the number of complex degrees of freedom is $L+J-N-M$. From the results of Appendix 1 it also follows that T is independent of S_{AB}. These facts are established in Appendix 1 first under conditioning on the B components, but the probability density function of T does not depend on the values of the conditioning variables. Therefore, the complex Wishart character of T, as well as its independence of S_{AB}, remains true when the conditioning is removed. By the same argument, T is proved to be unconditionally independent of $S_{AB}S_{BB}^{-1}$, because the second factor of this product is constant under the conditioning. Since T is formed from W_A and W_B, it is clearly independent of the components Z_A and Z_B. Thus, T is unconditionally independent of Y, as defined by Equation (3-11), and also of C_M and V, defined in Equations (3-14). T can be expressed in terms of a Gaussian array, say W, of dimension $J \times (L+J-N-M)$, as follows:

$$T = WW^H . \tag{3-31}$$

The mean of W is zero, and its covariance is

$$\mathrm{Cov}(W) = (\Sigma^{AA})^{-1} \otimes I_{L+J-N-M} , \tag{3-32}$$

a property established in Appendix 1.

The last step in the statistical characterization of the test statistic is a "whitening" operation. With the conditioning on the B-components still in effect, we define the new arrays

$$V_0 \equiv (\Sigma^{AA})^{1/2} V$$

$$T_0 \equiv (\Sigma^{AA})^{1/2} T (\Sigma^{AA})^{1/2} , \tag{3-33}$$

using the subscript zero to indicate the whitening. The matrix T_0 is also a complex Wishart matrix, and it can be expressed in terms of a new zero-mean complex Gaussian array W_0 (unrelated to W_A and W_B):

$$T_0 \equiv W_0 W_0^H . \tag{3-34}$$

These new arrays have identity matrices for their covariances:

$$\mathrm{Cov}_B(V_0) = I_J \otimes I_M$$

$$\mathrm{Cov}_B(W_0) = I_J \otimes I_{L+J-N-M} . \tag{3-35}$$

Thus, all the elements of these arrays are conditionally independent. The whitened array V_0 is made up of the components

$$V_0 \equiv V_{0s} + V_{0n} , \tag{3-36}$$

where V_{0n} is a complex Gaussian array, with zero mean and covariance equal to the identity, and where

$$V_{0s} \equiv (\Sigma^{AA})^{1/2}\, b\, C_M^{-1/2} . \tag{3-37}$$

The columns of the conditioning arrays Z_B and W_B share the covariance matrix Σ_{BB}. The marginal probability density functions of these arrays are direct analogs of Equation (A1-79) of Appendix 1. These arrays are now also whitened, with the introduction of the new quantities

$$Z_{B0} \equiv (\Sigma_{BB})^{-1/2} Z_B$$

$$W_{B0} \equiv (\Sigma_{BB})^{-1/2} W_B . \tag{3-38}$$

The whitened arrays have zero means; their covariance matrices are given by

$$\mathrm{Cov}(Z_{B0}) = I_{N-J} \otimes I_M$$

$$\mathrm{Cov}(W_{B0}) = I_{N-J} \otimes I_{L-M} . \tag{3-39}$$

The whitening matrix cancels out in the formation of C_M, which has the same structure in terms of Z_{B0} and W_{B0}:

$$C_M = I_M + Z_{B0}^H (W_{B0} W_{B0}^H)^{-1} Z_{B0} . \tag{3-40}$$

Finally, the test statistic also retains its form when expressed in terms of the whitened arrays V_0 and W_0:

$$\ell = |I_M + V_0^H T_0^{-1} V_0| , \tag{3-41}$$

and this is the form which is analyzed in later sections. We note that the original covariance matrix Σ survives only in the "signal" array V_{0s}. The conditioning variables Z_{B0} and W_{B0} are also confined to that component, entering through its dependence on the C_M array.

We can therefore state that V_{0n} and W_0 are (unconditionally) independent complex Gaussian arrays, with zero means and covariance matrices given by the right sides of Equations (3-35), and that T_0 is subject to a complex Wishart distribution of dimension J, with $L+J-N-M$ complex degrees of freedom. T_0 is expressed in terms of W_0 by Equation (3-34). From this point forward, unless explicitly stated otherwise, when we say that a matrix is complex Wishart we mean that it has a form corresponding to Equation (3-31), and that the covariance matrix of the underlying Gaussian array is a Kronecker product of identity matrices.

The test statistic is expressed by Equation (3-41) and V_0 is given by Equation (3-36). Moreover, V_{0s} is independent of V_{0n} and W_0. To compute the probability of detection (PD) one can, in principle, begin by conditioning on V_{0s} itself, determine the conditional PD, and remove the conditioning at the end. The statistical character of V_{0s} is required, of course, and this is discussed in Section 5 . For the probability of false alarm (PFA), however, V_{0s} vanishes and our statistical analysis is formally complete. The statistical properties of the test can depend only on the dimensional parameters of the problem (in the absence of signal); hence, the GLR test is a CFAR decision rule. A more explicit statistical characterization will be obtained in Section 4.

The possibility of "mismatch" was mentioned earlier, and we introduce an example of it here. The departure from the modeled situation relates only to the signal component; hence, it will have no effect on the discussion of false alarm probability in the next section. We suppose that the true mean of the data array is not given by Equation (3-18), but instead has the more general form

$$E Z = D\tau \equiv dp \, . \tag{3-42}$$

The case of a completely arbitrary mean value of Z is certainly interesting, but its analysis appears to present considerable difficulties. With the new model, Equation (3-20) is replaced by

$$E\, U_N^H Z U_L^H = \begin{bmatrix} e^H \\ f^H \end{bmatrix} dp \, [p^H \; q^H] = \begin{bmatrix} b_A & 0 \\ b_B & 0 \end{bmatrix} , \tag{3-43}$$

where

$$b_A \equiv e^H d$$

$$b_B \equiv f^H d$$

$$d = D(\tau\tau^H)^{1/2} . \qquad (3\text{-}44)$$

According to Equation (3-43), the components W_A and W_B retain their zero means, but now

$$E Z_A = b_A$$

and

$$E Z_B = b_B .$$

In the analysis of the matched problem, we began by conditioning on the B components Z_P and W_B. Formula (3-22) remains valid for the conditional mean of W_A, but Equation (3-21) must now be replaced by

$$E_B Z_A = b_A + \Sigma_{AB} \Sigma_{BB}^{-1} (Z_B - b_B) . \qquad (3\text{-}45)$$

This is a direct analog of Equation (A1-81) in Appendix 1. The conditional mean of Y is evaluated as before, but now with the result

$$E_B Y = b_A - \Sigma_{AB} \Sigma_{BB}^{-1} b_B . \qquad (3\text{-}46)$$

The conditional covariance of Y is still correctly expressed by Equation (3-27), with C_M as defined by Equation (3-14). The effect of the non-vanishing mean value of Z_B which enters this definition will be felt when the conditioning is removed later.

After the transition to whitened arrays, Equation (3-37) becomes

$$V_{0s} = (\Sigma^{AA})^{1/2} (b_A - \Sigma_{AB} \Sigma_{BB}^{-1} b_B) C_M^{-1/2} . \qquad (3\text{-}47)$$

Expression (3-40), which defines C_M in terms of these whitened arrays, remains correct, but now

$$E Z_{B0} = (\Sigma_{BB})^{-1/2} b_B . \qquad (3\text{-}48)$$

These results will be utilized in Sections 5 and 6, where the effects of this kind of mismatch are studied in terms of signal parameter estimation and probability of detection.

4. THE PROBABILITY OF FALSE ALARM

The fundamental problem of performance analysis is the computation of the probability of accepting hypothesis H_1 by means of the GLR test: $\ell \geq \ell_0$. The general case is discussed in Section 6. We devote this section to the evaluation of the probability of false alarm (PFA), i.e., the probability of accepting H_1 when $B=0$. We simplify the notation of Section 3 by dropping the subscript 0 which was used to indicate the whitening of various arrays. The GLR test statistic, given by Equation (3-41), again assumes the form

$$\ell = |I_M + V^H T^{-1} V| , \tag{4-1}$$

where

$$T = WW^H . \tag{4-2}$$

The arrays V and W are Gaussian, independent of one another, and they both have mean value zero. We introduce the new parameter

$$K \equiv L - N - M , \tag{4-3}$$

and recall that $K \geq 0$ by the constraint first expressed as Equation (1-9). The dimension of V is $J \times M$, W is $J \times (J+K)$, and the covariances of these arrays are given by

$$\begin{aligned} \mathrm{Cov}(V) &= I_J \otimes I_M \\ \mathrm{Cov}(W) &= I_J \otimes I_{J+K} . \end{aligned} \tag{4-4}$$

The PFA will depend only on J, M, and K, and not on the actual covariance matrix Σ; hence, the GLR test has the CFAR property. The only change when signals are added will be the addition of a non-zero mean value for V.

Using Equation (A1-2), the test statistic can also be expressed in the form

$$\ell = \frac{|T + VV^H|}{|T|} . \tag{4-5}$$

The inverse of ℓ is the complex analog of Wilks' Lambda statistic,[2,9] which often arises in multivariate statistical analysis. It is usual in that context to test *for* the validity of H_0 against H_1, as we have defined the hypotheses, which accounts for the inversion of the test statistic. We note that T is a complex Wishart matrix and is non-singular (with probability one), but that VV^H is non-singular only when $M \geq J$.

It is useful to consider some special cases, and we begin with the simplest, namely $J=1$, with arbitrary M and non-negative K. Then, V and W are row vectors and T is a scalar:

$$T = WW^H = \sum_{j=1}^{K+1} |w_j|^2 ,$$

where the w_j are the elements of W. Thus, T is a complex chi-squared variable, with $K+1$ complex degrees of freedom. This terminology is introduced in Appendix 2, where a discussion of the complex chi-squared and other related distributions will be found.

Using Equation (4-5), the test statistic takes the form

$$\ell = 1 + \frac{VV^H}{T} , \tag{4-6}$$

and VV^H is also a complex chi-squared variable:

$$VV^H = \sum_{i=1}^{M} |v_i|^2 ,$$

with M complex degrees of freedom. The ratio of complex chi-squared variables which appears in Equation (4-6) is subject to a complex central F distribution, but we prefer to express the test statistic in the form

$$1/\ell = \frac{\sum_{j=1}^{K+1} |w_j|^2}{\sum_{j=1}^{K+1} |w_j|^2 + \sum_{i=1}^{M} |v_i|^2} \equiv x_\beta(K+1,M) . \tag{4-7}$$

In this formula, the notation $x_\beta(n,m)$ is used in a generic sense to denote a random variable which obeys the complex central Beta distribution, whose probability density function (pdf) is given by Equation (A2-12) of Appendix 2. The PFA, defined by the equation

$$\mathrm{PFA} = \mathrm{Prob}(\ell \geq \ell_0) = \mathrm{Prob}\left[x_\beta(K+1,M) \leq 1/\ell_0\right] , \tag{4-8}$$

is just the cumulative of the complex central Beta distribution, also presented in Appendix 2. Substituting the appropriate parameter values in Equation (A2-14), we obtain

$$\mathrm{PFA} = \frac{1}{\ell_0^{M+K}} \sum_{m=0}^{M-1} \binom{M+K}{m} (\ell_0 - 1)^m . \tag{4-9}$$

With the further specialization $M=1$, this formula reproduces the simple result found in Reference 3:

$$\mathrm{PFA} = \frac{1}{\ell_0^{K+1}} = \frac{1}{\ell_0^{L-N}} . \tag{4-10}$$

The other special case we wish to discuss is the dual version in which $M=1$, J is arbitrary J, and K is non-negative. V is now a column vector, and

$$\ell = 1 + V^H T^{-1} V .$$

The T matrix is of order J and satisfies a complex Wishart distribution with $J+K$ complex degrees of freedom. T is expressed in terms of a zero-mean Gaussian array in Equation (4-2). The covariance matrix of this array is the identity, as stated in Equation (4-4). As noted in the previous section, these properties of the underlying Gaussian array will be tacitly assumed for Wishart matrices in the following.

We define the unit vector

$$g \equiv V (V^H V)^{-1/2} ,$$

and write

$$\ell = 1 + (V^H V)(g^H T^{-1} g) . \tag{4-11}$$

Obviously, the quantity

$$V^H V = \sum_{i=1}^{J} |v_i|^2$$

is a complex chi-squared variable, with J complex degrees of freedom.

The unit vector g may be considered to form the basis for a one-dimensional subspace of $\mathbb{C}^J$, and thus, according to the general property of Wishart matrices established in Appendix 1, the inverse of $g^H T^{-1} g$ is also subject to a complex Wishart distribution. This latter Wishart matrix is simply a complex chi-squared variable in the present case, since its dimension is unity. This dimension is smaller than that of T by $J-1$, hence the complex chi-squared variable has $K+1$ complex degrees of freedom, according to the rule derived in Appendix 1. It is therefore statistically equivalent to the sum

$$(g^H T^{-1} g)^{-1} = \sum_{j=1}^{K+1} |w_j|^2 .$$

where the w_j are complex Gaussian variables of zero mean and unit variance. These properties are independent of the conditioning variables, hence they remain true without the conditioning which is now removed. Then, Equation (4-11) can be written

$$\ell = 1 + \frac{\sum_{i=1}^{J} |v_i|^2}{\sum_{j=1}^{K+1} |w_j|^2} , \qquad (4\text{-}12)$$

where the v_i are independent of the w_j. In other words,

$$1/\ell = x_\beta(K+1, J) . \qquad (4\text{-}13)$$

For the special case $M = 1$, we have therefore found:

$$\mathrm{PFA} = \frac{1}{\ell_0^{J+K}} \sum_{j=0}^{J-1} \binom{J+K}{j} (\ell_0 - 1)^j . \qquad (4\text{-}14)$$

This expression is in agreement with the corresponding result given in Reference 5.

We return to the general case and introduce a generic notation for the random matrix which appears in Equation (4-1):

$$\mathscr{C}(J,M,K) \equiv I_M + V^H T^{-1} V \quad (4\text{-}15)$$

The GLR test statistic itself is given a more specific notation, indicating the dimensional parameters to which it relates:

$$\ell(J,M,K) = |\mathscr{C}(J,M,K)| = |\mathscr{C}(J,M,L-N-M)| \quad (4\text{-}16)$$

It is useful to study some of the properties of these quantities, under the assumption that V and W are independent, zero-mean Gaussian arrays, with covariances given by Equation (4-4). By its very structure, the $\mathscr{C}$ matrix is always positive definite, and, when $M=1$, it reduces to a scalar. In the latter case, according to Equation (4-13),

$$\ell(J,1,K) = 1/x_\beta(K+1,J) \quad (4\text{-}17)$$

Similarly, when $J=1$, Equation (4-7) yields

$$\ell(1,M,K) = 1/x_\beta(K+1,M) \quad (4\text{-}18)$$

Equalities such as these are meant to indicate statistical identity, i.e., the equality of the probability density functions of the random variables which enter the equation. These two results constitute a particular example of a general duality property which will be derived later. We also note that the matrix C_M, defined by Equation (3-40), is of the same form, namely,

$$C_M = \mathscr{C}(N-J,M,L+J-N-M) = \mathscr{C}(N-J,M,J+K) \quad (4\text{-}19)$$

This matrix plays a central role in the analysis of performance under hypothesis H_1.

Let us introduce a decomposition of the vector space $\mathbb{C}^J$ into a subspace of dimension J_1 and its orthogonal complement, whose dimension will be $J_2 = J - J_1$. The arrays V and W are partitioned as follows:

$$V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

$$W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} . \qquad (4\text{-}20)$$

and we have

$$T = \begin{bmatrix} W_1 W_1^H & W_1 W_2^H \\ W_2 W_1^H & W_2 W_2^H \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} . \qquad (4\text{-}21)$$

We recall that T is a complex Wishart matrix of dimension J, with J + K degrees of freedom, and that the covariance of W is given by Equation (4-4).

We also define

$$T^{-1} = \begin{bmatrix} T^{11} & T^{12} \\ T^{21} & T^{22} \end{bmatrix} , \qquad (4\text{-}22)$$

and then substitute:

$$V^H T^{-1} V = \begin{bmatrix} V_1^H & V_2^H \end{bmatrix} \begin{bmatrix} T^{11} & T^{12} \\ T^{21} & T^{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} . \qquad (4\text{-}23)$$

Making use of identity (A1-9) of Appendix 1, we obtain

$$V^H T^{-1} V = (V_1 - T_{12} T_{22}^{-1} V_2)^H T^{11} (V_1 - T_{12} T_{22}^{-1} V_2) + V_2^H T_{22}^{-1} V_2 . \qquad (4\text{-}24)$$

By adding the identity matrix to Equation (4-24), we can express $\mathcal{C}(J,M,K)$ in the form of a product:

$$\mathcal{C}(J,M,K) = (I_M + V_2^H T_{22}^{-1} V_2)^{1/2} (I_M + \mathcal{V}^H \mathcal{T}^{-1} \mathcal{V}) (I_M + V_2^H T_{22}^{-1} V_2)^{1/2} , \qquad (4\text{-}25)$$

where

$$\mathcal{V} \equiv (V_1 - T_{12}T_{22}^{-1}V_2)(I_M + V_2^H T_{22}^{-1} V_2)^{-1/2} \tag{4-26}$$

and

$$\mathcal{T} \equiv (T^{11})^{-1} = T_{11} - T_{12}T_{22}^{-1}T_{21} . \tag{4-27}$$

The arrays V_2 and W_2 are Gaussian, independent of one another, and have mean values zero. Their covariances are given by

$$\mathrm{Cov}(V_2) = I_{J_2} \otimes I_M$$

$$\mathrm{Cov}(W_2) = I_{J_2} \otimes I_{J+K} .$$

We can therefore write

$$I_M + V_2^H T_{22}^{-1} V_2 = \mathscr{C}(J_2, M, J_1 + K) , \tag{4-28}$$

indicating thereby the statistical character of this matrix as an example of the family defined by Equation (4-15). It is directly analogous to the matrix C_M of the previous section.

We again recall the analysis of Section 3, which may be applied directly to the study of $\mathcal{V}$ and $\mathcal{T}$. These quantities correspond to V and T of that section. Conditioning on V_2 and W_2, it follows that $\mathcal{V}$ is a zero-mean Gaussian array with covariance

$$\mathrm{Cov}(\mathcal{V}) = I_{J_1} \otimes I_M ,$$

and that $\mathcal{V}$ and $\mathcal{T}$ are independent. According to a property of complex Wishart matrices, established in Appendix 1, it follows that $\mathcal{T}$ is a complex Wishart matrix, of dimension J_1, and with $J_1 + K$ complex degrees of freedom. Thus, $\mathcal{T}$ may be expressed in the form

$$\mathcal{T} = \mathcal{W}\mathcal{W}^H ,$$

where $\mathcal{W}$ is a zero-mean Gaussian array, with covariance

$$\mathrm{Cov}(\mathcal{W}) = I_{J_1} \otimes I_{J_1+K} .$$

All these statements are valid under the conditioning, but they do not involve the conditioning variables. In particular, the pdf of $\mathcal{I}$ does not depend on these variables. Hence, these statements remain true without the conditioning, which we now remove. We have therefore shown that

$$I_M + \mathcal{V}^H \mathcal{I}^{-1} \mathcal{V} = \mathscr{C}(J_1, M, K) , \tag{4-29}$$

again using this notation to identify the statistical character of this matrix. In addition, since the statistical properties of this array do not depend on the conditioning variables, it follows that the matrices expressed by Equations (4-28) and (4-29) are themselves independent.

From these results, we obtain the basic matrix factorization identity

$$\mathscr{C}(J,M,K) = \left[\mathscr{C}(J_2,M,J_1+K)\right]^{1/2} \mathscr{C}(J_1,M,K)\left[\mathscr{C}(J_2,M,J_1+K)\right]^{1/2} , \tag{4-30}$$

and, from it, the recursion relation

$$\ell(J,M,K) = \ell(J-J_1,M,J_1+K)\; \ell(J_1,M,K) . \tag{4-31}$$

The factors on the right are independent, and the recursion holds for any $J_1 < J$. Choosing $J_1 = 1$ and iterating, we obtain a representation in terms of independent factors:

$$\ell(J,M,K) = \prod_{j=0}^{J-1} \ell(1,M,K+j) . \tag{4-32}$$

The factors on the right side of this equation correspond to the special case $J = 1$ which we have already studied. Thus, using Equation (4-18), we have

$$1/\ell(J,M,K) = \prod_{j=1}^{J} x_\beta(K+j,M) . \tag{4-33}$$

The inverse of the test statistic is therefore the product of a set of independent complex central Beta variables. In the case of real data, Wilks' Lambda statistic is expressible as a product of independent real central Beta variables, with a sequence of parameters increasing in half-integral steps. Equation (4-33), which refers to the complex version of Wilks' statistic, is a direct analog. (See also Reference 20, where this result and the complex analogs of a number of other statistical theorems concerning real Gaussian variables are stated.)

We have also shown that

$$\ell(1,M,K) = \ell(M,1,K) \ , \tag{4-34}$$

by our discussion of the two special cases at the start of this section. As a special case of Equation (4-33), we have

$$1/\ell(M,1,K) = \prod_{m=1}^{M} x_\beta(K+m,1) \ ,$$

which, together with Equations (4-18) and (4-34), yields the following identity among complex central Beta variables:

$$x_\beta(K+1,M) = \prod_{m=1}^{M} x_\beta(K+m,1) \ . \tag{4-35}$$

The factors on the right are, of course, independent, and this identity can easily be verified by other means. Combining these results, we obtain the desired representation of the GLR test statistic as a double product of JM independent factors:

$$1/\ell(J,M,K) = \prod_{j=1}^{J} \prod_{m=1}^{M} x_\beta(K+j+m-1,1) \ . \tag{4-36}$$

The notation indicates the statistical character of each factor, their independence being understood. From this expression, it is clear that J and M may be interchanged without change to the PFA, provided only that K remains the same. This generalizes the duality noted earlier in this section.

Equation (4-36) provides a formally complete statistical characterization of the GLR test statistic under the null hypothesis. Except in the special cases already evaluated it is, however, of limited utility as a ' for numerical evaluation. This is particularly true in radar applications, whe FA values as small as 10^{-8} commonly occur. Similar difficulties are encountered in the evaluation of the real Wilks' statistic.[21] The double-product representation is, on the other hand, well suited to evaluation by the technique of numerical integration in the complex plane, following a contour of steepest descent. This procedure has been developed and successfully applied to a number of detection probability evaluations by Helstrom,[22] building on earlier work by Rice.[23] The analytical techniques involved in this procedure are quite unrelated to those used elsewhere in this study, and the entire topic is relegated to Appendix 6.

At this point it is useful to derive a result that will be needed in the next Section. We return to the definition of the $\mathscr{C}$ matrix and apply a unitary transformation to both sides, writing

$$U^H \mathscr{C}(J,M,K) U = I_M + \Phi^H T^{-1} \Phi , \tag{4-37}$$

where

$$\Phi \equiv V U ,$$

and U is an arbitrary unitary matrix of order M. Since Φ is statistically indistinguishable from V, the joint pdf of the elements of $\mathscr{C}(J,K,M)$ must also be invariant to the transformation expressed by Equation (4-37). It then follows that

$$E \mathscr{C}^n = E(U^H \mathscr{C} U)^n = U^H E \mathscr{C}^n U , \tag{4-38}$$

for any positive or negative integer n. Since Equation (4-38) holds for all unitary matrices, $E\mathscr{C}^n$ must be a multiple of the identity matrix.

We are particularly interested in the first moment of $\mathscr{C}$, and we make the definition

$$E \mathscr{C}(J,M,K) \equiv \mu(J,M,K) I_M . \tag{4-39}$$

Taking the trace of both sides of this equation, we have

$$\mu(J,M,K) = M^{-1}\, E\, \mathrm{Tr}\, \mathscr{C}(J,M,K)$$
$$= 1 + M^{-1}\, E\, \mathrm{Tr}\,(T^{-1} VV^{H})\,. \tag{4-40}$$

But T and V are independent, and

$$E\, VV^{H} = M\, I_J\,,$$

according to Equation (A1-42). The dependence on M therefore disappears, and

$$\mu(J,M,K) = 1 + E\, \mathrm{Tr}(T^{-1})\,.$$

If we take the trace of both sides of the factorization formula [Equation (4-30)] and recall the independence of the factors, we obtain the recursion

$$\mu(J,M,K) = \mu(J-J_1, M, J_1+K)\, \mu(J_1, M, K)\,.$$

This is just like Equation (4-31), and by iteration we find

$$\mu(J,M,K) = \prod_{j=0}^{J-1} \mu(1,M,K+j)\,. \tag{4-41}$$

When $J=1$, Equation (4-40) yields

$$\mu(1,M,K) = 1 + M^{-1} E \left(\frac{\sum_{i=1}^{M} |v_i|^2}{\sum_{j=1}^{K+1} |w_j|^2} \right).$$

As noted earlier, the ratio of complex chi-squared variables which enters here is subject to the complex F distribution [Equation (A2-9) of Appendix 2], and the required expectation value is just M/K. Thus,

$$\mu(1,M,K) = \frac{K+1}{K}\,. \tag{4-42}$$

from which we obtain

$$\mu(J,M,K) = \prod_{j=0}^{J-1} \frac{K+j+1}{K+j} = 1 + J/K\ , \tag{4-43}$$

and

$$E\,\mathscr{C}(J,M,K) = (1 + J/K)\, I_M\ . \tag{4-44}$$

This is the result we need later, and we note that the evaluation has also yielded the expected value of the trace of the inverse of a complex Wishart matrix, of dimension J and with J + K complex degrees of freedom:

$$E\,\mathrm{Tr}(T^{-1}) = J/K\ . \tag{4-45}$$

It is worth noting that, by a completely analogous argument, the following result may be obtained:

$$E[\,\mathscr{C}(J,M,K)\,]^{-1} = [\,1 - \frac{J}{M+K+J}\,]\, I_M\ . \tag{4-46}$$

5. THE ESTIMATION OF SIGNAL PARAMETERS

We begin this section by returning to the non-adaptive version of the problem and complete the analysis of its performance, both in terms of signal parameter estimation and detection probability. This exercise provides useful background for the adaptive version, and also serves to introduce some relevant notation. We recall that only the component Z_p of the data array enters the results in this case, since the covariance matrix Σ is assumed to be known.

The non-adaptive signal parameter array estimator, derived in Section 2, is

$$\hat{b}_\Sigma = (e^H \Sigma^{-1} e)^{-1} e^H \Sigma^{-1} Z_p \; . \tag{5-1}$$

In Section 3 [Equation (3-16)], it was expressed in terms of the A and B components of Z_p, as follows:

$$\hat{b}_\Sigma = Z_A - \Sigma_{AB} \Sigma_{BB}^{-1} Z_B \; .$$

This estimator is completely characterized as a Gaussian array, whose mean and covariance are

$$E \hat{b}_\Sigma = b$$

$$\mathrm{Cov}(\hat{b}_\Sigma) = (\Sigma^{AA})^{-1} \otimes I_M \; . \tag{5-2}$$

The first of these equations, which states that the estimator is unbiased, follows from Equation (3-20). The second equation is a direct analog of Equation (A1-82) of Appendix 1, since the estimator has the form of a prediction error.

A whitened estimator may be defined as follows:

$$\hat{b}_{\Sigma 0} \equiv (\Sigma^{AA})^{1/2} \hat{b}_\Sigma \; . \tag{5-3}$$

Its expected value is

$$E \hat{b}_{\Sigma 0} = (\Sigma^{AA})^{1/2} b \equiv b_0 \; . \tag{5-4}$$

which we will call the whitened true signal parameter array. The covariance of this whitened estimator is

$$\mathrm{Cov}(\hat{b}_{\Sigma 0}) = I_J \otimes I_M \, . \tag{5-5}$$

and its pdf is equal to

$$f(\hat{b}_{\Sigma 0}) = \frac{1}{\pi^{JM}} e^{-\mathrm{Tr}(\hat{b}_{\Sigma 0} - b_0)(\hat{b}_{\Sigma 0} - b_0)^H} \, . \tag{5-6}$$

The components of the whitened estimator array are independent, and all have variance unity.

The non-adaptive decision rule, given by Equation (3-17), assumes the simple form

$$\lambda = \mathrm{Tr}(\hat{b}_{\Sigma 0}^H \hat{b}_{\Sigma 0}) \geq \mathrm{Const} \, ,$$

in terms of the whitened signal parameter estimator. The test statistic is thus equal to the sum of the squared magnitudes of the elements of this matrix. Statistically, λ is a non-central complex chi-squared random variable, with JM complex degrees of freedom, according to the usage introduced in Appendix 2.

The "non-centrality" parameter of this distribution is

$$a_0 \equiv \mathrm{Tr}(b_0^H b_0) \, . \tag{5-7}$$

We call this quantity the non-adaptive signal-to-noise ratio. To express it in terms of the original variables of the problem, we write

$$b_0^H b_0 = b^H \Sigma^{AA} b = b^H e^H \Sigma^{-1} e b \, ,$$

and note that

$$e b = \sigma B \tau p^H = \sigma B (\tau \tau^H)^{1/2} \, .$$

Then, we have

$$b_0^H b_0 = (\tau \tau^H)^{1/2} B^H \sigma^H \Sigma^{-1} \sigma B (\tau \tau^H)^{1/2} \, . \tag{5-8}$$

and, finally,

$$a_0 = \mathrm{Tr}\left[(\sigma B \tau)^H \Sigma^{-1} (\sigma B \tau)\right] . \tag{5-9}$$

In the special case $M=1$, $\tau\tau^H$ is a scalar, the squared norm of the τ vector. As noted in Section 1, this vector can be normalized to unity by a redefinition of the B array. If this is done, we will have

$$a_0 = b_0^H b_0 = B^H \sigma^H \Sigma^{-1} \sigma B .$$

Moreover, if $J=1$, then B itself is a scalar, and the signal-to-noise ratio reduces to the familiar form

$$a_0 = |B|^2 \sigma^H \Sigma^{-1} \sigma .$$

In radar terms, the test statistic is a non-coherent integrator of JM complex samples, and its pdf is the non-central complex chi-squared distribution, which is discussed in Appendix 2. The detection probability is given by the Marcum Q-function:[24]

$$P_D \equiv \mathrm{Prob}(\lambda \geq \lambda_0) = \int_{\lambda_0}^{\infty} e^{-a_0-\lambda} (\lambda/a_0)^{(JM-1)/2} I_{JM-1}(2\sqrt{a_0\lambda})\, d\lambda . \tag{5-10}$$

The corresponding probability of false alarm is

$$PFA = G_{JM}(\lambda_0) , \tag{5-11}$$

where

$$G_m(y) \equiv e^{-y} \sum_{k=0}^{m-1} \frac{y^k}{k!} \tag{5-12}$$

is the incomplete Gamma function, introduced in Appendix 2.

We return to the adaptive problem and recall that the adaptive parameter array estimator, found in Section 2, has the form:

$$\hat{b} \equiv (e^H S^{-1} e)^{-1} e^H S^{-1} Z_p , \tag{5-13}$$

which is just like Equation (5-1), with S replacing Σ. The matrix S, of course, is (L − M) times the ML estimator of Σ, based on the Z_q component of the data array alone, as expressed by Equation (2-48) of Section 2. The proportionality constant will cancel out in the above expression for the amplitude parameter estimator. This estimator was later shown to be identical to the array Y, introduced in Section 3 [see Equation (3-13)]. Under conditioning on the B components of the data array, we found that this array is Gaussian, with conditional mean and covariance matrices given by Equations (3-23) and (3-27), respectively:

$$E_B \hat{b} = b$$

$$\mathrm{Cov}_B(\hat{b}) = (\Sigma^{AA})^{-1} \otimes C_M^* .$$

We introduce the whitened estimator

$$\hat{b}_0 \equiv (\Sigma^{AA})^{1/2} \hat{b} , \tag{5-14}$$

as in the non-adaptive case. Its conditional mean is

$$E_B \hat{b}_0 = b_0 , \tag{5-15}$$

and the corresponding conditional covariance matrix is

$$\mathrm{Cov}_B(\hat{b}_0) = I_J \otimes C_M^* . \tag{5-16}$$

The b_0 array is the whitened signal parameter array defined in Equation (5-4), and the matrix C_M (defined in Section 3) is

$$C_M = I_M + Z_B^H S_{BB}^{-1} Z_B . \tag{5-17}$$

In accordance with the usage begun in Section 4, we have dropped the subscript zero (which indicated whitening) on the B arrays in this definition. In the notation of Section 4, we have

$$C_M = \mathscr{C}(N-J, M, J+K) , \tag{5-18}$$

as noted there. It will be recalled that $K = L - N - M$. The conditional mean of our estimator is independent of the conditioning variables, hence it remains an unbiased estimator (like its non-adaptive counterpart) when the conditioning is removed:

$$E\hat{b}_0 = b_0 . \tag{5-19}$$

The unconditioned covariance matrix may be evaluated from the equation

$$\mathrm{Cov}(\hat{b}_0) = I_J \otimes (EC_M)^* ,$$

obtained by taking the expected value of both sides of Equation (5-16). The required expected value of C_M was found in Section 4, and Equation (4-44) [together with Equation (5-18) above] yields

$$EC_M = \left(1 + \frac{N-J}{J+K}\right) I_M .$$

Finally, we obtain

$$\mathrm{Cov}(\hat{b}_0) = \frac{K+N}{J+K} I_J \otimes I_M . \tag{5-20}$$

The removal of the conditioning has left us with uncorrelated columns for the parameter array estimator, but it is no longer Gaussian; hence, we cannot infer independence, as in the non-adaptive case. The relation between the covariance matrices in these cases is interesting. We have

$$\mathrm{Cov}(\hat{b}_0) = \frac{K+N}{J+K} \mathrm{Cov}(\hat{b}_{\Sigma 0}) ,$$

and the factor which connects them is generally greater than unity:

$$\frac{K+N}{J+K} = \frac{L-M}{L+J-M-N} \geq 1 .$$

Equality is attained when $J = N$, as we should expect, because in this special case the e array is unitary, and definitions (2-34) and (2-51) tell us that the estimators coincide in this case:

$$J = N: \quad \hat{b} = \hat{b}_\Sigma = e^H Z_p . \tag{5-21}$$

At the end of Section 3 we introduced a particular form of "mismatch," in which the signal component present in the actual data differs from the model on which the GLR detector and parameter estimator are based. In this model, the mean value of the data array is the form

$$E Z = D\tau = dp \, , \tag{5-22}$$

where D and d are $N \times M$ arrays, and τ and p have their usual meanings. The true parameter array now has a component b_A which is in the subspace defined by the signal model, and a component b_B in its orthogonal complement. These arrays, originally defined by Equations (3-44), are given by

$$b_A \equiv e^H d \, , \qquad b_B \equiv f^H d \, . \tag{5-23}$$

In order to assess the effects of this mismatch on parameter estimation, we introduce whitened versions of these signal components, as follows:

$$b_{A0} \equiv (\Sigma^{AA})^{1/2} (b_A - \Sigma_{AB} \Sigma_{BB}^{-1} b_B)$$

$$b_{B0} \equiv (\Sigma_{BB})^{-1/2} \, b_B \, . \tag{5-24}$$

These definitions are motivated by Equations (3-47) and (3-48) of Section 3, [illegible] become

$$V_{0s} = b_{A0} \, C_M^{-1/2} \tag{5-25}$$

and, again dropping the zero subscript on Z_B,

$$E Z_B = b_{B0} \, . \tag{5-26}$$

Recalling Equations (3-9) and (3-44) of Section 3, together with Equation (A1-8) of Appendix 1, it can be seen that

$$b_A - \Sigma_{AB} \Sigma_{BB}^{-1} b_B = b_A + (\Sigma^{AA})^{-1} \Sigma^{AB} b_B = (\Sigma^{AA})^{-1} e^H \Sigma^{-1} d \, ,$$

hence, we may write the first of Equations (5-24) in the form

$$b_{A0} = (\Sigma^{AA})^{-1/2} e^H \Sigma^{-1} d \, . \tag{5-27}$$

The conditional mean of the whitened parameter array estimator is

$$E_B \hat{b}_0 = b_{A0} . \tag{5-28}$$

which follows directly from Equation (3-46). Since this result does not depend on the values of the conditioning variables, Equation (5-28) expresses the unconditioned mean value array as well. The mean value of the original (unwhitened) estimator array is therefore given by

$$E \hat{b} = (\Sigma^{AA})^{-1/2} E \hat{b}_0 = (\Sigma^{AA})^{-1/2} b_{A0} . \tag{5-29}$$

Using Equation (5-27), together with the definition of Σ^{AA}, we obtain

$$E \hat{b} = (e^H \Sigma^{-1} e)^{-1} e^H \Sigma^{-1} d . \tag{5-30}$$

By way of comparison, we can evaluate the expected value of the non-adaptive parameter array estimator directly from Equation (5-1), using the fact that

$$Z_p = d p p^H = d .$$

We obtain

$$E \hat{b}_\Sigma = (e^H \Sigma^{-1} e)^{-1} e^H \Sigma^{-1} E Z_p = (e^H \Sigma^{-1} e)^{-1} e^H \Sigma^{-1} d , \tag{5-31}$$

which expresses the remarkable fact that the adaptive and non-adaptive parameter array estimators have the same expected values, even when the signals are not matched to the model in our original formulation.

Equation (5-16), which expresses the conditional covariance of the parameter estimator, is still valid in the presence of mismatch, whose effects will become apparent only when the conditioning is removed. To evaluate the expected value of C_M, we recall that Z_B and W_B are independent, and that their covariance matrices are

$$\begin{aligned} \mathrm{Cov}(Z_B) &= I_{N-J} \otimes I_M \\ \mathrm{Cov}(W_B) &= I_{N-J} \otimes I_{L-M} . \end{aligned} \tag{5-32}$$

Since

$$S_{BB} = W_B W_B^H .$$

it follows that S_{BB} is a complex Wishart matrix, of order $N-J$, with $L-M$ complex degrees of freedom. Following the convention established near the end of Section 3, it is understood that the covariance matrix of the underlying Gaussian array of the Wishart matrix is the identity. In the present case, this is expressed by Equation (5-32). Using Equation (4-45), we evaluate the mean of the trace of its inverse:

$$E\,\mathrm{Tr}\,(S_{BB}^{-1}) = \frac{N-J}{L+J-N-M} = \frac{N-J}{J+K} .$$

It is clear from the complex Wishart pdf, Equation (A3-10) of Appendix 3, that the expected value of any power of S_{BB} is proportional to the identity matrix. The argument is the same as that used in Section 4 to establish Equation (4-38), and we conclude that

$$E S_{BB}^{-1} = \frac{1}{J+K}\, I_{N-J} .$$

We can now evaluate the required expectation of both sides of Equation (5-16) when mismatch is present. First, we condition on Z_B in Equation (5-17), and then average over this array, to obtain

$$E C_M = I_M + \frac{1}{J+K}\, E(Z_B^H Z_B) .$$

But,

$$E(Z_B^H Z_B) = (E Z_B)^H (E Z_B) + E(Z_B - E Z_B)^H (Z_B - E Z_B) , \tag{5-33}$$

and EZ_B is given by Equation (5-26) above. The second term on the right of Equation (5-33) is evaluated as a special case of Equation (A1-42) of Appendix 1. In view of the covariance matrix, given in Equation (5-32), the result is

$$E(Z_B - E Z_B)^H (Z_B - E Z_B) = (N-J)\, I_M .$$

Combining these facts, we have the properties

$$E\hat{b}_0 = b_{A0}$$

$$\mathrm{Cov}(\hat{b}_0) = \frac{1}{J+K}\, I_J \otimes \left[(K+N)I_M + b_{B0}^{H} b_{B0}\right]^{*}, \tag{5-34}$$

which characterize the parameter estimator in the mismatched case. The estimator attempts to produce the component of the actual signal array which lies in the modeled subspace, and its performance is degraded by the effect of the orthogonal component of the signal array which increases its variance.

It is interesting to note that

$$b_{A0}^{H} b_{A0} + b_{B0}^{H} b_{B0} = (b_A - \Sigma_{AB}\Sigma_{BB}^{-1} b_B)^{H}\, \Sigma^{AA}\, (b_A - \Sigma_{AB}\Sigma_{BB}^{-1} b_B) + b_B^{H}\, \Sigma_{BB}^{-1}\, b_B$$

$$= \left[b_A^{H}\; b_B^{H}\right] \begin{bmatrix} \Sigma^{AA} & \Sigma^{AB} \\ \Sigma^{BA} & \Sigma^{BB} \end{bmatrix} \begin{bmatrix} b_A \\ b_B \end{bmatrix},$$

by application of Equation (A1-9). Moreover, we can write definitions (3-44) in the form

$$\begin{bmatrix} b_A \\ b_B \end{bmatrix} = U_N^{H} d, \tag{5-35}$$

where U_N is the unitary matrix defined by Equation (2-20). Then, recalling definition (3-9), we obtain

$$b_{A0}^{H} b_{A0} + b_{B0}^{H} b_{B0} = d^{H}\Sigma^{-1} d = (\tau\tau^{H})^{1/2}\, D^{H}\Sigma^{-1} D\, (\tau\tau^{H})^{1/2}. \tag{5-36}$$

In the matched case we have $D = \sigma B$, and Equation (5-36) then passes over into $b_0^{H} b_0$, as expressed by Equation (5-8) above. We return now to the matched problem, and its postulates are to be assumed throughout the ensuing discussion, except where the contrary is explicitly noted.

Before discussing the pdf of the amplitude parameter estimator, we recall the definition of the general $\mathscr{C}$ matrix [Equation (4-15)] and introduce the notation $\mathscr{R}$ for its inverse:

$$\mathcal{R}(J,M,K) \equiv \mathcal{C}(J,M,K)^{-1} = (I_M + V^H T^{-1} V)^{-1}$$

$$= I_M - V^H (T + V V^H)^{-1} V . \tag{5-37}$$

As in the definition of the $\mathcal{C}$ matrices, $\mathcal{R}$ is often used as a "generic" designator for a random quantity, not always a specific example. In the above definition, V is a zero-mean complex Gaussian array of dimension $J \times M$, with covariance

$$\mathrm{Cov}(V) = I_J \otimes I_M , \tag{5-38}$$

and T is a complex Wishart matrix of order J, with $J+K$ complex degrees of freedom. By analogy with C_M, we will write

$$R_M \equiv C_M^{-1} = \mathcal{R}(N-J, M, J+K) . \tag{5-39}$$

The general $\mathcal{R}$ matrix is a complex multivariate generalization of the complex central Beta random variable, and the joint pdf of its elements is derived in Appendix 3. We use the notation f_B for the probability density function of an $\mathcal{R}$ matrix, and $d_0(R)$ for the corresponding volume element. This pdf depends only on the dimensional parameters J, M, and K, and, when $M=1$, it reduces to the ordinary scalar complex Beta pdf (see Appendix 3 for details). The volume element is specific to positive-definite matrices, and it is the same as the volume element for the complex Wishart pdf. The notation is defined in Appendix 3. If Φ is a function of the random matrix $\mathcal{R}$, then we can evaluate its expected value by integrating over the appropriate pdf:

$$E\,\Phi[\mathcal{R}(J,M,K)] = \int \Phi(R)\, f_B(R; M, K+M, J)\, d_0(R) . \tag{5-40}$$

In the special cases to be discussed later, this Beta matrix will reduce to a complex scalar Beta variable, and the integration will be a simple, one-dimensional integral over the complex (scalar) Beta density.

The $\mathcal{R}$ matrices have some interesting properties, two of which will be established here and used presently. Let U_M be a unitary matrix of order M, which is partitioned as follows:

$$U_M = \begin{bmatrix} r \\ s \end{bmatrix} .$$

We assume that r is $M_1 \times M$, s is $M_2 \times M$, and that the sum of M_1 and M_2 is M. The V array is also partitioned, using U_M:

$$V U_M^H = [V_1 \ V_2] ,$$

where

$$V_1 \equiv V r^H$$

$$V_2 \equiv V s^H . \qquad (5\text{-}41)$$

The new components are complex Gaussian arrays with zero means, and with covariances

$$\mathrm{Cov}(V_1) = I_J \otimes I_{M_1}$$

$$\mathrm{Cov}(V_2) = I_J \otimes I_{M_2} . \qquad (5\text{-}42)$$

We note that

$$V V^H = V_1 V_1^H + V_2 V_2^H , \qquad (5\text{-}43)$$

and consider the matrix

$$s\,\mathcal{R}(J,M,K)\,s^H = s\,s^H - V_2^H (T + V V^H)^{-1} V_2 . \qquad (5\text{-}44)$$

From the unitary character of U_M, we have

$$s\,s^H = I_{M_2} .$$

Then, using Equation (5-43), we can write

$$s\mathcal{R}(J,M,K)s^H = I_{M_2} - V_2^H(T+V_1V_1^H+V_2V_2^H)^{-1}V_2$$

$$= [I_{M_2}+V_2^H(T+V_1V_1^H)^{-1}V_2]^{-1} .$$

The complex Wishart matrix T can be expressed in terms of a zero-mean complex Gaussian array $\mathcal{W}$:

$$T = \mathcal{W}\mathcal{W}^H ,$$

where

$$\mathrm{Cov}(\mathcal{W}) = I_J \otimes I_{J+K} .$$

It follows that

$$T+V_1V_1^H = [\,\mathcal{W} \;\; V_1\,]\begin{bmatrix}\mathcal{W}^H\\ V_1^H\end{bmatrix}$$

is also a complex Wishart matrix, of order J, and with $J+K+M_1$ complex degrees of freedom. Since the covariance of the V_2 component is given in Equation (5-42), we have therefore shown that

$$s\mathcal{R}(J,M,K)s^H = \mathcal{R}(J,M_2,K+M_1) . \tag{5-45}$$

Recall that s is $M_2 \times M$ in dimension, and that $M_1 = M - M_2$. In this equation, as in others which relate generic random variables, the equality sign refers to statistical identity, or equality of the corresponding probability density functions.

The second property concerns the determinant of an $\mathcal{R}$ matrix, which has the form of the inverse of the GLR test statistic in the signal-free case, as discussed in Section 4:

$$|\mathcal{R}(J,M,K)| = 1/\ell(J,M,K) . \tag{5-46}$$

As shown in Appendix 3, by a simple factoring of the determinants,

$$\ell(J,M,K) = \ell(J,M_2,K+M_1)\,\ell(J,M_1,K) . \tag{5-47}$$

This is Equation (A3-63) of Appendix 3, where it is further established that the two factors on the right side of this equation are statistically independent. The same applies, obviously, to their inverses, and we can therefore write the determinant of a general $\mathcal{R}$ matrix as a product of independent factors:

$$|\mathcal{R}(J,M,K)| = |\mathcal{R}(J,M_2,K+M_1)|\,|\mathcal{R}(J,M_1,K)| \,. \tag{5-48}$$

We resume our discussion of the parameter array estimator, in its whitened form, and define the estimation error array:

$$\xi \equiv \hat{b}_0 - E\,\hat{b}_0 = \hat{b}_0 - b_0 \,. \tag{5-49}$$

We exclude the special case $J = N$, because in this situation the adaptive estimator coincides with the non-adaptive one, as we have already noted. There are no B components when $J = N$, the C_M matrix reduces to the identity, and the pdf of the estimation error [see Equation (5-6)] takes the simple form

$$f(\xi) = \frac{1}{\pi^{NM}}\, e^{-\mathrm{Tr}(\xi^H \xi)} \tag{5-50}$$

in this case.

In general, the expected value of ξ is zero, and its conditional covariance is given by the right side of Equation (5-16). In terms of R_M, we may write it as

$$\mathrm{Cov}_B(\xi) = I_J \otimes (R_M^{-1})^* \,,$$

and then the conditional pdf of ξ becomes

$$f(\xi|R_M) = \frac{1}{\pi^{JM}}\, |R_M|^J\, e^{-\mathrm{Tr}(R_M \xi^H \xi)} \,. \tag{5-51}$$

This form of the multivariate Gaussian distribution is a special case of Equation (A1-62) of Appendix 1, and we have indicated the conditioning variables as the components of R_M itself, since it is only through them that the B components survive. The unconditioned pdf of ξ can therefore be expressed as the integral over the appropriate density of R_M:

$$f(\xi) = \frac{1}{\pi^{JM}} \int |R|^J e^{-\mathrm{Tr}(R\xi^H\xi)} f_B(R; M, J+M+K, N-J)\, d_0(R) . \qquad (5\text{-}52)$$

This is, of course, the pdf of the whitened parameter estimator array, and it can depend only on the dimensional parameters of our model.

It is also clear that $f(\xi)$ depends on the estimation error only through the product $\xi^H\xi$. In fact, it can depend only on the non-zero eigenvalues of this matrix, and these, of course, are the squares of the singular values of ξ itself. To prove this assertion, we express $\xi^H\xi$ in terms of its eigenvalues λ_m as follows:

$$\xi^H\xi = \mathcal{U}\Lambda\mathcal{U}^H ,$$

where $\mathcal{U}$ is unitary, of order M, and

$$\Lambda = \mathrm{Diag}[\lambda_1, \ldots, \lambda_M] .$$

In the conditional pdf we have

$$\mathrm{Tr}(R_M \xi^H \xi) = \mathrm{Tr}(\mathcal{U}^H R_M \mathcal{U} \Lambda) ,$$

and, of course,

$$|R_M| = |\mathcal{U}^H R_M \mathcal{U}| .$$

From its definition, we see that $\mathcal{R}$ is statistically indistinguishable from $\mathcal{U}^H\mathcal{R}\mathcal{U}$, since the latter is expressible as an $\mathcal{R}$ matrix in terms of $V\mathcal{U}$, which is statistically identical to V. Thus, the pdf of ξ depends on ξ only through Λ.

If signal mismatch is present, the ξ array is defined by the equation

$$\xi = \hat{b}_0 - E\,\hat{b}_0 = \hat{b}_0 - b_{A0} , \qquad (5\text{-}53)$$

so that it still has zero mean. In addition, R_M is now a particular example of the non-central generalization of the $\mathcal{R}$ matrix. R_M is the inverse of C_M, defined in Equation (5-17), and the non-centrality arises from the non-vanishing mean of the Z_B

array, which is now given by Equation (5-26). The effect of mismatch on the pdf of the estimator will be discussed later, in connection with a special case in which R_M reduces to a non-central complex (scalar) Beta variable.

If $J \geq M$, the matrix $\xi^H \xi$ will have full rank, except for a set of measure zero in the ordinary Euclidean sense represented by the volume element $d(\xi)$. Equation (5-52) provides a convenient starting point for the study of the unconditioned pdf of ξ in this situation. On the other hand, if $J < M$ the product $\xi\xi^H$ will have full rank, in the sense described above, and an alternative form of the conditional pdf of ξ can then be obtained. This form will be more convenient because it will involve an $\mathcal{R}$ matrix of lower order. To obtain this form, we introduce the array

$$s \equiv (\xi\xi^H)^{-1/2} \xi \,, \tag{5-54}$$

which has the familiar properties

$$\begin{aligned} s s^H &= I_J \\ s^H s &= \xi^H (\xi\xi^H)^{-1} \xi \\ \xi &= (\xi\xi^H)^{1/2} s \,. \end{aligned} \tag{5-55}$$

The orthonormal rows of s form a basis in the row space of ξ. The orthogonal complement of this space, which has dimension $M - J$, is given a basis array r which, together with s, forms a unitary matrix:

$$\begin{bmatrix} r \\ s \end{bmatrix} = U_M$$

in the standard way. Expressing ξ in terms of s, we have

$$\mathrm{Tr}(R_M \xi^H \xi) = \mathrm{Tr}(s R_M s^H \xi\xi^H) \,, \tag{5-56}$$

and the first property of the $\mathcal{R}$ matrices, derived above, may be applied. In the present application, $M_2 = J$ and $M_1 = M - J$; therefore,

$$s R_M s^H = s \mathcal{R}(N-J, M, J+K) s^H = \mathcal{R}(N-J, J, M+K) \,. \tag{5-57}$$

Comparing this form with Equation (5-39), we note that J and M have been interchanged in the second and third arguments of the $\mathcal{R}$ matrices here. We make the definitions

$$R_\alpha \equiv \mathcal{R}(N-J,J,M+K)$$
$$R_\beta \equiv \mathcal{R}(N-J,M-J,J+K)\ , \tag{5-58}$$

to simplify the writing. Thus,

$$sR_M s^H = R_\alpha$$
$$|R_M| = |R_\alpha||R_\beta|\ , \tag{5-59}$$

and

$$f(\xi|R_\alpha,R_\beta) = \frac{1}{\pi^{JM}}|R_\beta|^J|R_\alpha|^J e^{-\mathrm{Tr}(R_\alpha \xi\xi^H)}\ . \tag{5-60}$$

Since the factors on the right side of the second of Equations (5-59) are independent, we can average over R_β to obtain a form of the pdf which is conditioned only on R_α. Using Equation (4-36), we have

$$|R_\beta| = 1/\ell(N-J,M-J,J+K)$$
$$= \prod_{j=1}^{N-J}\prod_{m=1}^{M-J} x_\beta(J+K+j+m-1,1)\ .$$

All the complex Beta variables in this double product are independent, and it is easily shown from the complex central Beta density [Equation (A2-12)] that

$$E[x_\beta(n,1)]^J = \frac{n}{n+J}\ .$$

When applied to our problem, we get

$$\Psi \equiv E|R_\beta|^J = \prod_{j=1}^{N-J}\prod_{m=1}^{M-J} \frac{J+K+j+m-1}{2J+K+j+m-1}$$

$$= \prod_{j=0}^{N-J-1} \frac{(M+K+j)!\,(2J+K+j)!}{(J+K+j)!\,(J+M+K+j)!} \,. \tag{5-61}$$

This evaluation has given us the following expression for the conditional pdf of the estimation error, valid when the indicated inequality is satisfied:

$$J < M: \qquad f(\xi|R_\alpha) = \frac{\Psi}{\pi^{JM}} |R_\alpha|^J e^{-\mathrm{Tr}(R_\alpha \xi\xi^H)} \,, \tag{5-62}$$

and the corresponding unconditioned pdf of ξ is then

$$f(\xi) = \frac{\Psi}{\pi^{JM}} \int |R|^J e^{-\mathrm{Tr}(R\xi\xi^H)} f_B(R;J,J+M+K,N-J)\, d_0(R) \,. \tag{5-63}$$

It is established in Appendix 3 [see Equation (A3-57)] that

$$|R|^n f_B(R;M,K,J) = \prod_{j=0}^{J-1} \frac{(K+j)!\,(K-M+n+j)!}{(K-M+j)!\,(K+n+j)!} f_B(R;M,K+n,J) \,, \tag{5-64}$$

which holds for negative values of n, so long as $K-M+n$ is non-negative. When this identity is applied to our example, we obtain

$$\Psi f_B(R;J,J+M+K,N-J) = |R|^{M-J} f_B(R;J,2J+K,N-J) \,,$$

and, consequently, Equation (5-63) can be written in the form

$$f(\xi) = \frac{1}{\pi^{JM}} \int |R|^M e^{-\mathrm{Tr}(R\xi\xi^H)} {}_B(R;J,2J+K,N-J)\, d_0(R) \,. \tag{5-65}$$

We have obtained this result under the assumption that $J < M$. However, it is also true when $J = M$, in which case it may be seen that Equations (5-52) and (5-65) differ only in the argument of the *trace* operator, which appears in the exponential factor. But when J and M are equal, ξ is square and invertible (except for a set of zero measure) in the sense referred to earlier. It follows from Equations (5-55) that the array s, now square, is unitary. We have already seen that such a unitary transformation may be applied to an R matrix with no effect on its statistical properties, and Equation (5-56) tells us that interchanging the order of the factors ξ and ξ^H in the argument of the *trace* is equivalent to subjecting R_M to such a unitary transformation. The determinant of R_M is also unaltered by this unitary transformation, as we have observed already. Equation (5-65) is therefore obtained directly, without the need to factor the R matrix explicitly, and this completes the proof of our assertion.

The analysis which has led us to Equations (5-52) and (5-65) made use of an intermediate stage of conditioning (on the B components of the data array) which was originally introduced in Section 3. This method is particularly appropriate for the analysis of the GLR test statistic itself. However, another technique can be employed to obtain a formula for the conditional pdf of the estimation error array. This approach leads directly to Equation (5-65), but without the restriction on the relative values of J and M, and it is presented here as an interesting alternative.

We start from Equation (5-13), as before, and write it in the form

$$\hat{b} = (e^H S^{-1} e)^{-1} e^H S^{-1} Z_p = w^H Z_p , \tag{5-66}$$

where w is a "weight array," given by

$$w \equiv S^{-1} e \, (e^H S^{-1} e)^{-1} . \tag{5-67}$$

This array is of dimension $N \times J$, and it has the property that

$$e^H w = I_J .$$

We recall that

$$Z_p = Z p^H ,$$

and that the mean and covariance of the original data array are

$$EZ = ebp$$
$$\mathrm{Cov}(Z) = \Sigma \otimes I_L . \qquad (5\text{-}68)$$

Z_p is a complex Gaussian array, of course, with mean and covariance given by

$$EZ_p = eb$$
$$\mathrm{Cov}(Z_p) = \Sigma \otimes I_M . \qquad (5\text{-}69)$$

The covariance has been evaluated using Equation (A1-44) of Appendix 1.

In the new technique, we condition on the Z_q array instead of the R components, and we indicate this by a subscript q. Since

$$S = Z_q Z_q^H ,$$

the S matrix and the weight array w are fixed under this conditioning. The form of Equation (5-66) makes this a natural step in the analysis of the statistical properties of the estimator of the b array. Under the new conditioning, this estimator is obviously a complex Gaussian array, with conditioned mean and covariance given by

$$E_q \hat{b} = w^H E Z_p = w^H eb = b$$
$$\mathrm{Cov}_q(\hat{b}) = C_b \otimes I_M ,$$

where

$$C_b \equiv w^H \Sigma w$$
$$= (e^H S^{-1} e)^{-1} e^H S^{-1} \Sigma S^{-1} e \, (e^H S^{-1} e)^{-1} . \qquad (5\text{-}70)$$

The conditional pdf of the estimator array is therefore

$$f_q(\hat{b}) = \frac{1}{\pi^{JM} |C_b|^M} e^{-\mathrm{Tr}[C_b^{-1}(\hat{b}-b)(\hat{b}-b)^H]} . \qquad (5\text{-}71)$$

The conditioning variables survive only in the matrix C_b, whose statistical properties we now examine. The first step is a whitening transformation, in which we introduce the array

$$Z_0 \equiv \Sigma^{-1/2} Z_q . \tag{5-72}$$

Like Z_q, this is a complex Gaussian array with zero mean, but with covariance matrix

$$\mathrm{Cov}(Z_0) = I_N \otimes I_{L-M} .$$

We also introduced the whitened version of the S matrix:

$$S_0 \equiv Z_0 Z_0^H , \tag{5-73}$$

which obeys a complex Wishart distribution, and which, like S, is invertible with probability one.

Let e_0 be a whitened version of the e array:

$$e_0 \equiv \Sigma^{-1/2} e . \tag{5-74}$$

This array is no longer a basis array, and its column space is different from that of the original e (or σ) array. In terms of e_0, we have

$$e^H S^{-1} e = e_0^H S_0^{-1} e_0 ,$$

and

$$C_b = (e_0^H S_0^{-1} e_0)^{-1} \; e_0^H S_0^{-2} e_0 \; (e_0^H S_0^{-1} e_0)^{-1} . \tag{5-75}$$

From the definition of e_0 we make the evaluation

$$e_0^H e_0 = e^H \Sigma^{-1} e = \Sigma^{AA} ,$$

which is a positive-definite matrix of order J. We can establish a basis array in the column space of e_0 by the standard procedure, introducing the array

$$e_1 \equiv e_0 (e_0^H e_0)^{-1/2} = e_0 (\Sigma^{AA})^{-1/2} . \tag{5-76}$$

This development parallels the introduction of e itself from the original array σ, and we obtain the following identities directly:

$$e_1^H e_1 = I_J$$

$$e_1 e_1^H = e_0 (e_0^H e_0)^{-1} e_0^H$$

$$e_0 = e_1 (\Sigma^{AA})^{1/2} . \tag{5-77}$$

Continuing the analogy, we let f_1 be a basis array in the orthogonal complement of the column space of e_0, and form the unitary matrix

$$\mathcal{U}_N \equiv [\, e_1 \;\; f_1 \,] . \tag{5-78}$$

We use this matrix to transform and partition the Z_0 array:

$$\mathcal{U}_N^H Z_0 = \begin{bmatrix} e_1^H Z_0 \\ f_1^H Z_0 \end{bmatrix} \equiv \begin{bmatrix} X_A \\ X_B \end{bmatrix} , \tag{5-79}$$

the matrix S_0:

$$\mathcal{U}_N^H S_0 \mathcal{U}_N \equiv \begin{bmatrix} \mathcal{S}_{AA} & \mathcal{S}_{AB} \\ \mathcal{S}_{BA} & \mathcal{S}_{BB} \end{bmatrix} = \begin{bmatrix} X_A X_A^H & X_A X_B^H \\ X_B X_A^H & X_B X_B^H \end{bmatrix} , \tag{5-80}$$

and its inverse:

$$\mathcal{U}_N^H S_0^{-1} \mathcal{U}_N \equiv \begin{bmatrix} \mathcal{S}^{AA} & \mathcal{S}^{AB} \\ \mathcal{S}^{BA} & \mathcal{S}^{BB} \end{bmatrix} . \tag{5-81}$$

According to the third of Equations (5-77), we have

$$e_0^H S_0^{-1} e_0 = (\Sigma^{AA})^{1/2}\ e_1^H S_0^{-1} e_1\ (\Sigma^{AA})^{1/2}\ ,$$

and

$$e_0^H S_0^{-2} e_0 = (\Sigma^{AA})^{1/2}\ e_1^H S_0^{-2} e_1\ (\Sigma^{AA})^{1/2}\ .$$

Then, substituting in Equation (5-75), we obtain

$$C_b = (\Sigma^{AA})^{-1/2}\ C_0\ (\Sigma^{AA})^{-1/2}\ , \tag{5-82}$$

where

$$C_0 \equiv (e_1^H S_0^{-1} e_1)^{-1}\ e_1^H S_0^{-2} e_1\ (e_1^H S_0^{-1} e_1)^{-1}\ . \tag{5-83}$$

We make use of Equation (5-81) to express C_0 in terms of the new partitioned components:

$$\begin{aligned} C_0 &= (\mathcal{S}^{AA})^{-1}\ (\mathcal{S}^{AA}\mathcal{S}^{AA} + \mathcal{S}^{AB}\mathcal{S}^{BA})\ (\mathcal{S}^{AA})^{-1} \\ &= I_J + (\mathcal{S}^{AA})^{-1}\ \mathcal{S}^{AB}\mathcal{S}^{BA}\ (\mathcal{S}^{AA})^{-1}\ . \end{aligned}$$

In view of the identities contained in Equation (A1-8), this expression is equivalent to

$$C_0 = I_J + \mathcal{S}_{AB}\ \mathcal{S}_{BB}^{-2}\ \mathcal{S}_{BA}\ .$$

The statistical properties of C_0 do not depend on the true covariance matrix Σ. In fact, they can depend only on the dimensional parameters of the problem. We will derive these properties shortly, but first we wish to express the conditional pdf, Equation (5-71), in terms of C_0. From Equation (5-82), it follows directly that

$$|C_b| = |C_0|\ |\Sigma^{AA}|^{-1}\ . \tag{5-84}$$

and

$$\mathrm{Tr}\left[C_b^{-1}(\hat{b} - b)(\hat{b} - b)^H\right] = \mathrm{Tr}\left[C_0^{-1}(\Sigma^{AA})^{1/2}(\hat{b} - b)(\hat{b} - b)^H(\Sigma^{AA})^{1/2}\right]\ . \tag{5-85}$$

But

$$(\Sigma^{AA})^{1/2}(\hat{b} - b) = \hat{b}_0 - b_0 = \xi \; , \tag{5-86}$$

according to definitions (5-4) and (5-14), hence we can obtain the conditional pdf of the ξ array itself. We can view Equation (5-86) as a change of the variables of integration, and use Equation (A1-66) of Appendix 1 to find the appropriate Jacobian:

$$d(\hat{b}) = |\Sigma^{AA}|^{-M} \, d(\xi) \; .$$

Combining these results, we get

$$f_q(\xi) = \frac{1}{\pi^{JM} |C_0|^M} \, e^{-\mathrm{Tr}(C_0^{-1} \xi \xi^H)} \; . \tag{5-87}$$

Returning to the C_0 array, we define

$$T \equiv \mathscr{S}_{BB} = X_B X_B^H \; , \tag{5-88}$$

which is a complex Wishart matrix of order $N - J$, since X_B is a zero-mean complex Gaussian array, whose covariance matrix is easily found to be

$$\mathrm{Cov}(X_B) = I_{N-J} \otimes I_{L-M} \; . \tag{5-89}$$

From this property it follows that T has $L - M = N + K$ complex degrees of freedom. We also define

$$\mathcal{V} \equiv \mathscr{S}_{AB} (\mathscr{S}_{BB})^{-1/2} \; , \tag{5-90}$$

which has dimension $J \times (N - J)$, and then we can write

$$C_0 = I_J + \mathcal{V} \, T^{-1} \mathcal{V}^H \; .$$

$\mathcal{V}$ is a function of the arrays X_A and X_B which are, of course, complex Gaussian arrays with zero means. The covariance matrix of X_A is

$$\mathrm{Cov}(X_A) = I_J \otimes I_{L-M} \; .$$

We write $\mathcal{V}$ in the form

$$\mathcal{V} = X_A Q ,$$

where

$$Q \equiv X_B^H (\mathcal{S}_{BB})^{-1/2} . \tag{5-91}$$

If we condition on the elements of X_B, Q will be a constant array and $\mathcal{V}$ will be conditionally Gaussian, with zero mean and conditional covariance matrix

$$\text{Cov}_B(\mathcal{V}) = I_J \otimes (Q^H Q)^* ,$$

using Equation (A1-44) of Appendix 1. The subscript B is intended to indicate conditioning on X_B. But

$$Q^H Q = (\mathcal{S}_{BB})^{-1/2} X_B X_B^H (\mathcal{S}_{BB})^{-1/2} = I_{N-J} . \tag{5-92}$$

hence,

$$\text{Cov}_B(\mathcal{V}) = I_J \otimes I_{N-J} . \tag{5-93}$$

The $\mathcal{V}$ array has been shown to be conditionally Gaussian, with a mean array and a covariance matrix which do not depend on the conditioning variables. Hence, $\mathcal{V}$ has the same statistical properties without the conditioning, and this is now removed.

Finally, we replace $\mathcal{V}$ by its Hermitian transpose, making the definition

$$V \equiv \mathcal{V}^H . \tag{5-94}$$

Then, V is a zero-mean complex Gaussian array, with covariance matrix

$$\text{Cov}(V) = I_{N-J} \otimes I_J , \tag{5-95}$$

and C_0 can be written

$$C_0 = I_J + V^H T^{-1} V .$$

This is clearly a random $\mathscr{C}$ matrix, of the kind defined in Section 4. Its parameters are determined from the definitions of that section, together with Equations (5-88), (5-89), and (5-95). Since $(N+K)-(N-J) = J+K$, we obtain

$$C_0 = \mathscr{C}(N-J, J, J+K) . \tag{5-96}$$

The inverse of C_0, which we call R_γ, is an $\mathscr{R}$ matrix:

$$R_\gamma \equiv C_0^{-1} = \mathscr{R}(N-J, J, J+K) . \tag{5-97}$$

The conditional pdf of ξ is therefore given by

$$f_q(\xi) = f(\xi | R_\gamma) = \frac{1}{\pi^{JM}} |R_\gamma|^M e^{-\mathrm{Tr}(R_\gamma \xi \xi^H)} , \tag{5-98}$$

and the unconditioned pdf is therefore

$$f(\xi) = \frac{1}{\pi^{JM}} \int |R|^M e^{-\mathrm{Tr}(R \xi \xi^H)} f_B(R; J, 2J+K, N-J) \, d_0(R) . \tag{5-99}$$

This is identical to Equation (5-65), but it is valid for all values of J and M which are permissible in the general formulation of Section 1.

Using the apparatus of Appendix 3, it is possible (when $J \geq M$) to integrate out the extraneous variables in Equation (5-51) in order to obtain a formula for the conditional pdf of the elements of $\xi^H \xi$ itself, which is positive definite under this assumption. A similar formula can be derived [from Equation (5-98)] for the conditional pdf of the elements of $\xi \xi^H$, which is positive definite when $J \leq M$. To give expression to these conditional densities, we define the matrices

$$A \equiv \xi^H \xi$$

$$A' \equiv \xi \xi^H . \tag{5-100}$$

The conditional pdf of A then assumes the form

$$J \geq M: \quad g(A | R_M) = \frac{1}{\Gamma_M(J)} |A|^{J-M} |R_M|^J e^{-\mathrm{Tr}(R_M A)} , \tag{5-101}$$

and that of A' becomes

$$J \le M: \qquad g(A'|R_\gamma) = \frac{1}{\Gamma_J(M)} |A'|^{M-J} |R_\gamma|^M e^{-\mathrm{Tr}(R_\gamma A')} . \tag{5-102}$$

The associated volume elements are $d_0(A)$ and $d_0(A')$, respectively. As noted earlier, this is the same volume element used in connection with the Wishart pdf. The normalization factor $\Gamma_n(m)$ is defined in Appendix 3 [Equation (A3-8)]; it is a multivariate generalization of the Gamma function. The unconditioned densities of A and A' are expressed as the following integrals:

$$J \ge M: \qquad g(A) = \frac{|A|^{J-M}}{\Gamma_M(J)} \int |R|^J e^{-\mathrm{Tr}(RA)} f_B(R;M,J+M+K,N-J)\, d_0(R)$$

$$J \le M: \qquad g(A') = \frac{|A'|^{M-J}}{\Gamma_J(M)} \int |R|^M e^{-\mathrm{Tr}(RA')} f_B(R;J,2J+K,N-J)\, d_0(R) . \tag{5-103}$$

To get explicit results for the unconditional pdf of the estimation error array, we must specialize to either of the cases: $J=1$, M arbitrary, or $M=1$, J arbitrary. We note that the original parameter array B has rank unity in these situations, and we anticipate that only in these special cases will we find explicit results for the probability of detection.

We consider the case $M=1$ first, and recall that J must be less than N, but is otherwise arbitrary. In this specialization of the signal model, τ becomes a row vector which distributes the signal among the columns of the data array Z with known relative amplitudes. If Z is post-multiplied by a suitable unitary matrix, τ can be converted into a vector all of whose components are zero except the first. The value of this first component can then be factored from τ, and incorporated into a redefined B array. The general problem with $M=1$ is thus equivalent to the special choice

$$\tau = [1,0,\ldots,0] .$$

In this model, the signal is confined to the first column of Z, which becomes synonymous with Z_p, defined in Section 2. The remaining components comprise the Z_q array. The signal itself is any vector in a given J-dimensional subspace of $\mathbb{C}^N$, and B is a column vector of dimension J. These specific transformations have been mentioned only to illuminate the special case at hand; in the following discussion, we do not assume that they have been made.

When $M=1$, the $\mathcal{R}$ matrix of definition (5-37) becomes a scalar. The relation expressed by Equation (5-46) is then simply

$$\mathcal{R}(J,1,K) = 1/\ell(J,1,K) \ . \tag{5-104}$$

We can therefore make use of Equation (4-17) of Section 4 to obtain the statistical character of R_M in this case:

$$\begin{aligned} R_1 &= \mathcal{R}(N-J,1,J+K) = 1/\ell(N-J,1,J+K) \\ &= x_\beta(J+K+1,N-J) \ . \end{aligned} \tag{5-105}$$

The same result can be obtained by specialization of the complex multivariate Beta distribution, given by Equation (A3-53), which becomes a complex scalar Beta variable as indicated in Equation (A3-54). From Equation (5-52) we now obtain the unconditioned pdf of ξ as the integral:

$$f(\xi) = \frac{1}{\pi^J}\int_0^1 \rho^J e^{-\rho\xi^H\xi} f_\beta(\rho;J+K+1,N-J)\,d\rho \ . \tag{5-106}$$

The complex central Beta pdf which enters this formula is defined in Equation (A2-12). Note that $N-J$ is positive, so there will be no difficulty at the upper limit of this integral.

The estimation error is a J vector in this case, and $A=\xi^H\xi$ is a scalar, the square of its norm. According to Equation (5-106), the pdf of ξ is a spherically symmetric function in $\mathbb{C}^J$, depending only on A. By setting $M=1$ in the first of Equations (5-103), we obtain the pdf of A directly:

$$g(A) = \frac{A^{J-1}}{(J-1)!}\int_0^1 \rho^J e^{-\rho A} f_\beta(\rho;J+K+1,N-J)\,d\rho \ . \tag{5-107}$$

Alternatively, one can introduce spherical coordinates in the 2J-dimensional real space corresponding to $\mathbb{C}^J$, and then integrate out the angle variables in Equation (5-106).

The result is a function of radial distance only, and the quantity A is the square of this radius. The procedure just described is exactly that to which the integration theorem, used in the derivation of Equation (5-101), reduces when $M=1$. It is also the starting point for our inductive proof of the theorem in Appendix 3.

The integration indicated in Equation (5-107) leads to a confluent hypergeometric function.[25] We introduce it here by means of an integral representation:[25,26]

$$\int_0^1 e^{\rho x} f_\beta(\rho;n,m)\,d\rho = {}_1F_1(n;n+m;x) \,,$$

which is valid when n and m are positive integers. More relevant to our needs is the formula obtained when the variable of integration is changed from ρ to $1-\rho$:

$$\int_0^1 e^{-\rho x} f_\beta(\rho;m,n)\,d\rho = e^{-x}\,{}_1F_1(n;n+m;x) \,. \qquad (5\text{-}108)$$

The effect of the change of variable on the complex Beta density function is to interchange its parameters, an obvious consequence of its definition. The process we have just carried out is equivalent to Kummer's first transformation of the confluent hypergeometric function, Equation (A2-21) of Appendix 2.

Another property of the complex Beta pdf is

$$\rho^k f_\beta(\rho;n,m) = \frac{(n+m-1)!\,(n+k-1)!}{(n-1)!\,(n+m+k-1)!}\, f_\beta(\rho;n+k,m) \,, \qquad (5\text{-}109)$$

which is easily verified from the definition of this function. This formula holds for negative integral k as well, as long as $n+k$ is positive, and it represents a special case of Equation (5-64).

Returning to integral (5-107), we apply Equation (5-109) to obtain

$$\rho^J f_\beta(\rho;J+K+1,N-J) = \frac{(K+N)!\,(2J+K)!}{(J+K)!\,(J+K+N)!}\, f_\beta(\rho;2J+K+1,N-J) \,,$$

and then make use of Equation (5-108). The result is

$$g(A) = \frac{(K+N)!(2J+K)!}{(J+K)!(J+K+N)!} \frac{A^{J-1}}{(J-1)!} e^{-A} \, {}_1F_1(N-J; J+K+N+1; A) \; . \tag{5-110}$$

The normalization of this pdf can be verified by using the formula

$$\int_0^\infty x^k e^{-x} \, {}_1F_1(n; n+m; x) dx = \frac{k!(n+m-1)!(m-k-2)!}{(m-1)!(n+m-k-2)!} \; ,$$

which holds when $m + k \geq 2$, and which follows from results already obtained.

If the first argument of a confluent hypergeometric function is $-k$, where k is a non-negative integer, then the function is expressible as a polynomial of order k. The general case is given as Equation (A2-22) of Appendix 2, and, in particular,

$${}_1F_1(0; m; x) = 1 \; . \tag{5-111}$$

If we formally put $J = N$ in Equation (5-110) and use this result, we obtain

$$g(A) = \frac{A^{N-1}}{(N-1)!} e^{-A} \; ,$$

which is the correct answer. It follows directly from Equation (5-50), with $M = 1$, when the integration theorem of Appendix 3 is applied to convert it to a density function for A.

When $M = 1$ and J is less than N, exact results can be obtained for the mismatch problem described earlier. (There can be no mismatch problem when $J = N$!) As noted earlier, the expected value of the parameter estimator is altered by the mismatch. ξ always refers to the difference between the estimator and its mean, as given by Equation (5-53). The quantity b_{A0} is expressed by Equation (5-27) and, in the present instance, the d array is an N vector.

We recall the definition of C_M and note that

$$C_1 = 1 + Z_B^H S_{BB}^{-1} Z_B \; ,$$

where Z_B is now a column vector, of dimension $N-J$, and S_{BB} is a complex Wishart matrix of order $N-J$, with $L-M=K+N$ complex degrees of freedom. We have noted the effect of signal mismatch on the expected value of the parameter array estimator and on the mean value of Z_B, given by Equations (5-28) and (5-26).

The method of analysis used to deal with the case $M=1$ in Section 4 may be applied directly to the study of C_1 and its inverse R_1. We write Z_B as the product of its norm and a unit vector, condition on Z_B, and then make use of the property of complex Wishart matrices established in Appendix 1. As a result, we may write

$$Z_B^H S_{BB}^{-1} Z_B = \frac{\sum_{i=1}^{N-J} |v_i|^2}{\sum_{j=1}^{J+K+1} |w_j|^2} , \tag{5-112}$$

where the v_i and w_j are independent circular complex Gaussian variables, all with variance unity. The w_j have zero means, but the v_i, which are the components of Z_B, have non-zero expected values, as noted above. Thus, the ratio expressed by formula (5-112) is subject to a complex non-central F distribution, with non-centrality parameter

$$c \equiv (E\,Z_B)^H(E\,Z_B) = b_{B0}^H b_{B0} . \tag{5-113}$$

It follows that R_1 is the corresponding complex non-central Beta variable:

$$R_1 = \frac{1}{1 + Z_B^H S_{BB}^{-1} Z_B} = x_\beta(J+K+1, N-J|c) . \tag{5-114}$$

This notation is defined in Appendix 2, and the pdf of the complex non-central Beta is given by Equation (A2-23). Thus, the generalization of Equation (5-106) is

$$f(\xi) = \frac{1}{\pi^J} \int_0^1 \rho^J e^{-\rho \xi^H \xi} f_\beta(\rho; J+K+1, N-J|c)\, d\rho . \tag{5-115}$$

Similarly, the generalization of Equation (5-107), the pdf of the squared norm of ξ, is

$$g(A) = \frac{A^{J-1}}{(J-1)!} \int_0^1 \rho^J e^{-\rho A} f_\beta(\rho;J+K+1,N-J|c)\,d\rho \; . \tag{5-116}$$

In the present case, we have

$$f_\beta(\rho;J+K+1,N-J|c) = e^{-c\rho} \sum_{k=0}^{J+K+1} \binom{J+K+1}{k} \frac{(K+N)!}{(K+N+k)!}$$

$$\times\, c^k f_\beta(\rho;J+K+1,N-J+k) \; . \tag{5-117}$$

The required integrations are carried out by the same methods used before. The exponential factor which occurs in the above formula combines with those already present in the integrands of Equations (5-115) and (5-116). In the latter case, the result is

$$g(A) = \frac{(K+N)!(2J+K)!}{(J-1)!(J+K)!} A^{J-1} e^{-A-c}$$

$$\times \sum_{k=0}^{J+K+1} \binom{J+K+1}{k} \frac{c^k}{(J+K+N+k)!} \, {}_1F_1(N-J+k;J+K+N+k+1;A+c) \; . \tag{5-118}$$

When c vanishes, this expression reduces to Equation (5-110).

The covariance of ξ in the general mismatched case is given by Equation (5-34). Putting M=1 in this expression and using definition (5-113), we obtain

$$\mathrm{Cov}(\xi) = \frac{1}{J+K} I_J \otimes (K+N+c) \; . \tag{5-119}$$

Since $A=\xi^H\xi$, we can apply Equation (A1-42) to compute

$$E\,A = \frac{J}{J+K} (K+N+c) \; . \tag{5-120}$$

It can be verified directly that this result is consistent with the pdf of A, as expressed by Equation (5-118). There is, however, a much simpler route in which we start from Equation (5-116) and write

$$E\,A = \int_0^\infty \frac{A^J}{(J-1)!} \int_0^1 \rho^J e^{-\rho A} f_\beta(\rho; J+K+1, N-J|c)\, d\rho\, dA \; .$$

The order of integration is now reversed, which gives us

$$E\,A = J \int_0^1 \rho^{-1} f_\beta(\rho; J+K+1, N-J|c)\, d\rho \; .$$

Next, we make use of the infinite series representation for the non-central complex Beta pdf, given by Equation (A2-20) of Appendix 2, which gives us the form:

$$E\,A = J e^{-c} \sum_{k=0}^{\infty} \frac{c^k}{k!} \int_0^1 \rho^{-1} f_\beta(\rho; J+K+1, N-J+k)\, d\rho \; . \tag{5-121}$$

From Equation (5-109), we obtain

$$\rho^{-1} f_\beta(\rho; J+K+1, N-J+k) = \frac{K+N+k}{J+K} f_\beta(\rho; J+K, N-J+k) \; ,$$

and, when this result is substituted in Equation (5-121), the integrals all evaluate to unity due to the normalization of the Beta densities. The result is therefore

$$E\,A = J e^{-c} \sum_{k=0}^{\infty} \frac{c^k}{k!} \frac{K+N+k}{J+K} = \frac{J}{J+K}(K+N+c) \; , \tag{5-122}$$

which agrees with Equation (5-120).

The other special case mentioned earlier corresponds to $J=1$, with arbitrary M. We exclude the case $M=J=1$, which is covered by our previous analysis. We also return to the matched version of the problem, the analysis of which cannot so easily be extended to mismatched signals in this instance.

A particular example of the special case now under study is described by a σ array, now an N vector, all of whose components vanish except the first, which is unity. This form can be attained by pre-multiplication by a suitable unitary matrix. Signals are now confined to the first row of the data array Z, whose signal component is an arbitrary row vector in an M-dimensional subspace of $\mathbb{C}^L$: the row space of τ. As before, we do not assume that a transformation to this special form has been carried out.

The estimation error is now a row vector, of dimension M, and its conditional pdf can be obtained from Equation (5-98) by putting $J=1$. This pdf is a spherically symmetric function in $\mathbb{C}^M$, hence it depends only on the squared norm of ξ. We could obtain the unconditioned pdf of this latter quantity (previously called A') from the second of Equations (5-103) by integrating over the conditioning R matrix, which is now a scalar complex Beta variable. However, we prefer to derive the unconditioned pdf of ξ itself in this case, because of its relevance to the adaptive nulling problem mentioned in Section 2.

Substituting $J=1$ in Equation (5-98), we observe that the quantity R_γ which enters there is a scalar in the present case. Using Equation (5-104) and Equation (4-17) once again, we obtain its explicit representation as a complex Beta variable:

$$\begin{aligned} R_\gamma &= \mathcal{R}(N-1,1,K+1) = 1/\ell(N-1,1,K+1) \\ &= x_\beta(K+2,N-1) . \end{aligned} \tag{5-123}$$

The unconditioned pdf of ξ is the integral of the conditional pdf over the density function of the Beta variable:

$$f(\xi) = \int_0^1 \left(\frac{\rho}{\pi}\right)^M e^{-\rho\xi\xi^H} f_\beta(\rho;K+2,N-1)\,d\rho . \tag{5-124}$$

This result also follows directly from Equation (5-99), of course, when the specialization to $J=1$ is carried out [see Equation (A3-54) of Appendix 3].

Although the special case $M=1$ was originally excluded to assure the validity of Equation (5-62), the result when we set M equal to unity in Equation (5-124) is correct, as may be seen from Equation (5-106) (with $J=1$), together with the fact that ξ is a scalar in this case.

The integral in Equation (5-124) can be evaluated as another confluent hypergeometric function, but it is much more useful to view it as the expected value of a conditional pdf of the row vector ξ. Under conditioning by the Beta variable ρ, this pdf is Gaussian, and the elements of ξ are conditionally independent with zero means and variances equal to ρ^{-1}. The pdf of the whitened signal parameter estimator itself is thus given by

$$f_0(\hat{b}_0) = \int_0^1 f_0(\hat{b}_0|\rho)\, f_\beta(\rho;K+2,N-1)\, d\rho \,, \tag{5-125}$$

where

$$f_0(\hat{b}_0|\rho) = \left(\frac{\rho}{\pi}\right)^M e^{-\rho(\hat{b}_0 - b_0)(\hat{b}_0 - b_0)^H} . \tag{5-126}$$

If we make the definition

$$\sigma_b^2 \equiv (\Sigma^{AA})^{-1} = \frac{1}{e^H \Sigma^{-1} e} \,, \tag{5-127}$$

we can express the original parameter array estimator as

$$\hat{b} = \sigma_b \hat{b}_0 .$$

Then, the joint pdf of the elements of this estimator is

$$f(\hat{b}) = \int_0^1 f(\hat{b}|\rho)\, f_\beta(\rho;K+2,N-1)\, d\rho \,, \tag{5-128}$$

where

$$f(\hat{b}|\rho) = \left(\frac{\rho}{\pi\sigma_b^2}\right)^M e^{-\rho(\hat{b} - b)(\hat{b} - b)^H/\sigma_b^2} . \tag{5-129}$$

Since $J = 1$, we can assume that the σ vector is normalized to unity with no loss of generality, in which case σ and e are identical. Moreover, let us now consider the special form of the τ array described by Equation (1-3), in which signals appear in the first M columns of the data array. Then, b and B are identical [see Equation (2-23)], and we have the same situation for which the connection with adaptive nulling was first discussed in Section 2. Equations (5-128) and (5-129) then describe the joint pdf of the M outputs of an adaptive nulling system which applies weights based on the Z_q array to the data vectors which comprise Z_p.

The marginal pdf of the m^{th} element of this output vector can be obtained by integrating out the other components under the integral sign in Equation (5-128). The result is an integral of the product of the same complex Beta density and a univariate complex Gaussian pdf. This conditional pdf describes a complex Gaussian variable with mean value b_m (the m^{th} component of b) and variance equal to σ_b^2 divided by ρ. A "conditional signal-to-noise ratio" can be defined for this variable, in the usual way, as the ratio of squared mean to variance. It is given by

$$SNR_\rho = \frac{|b_m|^2}{\sigma_b^2} \rho ,$$

which reproduces the well-known result of Reed, Mallett, and Brennan,[27] in which the Beta variable plays the role of a loss factor.

Quite apart from the detection problem which has been the focus of our attention in this study, one can use these formulas to analyze the performance of various algorithms for processing the output sequence of such an adaptive nulling system. The procedure is first to use the conditional pdf (which describes simple, independent Gaussian variables) and later average over the complex Beta pdf according to Equation (5-128). It has been tacitly assumed that the adaptive weights based on the Z_q array are not changed as they are applied to the sample vectors of Z_p. In practice, such weights are often "frozen" in this way for a brief interval of time, after which new weights, based on a new array like Z_q, are found and applied to a new block of data vectors. If the "new" Z_q and Z_p arrays are independent of all the "old" vectors, then the new adaptively nulled outputs are statistically just like those of the first block and independent of them. In our model, the true covariance matrix is the same for all the sample vectors in the data array, which now constitutes only one of many such blocks of data. If we allow this covariance matrix (always unknown) to be different from block to block, the only effect on the adaptively nulled outputs will be a changing value of σ_b from block to block. This extension of our original model begins to accommodate the non-stationarity typical of situations ordinarily met in practical applications.

6. THE PROBABILITY OF DETECTION FOR THE GLR TEST

We proceed now to a discussion of the probability of detection (PD) of the GLR test, beginning with the same special cases for which the pdf of the amplitude array estimator was analyzed in Section 5. The general method will be to formulate the conditional PD, given the B components of the data array, and then to remove the conditioning by averaging over these components. As noted at the end of Section 3, the conditioning variables survive only through the matrix C_M, which enters the "signal component" V_{0s} of the V array. For the special cases to be considered first, we can build on the analysis of Section 4, making suitable modifications to account for the presence of signals, in order to derive the conditional probabilities of detection. As we have already seen, when $J = N$ the C_M matrix reduces to the identity and there are no conditioning variables. This case is relatively simple, and it will be therefore be considered separately.

Let $M = 1$ and J be less than N, but otherwise arbitrary. In Section 4, the following expression was obtained for the test statistic:

$$\ell = 1 + \frac{\sum_{i=1}^{J} |v_i|^2}{\sum_{j=1}^{K+1} |w_j|^2} . \tag{6-1}$$

The v_i are the components of the original V array, which is a J vector in this case. The argument which led to this formula remains valid when V contains a signal component, but the numerator of the fraction here is now a non-central complex chi-squared variable under the conditioning. In Section 3, we wrote V as the sum of a "signal component" and a "noise component." After whitening, this representation took the form of Equation (3-36):

$$V = V_{0s} + V_{0n} .$$

The subscript zero has been dropped from V itself, but retained on the components.

The noise component has zero mean and, in the present special case, the signal component is

$$V_{0s} = b_0 C_1^{-1/2} = b_0 R_1^{1/2} . \tag{6-2}$$

In these expressions b_0 is the whitened signal parameter array (a J vector) and R_1 is a scalar, given by Equation (5-105). It follows that

$$V_{0s}^H V_{0s} = b_0^H b_0 R_1$$

is the non-centrality parameter of the complex chi-squared variable in the numerator of Equation (6-1). We write

$$\rho \equiv R_1 = x_\beta(J+K+1, N-J) \ , \tag{6-3}$$

and also make the definition

$$a \equiv a_0 \rho \ . \tag{6-4}$$

where a_0 is again the non-adaptive signal-to-noise ratio (SNR). This quantity was expressed in terms of the arrays in which the detection problem was originally formulated by means of Equation (5-8), which takes the form

$$a_0 = b_0^H b_0 = \tau\tau^H (\sigma B)^H \Sigma^{-1} (\sigma B) \ . \tag{6-5}$$

in the present case. Since $M=1$, σB is an N vector, while $\tau\tau^H$ and a_0 are scalars. The new quantity "a" will play the role of a SNR for the conditional detection problem, and ρ will act as a "loss factor." When $J=N$, the same reasoning is valid, except that $a=a_0$. Hence, this special case can be recovered by replacing ρ by unity in the following analysis.

Under the conditioning, the inverse of the test statistic is a complex Beta variable, but now it is a non-central one, and we may write

$$1/\ell = x_\beta(K+1, J|a) \tag{6-6}$$

which reduces to Equation (4-13) when a vanishes. The conditional detection probability is a cumulative non-central complex Beta distribution, and we can make use of Equation (A2-26) of Appendix 2 to write it in the form

$$\text{Prob}_B(\ell \geq \ell_0) = F_\beta(1/\ell_0;K+1,J|a_0\rho)$$

$$= 1 - \frac{(\ell_0 - 1)^J}{\ell_0^{J+K}} \sum_{k=0}^{K} \binom{J+K}{J+k} (\ell_0 - 1)^k G_{k+1}\left(\frac{a_0\rho}{\ell_0}\right) . \tag{6-7}$$

Considering again the case $J = N$, we see that Equation (6-7), with ρ replaced by unity, provides the final detection probability for the GLR test in that specialization.

In general, we must still average over ρ, which gives us the formula

$$PD = 1 - \frac{(\ell_0 - 1)^J}{\ell_0^{J+K}} \sum_{k=0}^{K} \binom{J+K}{J+k} (\ell_0 - 1)^k H_{k+1}\left(\frac{a_0}{\ell_0}\right) , \tag{6-8}$$

where

$$H_k(y) \equiv E\, G_k(y\rho) = \int_0^1 G_k(y\rho)\, f_\beta(\rho;J+K+1,N-J)\, d\rho . \tag{6-9}$$

Substituting for the incomplete Gamma function [Equation (5-12)] and using Equation (5-109), we obtain

$$H_k(y) = \sum_{m=0}^{k-1} \frac{y^m}{m!} \int_0^1 e^{-y\rho} \rho^m f_\beta(\rho;J+K+1,N-J)\, d\rho$$

$$= \sum_{m=0}^{k-1} \frac{(K+N)!(J+K+m)!}{(J+K)!(K+N+m)!} \frac{y^m}{m!} \int_0^1 e^{-y\rho} f_\beta(\rho;J+K+m+1,N-J)\, d\rho . \tag{6-10}$$

From Equation (5-108), we obtain the final result

$$H_k(y) = \frac{(K+N)!}{(J+K)!} e^{-y} \sum_{m=0}^{k-1} \frac{(J+K+m)!}{(K+N+m)!} \frac{y^m}{m!} \, {}_1F_1(N-J;K+N+m+1;y) . \tag{6-11}$$

Once again, the formula derived for $J<N$ gives the correct answer when $J=N$. As can be seen from Equation (5-111), the confluent hypergeometric function in Equation (6-11) is simply unity in this case, hence H_k reduces to G_k.

Equations (6-8) and (6-11) provide a complete solution for the probability of detection of the GLR test in the special case when $M=1$. These formulas depend only on the non-adaptive SNR, the detection threshold, and the dimensional parameters of the problem. The threshold, in turn, is related to the probability of false alarm, which is given by the cumulative complex central Beta distribution:

$$\mathrm{PFA} = F_\beta(1/\ell_0;K+1,J) = \frac{1}{\ell_0^{J+K}} \sum_{k=0}^{J-1} \binom{J+K}{k} (\ell_0-1)^k , \tag{6-12}$$

which otherwise depends only on the same dimensional parameters. This is the result previously obtained in Section 4. When a_0 vanishes, the H_k functions of Equation (6-8) all reduce to unity, and that equation becomes identical to Equation (6-12), as is easily verified.

These equations are the basis of the numerical analysis and results of Reference 4, in which the performance of the GLR test (in this specific case) is compared with that of a conventional non-adaptive test for the same problem, but assuming that the covariance is known. It may be seen from Equation (6-11) that the function H_k depends on k only through the upper limit of the summation, hence these functions can be computed recursively. The confluent hypergeometric functions are well behaved, since the second argument always exceeds the first as they occur in this formula. The terms of their series are positive, and they decrease faster than those of exp(y). The error caused by truncation of these series is easily bounded by the tail of the series for this exponential. The bound becomes tighter as one progresses along in the series. Once these functions are obtained, the remainder of the computation of PD, from Equation (6-8), is quite straightforward.

From Equation (6-12), we can evaluate the derivative:

$$\frac{d}{d\ell_0}\mathrm{PFA} = -\frac{1}{\ell_0^2} f_\beta(1/\ell_0;K+1,J) = -\frac{(J+K)!}{(J-1)!\,K!} \frac{(\ell_0-1)^{J-1}}{\ell_0^{J+K+1}} ,$$

which may be used to carry out an iterative solution for threshold in terms of PFA, by the Newton-Raphson technique. The threshold that is obtained by approximating

Equation (6-12) by its first term has been successfully used as a starting point for this procedure.

As long as $M=1$, we can evaluate the detection performance in the case of signal mismatch, paralleling our discussion of the estimation error, from which many of the results we need can be obtained. The signal component of the V array, given by Equation (5-25), becomes

$$V_{0s} = b_{A0} C_1^{-1/2} = b_{A0} R_1^{1/2}$$

in the present case. Thus, the non-centrality parameter of the numerator of Equation (6-1) is changed to

$$a = V_{0s}^H V_{0s} = b_{A0}^H b_{A0} R_1 \equiv b_{A0}^H b_{A0} \rho \ .$$

The non-adaptive SNR of the matched case is now replaced by the scalar

$$b_{A0}^H b_{A0} = (b_A - \Sigma_{AB}\Sigma_{BB}^{-1} b_B)^H \Sigma^{AA} (b_A - \Sigma_{AB}\Sigma_{BB}^{-1} b_B) \ .$$

According to Equation (5-27), this quantity can also be written

$$b_{A0}^H b_{A0} = d^H \Sigma^{-1} e \, (e^H \Sigma^{-1} e)^{-1} \, e^H \Sigma^{-1} d \ . \tag{6-13}$$

The "loss factor" ρ is now a non-central complex Beta variable, given by Equation (5-114). The pdf of ρ is now $f_\beta(\rho;J+K+1,N-J|c)$, given explicitly by Equation (5-117), and the appropriate non-centrality parameter for this distribution is expressed by Equation (5-113):

$$c = b_{B0}^H b_{B0}$$

The sum of these non-centrality parameters was evaluated in Equation (5-36), which is a scalar in the present case. We define

$$a_1 \equiv \tau\tau^H \, D^H \Sigma^{-1} D = \mathrm{Tr}\left[(D\tau)^H \Sigma^{-1} (D\tau)\right] , \tag{6-14}$$

and then Equation (5-36) becomes

$$b_{A0}^H b_{A0} + b_{B0}^H b_{B0} = a_1 \tag{6-15}$$

Had we modeled our problem differently, so that signal arrays of the form $D\tau$ were expected, then a_1 would play the role of the non-adaptive SNR. In fact, it would be the actual non-adaptive SNR for a processor designed to anticipate such signals. For this reason, a_1 may be called the "available SNR" associated with this signal and interference environment.

Since the non-centrality parameters are non-negative, we can make the definitions

$$b_{A0}^H b_{A0} = a_1 \cos^2\Theta$$

$$b_{B0}^H b_{B0} = c = a_1 \sin^2\Theta \ , \tag{6-16}$$

and then Θ characterizes the degree of mismatch in a simple way. Thus,

$$a = a_1 \cos^2\Theta \, \rho \ ,$$

and the detection probability becomes

$$PD = 1 - \frac{(\ell_0 - 1)^J}{\ell_0^{J+K}} \sum_{k=0}^{K} \binom{J+K}{J+k} (\ell_0 - 1)^k \, H_{k+1}\left(\frac{a_1 \cos^2\Theta}{\ell_0} \right) . \tag{6-17}$$

where

$$H_k(y) = \int_0^1 G_k(y\rho) \, f_\beta(\rho; J+K+1, N-J|c) \, d\rho \ . \tag{6-18}$$

Substituting for the complex Beta density, we obtain

$$H_k(y) = \sum_{j=0}^{J+K+1} \binom{J+K+1}{j} \frac{(K+N)!}{(K+N+j)!} c^j$$

$$\times \int_0^1 e^{-c\rho} \, G_k(y\rho) \, f_\beta(\rho; J+K+1, N-J+j) \, d\rho \ .$$

These integrals have the same form as those evaluated before, since the exponential factor combines with a similar one contained in the G_k function. When these integrations are carried out, and the order of summation reversed, we obtain the generalization of Equation (6-11):

$$H_k(y) = \frac{(K+N)!}{(J+K)!} e^{-y-c} \sum_{m=0}^{k-1} \frac{(J+K+m)!}{(K+N+m)!} \frac{y^m}{m!}$$

$$\times \sum_{j=0}^{J+K+1} \binom{J+K+1}{j} \frac{(K+N+m)!}{(K+N+m+j)!} c^j \, {}_1F_1(N-J+j;K+N+m+j+1;y+c) \ . \qquad (6\text{-}19)$$

When c vanishes, or $\Theta = 0$, this formula reduces directly to Equation (6-11). The PFA for the mismatched case is, of course, unchanged and is given by Equation (6-12).

Numerical evaluation of the PD from Equation (6-19) presents no new difficulties, relative to the use of Equation (6-11). The problem of the detection of mismatched signals using the GLR decision rule has been discussed for the special case $J = M = 1$ in Reference 5 where numerical results are presented, together with an interpretative analysis of the behavior of this detector. The parameter Θ plays a central role in that analysis.

The other special case considered in connection with the estimation error is characterized by $J = 1$ and arbitrary M. We exclude the case $J = N$ by requiring N to exceed the value unity. The form taken by the test statistic was found in Section 4. Equation (4-6) may be written

$$\ell = 1 + \frac{\sum_{i=1}^{M} |v_i|^2}{\sum_{j=1}^{K+1} |w_j|^2} \ , \qquad (6\text{-}20)$$

which is analogous to Equation (6-1). The v_i are the components of the V array, which is now a row vector of M elements. The denominator here is a complex central chi-squared variable, just as before, and the numerator is again a non-central complex chi-squared variable. In the present case, the (scalar) non-centrality parameter of this variable is

$$a = (EV)(EV)^H \ .$$

The general expression for the signal component, Equation (3-37), takes the form

$$EV = V_{0s} = b_0 C_M^{-1/2} = b_0 R_M^{1/2} , \tag{6-21}$$

in terms of the whitened signal array, which is also a row vector in this case. Thus,

$$a = b_0 R_M b_0^H . \tag{6-22}$$

We define the row vector

$$t \equiv (b_0 b_0^H)^{-1/2} b_0 ,$$

which is always possible unless b_0 itself is identically zero. We expressly exclude this case, since we are dealing here with the probability of detection. It follows that

$$b_0 = (b_0 b_0^H)^{-1/2} t$$

and

$$t t^H = 1 .$$

Then,

$$a = b_0 b_0^H \, t R_M t^H \equiv a_0 \rho , \tag{6-23}$$

where

$$a_0 = b_0 b_0^H = \mathrm{Tr}\,(b_0^H b_0) = \sigma^H \Sigma^{-1} \sigma (B\tau)(B\tau)^H \tag{6-24}$$

is again the non-adaptive SNR, and

$$\rho = t R_M t^H . \tag{6-25}$$

Using identity (5-45), applied to the present situation, we obtain

$$\rho = t \mathcal{R}(N-1, M, K+1) t^H = \mathcal{R}(N-1, 1, K+M) = x_\beta(K+M+1, N-1) . \tag{6-26}$$

The identification of this one-dimensional $\mathcal{R}$ matrix with a complex central Beta variable is exactly the same as in the study of the estimation error in Section 5.

The remainder of the evaluation is a direct parallel to the previous special case, but without mismatch. The test statistic, Equation (6-20), is the inverse of a non-central complex Beta variable, and the conditional probability of detection is given by

$$\mathrm{Prob}_B(\ell \geq \ell_0) = F_\beta(1/\ell_0; K+1, M|a_0\rho)$$

$$= 1 - \frac{(\ell_0 - 1)^M}{\ell_0^{K+M}} \sum_{k=0}^{K} \binom{M+K}{M+k} (\ell_0 - 1)^k G_{k+1}\left(\frac{a_0\rho}{\ell_0}\right). \qquad (6\text{-}27)$$

Note that this formula is the same as Equation (6-7), but with J and M interchanged and K held constant. Similarly, the probability of false alarm is given by

$$\mathrm{PFA} = \frac{1}{\ell_0^{M+K}} \sum_{k=0}^{M-1} \binom{M+K}{k} (\ell_0 - 1)^k \qquad (6\text{-}28)$$

This is formula (4-9) of Section 4, and it is also the limiting form of Equation (6-27) when the SNR tends to zero.

The unconditional PD is therefore

$$\mathrm{PD} = 1 - \frac{(\ell_0 - 1)^M}{\ell_0^{K+M}} \sum_{k=0}^{K} \binom{M+K}{M+k} (\ell_0 - 1)^k H_{k+1}\left(\frac{a_0}{\ell_0}\right), \qquad (6\text{-}29)$$

where

$$H_k(y) = \int_0^1 G_k(y\rho)\, f_\beta(\rho; K+M+1, N-1)\, d\rho$$

$$= \frac{(K+N+M-1)!}{(K+M)!} e^{-y} \sum_{m=0}^{k-1} \frac{(K+M+m)!}{(K+N+M+m-1)!} \frac{y^m}{m!} \, {}_1F_1(N-1; K+N+M+m; y) \;. \qquad (6\text{-}30)$$

Recalling the definition of K, it is seen that Equation (6-30) takes a somewhat simpler form in terms of the original parameter L:

$$H_k(y) = \frac{(L-1)!}{(L-N)!} e^{-y} \sum_{m=0}^{k-1} \frac{(L-N+m)!}{(L+m-1)!} \frac{y^m}{m!} \, {}_1F_1(N-1; L+m; y) .$$

If we put $N=1$ formally in this expression, and use Equation (5-111) again, we see that the H_k functions reduce to the corresponding G_k functions, and the PD formula reverts to the conditional PD expression, which we have seen to be correct whenever $J=N$.

The behavior of the GLR test in the special cases just discussed can be interpreted in a simple way in terms of familiar radar concepts. If we express the decision threshold in the form

$$\ell_0 \equiv 1 + \mu , \tag{6-31}$$

then for $M=1$ (and $J<N$) the decision rule based on the test statistic of Equation (6-1) can be written as

$$\sum_{i=1}^{J} |v_i|^2 \geq \mu \sum_{j=1}^{K+1} |w_j|^2 . \tag{6-32}$$

In this criterion, the v_i and w_j are mutually independent complex Gaussian variables of variance unity. The w_j have zero means, and

$$\sum_{i=1}^{J} |Ev_i|^2 = a = a_0 \rho .$$

Equation (6-32) may be interpreted as the detection criterion of a conventional CFAR detector, based on $K+1=L-N$ samples of "noise," and using non-coherent integration of J samples of "signal plus noise." The effective SNR for this equivalent detector is the product of a_0 and the loss factor ρ, which appears in the place of a more conventional random target fluctuation variable, as these flucttutation models are frequently used in radar analysis. Unlike the conventional models, our loss factor is always less than or equal to unity. Due to this effect, the average value of the effective SNR is reduced and is given by

Recalling the definition of K, it is seen that Equation (6-30) takes a somewhat simpler form in terms of the original parameter L:

$$H_k(y) = \frac{(L-1)!}{(L-N)!} e^{-y} \sum_{m=0}^{k-1} \frac{(L-N+m)!}{(L+m-1)!} \frac{y^m}{m!} \, {}_1F_1(N-1; L+m; y) \,.$$

If we put $N=1$ formally in this expression, and use Equation (5-111) again, we see that the H_k functions reduce to the corresponding G_k functions, and the PD formula reverts to the conditional PD expression, which we have seen to be correct whenever $J=N$.

The behavior of the GLR test in the special cases just discussed can be interpreted in a simple way in terms of familiar radar concepts. If we express the decision threshold in the form

$$\ell_0 \equiv 1 + \mu \,, \tag{6-31}$$

then for $M=1$ (and $J<N$) the decision rule based on the test statistic of Equation (6-1) can be written as

$$\sum_{i=1}^{J} |v_i|^2 \geq \mu \sum_{j=1}^{K+1} |w_j|^2 \,. \tag{6-32}$$

In this criterion, the v_i and w_j are mutually independent complex Gaussian variables of variance unity. The w_j have zero means, and

$$\sum_{i=1}^{J} |Ev_i|^2 = a = a_0 \rho \,.$$

Equation (6-32) may be interpreted as the detection criterion of a conventional CFAR detector, based on $K+1=L-N$ samples of "noise," and using non-coherent integration of J samples of "signal plus noise." The effective SNR for this equivalent detector is the product of a_0 and the loss factor ρ, which appears in the place of a more conventional random target fluctuation variable, as these fluctutation models are frequently used in radar analysis. Unlike the conventional models, our loss factor is always less than or equal to unity. Due to this effect, the average value of the effective SNR is reduced and is given by

$$E\,a = a_0\, E\,\rho$$

The mean value of a complex Beta variable is easily derived from the complex Beta pdf [Equation (A2-12) of Appendix 2]:

$$E\,x_\beta(n,m) = \frac{n}{n+m} \;,$$

and in the present case, which is characterized by Equation (6-3), we obtain

$$E\,a = a_0 \frac{J+K+1}{N+K+1} = a_0 \frac{L+J-N}{L} \;. \tag{6-33}$$

There is, of course, no loss when $J = N$ and ρ is replaced by unity.

Formulas (6-7) and (6-12) are well known in connection with the performance of conventional CFAR radar detectors. The loss factor is, of course, directly associated with adaptive detection and its inevitable covariance estimation. It is easy to insert a target fluctuation model, such as one of the Swerling models, into the analysis at this point. The procedure is to replace a_0 by ua_0 in the formula for the conditional detection probability. The new factor u is a random variable, independent of everything else, and subject to a pdf which represents the desired target fluctuation model. (In effect, every element of the true signal parameter array has been multiplied by the square root of u.) In the Swerling models, u is a complex chi-squared variable, and the number of its complex degrees of freedom can be related to J, the dimensionality of the signal subspace, so as to achieve the desired effect in the model. This is analogous to choosing the number of degrees of freedom in relation to the number of pulses which are subjected to non-coherent integration in the ordinary application of the fluctuation models.

To compute the probability of detection using one of the Swerling models, it is best to average first over the target fluctuation parameter, since this will usually lead to a simpler formula than Equation (6-7) for the conditional PD. A collection of such detection formulas, for various fluctuation models, may be found in Reference 28. The resulting expression is then averaged over the complex Beta pdf to obtain the final result. The probability of false alarm is, of course, unaffected by the addition of a target fluctuation model.

The other special case studied earlier can be interpreted in an analogous fashion, and a target fluctuation factor can be added to the model. Our starting point will be

Equation (6-20), which describes the performance of an equivalent CFAR detector based on K+1 samples of "noise" and M samples of "signal plus noise." The effective SNR has the same form as before, namely the product of a_0 and a loss factor ρ, whose statistical characterization is expressed by Equation (6-26). The average SNR is now given by

$$Ea = a_0 \frac{K+M+1}{K+M+N} = a_0 \frac{L+1-N}{L} .$$

In terms of L, this is the same as Equation (6-33), with $J=1$. Target fluctuation can be added to the formulation exactly as before, and now the number of complex degrees of freedom of the variable u must be related to M. In the special case described by Equation (1-3), so often invoked here for illustrative purposes, M is just the number of sample vectors for which signal components may be present, and the correspondence with ordinary non-coherent integration is quite precise.

In Section 3 we discussed the transition from the adaptive test to the non-adaptive one in a heuristic way. Now, with explicit formulas before us, we can sharpen that discussion, at least for those special cases for which we have obtained explicit results. We consider only the first of the special cases, namely $M=1$, since the other can be obtained by a trivial interchange of parameters. If we put

$$\mu \equiv \frac{\lambda_0}{K+1} .$$

Equation (6-32) becomes

$$\sum_{i=1}^{J} |v_i|^2 \geq \frac{\lambda_0}{K+1} \sum_{j=1}^{K+1} |w_j|^2 .$$

The expected value of the right side of this equation is just λ_0, and its variance will tend to zero as K is allowed to increase indefinitely. The test will then correspond to a non-adaptive decision rule which takes the form of non-coherent integration of J samples of "signal plus noise." Making the same substitution in Equation (6-12), and letting K tend to infinity, we obtain

$$\text{PFA} = \left(1 + \frac{\lambda_0}{K+1}\right)^{-(J+K)} \sum_{k=0}^{J-1} \binom{J+K}{k} \left(\frac{\lambda_0}{K+1}\right)^k$$

$$\to e^{-\lambda_0} \sum_{k=0}^{J-1} \frac{\lambda_0^k}{k!} = G_J(\lambda_0) \, ,$$

which is the standard result for the PFA of such a test, and it agrees with Equation (5-11), when the substitution $M = 1$ is made.

When K tends to infinity, the pdf of the loss factor becomes more and more concentrated near the value $\rho = 1$. Formula (6-9) suggests that we should have

$$H_k(y) \underset{K \to \infty}{\to} G_k(y)$$

in this case, and this is confirmed by an analysis of Equation (6-11) as K goes to infinity. The detection probability can thus be obtained from Equation (6-7) by replacing ρ by unity and substituting for μ. The result is

$$\text{PD} \to 1 - \left(1 + \frac{\lambda_0}{K+1}\right)^{-(J+K)} \sum_{k=0}^{K} \binom{J+K}{J+k} \left(\frac{\lambda_0}{K+1}\right)^{J+k} G_{k+1}\left(\frac{a_0(K+1)}{K+1+\lambda_0}\right) .$$

Passing to the limit on K, the final result may be written

$$\lim_{K \to \infty} \text{PD} = 1 - e^{-\lambda_0} \sum_{k=0}^{\infty} \frac{\lambda_0^{J+k}}{(J+k)!} G_{k+1}(a_0)$$

$$= G_J(\lambda_0) + e^{-\lambda_0} \sum_{k=J}^{\infty} \frac{\lambda_0^k}{k!} \left[1 - G_{k-J+1}(a_0)\right] .$$

This is a well-known[29] series representation for the Marcum Q-function, and it is in agreement with our earlier result for the non-adaptive problem, Equation (5-10), again with $M = 1$. It follows that the performance of the GLR test will tend to that of a non-adaptive decision rule as K tends to infinity. This is the same limit, of course, in which the sample covariance matrix tends to the true covariance.

In the general case, the evaluation of the probability of detection presents formidable difficulties. It is not evaluated explicitly here, but some general properties of the exact solution will be derived. We will then review the analysis of Section 4, taking account of the presence of signal components in the data. This exercise will illustrate the difficulties of the general problem, and will also provide the basis for a proof of another useful property of the exact probability of detection.

To deal with this generalization effectively, some new notation is required. As before, let T be a complex Wishart matrix of order J, with $J+K$ complex degrees of freedom. This matrix can be expressed in the form $T=WW^H$, where W is a complex Gaussian array with zero mean. We also let V be a complex Gaussian array of dimension $J \times M$, independent of W, whose mean value is given by a constant array A, and whose covariance matrix is the identity. The complete set of definitions is:

$$\begin{aligned} E\,V &= A\,, & \mathrm{Cov}(V) &= I_J \otimes I_M \\ E\,W &= 0\,, & \mathrm{Cov}(W) &= I_J \otimes I_{J+K}\,. \end{aligned} \tag{6-34}$$

We now introduce the "non-central" $\mathscr{C}$ matrix, extending the notation used earlier:

$$\mathscr{C}(J,M,K|A) \equiv I_M + V^H T^{-1} V\,. \tag{6-35}$$

Continuing the analogy, we define

$$\mathscr{R}(J,M,K|A) \equiv \mathscr{C}(J,M,K|A)^{-1} \tag{6-36}$$

and

$$\ell(J,M,K|A) \equiv |\mathscr{C}(J,M,K|A)|\,. \tag{6-37}$$

The matrix A can actually be a function of different random quantities, as long as these are completely independent of the random variables which appear in the definition of the $\mathscr{C}$ matrix. A is then the conditional mean value of V, with these "different" random quantities held fixed. More precisely, we can say that $V-A$ is a zero-mean complex Gaussian array, whose covariance is the identity matrix given above for the covariance of V itself. This extension of the significance of the notation is needed in the discussion of the GLR test in the general case.

The first property we wish to establish is a generalization of the duality between the parameters J and M, observed first in connection with the PFA, and noted again in the study of the two special cases of the present section. To establish this property, we assume that J is less than M and note that VV^H will then be positive definite (with probability one). We fix the arrays V and T, and introduce the array

$$\mathcal{E} \equiv (V V^H)^{-1/2} V . \tag{6-38}$$

The properties

$$\mathcal{E}\mathcal{E}^H = I_J$$

$$V = (V V^H)^{1/2} \mathcal{E}$$

follow directly. Now let $\mathcal{I}$ be a complex Wishart matrix, of order M, with M + K complex degrees of freedom. Like T, the new matrix can be expressed in terms of a complex Gaussian array with zero mean. According to the property established in Appendix 1, the matrix

$$(\mathcal{E} \mathcal{I}^{-1} \mathcal{E}^H)^{-1}$$

is also complex Wishart, of order J, and with J + K complex degrees of freedom. It is statistically identical to T, hence we can write

$$|I_M + V^H T^{-1} V| = |I_M + V^H \mathcal{E} \mathcal{I}^{-1} \mathcal{E}^H V| .$$

The factors in this determinant may be permuted cyclically, as shown in Appendix 1, so that

$$|I_M + V^H T^{-1} V| = |I_J + \mathcal{E} \mathcal{I}^{-1} \mathcal{E}^H V V^H|$$

$$= |I_J + (V V^H)^{1/2} \mathcal{E} \mathcal{I}^{-1} \mathcal{E}^H (V V^H)^{1/2}| = |I_J + V \mathcal{I}^{-1} V^H| .$$

Finally, if we define

$$\mathcal{V} \equiv V^H . \tag{6-39}$$

we obtain the form:

$$|I_M + V^H T^{-1} V| = |I_J + \mathcal{V}^H \mathcal{T}^{-1} \mathcal{V}| ,$$

where now $\mathcal{V}$ is $M \times J$ and $\mathcal{T}$ is of order M.

This is the desired duality property, and it may be expressed by the relation

$$\ell(J,M,K|A) = \ell(M,J,K|A^H) . \tag{6-40}$$

As in other similar situations, the equality here refers to statistical identity, or equality of the corresponding distribution functions. The form of the result itself shows that it is valid regardless of the relationship between J and M. The symmetry between J and M will be lost when this identity is applied to the conditional detection probability and the conditioning is subsequently removed, as we have seen in the two special cases already worked out.

The non-central $\mathscr{C}$ matrices exhibit another feature, which will lead us to a useful general property of the unconditioned probability of detection. Let U_J and U_M be arbitrary unitary matrices, whose orders are indicated by their subscripts, and let

$$\tilde{V} \equiv U_J^H V U_M^H$$

$$\tilde{T} \equiv U_J^H T U_J . \tag{6-41}$$

It follows that

$$|I_M + V^H T^{-1} V| = |I_M + \tilde{V}^H \tilde{T}^{-1} \tilde{V}| , \tag{6-42}$$

and also that

$$E\tilde{V} = U_J^H A U_M^H \equiv \tilde{A} . \tag{6-43}$$

The unitary transformation has no effect on the statistical character of the T matrix, and the transformed V array is still complex Gaussian, with the same covariance matrix as V itself. Only its mean value is changed, according to Equation (6-43). This yields another statistical equivalence, expressed by the relation

$$\ell(J,M,K|A) = \ell(J,M,K|\tilde{A}) . \tag{6-44}$$

We now introduce the singular value decomposition of the A array, writing

$$A = U_J A_0 U_M . \tag{6-45}$$

where U_J and U_M are unitary matrices of orders J and M, respectively, and A_0 is a diagonal array. The diagonal elements of A_0 are the singular values of A, ordered in an arbitrary way. If we identify U_J and U_M with the unitary matrices of Equations (6-41), we see that

$$\ell(J,M,K|A) = \ell(J,M,K|A_0) , \tag{6-46}$$

in the sense of statistical equivalence. Thus, the probability distribution function of the random variable $\ell(J,M,K|A)$ depends only on the singular values of the A array. It is, in fact, a symmetric function of these numbers, since they may be permuted arbitrarily by a transformation of the kind described by Equations (6-41). The singular values of A are, in turn, the non-negative square roots of the eigenvalues of AA^H. (If $J > M$, this matrix will be rank-deficient, and it will have $J - M$ zero eigenvalues, in addition to the squares of the singular values of A.) In any case, we can say that the statistical properties of $\ell(J,M,K|A)$ depend only on these eigenvalues. In particular, we can write

$$\text{Prob}\left[\ell(J,M,K|A) \geq \ell_0\right] \equiv \Phi(J,M,K;\ell_0;AA^H) , \tag{6-47}$$

where $\Phi(J,M,K,x,X)$ is a real-valued function of the scalar parameters J, M, K, and x, and of the square $J \times J$ matrix X. Φ depends only on the eigenvalues of X, hence it is unaffected if X undergoes a similarity transformation:

$$X \rightarrow U_J X U_J^H .$$

Now let us apply these results to the GLR test statistic, by identifying A with the signal component of the V array, V_{0s}, first defined in Equation (3-37). Then we will have

$$A = V_{0s} = b_0 R_M^{1/2} ,$$

where b_0 is the whitened true signal amplitude parameter array, and R_M is statistically described in Equation (5-39) as a central $\mathcal{R}$ matrix:

$$R_M = \mathscr{R}(N-J, M, J+K) ,$$

which is completely independent of $V_{0n} = V - A$ and T. With this substitution for A, Equation (6-47) expresses the conditional probability of detection of the GLR test. The unconditioned PD is obtained, formally, by averaging over the $\mathscr{R}$ matrix:

$$PD = \int \Phi(J, M, K; \ell_0; b_0 R b_0^H) \, f_B(R; M, J+M+K, N-J) \, d_0(R) , \tag{6-48}$$

where f_B is the pdf of the multivariate Beta matrix. This integral is an example of a general type discussed in Appendix 3 [see Equation (A3-52)].

We now introduce the singular value decomposition of b_0:

$$b_0 = u_J \beta u_M , \tag{6-49}$$

where u_J and u_M are unitary, and β is a diagonal $J \times M$ array, whose diagonal elements are the singular values of b_0. In terms of β, we have

$$b_0 R b_0^H = u_J \beta u_M R u_M^H \beta^H u_J^H .$$

We can now make a change of variables in the integral, defining the new matrix

$$R' \equiv u_M R u_M^H . \tag{6-50}$$

The Jacobian of this transformation is unity [it is a special case of Equation (A3-14) of Appendix 3], and it also leaves the pdf of the matrix R unchanged, a fact we used repeatedly in Section 5. Finally, the function Φ is unaffected by the application of the similarity transformation described by u_J, and we conclude that

$$PD = \int \Phi(J, M, K; \ell_0; \beta R \beta^H) \, f_B(R; M, J+M+K, N-J) \, d_0(R) . \tag{6-51}$$

This shows that the final probability of detection depends only on the singular values of b_0, the whitened signal parameter array. The b_0 array depends, in turn, on the true covariance matrix Σ and the true signal parameter array b (or the original array B). The singular values of b_0 are the non-negative square roots of the eigenvalues of the matrix $b_0^H b_0$, which we have encountered already in Equation (5-8) of Section 5. We may call it the "signal-to-noise-ratio matrix," and we recall that the non-adaptive SNR is its trace. According to Equation (5-8), the SNR matrix depends on τ only through the product $\tau\tau^H$, and it is therefore unchanged if τ is post-multiplied by any unitary matrix of order L. This fact confirms the invariance property of the GLR detection probability already observed at the end of Section 2. In the two special cases for which we have obtained complete performance results, the SNR matrix has rank unity. The extension of our results to cases for which this matrix has higher rank remains an interesting challenge.

In Section 4 we derived a formula which expresses the test statistic as a product of two factors which proved to be statistically independent of one another. This factorization was then iterated, to obtain a double-product representation which provides the basis for the evaluation of the PFA in the general case. When signal components are present, the factorization is still valid, but the factors are no longer independent, and the conditional detection probability (conditioned on R_M) cannot be obtained by the methods used for the evaluation of the PFA. The factorization is useful, however, for the proof of a monotonicity property of the exact solution which will now be derived.

Following closely the analysis of Section 4, we introduce a subspace of the vector space $\mathbb{C}^J$ by separating all column vectors into two components of dimension J_1 and J_2, where $J_1 + J_2 = J$. We write

$$V \equiv \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}, \qquad A \equiv \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \qquad W \equiv \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}, \tag{6-52}$$

which extends Equations (4-20) to include the mean value array A, introduced in Equation (6-34). Components of the T matrix and its inverse are introduced, using definitions (4-21) and (4-22), and then Equations (4-23) and (4-24) are still valid. We can write

$$I_M + V_2^H T_{22}^{-1} V_2 = \mathscr{C}(J_2, M, J_1 + K \mid A_2) . \tag{6-53}$$

applying our new notation to the problem, and this equation replaces Equation (4-28) as a statement of the statistical character of the quantity on the left side. As before, we put

$$\mathcal{V} \equiv (V_1 - T_{12} T_{22}^{-1} V_2)(I_M + V_2^H T_{22}^{-1} V_2)^{-1/2} . \tag{6-54}$$

and

$$\mathcal{G} \equiv (T^{11})^{-1} , \tag{6-55}$$

and then we have

$$\ell(J,M,K|A) = |I_M + V^H T^{-1} V| = |I_M + \mathcal{V}^H \mathcal{G}^{-1} \mathcal{V}|\,|I_M + V_2^H T_{22}^{-1} V_2| . \tag{6-56}$$

If we condition on the 2-components, and recall that we are dealing with whitened quantities in the present case, we can compute

$$E_2 \mathcal{V} \equiv \mathcal{A}_1 = A_1 (I_M + V_2^H T_{22}^{-1} V_2)^{-1/2} , \tag{6-57}$$

since $T_{12} = W_1 W_2^H$, and W_1 has zero mean. Stretching the notation slightly, we can express the statistical character of the left side of Equation (6-56) by writing

$$\ell(J,M,K|A) = \ell(J_1,M,K|\mathcal{A}_1)\,\ell(J_2,M,J_1+K|A_2) . \tag{6-58}$$

Because $\mathcal{A}_1$ depends on the 2-components of V and W, the factors in this expression are not independent, and this fact is the main impediment to the derivation of an explicit formula for the conditional probability of detection. Of course, if we had such an expression, we would then be faced with the evaluation of the integral in Equation (6-48)!

To obtain the monotonicity property referred to above, we specialize our factorization to the case $J_1 = 1$, so that A_1 becomes a row vector of M components. Conditioned on the 2-components, $\mathcal{A}_1$ is fixed, and we can write

$$\text{Prob}_2[\,\ell(1,M,K|\mathcal{A}_1) \geq \mu] = \Phi(1,M,K;\mu;G) , \tag{6-59}$$

where μ is a constant, and G is given by

$$G \equiv \mathcal{A}_1 \mathcal{A}_1^H = A_1 (I_M + V_2^H T_{22}^{-1} V_2)^{-1} A_1^H . \tag{6-60}$$

When $J_1 = 1$, we have

$$\ell(1, M, K | \mathcal{A}_1) = 1/x_\beta(K+1, M|G) , \tag{6-61}$$

which is a direct generalization of Equation (4-18) of Section 4. The extension to a non-central Beta variable made here is very much like the extension discussed in detail in Section 5, in connection with the mismatched signal problem. Reference may be made to Equations (5-112) and (5-114) for details of that discussion. Finally, using the notation of Appendix 2, we can write

$$\Phi(1, M, K; \mu; G) = F_\beta(1/\mu; K+1, M|G) . \tag{6-62}$$

From the explicit form of the cumulative complex non-central Beta distribution, given by Equation (A2-27) of Appendix 2, it may be seen that the right side of Equation (6-62) is an increasing function of G (we will use the term "increasing" here as shorthand for "monotone non-decreasing").

Now we let

$$\mu \equiv \frac{\ell_0}{|I_M + V_2^H T_{22}^{-1} V_2|} = \frac{\ell_0}{\ell(J_2, M, J_1 + K | A_2)} , \tag{6-63}$$

which makes μ a function of the 2-components. Then, in view of Equation (6-58), we can express the right side of Equation (6-47) in the form of an expectation value over the 2-component variables implicit in μ and G:

$$\Phi(J, M, K; \ell_0; AA^H) = E\,\Phi(J_1, M, K; \mu; G) . \tag{6-64}$$

We have seen that this probability depends only on the singular values of A. We can therefore assume that A is already in diagonal form, since this can be accomplished by the transformation indicated in Equation (6-45). Now suppose the two unitary matrices which appear in that equation are fixed, and that one of the singular values is allowed to vary, all the others being held constant. Since the order of the singular values was, in any case, immaterial, we can take the variable one to be the first entry in the diagonal form of A. When we apply the factorization described above, with $J_1 = 1$, the row vector A_1 will then have all zero entries except the first, which we may call a_1:

$$A_1 = [a_1\, 0 \ldots 0] \,.$$

Then G, defined by Equation (6-60), will take the form

$$G = (a_1)^2\, [1\, 0 \ldots 0]\, (I_M + V_2^H T_{22}^{-1} V_2)^{-1}\, [1\, 0 \ldots 0]^H \,.$$

This matrix product is necessarily positive; hence, the left side of Equation (6-62) is an increasing function of a_1. This property is preserved when the expectation indicated in Equation (6-64) is carried out. We have therefore shown that the left side of that equation is an increasing function of a_1 which was an arbitrary singular value of A.

Let A and B be two $J \times M$ arrays which have identical singular values except for one, say a and b. Then, if $a \leq b$, we will have

$$\Phi(J,M,K;\ell_0;AA^H) \leq \Phi(J,M,K;\ell_0;BB^H) \,, \tag{6-65}$$

since A and B can be put into diagonal form, with a and b as the first entries in the respective diagonals, and the result proved above can then be applied. More generally, let the ordered singular values of A and B be related as follows:

$$a_i \leq b_i \,, \qquad 1 \leq i \leq \mathrm{Min}(J,M) \,. \tag{6-66}$$

Then, Equation (6-65) is again correct since the singular values can be increased one by one, changing from the A values to those of B, and the corresponding probability is always increasing. Inequality (6-66) defines an ordering of $J \times M$ arrays, and, in terms of this ordering, the probability function on the left side of Equation (6-65) is an increasing function of the A array.

Let b_0 and b_0' be two whitened signal parameter arrays, and suppose that

$$b_0 \leq b_0' \,, \tag{6-67}$$

in the sense of the ordering defined above. Let the singular value decompositions of these arrays be given by the equations

$$b_0 = u_J\, \beta\, u_M \,, \qquad b_0' = u_J'\, \gamma\, u_M' \,.$$

and let us assume that in both cases the singular values are ordered, say from the largest to the least. Then, according to the ordering of the arrays, we have

$$\beta_i \le \gamma_i \,, \qquad 1 \le i \le \mathrm{Min}(J,M) \,.$$

From our previous discussion, it follows that we can replace the original signal parameter arrays by the diagonal arrays β and γ in the expressions for the unconditioned probability of detection for these two cases. This probability is given by Equation (6-51) for b_0, and by the same formula (with γ replacing β) for the other case. The two probabilities are therefore expressible as integrals of appropriate conditional probabilities over the same complex multivariate Beta distribution.

The conditional probabilities depend, in turn, on the eigenvalues of the matrices $\beta R \beta^H$ and $\gamma R \gamma^H$. Let ν stand for the smaller of the parameters J and M. Then, the ν largest eigenvalues of these $J \times J$ matrices coincide with the ν largest eigenvalues of the respective $M \times M$ matrices, X and Y, which are defined by the equations

$$X \equiv R^{1/2} \beta^H \beta R^{1/2} \,, \qquad Y \equiv R^{1/2} \gamma^H \gamma R^{1/2} \,.$$

If $\nu = J$, then the eigenvalues of these new matrices will be augmented by one or more zero values. The difference

$$Y - X = R^{1/2} (\gamma^H \gamma - \beta^H \beta) R^{1/2}$$

is clearly a non-negative definite matrix. In Appendix 1, by an application of the Courant-Fisher theorem, it is shown that the ordered eigenvalues of X are less than or equal to their counterparts in the list of ordered eigenvalues of Y. We may conclude that the ordered eigenvalues of $\beta R \beta^H$ are less than or equal to their counterparts in the list of ordered eigenvalues of $\gamma R \gamma^H$. From this relation it follows that the unconditioned probability of detection for the signal parameter array b_0 is less than or equal to that corresponding to the other parameter array b_0'. Thus, the probability of detection is an increasing function of the singular values of the whitened signal parameter array, or, equivalently, of the eigenvalues of the SNR matrix $b_0^H b_0$, and this is the monotonicity property we set out to establish.

7. A GENERALIZATION OF THE MODEL

In Section 1 we mentioned a generalization of the basic model of the hypothesis testing problem. The null hypothesis, which previously corresponded to the complete absence of signal components, is replaced by the hypothesis that a particular component of the signal parameter array is zero, the rest being arbitrary. More precisely, this model takes the form

$$H_0 : \quad \alpha B \gamma = 0$$

$$H_1 : \quad B \text{ is arbitrary .} \qquad (7\text{-}1)$$

The fixed arrays α $(r \times J)$ and γ $(M \times t)$ determine the component of the B array whose presence or absence constitutes the purpose of the test. We postulate that the rank of α is $r \leq J$, while that of γ is $t \leq M$, and anticipate that these arrays will determine subspaces in $\mathbb{C}^J$ and $\mathbb{C}^M$, respectively.

The significance of the model is illustrated by the specific example

$$\alpha = \begin{bmatrix} 0 & I_r \end{bmatrix}$$

$$\gamma = \begin{bmatrix} 0 \\ I_t \end{bmatrix} ,$$

in which α and γ provide direct decompositions of $\mathbb{C}^J$ and $\mathbb{C}^M$. In accordance with these decompositions, we may partition B as follows:

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} .$$

Then, the test becomes a decision on whether or not B_{22} is zero, while the other three components of B may have any values on either hypothesis. These latter components may be considered to describe "nuisance signals," while B_{22} describes the "desired signal" component which may be present in the data array.

To specialize further, suppose that both Equations (1-3) and (1-4) hold, so that the signal structure itself corresponds to the "canonical form" discussed in Section 1. As shown by Equation (1-8), the signal components are then confined to the upper left corner of the data array, and the first $M - t$ of these columns contain only nuisance signals. The remaining t columns which are allowed to contain signals are further divided into two subspaces (corresponding to B_{12} and B_{22}), of which one contains desired signals and the other only more nuisance components.

The task of the decision rule in the general case is to detect the desired signals in the presence of the others, against a background of unknown noise and interference. A GLR test will now be derived which accomplishes this goal and which turns out to have very similar structure to the test studied in the earlier sections of this study. In particular, this test will have the same extended CFAR property as the former one, and, in addition, its performance will not be influenced by the presence of nuisance signal components.

We begin by expressing the null hypothesis in terms of the "normalized" signal parameter array b, defined in Equation (2-23), writing

$$\alpha B \gamma = abc \ , \tag{7-2}$$

where

$$a \equiv \alpha (\sigma^H \sigma)^{-1/2}$$

$$c \equiv (\tau \tau^H)^{-1/2} \gamma \ . \tag{7-3}$$

To set up the subspace projections, we introduce the basis arrays

$$a_2 \equiv (a a^H)^{-1/2} a$$

$$c_2 \equiv c (c^H c)^{-1/2} \tag{7-4}$$

in the usual way, and note that the null hypothesis now corresponds to the condition

$$a_2 b c_2 = 0 \ . \tag{7-5}$$

The relations

$$a_2 a_2^H = I_r$$

$$a = (a a^H)^{1/2} a_2$$

$$c_2^H c_2 = I_t$$

$$c = c_2 (c^H c)^{1/2}$$

follow directly, and we work with a_2 and c_2 from here on, instead of with a and c.

The row space of a_2 is an r-dimensional subspace of $\mathbb{C}^J$, and we introduce an orthonormal basis a_1 for its complementary subspace. (This nomenclature, which uses the subscript 2 for the subspaces representing desired signals, is arbitrary, but proves convenient in the later analysis.) Similarly, let c_1 be an array of basis vectors in the space complementary to the column space of c_2, so that

$$a_1 a_1^H = I_{J-r}$$

$$a_1^H a_1 + a_2^H a_2 = I_J$$

$$c_1^H c_1 = I_{M-t}$$

$$c_1 c_1^H + c_2 c_2^H = I_M \ .$$

Finally we introduce unitary matrices

$$U_J \equiv \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad , \quad U_M \equiv \begin{bmatrix} c_1 & c_2 \end{bmatrix} \qquad (7\text{-}6)$$

in analogy to the matrices

$$U_N = \begin{bmatrix} e & f \end{bmatrix} \quad , \quad U_L = \begin{bmatrix} p \\ q \end{bmatrix} ,$$

which we will also need.

As before, the data array is first decomposed using U_L:

$$Z\,U_L^H = [Z_p\ Z_q]\ ,$$

where

$$Z_p \equiv Zp^H$$

$$Z_q \equiv Zq^H\ . \qquad (7\text{-}7)$$

The Z_p component is further decomposed by means of U_M:

$$Z_p U_M = [Z_{p1}\ Z_{p2}]\ ,$$

where

$$Z_{p1} \equiv Z\,p^H c_1$$

$$Z_{p2} \equiv Z\,p^H c_2\ . \qquad (7\text{-}8)$$

Together, a threefold decomposition of $\mathbb{C}^L$ is produced, based on the unitary matrix

$$\tilde{U}_L \equiv \begin{bmatrix} U_M^H & 0 \\ 0 & I_{L-M} \end{bmatrix} U_L = \begin{bmatrix} c_1^H p \\ c_2^H p \\ q \end{bmatrix}\ . \qquad (7\text{-}9)$$

When applied to the data array, this decomposition gives us the equation

$$Z\,\tilde{U}_L^H = [\,Z_{p1}\ \ Z_{p2}\ \ Z_q\,]\ . \qquad (7\text{-}10)$$

In a similar way, U_J and U_N are combined to form a threefold decomposition of $\mathbb{C}^N$:

$$\tilde{U}_N \equiv U_N \begin{bmatrix} U_J^H & 0 \\ 0 & I_{N-J} \end{bmatrix} = [\,e_1\ \ e_2\ \ f\,]\ . \qquad (7\text{-}11)$$

where

$$e_1 \equiv e a_1^H$$

$$e_2 \equiv e a_2^H \, . \tag{7-12}$$

The derivation of the GLR test begins, as in Section 2, with the maximization of the probability density functions over the unknown covariance matrix. The test statistic can then be expressed in the form

$$l = \frac{\operatorname*{Min}_{H_0} |F(b)|}{\operatorname*{Min}_{H_1} |F(b)|} ,$$

where F(b) is still given by

$$F(b) = (Z - ebp)(Z - ebp)^H \, .$$

Under H_1 the array b is unconstrained, while under H_0 it is subject to the linear constraint (7-5). We begin with the null hypothesis and introduce some notation in order to accommodate the constraint. Consider the matrix product

$$U_J \, b \, U_M = \begin{bmatrix} a_1 b c_1 & a_1 b c_2 \\ a_2 b c_1 & a_2 b c_2 \end{bmatrix} \equiv \begin{bmatrix} \delta_1 & \beta \\ \delta_2 & 0 \end{bmatrix} . \tag{7-13}$$

by which β, δ_1, and δ_2 are defined. The zero component is the result of the constraint, as expressed by Equation (7-5). We use the new parameters to express b in the form

$$b = U_J^H \begin{bmatrix} \delta_1 & \beta \\ \delta_2 & 0 \end{bmatrix} U_M^H \equiv \delta c_1^H + a_1^H \beta c_2^H \, , \tag{7-14}$$

where

$$\delta \equiv U_J^H \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} .$$

and, of course,

$$U_J^H \begin{bmatrix} \beta \\ 0 \end{bmatrix} = a_1^H \beta \ .$$

The $(J-r)\times t$ array β is the analog of B_{12} in the special example described above, while δ, which is of dimension $J\times(M-t)$, represents the components analogous to both B_{11} and B_{21}. The minimization required under H_0 is the same as an unconstrained minimization over δ and β.

To bring the new arrays into play, we separate F(b) into terms corresponding to the decomposition of $\mathbb{C}^L$ by writing

$$\begin{aligned} F(b) &= (Z - ebp)\,\tilde{U}_L^H \tilde{U}_L (Z - ebp)^H \\ &= (Z_{p1} - ebc_1)(Z_{p1} - ebc_1)^H + (Z_{p2} - ebc_2)(Z_{p2} - ebc_2)^H + S \ , \end{aligned}$$

where, as in Section 2,

$$S \equiv Z_q Z_q^H \ .$$

Using the representation (7-14), we have

$$\begin{aligned} bc_1 &= \delta \\ bc_2 &= a_1^H \beta \ , \end{aligned} \qquad (7\text{-}15)$$

and, therefore,

$$F(b) = (Z_{p1} - e\delta)(Z_{p1} - e\delta)^H + (Z_{p2} - e_1\beta)(Z_{p2} - e_1\beta)^H + S \ . \qquad (7\text{-}16)$$

We make the definition

$$\tilde{S}(\beta) \equiv (Z_{p2} - e_1\beta)(Z_{p2} - e_1\beta)^H + S \qquad (7\text{-}17)$$

and proceed to carry out the minimization over δ. This follows precisely the procedure of Section 2, with a result analogous to Equation (2-41):

$$\operatorname*{Min}_{\delta} |F(b)| = |\tilde{S}(\beta)|\,|I_{M-t} + Z_{p1}^{H}\tilde{P}Z_{p1}|\,, \tag{7-18}$$

where

$$\tilde{P} \equiv \tilde{S}^{-1} - \tilde{S}^{-1}e\,(e^{H}\tilde{S}^{-1}e)^{-1}\,e^{H}\tilde{S}^{-1}\,. \tag{7-19}$$

This quantity appears to depend upon β, but it is actually independent of that array; hence, the right side of Equation (7-18) will depend on β only through the first of the two factors. In analogy to Equation (3-12), only the component

$$f^{H}\tilde{P}f = (f^{H}\tilde{S}f)^{-1}$$

is non-vanishing, and the evaluation

$$f^{H}\tilde{S}f = f^{H}(Z_{p2}Z_{p2}^{H} + S)f$$

shows the claimed independence of β. It follows that

$$\operatorname*{Min}_{H_0} |F(b)| = |I_{M-t} + Z_{p1}^{H}\tilde{P}Z_{p1}|\,\operatorname*{Min}_{\beta}|\tilde{S}(\beta)|\,.$$

The minimization over β is the same problem over again, and we can immediately write

$$\operatorname*{Min}_{\beta}|\tilde{S}(\beta)| = |S|\,|I_t + Z_{p2}^{H}\Pi Z_{p2}|\,,$$

where Π is defined by

$$\Pi \equiv S^{-1} - S^{-1}e_1(e_1^{H}S^{-1}e_1)^{-1}e_1^{H}S^{-1}\,. \tag{7-20}$$

Combining our results, we obtain

$$\operatorname*{Min}_{H_0}|F(b)| = |S|\,|I_{M-t} + Z_{p1}^{H}\tilde{P}Z_{p1}|\,|I_t + Z_{p2}^{H}\Pi Z_{p2}|\,. \tag{7-21}$$

The minimization of $F(b)$ under H_1 has, of course, been carried out in Section 2, but it is useful to derive the result again, in a slightly different w[illegible]h parallels the analysis just given. Specifically, we represent b in terms of tv[illegible]ays, as follows:

$$b = \delta' c_1^H + \beta' c_2^H \, . \tag{7-22}$$

These arrays are unconstrained, and their role is to allow the minimization to be carried out in two steps, as was done under H_0. The new expression for $F(b)$ is the same as Equation (7-16), but with the array e_1 replaced by e itself. The final result is then

$$\underset{H_1}{\mathrm{Min}} \, |F(b)| = |S| \, |I_{M-t} + Z_{p1}^H \tilde{P} \, Z_{p1}| \, |I_t + Z_{p2}^H P \, Z_{p2}| \, ,$$

where P is the same array which appeared in Section 2:

$$P \equiv S^{-1} - S^{-1} e \, (e^H S^{-1} e)^{-1} \, e^H S^{-1} \, . \tag{7-23}$$

The two versions of the minimization under H_1 yield the equation

$$|I_M + Z_p^H P \, Z_p| = |I_{M-t} + Z_{p1}^H \tilde{P} \, Z_{p1}| \, |I_t + Z_{p2}^H P \, Z_{p2}| \, , \tag{7-24}$$

which can also be verified directly as an identity involving determinants.

The GLR test statistic now assumes the form

$$\ell = \frac{|I_t + Z_{p2}^H \Pi \, Z_{p2}|}{|I_t + Z_{p2}^H P \, Z_{p2}|} \, , \tag{7-25}$$

which corresponds to Equation (2-42). We note that the component Z_{p1} has dropped out of the test completely. In the case of the special example described at the beginning of this section, the first $M - t$ columns of the data array would be discarded in forming the GLR test statistic. The remaining data array components, Z_{p2} and Z_q, are partitioned as follows:

$$\tilde{U}_N^H Z_{p2} \equiv \begin{bmatrix} Z_N \\ Z_A \\ Z_B \end{bmatrix} \quad , \quad \tilde{U}_N^H Z_q \equiv \begin{bmatrix} W_N \\ W_A \\ W_B \end{bmatrix} . \tag{7-26}$$

The subscript N refers to the "nuisance" components, while the A and B portions are directly analogous to the corresponding components employed in Section 3. In analogy to Equation (3-5), the S matrix is also expressed in component form:

$$\tilde{U}_N^H S \tilde{U}_N \equiv \begin{bmatrix} S_{NN} & S_{NA} & S_{NB} \\ S_{AN} & S_{AA} & S_{AB} \\ S_{BN} & S_{BA} & S_{BB} \end{bmatrix} . \tag{7-27}$$

By repeating the analysis of Section 3, using appropriate partitionings of this S matrix, we obtain the evaluations

$$Z_{p2}^H P Z_{p2} = Z_B^H S_{BB}^{-1} Z_B$$

and

$$Z_{p2}^H \Pi Z_{p2} = [\, Z_A^H \;\; Z_B^H \,] \begin{bmatrix} S_{AA} & S_{AB} \\ S_{BA} & S_{BB} \end{bmatrix}^{-1} \begin{bmatrix} Z_A \\ Z_B \end{bmatrix} .$$

Again, using Equation (A1-9) of Appendix 1, we have

$$Z_{p2}^H \Pi Z_{p2} = Y^H T^{-1} Y + Z_B^H S_{BB}^{-1} Z_B , \tag{7-28}$$

where Y and T are given by

$$Y \equiv Z_A - S_{AB} S_{BB}^{-1} Z_B$$

$$T \equiv S_{AA} - S_{AB} S_{BB}^{-1} S_{BA} . \tag{7-29}$$

Substituting these results, we find that

$$\ell = \frac{|I_t + Z_B^H S_{BB}^{-1} Z_B + Y^H T^{-1} Y|}{|I_t + Z_B^H S_{BB}^{-1} Z_B|} .$$

By introducing the definitions

$$C_t \equiv I_t + Z_B^H S_{BB}^{-1} Z_B$$

$$V \equiv Y C_t^{-1/2} , \qquad (7\text{-}30)$$

we obtain the final result

$$\ell = |I_t + V^H T^{-1} V| , \qquad (7\text{-}31)$$

all in direct correspondence with the analysis of the original model of the hypothesis testing problem. We note that the components Z_N and W_N have also dropped out of the test, so that in the special example mentioned earlier, the first $J - r$ rows of the data array would also be discarded.

The performance of the GLR test in the more general context of the present section is exactly the same as in the original problem, when the appropriate parameter correspondences are made. To establish these correspondences, we retrace the steps through the various transformations which have been made, evaluating their statistical consequences. The quantities B (or b) and Σ now represent the actual values of these arrays, hence the expected value of the original data array is

$$E Z = ebp .$$

Recalling the definition (7-9), we have

$$E Z \tilde{U}_L^H = eb \left[c_1 \;\; c_2 \;\; 0 \right] ,$$

and, therefore, in view of Equation (7-10), the component Z_q has zero mean, while

$$E Z_{p2} = ebc_2 . \qquad (7\text{-}32)$$

Similarly, from the original covariance property

$$\mathrm{Cov}(Z) = \Sigma \otimes I_L$$

and Equation (A1-42) of Appendix 1, we obtain the results

$$\mathrm{Cov}(Z_{p2}) = \Sigma \otimes I_t$$

$$\mathrm{Cov}(Z_q) = \Sigma \otimes I_{L-M} .$$

In addition, the components Z_q and Z_{p2} are independent.

The components W_A and W_B obviously have zero mean, and from definitions (7-11) and (7-12), together with Equation (7-32), we obtain

$$E\, \tilde{U}_N^H Z_{p2} = \begin{bmatrix} a_1 \\ a_2 \\ 0 \end{bmatrix} bc_2 .$$

and, consequently,

$$E\, Z_A = a_2 bc_2$$

$$E\, Z_B = 0 .$$

The only component of the actual signal parameter array which can have any effect on the GLR test is a_2bc_2, which is just the component whose presence is being tested. The fact that nuisance signals enter into the hypotheses has the consequence that, in general, only a portion of any signal of the original postulated form $\sigma B\tau$ will contribute to the decision to accept H_1.

In analogy to Equation (7-27), we introduce the components of the transformed true covariance array:

$$\tilde{\Sigma} \equiv \tilde{U}_N^H \Sigma \tilde{U}_N = \begin{bmatrix} \Sigma_{NN} & \Sigma_{NA} & \Sigma_{NB} \\ \Sigma_{AN} & \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BN} & \Sigma_{BA} & \Sigma_{BB} \end{bmatrix} . \tag{7-33}$$

It follows that

$$\mathrm{Cov}(\tilde{U}_N^H Z_{p2}) = \tilde{\Sigma} \otimes I_t$$

$$\mathrm{Cov}(\tilde{U}_N^H Z_q) = \tilde{\Sigma} \otimes I_{L-M} .$$

If we introduce the notations

$$Z_c \equiv \begin{bmatrix} Z_A \\ Z_B \end{bmatrix} \quad , \quad W_c \equiv \begin{bmatrix} W_A \\ W_B \end{bmatrix} ,$$

for the surviving components of Z_{p2} and Z_q, we can write

$$E Z_c = \begin{bmatrix} a_2 b c_2 \\ 0 \end{bmatrix} \quad , \quad E W_c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and

$$\mathrm{Cov}(Z_c) = \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{bmatrix} \otimes I_t \quad , \quad \mathrm{Cov}(W_c) = \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{bmatrix} \otimes I_{L-M} .$$

Next, we define the components of the inverse matrix:

$$\begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{bmatrix}^{-1} \equiv \begin{bmatrix} \Sigma^{AA} & \Sigma^{AB} \\ \Sigma^{BA} & \Sigma^{BB} \end{bmatrix} \tag{7-34}$$

to complete the parallel with the original problem. Note that the components defined on the right side of Equation (7-34) are not partitions of the inverse of the full Σ matrix.

Finally, a whitened array is defined:

$$V_0 \equiv (\Sigma^{AA})^{1/2} V = V_{0s} + V_{0n} , \tag{7-35}$$

in which the "signal component" is given by

$$V_{0s} = (\Sigma^{AA})^{1/2} a_2 b c_2 C_t^{-1/2} . \tag{7-36}$$

Note that the dimension of V_0 is $r \times t$. The whitened T array in the present case obeys a complex Wishart distribution of dimension r (and with $L + J - N - M$ complex degrees of freedom) as it did in the original problem.

The rest of the analysis is identical to that of Section 3, whose results apply directly to the present case with the replacements

$$\begin{aligned} b &\rightarrow a_2 b c_2 \\ J &\rightarrow r \\ M &\rightarrow t \\ L &\rightarrow L + t - M \\ N &\rightarrow N + r - J . \end{aligned} \tag{7-37}$$

With these correspondences, the results obtained in Sections 4, 5, and 6 are also directly applicable.

APPENDIX 1
MATHEMATICAL BACKGROUND

Several groups of related mathematical results, most of them well known, are collected here for reference; they are used freely in the text.

A. LEMMAS INVOLVING PARTITIONED MATRICES

Partitioned matrices occur frequently in the analysis, and we begin with a derivation of some indispensible identities. If A and D are square non-singular matrices, where A is of order K and B is of order L, then the partitioned array whose blocks are A, B, C, and D can be factored in two ways, as follows:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_K & 0 \\ CA^{-1} & I_L \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D-CA^{-1}B \end{bmatrix} \begin{bmatrix} I_K & A^{-1}B \\ 0 & I_L \end{bmatrix}$$

$$= \begin{bmatrix} I_K & BD^{-1} \\ 0 & I_L \end{bmatrix} \begin{bmatrix} A-BD^{-1}C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I_K & 0 \\ D^{-1}C & I_L \end{bmatrix} . \qquad \text{(A1-1)}$$

As a direct consequence, we obtain the useful determinant identity

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A|\,|D-CA^{-1}B| = |D|\,|A-BD^{-1}C| . \qquad \text{(A1-2)}$$

The special case

$$|I+BC| = |I+CB| \qquad \text{(A1-3)}$$

is frequently applied in the text.

By inverting the factors in Equation (A1-1), which is a straightforward process, and then multiplying out the results, we obtain the standard inversion formulas

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D-CA^{-1}B)^{-1} \\ -(D-CA^{-1}B)^{-1}CA^{-1} & (D-CA^{-1}B)^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} (A-BD^{-1}C)^{-1} & -(A-BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A-BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A-BD^{-1}C)^{-1}BD^{-1} \end{bmatrix} . \tag{A1-4}$$

Further, by comparing these expressions, we obtain the generalized Woodbury[7] formula

$$(A-BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1} . \tag{A1-5}$$

Another useful identity may be obtained, using the first of Equations (A1-1), as follows:

$$\begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = (U + VCA^{-1})\,A\,(X + A^{-1}BY) + V(D-CA^{-1}B)Y . \tag{A1-6}$$

We often use the notation

$$M \equiv \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} , \qquad M^{-1} \equiv \begin{bmatrix} M^{11} & M^{12} \\ M^{21} & M^{22} \end{bmatrix} , \tag{A1-7}$$

as a convenient way of identifying the blocks of a partitioned matrix and its inverse. By applying Equation (A1-4) to M and also to its inverse, we obtain the relations

$$\begin{aligned} M^{11} &= \left(M_{11} - M_{12}M_{22}^{-1}M_{21}\right)^{-1} \\ M_{11}^{-1} &= M^{11} - M^{12}(M^{22})^{-1}M^{21} \\ (M^{11})^{-1}M^{12} &= -M_{12}M_{22}^{-1} \\ M_{11}^{-1}M_{12} &= -M^{12}(M^{22})^{-1} \end{aligned} \tag{A1-8}$$

and so on. A special case of Equation (A1-6) is frequently encountered:

$$[U^H\ V^H]\ M^{-1}\begin{bmatrix} U \\ V \end{bmatrix} = (U - M_{12}M_{22}^{-1}V)^H M^{11}(U - M_{12}M_{22}^{-1}V) + V^H M_{22}^{-1} V\ , \quad \text{(A1-9)}$$

in which we have also made use of some of the relations expressed in Equation (A1-8).

B. MATRIX LEMMAS INVOLVING EIGENVALUES

Suppose the product $AB \ldots YZ$ of some number of arrays is square, although some or all of the factors may be rectangular. Then $ZAB \ldots Y$ is also square, and generally of a different order than the original matrix, as is every other product formed by cyclic permutation. Suppose the original product has a non-zero eigenvalue λ. There will then be a normalized eigenvector ψ which satisfies the eigenvalue equation

$$AB \ldots YZ\psi = \lambda\psi\ .$$

Since $\lambda\psi$ is not zero, the vector $Z\psi$ cannot vanish.

Multiplying on the left by Z, we obtain

$$ZAB \ldots YZ\psi = \lambda Z\psi\ ,$$

which shows that λ is also an eigenvalue of $ZAB \ldots Y$. Thus, λ is an eigenvalue of every cyclic permutation of the original product. Many (perhaps all) of these products will be rank-deficient, with null eigenvalues supplementing the shared non-vanishing ones. We may say that these products are "eigenvalue-equivalent" matrices, since every non-vanishing eigenvalue of one of them is an eigenvalue of every other.

The sum of all the non-zero eigenvalues of each of these products is the same, which is consistent with the equality of their traces. If we add the appropriate identity matrix to each cyclic product and form the determinants of the resulting sums, then all these determinants will be equal, a fact which also follows from Equation (A1-3).

We consider the maximization problem posed in Section 2. We are given a pair of positive-definite matrices A_1 and A_2, of order N, and we are to evaluate

$$y \equiv \underset{\sigma}{\text{Max}}\ \frac{|\sigma^H A_1 \sigma|}{|\sigma^H A_2 \sigma|}\ . \quad \text{(A1-10)}$$

the maximization being carried out over all full-rank arrays σ of dimension $N \times J$. We introduce the positive-definite square root of A_2 and define

$$u \equiv A_2^{1/2} \sigma \ . \tag{A1-11}$$

Then,

$$y = \underset{u}{\mathrm{Max}} \frac{|u^H B u|}{|u^H u|} \ . \tag{A1-12}$$

the maximization being over all $N \times J$ arrays u, of rank J, where

$$B \equiv A_2^{-1/2} A_1 A_2^{-1/2} \ . \tag{A1-13}$$

The matrix $u^H u$ is positive definite, as a result of our rank assumption; hence, we can introduce the array

$$\mu \equiv u(u^H u)^{-1/2} \ , \tag{A1-14}$$

which satisfies the relation

$$\mu^H \mu = 1_J \ . \tag{A1-15}$$

Since

$$u = \mu (u^H u)^{1/2} \ ,$$

we have

$$|u^H B u| = |u^H u| \, |\mu^H B \mu| \ .$$

Thus,

$$y = \underset{\mu}{\mathrm{Max}} \, |\mu^H B \mu| \ , \tag{A1-16}$$

subject to the validity of Equation (A1-15), now viewed as a constraint.

If the eigenvalues of the positive-definite matrix B are called λ_n, placed in decreasing (or non-increasing) order from λ_1 through λ_N, then

$$y = \lambda_1 \ldots \lambda_J \ . \qquad \text{(A1-17)}$$

as will be proved below. By the cyclic permutation lemma, the λ_n are also the eigenvalues of $A_1(A_2)^{-1}$, and this is the property which was used in Section 2.

To prove the assertion made above, let the eigenvectors of B be ψ_n, properly orthogonalized in case of the degeneracy of any of the eigenvalues, and also normalized. If we take for μ the array whose columns are the first J of these eigenvectors, the constraint will automatically be satisfied and the result claimed for the maximum will be attained.

Now suppose that μ is an array which satisfies Equation (A1-15), and such that

$$|\mu^H B \mu| > \lambda_1 \ldots \lambda_J \ . \qquad \text{(A1-18)}$$

We define

$$M \equiv \mu^H B \mu \ , \qquad \text{(A1-19)}$$

and note that M is a positive-definite matrix of order J. Let its ordered eigenvalues be μ_m, and let U_J be a unitary matrix which diagonalizes M, placing the eigenvalues in decreasing order, according to

$$U_J^H M U_J = \text{Diag}[\mu_1, \ldots, \mu_J] \ ,$$

or

$$\nu^H B \nu = \text{Diag}[\mu_1, \ldots, \mu_J] \ ,$$

where

$$\nu \equiv \mu U_J \ . \qquad \text{(A1-20)}$$

Then,

$$|M| = |\nu^H B \nu| = \mu_1 \ldots \mu_J > \lambda_1 \ldots \lambda_J \ . \qquad \text{(A1-21)}$$

Since the μ's and the λ's are positive and similarly ordered, we must have

$$\mu_k > \lambda_k \tag{A1-22}$$

for at least one value of k between unity and J. Fixing this value of k, we form an array η, of dimension $N \times k$, which consists of the first k columns of ν. Then,

$$\eta^H M \eta = \mathrm{Diag}[\mu_1, \ldots, \mu_k] , \tag{A1-23}$$

and, since U_J is unitary,

$$\eta^H \eta = I_k . \tag{A1-24}$$

Let S be the subspace of $\mathbb{C}^N$ for which the columns of η form a basis, and let Θ be an arbitrary vector in S. Then,

$$x(\Theta) \equiv \frac{\Theta^H M \Theta}{\Theta^H \Theta} = \frac{\sum_{m=1}^{k} \mu_m |\Theta_m|^2}{\sum_{m=1}^{k} |\Theta_m|^2} , \tag{A1-25}$$

where the Θ_m are the coefficients of Θ in the basis defined by η:

$$\Theta = \sum_{m=1}^{k} \Theta_m \eta_m .$$

Equation (A1-25) follows directly from the properties of η, as expressed by Equations (A1-23) and (A1-24). From Equation (A1-25), we conclude that

$$\min_{\Theta} x(\Theta) = \mu_k > \lambda_k , \tag{A1-26}$$

because the μ_m are positive and in decreasing order. But Equation (A1-26) contradicts the Courant-Fisher theorem,[30] according to which

$$\operatorname*{Max}_{S} \operatorname*{Min}_{\Theta \epsilon S} \frac{\Theta^H M \Theta}{\Theta^H \Theta} = \mu_k . \qquad \text{(A1-27)}$$

the maximization being carried out over all subspaces of dimension k, and this completes the proof.

In Section 6 of the text, another relationship between eigenvalues was used which is a direct consequence of the Courant-Fisher theorem itself. Suppose that A and B are Hermitian matrices, of order N, and that the difference $B - A$ is non-negative definite. We can write $A \leq B$ to indicate the ordering of these matrices. If the ordered eigenvalues of A and B are a_k and b_k, respectively, then it follows that $a_k \leq b_k$ for all k from 1 to N.

To prove this claim, we let w be any N vector and observe that

$$w^H A w \leq w^H B w .$$

This inequality is fully equivalent to the statement that $B - A$ is non-negative definite. If S_k is any k-dimensional subspace of $\mathfrak{C}^N$, then we can certainly say that

$$\operatorname*{Min}_{w \epsilon S_k} \frac{w^H A w}{w^H w} \leq \operatorname*{Min}_{w \epsilon S_k} \frac{w^H B w}{w^H w} .$$

But, according to the Courant-Fisher theorem, we have

$$\operatorname*{Min}_{w \epsilon S_k} \frac{w^H B w}{w^H w} \leq \operatorname*{Max}_{S_k} \operatorname*{Min}_{w \epsilon S_k} \frac{w^H B w}{w^H w} = b_k ,$$

where the Max is taken over all k-dimensional subspaces of $\mathfrak{C}^N$. Thus,

$$\operatorname*{Min}_{w \epsilon S_k} \frac{w^H A w}{w^H w} \leq b_k ,$$

and the desired result follows immediately:

$$a_k = \operatorname*{Max}_{S_k} \operatorname*{Min}_{w \epsilon S_k} \frac{w^H A w}{w^H w} \leq b_k . \qquad \text{(A1-28)}$$

C. THE KRONECKER PRODUCT

In the main text, we dealt with collections of random variables which are arranged as rectangular arrays. Such a collection may also be viewed as a vector, by mapping the pair of indices of the array into a single index in some definite way. The covariance matrix of a rectangular array of random variables will be an array which is characterized by a pair of double indices, and the use of this mapping will allow us to establish a consistent notation for such matrices and their products with vectors.

Let Z be an array with components $Z_{i,j}$, and let the single index α correspond to the pair (i,j), according to some one-to-one mapping such as lexigraphical ordering. Then, the Z array can be written as a vector, as follows:

$$z_\alpha = Z_{i,j} \quad , \quad \alpha \longleftrightarrow (i,j) \ . \tag{A1-29}$$

We use a lowercase symbol to indicate the vector which corresponds to an array identified by the same letter in uppercase. The inner product of a pair of such vectors can then be expressed in terms of the original arrays, according to the evaluation

$$x^H y = \sum_\alpha x^*_\alpha y_\alpha = \sum_{i,j} X^*_{i,j} Y_{i,j} = \mathrm{Tr}(X^H Y) \ . \tag{A1-30}$$

The notation is extended in a natural way to matrices whose rows and columns are each designated by index pairs. An element of such a matrix may be written in the form $A_{(i,j);(k,l)}$, or, equivalently, as

$$a_{\alpha,\beta} = A_{(i,j);(k,l)} \ ,$$

where

$$\alpha \longleftrightarrow (i,j) \quad , \quad \beta \longleftrightarrow (k,l) \ .$$

A general bilinear form in this notation is evaluated as follows:

$$x^H a y = \sum_{\alpha,\beta} x^*_\alpha a_{\alpha,\beta} y_\beta = \sum_{i,j} \sum_{k,l} X^*_{i,j} A_{(i,j);(k,l)} Y_{k,l} \ . \tag{A1-31}$$

If the elements of such an array can be expressed as products of the elements of two other arrays, indexed in the ordinary way, according to the rule

$$A_{(i,j);(k,l)} = B_{i,k}\,C_{j,l} \ . \tag{A1-32}$$

then A is called the Kronecker product of B and C, and we write

$$A = B \otimes C \ . \tag{A1-33}$$

If B is $J \times K$ and C is $M \times N$, then A is $JM \times KN$ in dimension. The algebraic properties of the Kronecker product, as an operator, follow easily from its definition. In particular, we note that

$$\begin{aligned} (B \otimes C)^H &= B^H \otimes C^H \\ \mathrm{Tr}(B \otimes C) &= \mathrm{Tr}(B)\,\mathrm{Tr}(C) \\ (B_1 \otimes C_1)(B_2 \otimes C_2) &= (B_1 B_2) \otimes (C_1 C_2) \ . \end{aligned} \tag{A1-34}$$

and, if B and C are square and non-singular,

$$(B \otimes C)^{-1} = B^{-1} \otimes C^{-1} \ . \tag{A1-35}$$

If the square matrices B and C are of orders J and M, respectively, then the Kronecker product is square and of order JM. Its determinant is given by

$$|B \otimes C| = |B|^M\,|C|^J \ . \tag{A1-36}$$

Finally, if A has the form of Equation (A1-33), the general bilinear form [Equation (A1-31)] becomes

$$x^H a y = \mathrm{Tr}(X^H B Y C^T) \ , \tag{A1-37}$$

and, as a special case, we obtain the multiplication rule

$$(ay)_\alpha = (BYC^T)_{i,j} \quad , \quad \alpha \longleftrightarrow (i,j) \ . \tag{A1-38}$$

D. RANDOM ARRAYS

Consider a complex random array Z, of dimension $J \times M$. For simplicity of writing, we assume that the mean value of Z is zero, since we are interested primarily in its covariance properties here. Since Z is a doubly indexed set of random variables, its covariance matrix is automatically of the doubly indexed type, and we make the definition:

$$[\mathrm{Cov}(Z)]_{(i,j);(k,l)} \equiv E\, Z_{i,j} Z^{*}_{k,l} \, . \tag{A1-39}$$

If this covariance has the form

$$E\, Z_{i,j} Z^{*}_{k,l} = B_{i,k}\, C^{*}_{j,l} \, , \tag{A1-40}$$

then we have

$$\mathrm{Cov}(Z) = B \otimes C^{*} \, . \tag{A1-41}$$

In this case, B is square and of order J, while C (also square) will be of order M. The paradigm for this choice of ordering of the indices is the array $Z_{i,j} = b_i c^{*}_j$, where b and c are independent random vectors whose covariance matrices are B and C, respectively. The full covariance matrix is, of course, Hermitian, and it can always be arranged that the factors B and C are individually Hermitian. Then, the identities

$$E\, Z Z^{H} = B\, \mathrm{Tr}\, C$$

$$E\, Z^{H} Z = C\, \mathrm{Tr}\, B \tag{A1-42}$$

follow directly from the definition. More generally, if X and Y are complex random arrays whose means are zero and whose elements satisfy the equation

$$E\, X_{i,j} Y^{*}_{k,l} = D_{i,k}\, E^{*}_{j,l} \, ,$$

then we write

$$\mathrm{Cov}(X, Y) = D \otimes E^{*} \, . \tag{A1-43}$$

Now suppose that U and V are fixed arrays, and that the product

$$Z' \equiv UZV$$

makes sense dimensionally. If the covariance of Z satisfies Equation (A1-41), it follows that

$$\mathrm{Cov}(Z') = (UBU^H) \odot (V^H C V)^* . \qquad \text{(A1-44)}$$

More generally, if X and Y satisfy Equation (A1-43) and if

$$X' = U_x X V_x$$

$$Y' = U_y Y V_y ,$$

where U_x, U_y, V_x, and V_y are fixed arrays, then we have

$$\mathrm{Cov}(X', Y') = (U_x D U_y^H) \odot (V_x^H E V_y)^* . \qquad \text{(A1-45)}$$

E. COMPLEX GAUSSIAN VECTORS

In the above discussion, and also throughout the main text, we encounter collections of complex random variables. In order to fix our ideas and our notation about such collections, especially about arrays of Gaussian random variables, we review here some of the basic facts concerning them, beginning with complex Gaussian vectors. Let z be a column vector of dimension J, whose elements are complex Gaussian random variables with zero means. Then, the joint probability density function of z takes the general form

$$f(z) = \frac{1}{\pi^J |\Gamma|} e^{-z^H \Gamma^{-1} z} , \qquad \text{(A1-46)}$$

where Γ is a complex positive-definite matrix. With the definition $z_k = x_k + iy_k$ for each of the elements of z, the volume element associated with this pdf is written

$$d(z) = dx_1 \ldots dx_J dy_1 \ldots dy_J . \qquad \text{(A1-47)}$$

The statistical significance of definition (A1-46) will follow from its expression in terms of the real component random variables themselves. To derive this form, we consider the one-to-one correspondence between z and the real vector u, of dimension 2J, defined by

$$z \equiv [z_1,\ldots,z_J]^T \longleftrightarrow u \equiv [x_1,\ldots,x_J,y_1,\ldots,y_J]^T .$$

Let Φ be a complex matrix, of order J, and let

$$z' \equiv \Phi z . \tag{A1-48}$$

Then, if the real vector corresponding to z' is called u', a linear relationship

$$u' = Fu \tag{A1-49}$$

will hold for a suitable real matrix F. We separate Φ into real and imaginary parts, making the definition

$$\Phi \equiv \Phi_R + i\Phi_I , \tag{A1-50}$$

where Φ_R and Φ_I are real matrices of order J. Then, applying our definitions, we find that F is expressible in block form, as follows:

$$F = \begin{bmatrix} \Phi_R & -\Phi_I \\ \Phi_I & \Phi_R \end{bmatrix} . \tag{A1-51}$$

This equation establishes a mapping between complex matrices of a given order and real matrices of twice that order. Under this mapping, the product $\Phi_2\Phi_1$ corresponds to F_2F_1, the inverse Φ^{-1} correponds to F^{-1}, and so on. If Φ is Hermitian, then F is symmetric, since Φ_R is symmetric and Φ_I is skew-symmetric in this case. It is also easily verified that

$$z^H \Phi z = u^T F u . \tag{A1-52}$$

Obviously, each vector z has the same quadratic norm as its real counterpart u. Finally, by elementary row and column operations, we evaluate the determinant

$$|F| = \begin{vmatrix} \Phi_R & -\Phi_I \\ \Phi_I & \Phi_R \end{vmatrix} = \begin{vmatrix} \Phi_R + i\Phi_I & -\Phi_I + i\Phi_R \\ \Phi_I & \Phi_R \end{vmatrix}$$

$$= \begin{vmatrix} \Phi_R + i\Phi_I & 0 \\ \Phi_I & \Phi_R - i\Phi_I \end{vmatrix} = |\Phi\Phi^H| . \qquad \text{(A1-53)}$$

If Equation (A1-48) is viewed as a linear transformation of variables, applied to a multiple integral over the volume element of Equation (A1-47), then Equation (A1-53) provides an evaluation of its Jacobian.

Returning to the Gaussian pdf, we put $\Gamma = \Gamma_R + i\Gamma_I$ and make the definition

$$\mathcal{M} \equiv \frac{1}{2} \begin{bmatrix} \Gamma_R & -\Gamma_I \\ \Gamma_I & \Gamma_R \end{bmatrix} . \qquad \text{(A1-54)}$$

Thus, Γ is associated with $2\mathcal{M}$, according to the mapping just discussed, and Γ^{-1} corresponds to $1/2\ \mathcal{M}^{-1}$. Then, from Equation (A1-52) we obtain

$$z^H \Gamma^{-1} z = \frac{1}{2} u^T \mathcal{M}^{-1} u .$$

Since Γ is Hermitian, Equation (A1-53) yields

$$|\Gamma| = |2\mathcal{M}|^{1/2} = 2^J |\mathcal{M}|^{1/2} .$$

Substituting in Equation (A1-46), we find the desired form

$$f(u) = \frac{1}{(2\pi)^J |\mathcal{M}|^{1/2}} e^{-\frac{1}{2} u^T \mathcal{M}^{-1} u} . \qquad \text{(A1-55)}$$

This represents a conventional Gaussian pdf for a real vector u with zero mean value and with covariance matrix

$$\mathcal{M} = E u u^T . \qquad \text{(A1-56)}$$

If we put

$$u \equiv \begin{bmatrix} x \\ y \end{bmatrix} , \tag{A1-57}$$

where

$$x \equiv [x_1, \ldots, x_J]^T \quad , \quad y \equiv [y_1, \ldots, y_J]^T . \tag{A1-58}$$

then we can write

$$\mathcal{M} = \begin{bmatrix} Exx^T & Exy^T \\ Eyx^T & Eyy^T \end{bmatrix} .$$

Comparison with Equation (A1-54) shows us that

$$Exx^T = Eyy^T = \frac{1}{2} \Gamma_R$$

and

$$Eyx^T = -Exy^T = \frac{1}{2} \Gamma_I .$$

Thus, the real variables corresponding to a set of complex Gaussian variables have a special covariance structure, expressed by the above equations. These relations, in turn, give us the basic covariance properties of the complex random vector itself:

$$\begin{aligned} Ezz^H &= E(xx^T + yy^T) + iE(yx^T - xy^T) \\ &= \Gamma_R + i\Gamma_I = \Gamma \end{aligned} \tag{A1-59}$$

and

$$Ezz^T = E(xx^T - yy^T) + iE(yx^T + xy^T) = 0 . \tag{A1-60}$$

Equation (A1-60) expresses the "circular symmetry property," which is a necessary and sufficient condition for the validity of the complex Gaussian probability density itself. For a complex scalar random variable, the joint pdf of the real and imaginary parts exhibits circular symmetry in the x-y plane.

F. COMPLEX GAUSSIAN ARRAYS

Now let us identify z with a $J \times M$-dimensional array of random variables Z, according to the correspondence (A1-29). We assume that the mean value of Z is not zero, but is given by an array $\bar{Z}$, and that the associated vector z has a corresponding mean value. The circularity condition will then be expressed by the relation

$$E(Z - \bar{Z})_{i,j}(Z - \bar{Z})_{k,l} = 0 , \tag{A1-61}$$

and the covariance matrix of Z will be given, in general, by an expression analogous to definition (A1-39). The Gaussian joint pdf of Z will be a direct generalization of Equation (A1-46).

We now assume that the covariance of Z has the special form given in Equation (A1-41), and we associate the covariance matrix Γ of the vector variable with the Kronecker product matrix $B \otimes C^*$ of the Z array. The determinant of this matrix is equal to the right side of Equation (A1-36), since C is Hermitian, and we make use of Equation (A1-35) for its inverse. Equation (A1-37) is then used to evaluate the exponent of the Gaussian distribution, completing the transition from the vector form of Equation (A1-46) to the desired expression in terms of the Z array itself. The resulting joint pdf of the elements of Z is

$$f(Z) = \frac{1}{\pi^{JM} |B|^M |C|^J} \, e^{-\mathrm{Tr}[B^{-1}(Z-\bar{Z})C^{-1}(Z-\bar{Z})^H]} . \tag{A1-62}$$

The corresponding volume element is written

$$d(Z) \equiv \prod_{j=1}^{J} \prod_{m=1}^{M} d[\mathrm{Re}(Z_{j,m})]\, d[\mathrm{Im}(Z_{j,m})] , \tag{A1-63}$$

which generalizes Equation (A1-47).

Consider the linear transformation

$$Z' = FZG , \tag{A1-64}$$

where F and G are square matrices of appropriate orders. Then, according to Equation (A1-38), this is the same as

$$z' = az \, , \tag{A1-65}$$

where z corresponds to Z, z' corresponds to Z', and a corresponds to the Kronecker product matrix $F \otimes G^T$. Identifying Equation (A1-65) with transformation (A1-48), we conclude that the Jacobian of transformation (A1-64) is given by

$$|aa^H| = |FF^H \otimes G^T G^*| \, .$$

Finally, the change of volume element corresponding to this transformation can be expressed in the form

$$d(Z') = |FF^H|^M \, |GG^H|^J \, d(Z) \, . \tag{A1-66}$$

As an example, suppose that Z is a Gaussian array, subject to the pdf given by Equation (A1-62), and consider the "whitening" transformation

$$Z' \equiv B^{-1/2} Z C^{-1/2} \, . \tag{A1-67}$$

Inverting this relation, we see that the volume elements are related according to the equation

$$d(Z) = |B|^M |C|^J \, d(Z') \, .$$

In terms of the expected value of the new random array,

$$\bar{Z}' \equiv EZ' = B^{-1/2} \bar{Z} C^{-1/2} \, ,$$

the joint pdf of Z' is

$$f(Z') = \frac{1}{\pi^{JM}} \, e^{-\mathrm{Tr}\left[(Z'-\bar{Z}')(Z'-\bar{Z}')^H\right]} \, . \tag{A1-68}$$

This pdf is, of course, consistent with the new covariance matrix

$$\mathrm{Cov}(Z') = I_J \otimes I_M \, .$$

G. THE MULTIVARIATE CONDITIONAL GAUSSIAN DISTRIBUTION

Let Z be a Gaussian array, of dimension $J \times M$, with expected value $\bar{Z}$ and covariance given by

$$\mathrm{Cov}(Z) = \Sigma \otimes I_M \,. \tag{A1-69}$$

This special case, in which the columns of Z are independent and share a common covariance matrix Σ, forms the setting for the entire analysis given in the main body of this study. It is also the usual setting for discussions of multivariate Gaussian statistics in the large literature of that subject. The covariance matrix Σ is, of course, a $J \times J$ positive-definite matrix, and, with these assumptions, the joint pdf of Z assumes the form

$$f(Z) = \frac{1}{\pi^{JM} |\Sigma|^M} e^{-\mathrm{Tr}[\Sigma^{-1}(Z-\bar{Z})(Z-\bar{Z})^H]} \,. \tag{A1-70}$$

Let U_J be a unitary matrix, of order J, which is partitioned as follows:

$$U_J \equiv [\, a \quad b \,] \,, \tag{A1-71}$$

where a has dimension $J \times j$, b is $J \times k$, and $j + k = J$. Then, a and b are basis arrays in orthogonal subspaces of $\mathbb{C}^J$. We apply this matrix to Z, viewing the result as a rotation, followed by a partitioning of the Z array. In analogy to the many similar transformations used in the main text, we write this operation in the form

$$U_J^H Z = \begin{bmatrix} a^H Z \\ b^H Z \end{bmatrix} \equiv \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} , \qquad U_J^H \bar{Z} = \begin{bmatrix} a^H \bar{Z} \\ b^H \bar{Z} \end{bmatrix} \equiv \begin{bmatrix} \bar{Z}_1 \\ \bar{Z}_2 \end{bmatrix} , \tag{A1-72}$$

where Z_1 has dimension $j \times M$ and Z_2 is $k \times M$. As indicated by this equation, the mean value array $\bar{Z}$ is also subjected to this rotation and partitioning. The same transformation is applied to both the rows and columns of the covariance matrix Σ:

$$U_J^H \Sigma U_J = \begin{bmatrix} a^H \Sigma a & a^H \Sigma b \\ b^H \Sigma b & b^H \Sigma b \end{bmatrix} \equiv \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} , \tag{A1-73}$$

and also to its inverse:

$$U_J^H \Sigma^{-1} U_J = (U_J^H \Sigma U_J)^{-1} = \begin{bmatrix} a^H\Sigma^{-1}a & a^H\Sigma^{-1}b \\ b^H\Sigma^{-1}a & b^H\Sigma^{-1}b \end{bmatrix} \equiv \begin{bmatrix} \Sigma^{11} & \Sigma^{12} \\ \Sigma^{21} & \Sigma^{22} \end{bmatrix} . \tag{A1-74}$$

These equations serve to define the components of Σ and its inverse relative to the pair of subspaces determined by a and b.

We now apply identity (A1-9) to obtain the formula

$$(Z-\bar{Z})^H \Sigma^{-1}(Z-\bar{Z}) = \left[(Z_1-\bar{Z}_1)^H \ (Z_2-\bar{Z}_2)^H \right] \begin{bmatrix} \Sigma^{11} & \Sigma^{12} \\ \Sigma^{21} & \Sigma^{22} \end{bmatrix} \begin{bmatrix} Z_1-\bar{Z}_1 \\ Z_2-\bar{Z}_2 \end{bmatrix}$$

$$= Y^H \Sigma^{11} Y + (Z_2-\bar{Z}_2)^H \Sigma_{22}^{-1}(Z_2-\bar{Z}_2) , \tag{A1-75}$$

where

$$Y \equiv Z_1-\bar{Z}_1 - \Sigma_{12}\Sigma_{22}^{-1}(Z_2-\bar{Z}_2) . \tag{A1-76}$$

We also note, using Equation (A1-8), that

$$\Sigma^{11} = (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} . \tag{A1-77}$$

Next, by taking the trace of Equation (A1-75), we obtain

$$\mathrm{Tr}\left[\Sigma^{-1}(Z-\bar{Z})(Z-\bar{Z})^H\right] = \mathrm{Tr}(\Sigma^{11} YY^H) + \mathrm{Tr}\left[\Sigma_{22}^{-1}(Z_2-\bar{Z}_2)(Z_2-\bar{Z}_2)^h\right] .$$

From this result, we obtain the formula

$$f(Z)\, d(Z) = f_1(Z_1|Z_2)\, f_2(Z_2)\, d(Z_1)\, d(Z_2) , \tag{A1-78}$$

where

$$f_2(Z_2) = \frac{1}{\pi^{kM}|\Sigma_{22}|^M} e^{-\mathrm{Tr}\left[\Sigma_{22}^{-1}(Z_2-\bar{Z}_2)(Z_2-\bar{Z}_2)^H\right]} \tag{A1-79}$$

and

$$f_1(Z_1|Z_2) = \frac{1}{\pi^{JM}|\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|^M} e^{-\mathrm{Tr}[(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}(Z_1-\bar{Z}_{12})(Z_1-\bar{Z}_{12})^H]} . \tag{A1-80}$$

The volume elements which appear in Equation (A1-78) are all of the kind defined by Equation (A1-63), and the Jacobian of the original unitary transformation is, of course, unity. Identity (A1-2) has been used to factor the determinant of Σ, and the conditional mean of Z_1 which appears in Equation (A1-80) is given by

$$\bar{Z}_{12} \equiv E(Z_1|Z_2) = \bar{Z}_1 + \Sigma_{12}\Sigma_{22}^{-1}(Z_2-\bar{Z}_2) . \tag{A1-81}$$

The corresponding conditional covariance of Z_1 is

$$\mathrm{Cov}(Z_1|Z_2) = (\Sigma^{11})^{-1} \otimes I_M . \tag{A1-82}$$

These formulas are straightforward generalizations of standard results for Gaussian vectors, expressing the pdf of Z as the product of the conditional pdf of Z_1 (given Z_2) and the marginal pdf of Z_2. The conditional expectation given by Equation (A1-81) is, of course, the least-squares predictor of Z_1 (given Z_2), and Y [defined in Equation (A1-76)] is the corresponding prediction error. The conditional expectation of Y is zero, and its conditional covariance matrix is the same as that of Z_1.

H. SOME PROPERTIES OF COMPLEX WISHART MATRICES

We return to the untransformed Gaussian array Z and assume that its mean value is zero. The object of our discussion is the $J \times J$ matrix

$$S \equiv ZZ^H . \tag{A1-83}$$

We also make the assumption that $J \leq M$, in which case S is a complex Wishart matrix of random variables. In accordance with the dimension of the Z array, we say that S is of order J, with M complex degrees of freedom. The notation $CW_J(M,\Sigma)$ is often used to describe the distribution of S. In addition to the dimensional parameters, it indicates the covariance matrix shared by the columns of the original Gaussian array from which S is formed. Whenever Wishart matrices are discussed, it should be understood that the actual covariance matrix of the underlying Z array has the form expressed by Equation (A1-69).

A derivation of the Wishart distribution function is given in Appendix 3. We note here that S is positive definite with probability one, according to this distribution. The S matrix we have defined here is a "central" complex Wishart matrix, because the mean value of the underlying Gaussian array is zero. If this Gaussian array has a non-zero mean value, the corresponding S matrix is subject to a non-central Wishart distribution. The latter distribution is not explicitly discussed in this study, but some of the consequences of a non-vanishing mean value for the underlying Gaussian array are derived later on.

We recall the transformation of Z described by Equation (A1-72) and apply it to the rows and columns of S. The result is a partitioning of S itself, according to the equation

$$U_J^H S U_J = \begin{bmatrix} Z_1 Z_1^H & Z_1 Z_2^H \\ Z_2 Z_1^H & Z_2 Z_2^H \end{bmatrix} \equiv \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} . \qquad \text{(A1-84)}$$

The diagonal blocks in this partitioned matrix are square: S_{11} is of order j and S_{22} is of order k, according to the definitions used previously. The transformation is also applied to the inverse of S, and we write

$$U_J^H S^{-1} U_J \equiv \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{bmatrix} . \qquad \text{(A1-85)}$$

We will now show that the matrix

$$T \equiv (S^{11})^{-1}$$

is a complex Wishart matrix, of order j, with $M-k$ complex degrees of freedom. In addition, we will show that T is independent of the matrix block S_{12}. These properties are indispensable to the analysis carried out in the main text. Making use of Equations (A1-8), we can write

$$\begin{aligned} T &= S_{11} - S_{12} S_{22}^{-1} S_{21} \\ &= Z_1 \left(I_M - Z_2^H (Z_2 Z_2^H)^{-1} Z_2 \right) Z_1^H . \end{aligned} \qquad \text{(A1-86)}$$

Recall that Z_1 is an array of dimension $j \times M$, and note that a projection matrix appears in the second line of the expression for T. This matrix is very similar to the one which occurs in Equation (2-43) of Section 2, and we deal with it in much the same way.

First, an array analogous to p is introduced:

$$\alpha \equiv (Z_2 Z_2^H)^{-1/2} Z_2 \, , \tag{A1-87}$$

which is possible because $Z_2 Z_2^H$ is a complex Wishart matrix of dimension k, with M complex degrees of freedom. Since M exceeds k, this matrix is positive definite (with probability one); hence, it has a positive-definite square-root matrix. The properties

$$\begin{aligned} \alpha\alpha^H &= I_k \\ \alpha^H\alpha &= Z_2^H (Z_2 Z_2^H)^{-1} Z_2 \\ Z_2 &= (Z_2 Z_2^H)^{1/2} \alpha \end{aligned} \tag{A1-88}$$

follow directly from the definition of α. The projection matrix $\alpha^H\alpha$ thus defines a subspace of dimension k of $\mathbb{C}^M$ which is, in fact, the row space of Z_2. Now, corresponding to the q array of Section 2, we introduce an array β, which provides a basis in the orthogonal complement of this subspace. This array has the properties

$$\begin{aligned} \beta\beta^H &= I_{M-k} \\ \alpha\beta^H &= 0 \\ \alpha^H\alpha + \beta^H\beta &= I_M \, . \end{aligned} \tag{A1-89}$$

The two sets of basis vectors form a unitary matrix, in analogy to Equation (2-12):

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \equiv \mathcal{U}_M \, . \tag{A1-90}$$

Finally, we decompose the array Z_1 into further components, according to the definition

$$Z_1 \mathcal{U}_M^H = [Z_{1\alpha} \; Z_{1\beta}] \,. \tag{A1-91}$$

where

$$Z_{1\alpha} \equiv Z_1 \alpha^H$$

$$Z_{1\beta} \equiv Z_1 \beta^H \,. \tag{A1-92}$$

Using this apparatus, we find that T has the form

$$T = Z_1 \beta^H \beta Z_1^H = Z_{1\beta} Z_{1\beta}^H \,. \tag{A1-93}$$

We now condition on the elements the Z_2 array so that the subspaces, as well as the bases introduced in them, become fixed. For brevity of notation, we will use the subscript "2" to indicate this conditioning. The conditional covariance of Z_1 (given Z_2) is expressed by Equation (A1-82), and a straightforward evaluation [using Equation (A1-44)] now gives us the conditional covariance matrix

$$\mathrm{Cov}_2(Z_{1\beta}) = (\Sigma^{11})^{-1} \otimes (\beta\beta^H)^* = (\Sigma^{11})^{-1} \otimes I_{M-k} \,. \tag{A1-94}$$

Thus, $Z_{1\beta}$ is a zero-mean complex Gaussian array with independent columns, when conditioned on Z_2. As the conditioning variables themselves do not appear in any way in this statistical characterization, we have shown that $Z_{1\beta}$ is a zero-mean complex Gaussian array, whose covariance is given by the right side of Equation (A1-94) when the conditioning is removed. Thus, T is a complex Wishart matrix of dimension j. The number of degrees of freedom of this distribution is $M - k$, which is the dimensionality of the subspace onto which $\beta\beta^H$ projects. Since $j = J - k$, we can say that *the number of degrees of freedom of T is smaller than that of S by the same amount that its dimension is less than that of S*. Taking cognizance of the covariance properties of $Z_{1\beta}$, we may say that T has the distribution $CW_j(M-k, \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$.

The array $Z_{1\alpha}$ also has independent columns, and the two components of Z_1 are conditionally independent. To show this, we restore the conditioning on Z_2 and use Equation (A1-45) to make the evaluation

$$\mathrm{Cov}_2(Z_{1\alpha}, Z_{1\beta}) = (\Sigma^{11})^{-1} \otimes (\alpha\beta^H)^* = 0 \,. \tag{A1-95}$$

Since

$$Z_{1\alpha} = Z_1 Z_2^H (Z_2 Z_2^H)^{-1/2} = S_{12} (Z_2 Z_2^H)^{-1/2} ,$$

we can write

$$S_{12} = Z_{1\alpha} (Z_2 Z_2^H)^{1/2} ,$$

from which it follows that

$$\mathrm{Cov}_2(S_{12}, Z_{1\beta}) = 0 . \tag{A1-96}$$

Under the conditioning, S_{12} and $Z_{1\beta}$ are zero-mean Gaussian arrays, and the vanishing of this covariance matrix implies that they are independent as well. Independence means that the joint pdf of both arrays is the product of the separate density functions. Since the conditional pdf of $Z_{1\beta}$ does not depend on the values of the conditioning variables, the joint pdf remains a product of factors when the conditioning is removed. The unconditioned pdf of S_{12} will, of course, be different from the conditional pdf of that array, but S_{12} and $Z_{1\beta}$ are still independent without the conditioning, and it follows that T is unconditionally independent of S_{12}.

If the Z array has a mean value $\bar{Z}$, then this array is transformed and partitioned, along with Z, and its component arrays are defined by Equation (A1-72). The matrix S, defined by Equation (A1-83), is now a non-central complex Wishart matrix. It can be transformed and partitioned as before, after which its components are described by Equation (A1-84) above. S is still positive definite (with probability one), and its inverse can also be transformed and partitioned according to Equation (A1-85). The T array is defined as before, the subspace basis arrays are again introduced, and the analysis up through Equation (A1-93) is valid without change.

When conditioned on the Z_2 array, the covariance of Z_1 is still expressed by Equation (A1-82), but the conditional mean value, no longer zero, is given by Equation (A1-81). Equation (A1-94) still correctly describes the conditional covariance matrix of $Z_{1\beta}$, but the conditional mean of this array is now given by

$$E_2 Z_{1\beta} = \left[\bar{Z}_1 + \Sigma_{12} \Sigma_{22}^{-1} (Z_2 - \bar{Z}_2)\right] \beta^H .$$

Since

$$Z_2 \beta^H = (Z_2 Z_2^H)^{1/2} \alpha \beta^H = 0 ,$$

we can write

$$E_2 Z_{1\beta} = \tilde{Z}_1 \beta^H . \tag{A1-97}$$

where

$$\tilde{Z}_1 \equiv \bar{Z}_1 - \Sigma_{12} \Sigma_{22}^{-1} \bar{Z}_2 . \tag{A1-98}$$

The conditional probability density function of $Z_{1\beta}$ is still Gaussian, but the mean value of this pdf depends on the conditioning variables through the basis array β which enters the conditional mean. This fact destroys the Wishart character of T when the conditioning is removed. It also precludes the independence of T and S_{12}, since we can no longer infer independence from the vanishing of the conditional covariance matrix, although Equation (A1-96) remains valid. In spite of these complications, the analysis just given is useful in connection with another property of the Wishart matrices, to which we now turn.

We assume that Z is a complex Gaussian array, with a non-zero mean value, which is partitioned into components Z_1 and Z_2, as discussed above. Let U_M be a unitary matrix of order M, partitioned as follows:

$$U_M = \begin{bmatrix} c \\ d \end{bmatrix} , \tag{A1-99}$$

where c is of dimension $m \times M$, d is $n \times M$, and $m + n = M$. Then, c and d form basis arrays in complementary orthogonal subspaces of $\mathbb{C}^M$. We post-multiply the arrays Z_1 and Z_2 by the Hermitian transpose of U_M, and use its partitioning to define new component arrays:

$$Z_1 U_M^H = \left[Z_1 c^H \;\; Z_1 d^H \right] \equiv [X_1 \;\; Y_1]$$

$$Z_2 U_M^H = \left[Z_2 c^H \;\; Z_2 d^H \right] \equiv [X_2 \;\; Y_2] . \tag{A1-100}$$

We also replace the restriction $J \leq M$ by the stronger condition $J \leq n$.

The transformed Z array is therefore partitioned into four components:

$$U_J^H Z U_M^H = \begin{bmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{bmatrix} . \tag{A1-101}$$

It is also useful to introduce the notation

$$Z U_M^H \equiv [\, X \quad Y \,] , \tag{A1-102}$$

so that

$$U_J^H X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} , \qquad U_J^H Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} . \tag{A1-103}$$

The covariance matrix of Z is given by Equation (A1-69), and the covariance matrices of the components X and Y are easily found to be

$$\begin{aligned} Cov(X) &= \Sigma \otimes I_m \\ Cov(Y) &= \Sigma \otimes I_n . \end{aligned} \tag{A1-104}$$

The mean values of the component arrays X_1 and X_2 are denoted by overbars, and we assume that the means of Y_1 and Y_2 are zero. Then, we can write

$$U_J^H \bar{Z} U_M^H = \begin{bmatrix} \bar{X}_1 & 0 \\ \bar{X}_2 & 0 \end{bmatrix} . \tag{A1-105}$$

This specialization is necessary for the results that follow, and it is also consistent with the situation which arises in the general problem formulated in the main text.

Making use of Equation (A1-102), we can express the S matrix in the form

$$S = Z Z^H = X X^H + S_Y , \tag{A1-106}$$

where

$$S_Y \equiv Y Y^H . \tag{A1-107}$$

Since Y has zero mean and $J \leq n$, S_Y is a (central) complex Wishart matrix whose distribution is $CW_J(n,\Sigma)$, and whose inverse exists with probability one. The components of S and its inverse, after transformation by U_J, are given by Equations (A1-84) and (A1-85). We make analogous definitions for the components of S_Y and its inverse, after the same transformation:

$$U_J^H S_Y U_J = \begin{bmatrix} Y_1 Y_1^H & Y_1 Y_2^H \\ Y_2 Y_1^H & Y_2 Y_2^H \end{bmatrix} \equiv \begin{bmatrix} S_{Y_{11}} & S_{Y_{12}} \\ S_{Y_{21}} & S_{Y_{22}} \end{bmatrix} , \tag{A1-108}$$

and

$$U_J^H S_Y^{-1} U_J \equiv \begin{bmatrix} S_Y^{11} & S_Y^{12} \\ S_Y^{21} & S_Y^{22} \end{bmatrix} . \tag{A1-109}$$

By our previous results, the $j \times j$ matrix

$$T_Y \equiv (S_Y^{11})^{-1} = Y_1 \left[I_n - Y_2^H (Y_2 Y_2^H)^{-1} Y_2 \right] Y_1^H$$

is a (central) complex Wishart matrix. We define

$$\alpha_Y \equiv (Y_2 Y_2^H)^{-1/2} Y_2 , \tag{A1-110}$$

which is the analog of α in the previous analysis, and which serves as a basis array in the k-dimensional row space of Y_2. We also introduce the array β_Y, analogous to β, which is a basis array in the $(n-k)$-dimensional orthogonal complement of this row space. It follows that

$$\alpha_Y \alpha_Y^H = I_k$$

$$\beta_Y \beta_Y^H = I_{n-k} ,$$

and that

$$\alpha_Y^H \alpha_Y + \beta_Y^H \beta_Y = I_n .$$

Finally, in analogy to Equation (A1-93), we have

$$T_Y = Y_{1\beta} Y_{1\beta}^H , \tag{A1-111}$$

where

$$Y_{1\beta} \equiv Y_1 \beta_Y^H . \tag{A1-112}$$

$Y_{1\beta}$ is a zero-mean complex Gaussian array, whose covariance matrix is

$$\mathrm{Cov}(Y_{1\beta}) = (\Sigma^{11})^{-1} \otimes I_{n-k} . \tag{A1-113}$$

This formula completes the characterization of T_Y as a complex Wishart matrix by showing that it has $n-k$ complex degrees of freedom, and by exhibiting the covariance matrix shared by the columns of the underlying Gaussian array $Y_{1\beta}$.

The matrix S, formed from the full Z array, is subject to a non-central complex Wishart distribution. As noted above, we can still introduce the matrix

$$T \equiv (S^{11})^{-1} = Z_1 \left[I_M - Z_2^H (Z_2 Z_2^H)^{-1} Z_2 \right] Z_1^H , \tag{A1-114}$$

and the basis array α of the row space of Z_2:

$$\alpha \equiv (Z_2 Z_2^H)^{-1/2} Z_2 . \tag{A1-115}$$

Then, we have

$$T = Z_1 (I_M - \alpha^H \alpha) Z_1^H . \tag{A1-116}$$

It will now be shown that T can be expressed in terms of T_Y, in the form

$$T = \xi \xi^H + T_Y , \tag{A1-117}$$

where ξ is a $j \times m$ array, independent of T_Y, whose statistical characteristics will subsequently be derived. Equation (A1-117) resembles Equation (A1-106), and ξ (like X) will have a non-vanishing mean value which is dependent on the components of the original mean array $\bar{Z}$.

To establish this result, we must find a link between the subspace decompositions described by α_Y and β_Y, which relate to the row space of Y_2, and that described by α, our basis in the row space of Z_2. We define the array

$$\beta_2 \equiv [\,0 \;\; \beta_Y\,]\, U_M \,. \tag{A1-118}$$

in which the null array is of dimension $(n-k)\times m$. We observe that the rows of β_2 are orthonormal:

$$\beta_2 \beta_2^H = [\,0 \;\; \beta_Y\,] \begin{bmatrix} 0 \\ \beta_Y^H \end{bmatrix} = \beta_Y \beta_Y^H = I_{n-k} \,.$$

Since β_Y is orthogonal to the row space of Y_2, the extended array $[0 \; \beta_Y]$ is orthogonal to the row space of $[X_2 \; Y_2]$. Post-multiplication by the unitary matrix U_M produces an array which is orthogonal to the row space of Z_2:

$$Z_2 \beta_2^H = Z_2 U_M^H \begin{bmatrix} 0 \\ \beta_Y^H \end{bmatrix} = [\,X_2 \;\; Y_2\,] \begin{bmatrix} 0 \\ \beta_Y^H \end{bmatrix}$$

$$= Y_2 \beta_Y^H = (Y_2 Y_2^H)^{1/2} \alpha_Y \beta_Y^H = 0 \,, \tag{A1-119}$$

and this relation provides the link we seek. We do not expect, however, that β_2 will provide a basis for the full orthogonal complement of this row space.

The span of α is k-dimensional, while that of β_2 is of dimension $(n-k)$. These spaces are orthogonal but they do not exhaust $\mathbb{C}^M$, and there is an m-dimensional subspace left over which is orthogonal to the spans of both α and β_2. Let β_1 be an orthonormal basis array in this remaining subspace, so that we have

$$\alpha \alpha^H = I_k$$

$$\beta_1 \beta_1^H = I_m$$

$$\beta_2 \beta_2^H = I_{n-k} \,, \tag{A1-120}$$

and

$$\alpha^H \alpha + \beta_1^H \beta_1 + \beta_2^H \beta_2 = I_M \,. \tag{A1-121}$$

From the latter relation, together with Equation (A1-116), we obtain

$$T = Z_1(\beta_1^H\beta_1 + \beta_2^H\beta_2)Z_1^H .$$

But

$$Z_1\beta_2^H = Z_1 U_M^H \begin{bmatrix} 0 \\ \beta_Y^H \end{bmatrix} = [X_1 \;\; Y_1] \begin{bmatrix} 0 \\ \beta_Y^H \end{bmatrix}$$

$$= Y_1\beta_Y^H = Y_{1\beta} , \tag{A1-122}$$

in direct analogy to the derivation of Equation (A1-119), and, therefore,

$$T = Z_1\beta_1^H\beta_1 Z_1^H + Y_{1\beta} Y_{1\beta}^H = X_{1\beta}X_{1\beta}^H + T_Y , \tag{A1-123}$$

where

$$X_{1\beta} \equiv Z_1\beta_1^H . \tag{A1-124}$$

A similar formula, expressing $Y_{1\beta}$ directly in terms of Z_1, is provided by Equation (A1-122).

We condition on the elements of the Z_2 array, which includes the array Y_2; thus, all the subspaces and the basis arrays introduced in them are now fixed. Under this conditioning, $X_{1\beta}$ and $Y_{1\beta}$ are complex Gaussian arrays, the latter with zero mean. Using definition (A1-124) and Equation (A1-122), we evaluate the conditional covariance matrices of these arrays:

$$\mathrm{Cov}_2(X_{1\beta}) = (\Sigma^{11})^{-1} \otimes (\beta_1\beta_1^H)^* = (\Sigma^{11})^{-1} \otimes I_m$$

$$\mathrm{Cov}_2(Y_{1\beta}) = (\Sigma^{11})^{-1} \otimes (\beta_2\beta_2^H)^* = (\Sigma^{11})^{-1} \otimes I_{n-k} .$$

These results are consequences of Equation (A1-82), of course, and, as they do not depend on the values of the conditioning variables, they remain valid when the conditioning is removed. Thus, Equation (A1-113) (which expresses the unconditioned covariance matrix of $Y_{1\beta}$) is recovered, and we also have

$$\mathrm{Cov}(X_{1\beta}) = (\Sigma^{11})^{-1} \otimes I_m . \tag{A1-125}$$

Since β_1 and β_2 are basis arrays of orthogonal subspaces, we see that $X_{1\beta}$ and $Y_{1\beta}$ are conditionally uncorrelated:

$$\mathrm{Cov}_2(X_{1\beta}, Y_{1\beta}) = (\Sigma^{11})^{-1} \otimes (\beta_1 \beta_2^H)^* = 0 \ . \tag{A1-126}$$

This equation implies independence when the conditioning is removed, since the conditional probability density function of $Y_{1\beta}$ (which has zero mean) does not depend in any way on the values of the conditioning variables. Thus, T_Y itself is independent of $X_{1\beta}$.

It remains only to discuss the mean value of the array $X_{1\beta}$ and to identify the array ξ to complete the proof of our assertion, expressed by Equation (A1-117). We begin with the conditioning on Z_2 in effect, and, from definition (A1-124), we obtain

$$E_2 X_{1\beta} = (E_2 Z_1) \beta_1^H \ .$$

Equation (A1-81), which is applicable to the present analysis, states that

$$E_2 Z_1 = \bar{Z}_1 + \Sigma_{12} \Sigma_{22}^{-1} (Z_2 - \bar{Z}_2) \ .$$

Since

$$Z_2 \beta_1^H = (Z_2 Z_2^H)^{1/2} \alpha \beta_1^H = 0 \ ,$$

we obtain

$$E_2 X_{1\beta} = (\bar{Z}_1 - \Sigma_{12} \Sigma_{22}^{-1} \bar{Z}_2) \beta_1^H \ , \tag{A1-127}$$

in direct analogy to our earlier discussion of the effects of a non-zero mean value on the properties of Wishart matrices. Following that discussion another step, we make the definition

$$\tilde{X}_1 \equiv \bar{X}_1 - \Sigma_{12} \Sigma_{22}^{-1} \bar{X}_2 \ . \tag{A1-128}$$

From Equations (A1-100), we deduce that

$$(\bar{Z}_1 - \Sigma_{12} \Sigma_{22}^{-1} \bar{Z}_2) U_M^H = [\, \tilde{X}_1 \quad 0 \,] \ ,$$

since the Y-components have zero means. Combining these results and recalling Equation (A1-99), we obtain

$$E_2 X_{1\beta} = [\, \tilde{X}_1 \;\; 0 \,]\, U_M \beta_1^H = \tilde{X}_1 c \beta_1^H . \tag{A1-129}$$

The $X_{1\beta}$ array is of dimension $j \times m$. We let $\mathcal{W}_m$ be a unitary matrix, of order m, which will be precisely defined later. This matrix will be a function of the conditioning variables, but it is constant under the conditioning. We also define ξ in terms of $\mathcal{W}_m$, as follows:

$$\xi \equiv X_{1\beta} \mathcal{W}_m^H . \tag{A1-130}$$

Obviously, we have

$$T = \xi \xi^H + T_Y , \tag{A1-131}$$

so that the form of this representation of T is not affected by the choice of $\mathcal{W}_m$. ξ is a Gaussian array under the conditioning, with the same covariance matrix as $X_{1\beta}$. The conditional mean of ξ is, of course,

$$E_2 \xi = \tilde{X}_1 c \beta_1^H \mathcal{W}_m^H .$$

Let us put

$$\psi \equiv c \beta_1^H , \tag{A1-132}$$

and observe that ψ is a square matrix, of order m, since c and β_1 are both of dimension $m \times M$. We now evaluate

$$\psi \psi^H = c \beta_1^H \beta_1 c^H .$$

Making use of Equation (A1-121), we have

$$\psi \psi^H = c\,(I_M - \beta_2^H \beta_2 - \alpha^H \alpha)\, c^H ,$$

and, from definitions (A1-99) and (A1-118), it follows that

$$c\beta_2^H = [I_m \;\; 0]\, U_M \beta_2^H = [I_m \;\; 0] \begin{bmatrix} 0 \\ \beta_Y^H \end{bmatrix} = 0 .$$

We have therefore found that

$$\psi\psi^H = c(I_M - \alpha^H\alpha)c^H$$

$$= I_m - c Z_2^H (Z_2 Z_2^H)^{-1} Z_2 c^H .$$

The fact that $cc^H = I_m$ follows from the unitary character of U_M. From Equation (A1-100), we now obtain

$$\psi\psi^H = I_m - X_2^H (X_2 X_2^H + Y_2 Y_2^H)^{-1} X_2$$

$$= \left[I_m + X_2^H (Y_2 Y_2^H)^{-1} X_2 \right]^{-1} , \qquad \text{(A1-133)}$$

the last step being an application of Equation (A1-5).

We define

$$C_m \equiv I_m + X_2^H (Y_2 Y_2^H)^{-1} X_2 , \qquad \text{(A1-134)}$$

so that

$$\psi\psi^H = C_m^{-1} ,$$

and observe that C_m is a positive-definite matrix, which is constant under the conditioning. It follows that ψ is non-singular and that

$$C_m^{1/2}\psi = (\psi\psi^H)^{-1/2}\psi$$

is unitary. We now make the deferred choice

$$\mathcal{W}_m = (\psi\psi^H)^{-1/2}\psi , \qquad \text{(A1-135)}$$

and we find that

$$\psi \mathcal{W}_m^H = (\psi \psi^H)^{1/2} = C_m^{-1/2} .$$

Finally, we obtain the desired form

$$E_2 \xi = \tilde{X}_1 \psi \mathcal{W}_m^H = \tilde{X}_1 C_m^{-1/2} . \tag{A1-136}$$

The conditioning variables survive only through the matrix C_m, whose statistical character (when the conditioning is removed) we now investigate. Y_2 is a zero-mean complex Gaussian array, whose covariance matrix is

$$\mathrm{Cov}(Y_2) = \Sigma_{22} \otimes I_n ,$$

in agreement with Equation (A1-79). Therefore,

$$S_{Y_{22}} = Y_2 Y_2^H \tag{A1-137}$$

is a complex Wishart matrix, of order k, and with n complex degrees of freedom. In the notation used earlier, its distribution is $CW_k(n, \Sigma_{22})$. The X_2 array is also complex Gaussian, independent of Y_2, with mean and covariance arrays given by

$$E X_2 = \bar{X}_2$$

$$\mathrm{Cov}(X_2) = \Sigma_{22} \otimes I_m .$$

We have shown that T can be expressed in the form given in Equation (A1-117), where the ξ array is statistically independent of the complex Wishart matrix T_Y. We have also seen that ξ is conditionally Gaussian, with conditional mean value

$$E(\xi | C_m) = \tilde{X}_1 C_m^{-1/2} , \tag{A1-138}$$

and with the unconditioned covariance matrix

$$\mathrm{Cov}(\xi) = (\Sigma^{11})^{-1} \otimes I_m . \tag{A1-139}$$

We can express these properties in a convenient way by making the definition

$$\xi = \xi_s + \xi_n . \tag{A1-140}$$

where

$$\xi_s \equiv \tilde{X}_1 C_m^{-1/2} . \tag{A1-141}$$

Then, ξ_n is a zero-mean complex Gaussian array, with covariance matrix

$$\mathrm{Cov}(\xi_n) = (\Sigma^{11})^{-1} \otimes I_m , \tag{A1-142}$$

and the three quantities T_Y, ξ_n, and ξ_s are statistically independent. The statistical characterization of ξ_s is provided by the definitions (A1-141), (A1-128), and (A1-134), together with the properties just established for the complex Gaussian arrays X_2 and Y_2.

The matrix C_m belongs to a family of complex random matrices which are generalizations of the $\mathscr{C}$ matrices introduced in Section 4. The generalization lies with the fact that the X_2 array has a non-zero mean. A special case of this generalized $\mathscr{C}$ matrix was discussed in Sections 5 and 6, in connection with the presence of "signal mismatch," a feature introduced in Section 3. The $\mathscr{C}$ matrices are also discussed in Appendix 3, where their relation to the complex multivariate F and Beta variables is established.

As an application of these results, consider the ratio

$$\ell \equiv \frac{|T|}{|T_Y|} = \frac{|a^H S_Y^{-1} a|}{|a^H S^{-1} a|} = \frac{|a^H S_Y^{-1} a|}{|a^H (S_Y + X X^H)^{-1} a|} . \tag{A1-143}$$

This quantity has exactly the same form as one of the versions of the GLR test statistic, obtained in Section 2 and expressed by Equation (2-56). Using Equation (A1-131), we can write

$$\ell = \frac{|\xi\xi^H + T_Y|}{|T_Y|} = |I_j + \xi^H T_Y^{-1} \xi| , \tag{A1-144}$$

which is directly analogous to Equation (3-15) of Section 3. With the appropriate identifications of terms, we can therefore use the results obtained here to derive the statistical properties of the GLR test, starting from Equation (2-56) and leading to Equation (3-15), with "signal mismatch" included.

APPENDIX 2

COMPLEX DISTRIBUTIONS RELATED TO THE GAUSSIAN

We introduce here the complex analogs of the chi-squared, F, and Beta distributions. In real-variable statistics these distributions are usually treated as a family, based on their definitions in terms of real Gaussian vector variables. The complex distributions bear the same relationship to one or more complex Gaussian vectors of the kind discussed in Appendix 1.

Let u be a complex Gaussian vector, of dimension n, with zero mean and covariance matrix I_n. The components of this vector are independent, with "complex variance" unity;

$$E|u_i|^2 = 1 \ .$$

Each component represents a pair of independent real variables, both of which have mean zero and variance one-half. The scalar

$$y \equiv u^H u = \mathrm{Tr}(uu^H) = \sum_{i=1}^{n} |u_i|^2 \tag{A2-1}$$

will be called a complex chi-squared random variable, with n complex degrees of freedom. This usage differs from that of real-variable statistics, where 2y would be called chi-squared, with 2n degrees of freedom.

The pdf of y is given by the familiar formula

$$f_\chi(y;n) = \frac{y^{n-1}}{(n-1)!} e^{-y} \ . \tag{A2-2}$$

The cumulative distribution function of y is $1 - G_n(y)$, where

$$G_n(y) = \int_y^\infty f_\chi(y';n)\,dy' = e^{-y} \sum_{k=0}^{n-1} \frac{y^k}{k!} \ . \tag{A2-3}$$

This function, which appears elsewhere in the analysis, is the incomplete Gamma function.[25]

When the means of the underlying Gaussian vectors of any of these distributions are zero, the corresponding distribution is called "central." The non- central complex chi-squared variable is still defined by Equation (A2-1), but the mean vector of u is no longer zero. The non-central complex chi-squared pdf depends on this mean only through the scalar "non-centrality parameter"

$$c \equiv \sum_{i=1}^{n} |Eu_i|^2 = (Eu)^H Eu \ . \tag{A2-4}$$

The corresponding pdf is

$$f_\chi(y;n|c) = e^{-y-c}\,(y/c)^{(n-1)/2}\, I_{n-1}(2\sqrt{cy}) \ , \tag{A2-5}$$

which is well known in radar detection theory. In this formula, I_n is the modified Bessel function, and the series obtained from its definition,

$$\frac{n!}{x^{n/2}}\, I_n(2\sqrt{x}) = \sum_{k=0}^{\infty} \frac{n!}{(n+k)!}\,\frac{x^k}{k!} = {}_0F_1(n+1;x) \ , \tag{A2-6}$$

is a hypergeometric function. Thus, Equation (A2-5) may be written in the form

$$f_\chi(y;n|c) = f_\chi(y;n)\, e^{-c}\, {}_0F_1(n;cy) \ . \tag{A2-7}$$

The cumulative non-central complex chi-squared distribution is, of course, directly related to the Marcum Q-function.[24]

The ratio of two complex chi-squared variables obeys the complex F distribution. Let u be a zero-mean complex Gaussian vector, as before, and let w be an independent complex Gaussian vector, of dimension m. The mean of w is also zero, and its covariance matrix is I_m. The ratio

$$x \equiv \frac{u^H u}{w^H w} = \frac{\sum_{i=1}^{n} |u_i|^2}{\sum_{j=1}^{m} |w_j|^2} \tag{A2-8}$$

will be called a complex central F random variable. We signify this by writing

$$x = x_F(n,m) .$$

The symbol on the right is a generic designator, rather than a specific random variable. The pdf of the complex central F variable follows easily from the standard formula for the pdf of a ratio of random variables:

$$f_F(x;n,m) = \int_0^\infty f_\chi(xy;n)\, f_\chi(y;m)\, y\, dy$$

$$= \frac{(n+m-1)!}{(n-1)!(m-1)!} \frac{x^{n-1}}{(1+x)^{n+m}} . \tag{A2-9}$$

The complex central Beta variable is closely related to the F variable. If u and w have the same meanings as before, then

$$\rho \equiv \frac{\sum_{i=1}^{n} |u_i|^2}{\sum_{i=1}^{n} |u_i|^2 + \sum_{j=1}^{m} |w_j|^2} = \frac{1}{1 + \frac{w^H w}{u^H u}} \tag{A2-10}$$

will be called a complex central Beta random variable. We use the generic notation

$$\rho = x_\beta(n,m)$$

to signify this statistical character. From Equation (A2-10), we obviously have

$$x_\beta(n,m) = \frac{1}{1 + x_F(m,n)} . \tag{A2-11}$$

Observe the transposition of parameters in this relationship, which occurs because we have retained some of the conventions[31] of real-variable statistics in making these definitions. The pdf of the complex central Beta is obtained from that of the complex central F by a simple change of variable:

$$f_\beta(\rho;n,m) = \frac{(n+m-1)!}{(n-1)!(m-1)!}\,\rho^{n-1}(1-\rho)^{m-1} . \qquad \text{(A2-12)}$$

The cumulative complex central Beta distribution is defined as

$$F_\beta(\rho;n,m) \equiv \int_0^\rho f_\beta(\rho';n,m)\,d\rho' , \qquad \text{(A2-13)}$$

and it is given by[26]

$$\begin{aligned} F_\beta(\rho;n,m) &= \rho^{n+m-1}\sum_{k=0}^{m-1}\binom{n+m-1}{k}\left(\frac{1-\rho}{\rho}\right)^k \\ &= \frac{1}{m+n}\sum_{k=0}^{m-1} f_\beta(\rho;n+m-k,k+1) \\ &= 1 - \frac{1}{m+n}\sum_{k=0}^{n-1} f_\beta(\rho;n-k;m+k+1) . \end{aligned} \qquad \text{(A2-14)}$$

This result is easily verified by repeated partial integration, proceeding directly from definition (A2-13).

The cumulative complex central F distribution is defined in a similar way:

$$F_F(x;n,m) \equiv \int_0^x f_F(x';n,m)\,dx' . \qquad \text{(A2-15)}$$

In view of Equation (A2-11), we have

$$F_F(x;n,m) = 1 - F_\beta(1/(1+x);m,n) ,$$

from which we obtain the analog of Equation (A2-14):

$$F_F(x;n,m) = \frac{x^n}{(1+x)^{n+m-1}} \sum_{k=0}^{m-1} \binom{n+m-1}{n+k} x^k . \qquad \text{(A2-16)}$$

The non-central complex F variable is still defined by Equation (A2-8), but the mean value vector of u is no longer zero. Being the ratio of a non-central complex chi-squared variable to a central one, the non-central complex F distribution can depend on the mean of u only through the non-centrality parameter c, defined in Equation (A2-4). We use the generic notation

$$x = x_F(n,m|c)$$

for this random variable. Its pdf is evaluated from the integral

$$f_F(x;n,m|c) = \int_0^\infty f_\chi(xy;n|c) f_\chi(y;m)\, y\, dy \ , \qquad \text{(A2-17)}$$

by substituting the series (A2-6) in the non-central complex chi-squared density, and performing the integration term by term. The resulting series is recognized as a confluent hypergeometric function:

$$f_F(x;n,m|c) = f_F(x;n,m)\, e^{-c}\, {}_1F_1[n+m;n;cx/(1+x)] . \qquad \text{(A2-18)}$$

The non-central complex Beta variable is defined by the generic relation

$$x_\beta(n,m|c) = \frac{1}{1+x_F(m,n|c)} \ ,$$

and its pdf follows directly from Equation (A2-18) by means of a change of variable:

$$f_\beta(\rho;n,m|c) = f_\beta(\rho;n,m)\, e^{-c}\, {}_1F_1[n+m;m;c(1-\rho)] . \qquad \text{(A2-19)}$$

In order to make connection with the notation of real-variable statistics,[31] we must recall that the real dimensional parameters corresponding to n and m are 2n and

2m, respectively, and that the real non-centrality parameter is 2c because of our convention for the variances of our complex Gaussian variables.

If the defining series for the confluent hypergeometric function[25] is substituted in Equation (A2-19), the non-central complex Beta pdf assumes the interesting form:

$$f_\beta(\rho;n,m|c) = e^{-c} \sum_{k=0}^{\infty} f_\beta(\rho;n;m+k) \frac{c^k}{k!} . \tag{A2-20}$$

This distribution can also be expressed in finite form, by making use of some well-known properties of the confluent hypergeometric function. First, the Kummer transformation[25,26]

$${}_1F_1(n;m;x) = e^x \, {}_1F_1(m-n;m;-x) \tag{A2-21}$$

is applied to Equation (A2-19), which results in a hypergeometric function whose first parameter is a non-positive integer. Functions of this kind reduce to polynomials, according to[25,26]

$${}_1F_1(-n;m;x) = \sum_{k=0}^{n} \frac{n!}{(n-k)!} \frac{(m-1)!}{(m+k-1)!} \frac{(-x)^k}{k!} , \tag{A2-22}$$

provided $n \geq 0$. Combining these facts, we obtain the result

$$f_\beta(\rho;n,m|c) = f_\beta(\rho;n,m) e^{-c\rho} \sum_{k=0}^{n} \binom{n}{k} \frac{(m-1)!}{(m+k-1)!} c^k (1-\rho)^k$$

$$= e^{-c\rho} \sum_{k=0}^{n} \binom{n}{k} \frac{(n+m-1)!}{(n+m+k-1)!} c^k f_\beta(\rho;n;m+k) . \tag{A2-23}$$

A similar expression can be derived for the non-central complex F distribution:

$$f_F(x;n,m|c) = f_F(x;n,m) e^{-c/(1+x)} \sum_{k=0}^{m} \binom{m}{k} \frac{(n-1)!}{(n+k-1)!} \left(\frac{cx}{1+x}\right)^k . \tag{A2-24}$$

The cumulative non-central complex Beta distribution is defined by

$$F_\beta(\rho;n,m|c) \equiv \int_0^\rho f_\beta(\rho';n,m|c)\,d\rho' . \qquad \text{(A2-25)}$$

We substitute Equation (A2-20) in the integral (A2-25) and use Equation (A2-14) to evaluate the typical term:

$$F_\beta(\rho;n,m+k) = \rho^{n+m+k-1}\sum_{j=0}^{m+k-1}\binom{n+m+k-1}{j}\left(\frac{1-\rho}{\rho}\right)^j$$

$$= 1 - \rho^{n-1}(1-\rho)^{m+k}\sum_{j=0}^{n-1}\binom{n+m+k-1}{j+m+k}\left(\frac{1-\rho}{\rho}\right)^j .$$

Combining these results, we find

$$F_\beta(\rho,n,m|c) = 1 - e^{-c}\rho^{n-1}(1-\rho)^m$$

$$\times\sum_{k=0}^{\infty}\frac{c^k(1-\rho)^k}{k!}\sum_{j=0}^{n-1}\binom{n+m+k-1}{j+m+k}\left(\frac{1-\rho}{\rho}\right)^j .$$

Reversing the order of summation, we again recognize the series as a confluent hypergeometric function, and thus

$$F_\beta(\rho;n,m|c) = 1 - \rho^{n-1}(1-\rho)^m\sum_{j=0}^{n-1}\frac{(n+m-1)!}{(n-j-1)!(j+m)!}\left(\frac{1-\rho}{\rho}\right)^j$$

$$\times\, e^{-c}\,{}_1F_1[n+m;j+m+1;c(1-\rho)] .$$

The Kummer transformation can be applied once more, and, with the help of Equations (A2-21) and (A2-22), we obtain

$$F_\beta(\rho;n,m|c) = 1 - e^{-c\rho}\rho^{n-1}(1-\rho)^m\sum_{j=0}^{n-1}\sum_{k=0}^{n-j-1}\binom{n+m-1}{j+m+k}\frac{(c\rho)^k}{k!}\left(\frac{1-\rho}{\rho}\right)^{j+k} .$$

The summation indices are now changed by introducing the sum $j+k$ as a new index in place of j. The new index is then called k, and the incomplete Gamma function [Equation (A2-3)] is introduced. The result is

$$F_\beta(\rho;n,m|c) = 1 - \rho^{n-1}(1-\rho)^m \sum_{k=0}^{n-1} \binom{n+m-1}{k+m} \left(\frac{1-\rho}{\rho}\right)^k G_{k+1}(c\rho) \ , \tag{A2-26}$$

or, finally,

$$F_\beta(\rho;n,m|c) = 1 - \frac{1}{n+m} \sum_{k=0}^{n-1} f_\beta(\rho;n-k,m+k+1)\, G_{k+1}(c\rho) \ . \tag{A2-27}$$

When c is zero the G_k functions are all equal to unity, hence Equation (A2-27) reduces to Equation (A2-14).

The cumulative non-central complex F distribution:

$$F_F(x;n,m|c) \equiv \int_0^x f_F(x';n,m|c)\,dx' \ . \tag{A2-28}$$

is obtained from that of the non-central complex Beta by the same procedure used in the central case. We have

$$F_F(x;n,m|c) = 1 - F_\beta[1/(1+x);m,n|c]$$

$$= \frac{1}{n+m} \sum_{k=0}^{m-1} f_\beta[1/(1+x);m-k,n+k+1]\, G_{k+1}\left(\frac{c}{1+x}\right) \ ,$$

and, finally

$$F_F(x;n,m|c) = \frac{x^n}{(1+x)^{n+m-1}} \sum_{k=0}^{m-1} \binom{n+m-1}{k+n} x^k\, G_{k+1}\left(\frac{c}{1+x}\right) \ . \tag{A2-29}$$

When $c=0$, Equation (A2-29) reverts immediately to Equation (A2-16). In Reference 5, formulas (A2-20), (A2-23), and (A2-29) are derived by a different technique, starting directly from the Gaussian distribution.

APPENDIX 3

INTEGRATION LEMMAS AND INTEGRAL REPRESENTATIONS

In this Appendix we discuss the properties of certain random matrices from a different point of view than the one employed in the text. Some results obtained already are re-derived, and some new ones (needed in the main analysis) are derived here. The approach is based on a general technique of multiple integration, which is applied to derive the multivariate generalizations of the complex F and Beta distributions. This technique also provides a very direct derivation of the Wishart pdf itself. The analysis is confined to the "central" case, in which all the Gaussian arrays which appear have mean values of zero. Specific applications are made to the GLR test statistic, in the special case in which no signal components are present.

In Appendix 1 we discussed some properties of multiple integration in which the variables of integration are the complex-valued elements of an array. This array is generally rectangular in shape, and the volume element is called d(Z). The dimensionality of the underlying real space is twice the number of elements in Z, and integration is carried out with respect to the ordinary Euclidean measure in this space. The fact that we describe the integration variables in terms of a complex array Z has no impact on the character of integration in this case. The integration technique we introduce here is based on another space, whose elements (points) are Hermitian matrices of order J.

Let A and B be $J \times J$ Hermitian matrices, and let x and y be real numbers. Then, the Hermitian matrix $xA + yB$ is also a point in our space, which is therefore shown to be a real vector space. We introduce an inner product in this space, as follows:

$$[A,B] \equiv \mathrm{Tr}(AB) \ . \qquad \text{(A3-1)}$$

It is easily verified that this definition satisfies the requirements of an inner product in a real vector space. In particular, it is a symmetric function f A and B as a result of an elementary property of the *trace* operator which we have frequently utilized. The squared norm of a vector in the space is given by

$$\|A\|^2 \equiv [A,A] = \sum_{i,j=1}^{J} |A_{i,j}|^2 \ . \qquad \text{(A3-2)}$$

which is one of the several norms commonly used in connection with matrices.

A Hermitian matrix of order J is described by J^2 real numbers, hence the new space is of dimension J^2. We can map its points onto a real space of J^2 dimensions, as follows. Let the real variables $a_1 \ldots a_J$ be equal to the diagonal elements of the Hermitian matrix A:

$$a_j \equiv A_{j,j} \qquad 1 \leq j \leq J \ . \tag{A3-3}$$

and let

$$a_{J+1} + i a_{J+2} \equiv \sqrt{2} A_{1,2} \ . \tag{A3-4}$$

Continuing in this way, pairs of real variables are defined in terms of the remaining complex elements of A which lie above the main diagonal. The reason the square root of 2 is included in these definitions will become apparent shortly.

Let A and B be Hermitian matrices, and let a and b stand for the real vectors, of dimension J^2, which correspond to them according to the mapping just defined:

$$A \longleftrightarrow a \ , \qquad B \longleftrightarrow b \ .$$

Then, we can evaluate the inner product of A and B in terms of a and b, as follows:

$$\begin{aligned} [A,B] &= \sum_{i,j=1}^{J} A_{i,j} B_{i,j}^{*} \\ &= \sum_{j=1}^{J} A_{j,j} B_{j,j} + \sum_{1 \leq i<j \leq J} (A_{i,j} B_{i,j}^{*} + A_{i,j}^{*} B_{i,j}) \\ &= \sum_{j=1}^{J^2} a_j b_j = (a,b) \ . \end{aligned} \tag{A3-5}$$

The last form is the conventional inner product in the real space which contains a and b. We have shown that the mapping defined above preserves inner products, and thus also norms, with our definitions of these quantities.

The mapping is now applied to sets of points in the two spaces, and then used to define a measure, i.e., a definition of integration, in the space of Hermitian matrices. The measure of a set in the latter space is defined to be proportional to the ordinary

Euclidean measure of the corresponding set in the real space of dimension J^2. In the latter space, the volume element of integration is given by

$$dV \equiv da_1 da_2 \ldots da_{J^2} ,$$

and in the new space it will be taken to be

$$d(A) \equiv \prod_{k=1}^{J} d(A_{k,k}) \prod_{1 \leq i<j \leq J} d[\mathrm{Re}(A_{i,j})]\, d[\mathrm{Im}(A_{i,j})] . \qquad \text{(A3-6)}$$

We therefore have

$$dV = 2^{J(J-1)/2} d(A) ,$$

and this relation establishes the proportionality constant between the two measures. In the analysis to follow, we will limit all integrals in the new space to the subspace of Hermitian matrices which are non-negative definite. This restriction will be indicated by the use of the notation $d_0(A)$ for the volume element of integration.

The two integration concepts are closely related, as shown by the following property. Let Z be an array of variables, of dimension $J \times M$, where $J \leq M$. Then, if $\mathcal{F}$ is any well-behaved function whose argument is a square matrix, the identity

$$\int \mathcal{F}(ZZ^H)\, d(Z) = \frac{\pi^{JM}}{\Gamma_J(M)} \int \mathcal{F}(S)\, |S|^{M-J}\, d_0(S) \qquad \text{(A3-7)}$$

holds, so long as the integrals themselves exist, where

$$\Gamma_J(K) \equiv \pi^{J(J-1)/2} \prod_{j=0}^{J-1} \Gamma(K-j) . \qquad \text{(A3-8)}$$

This quantity, which is a generalization of the Gamma function, will appear frequently in the following discussion, and we note that $\Gamma_1(K) = \Gamma(K)$.

The integration identity can be derived directly from geometric considerations, and a detailed exposition of the theorem (for the case of real variables) may be found in Chapter 2 of Reference 10 which contains further references to the literature. We give an inductive proof for the complex case later in this Appendix, using only elementary matrix methods. These are, in fact, the same methods of projection and partitioning which are utilized repeatedly in the main body of this study. Before proceeding with this proof, we first show some of the consequences of Equation (A3-7), beginning with a derivation of the Wishart pdf which is simpler than the conventional procedure.[15]

Let Z be a complex Gaussian array, of dimension $J \times (J+K)$, with mean value zero, and with covariance matrix

$$\mathrm{Cov}(Z) = I_J \otimes I_{J+K} \,,$$

where $K \geq 0$. Then, the expected value of an arbitrary function of the product

$$T \equiv ZZ^H$$

can be evaluated as the integral

$$E\mathcal{F}(T) = \frac{1}{\pi^{J(J+K)}} \int \mathcal{F}(ZZ^H)\, e^{-\mathrm{Tr}(ZZ^H)}\, d(Z) \,, \tag{A3-9}$$

taken over the pdf of Z. The latter is a special case of Equation (A1-68) of Appendix 1, with the mean value replaced by zero. Applying Equation (A3-7) to this integral, we obtain

$$E\mathcal{F}(T) = \frac{1}{\Gamma_J(J+K)} \int \mathcal{F}(S)\, |S|^K\, e^{-\mathrm{Tr}(S)}\, d_0(S) \,.$$

It follows immediately that the joint pdf of the elements of T is the complex Wishart density

$$f_W(T; J, K | I) \equiv \frac{1}{\Gamma_J(J+K)} |T|^K e^{-\mathrm{Tr}(T)} \,. \tag{A3-10}$$

This notation (which is not standard) is chosen to exhibit the complex Wishart pdf as a direct generalization of the complex chi-squared distribution. In the present case, the matrix dimension is J and the Wishart density has $J + K$ complex degrees of freedom. If Z has the more general covariance matrix

$$\mathrm{Cov}(Z) = \Sigma \otimes I_{J+K} ,$$

then Equation (A3-9) is replaced by

$$E\mathcal{F}(T) = \frac{1}{\pi^{J(J+K)} |\Sigma|^{J+K}} \int \mathcal{F}(ZZ^H)\, e^{-\mathrm{Tr}(\Sigma^{-1}ZZ^H)}\, d(Z) .$$

Applying Equation (A3-7) again, we obtain

$$E\mathcal{F}(T) = \int \mathcal{F}(S)\, f_W(S;J,K|\Sigma)\, d_0(S) ,$$

where

$$f_W(T;J,K|\Sigma) \equiv \frac{1}{\Gamma_J(J+K)} \frac{|T|^K}{|\Sigma|^{J+K}}\, e^{-\mathrm{Tr}(\Sigma^{-1}T)} , \tag{A3-11}$$

which is the general case of the complex Wishart density.

As another application of Equation (A3-7), we derive the Jacobian for the linear transformation of variables

$$S = G\tilde{S}G^H , \tag{A3-12}$$

where S is a matrix of complex variables of integration, and the volume element is defined by Equation (A3-6). The matrix G is, of course, non-singular. Any integral over S can be expressed as an integral over a $J \times J$ array Z of unconstrained complex variables, as follows:

$$\int \mathcal{F}(S)\, d_0(S) = \frac{\Gamma_J(J)}{\pi^{J^2}} \int \mathcal{F}(ZZ^H)\, d(Z) .$$

The validity of this representation is a special case of Equation (A3-7). Now let us introduce the change of variables

$$Z = G\tilde{Z} \quad , \qquad d(Z) = |GG^H|^J\, d(\tilde{Z}) , \tag{A3-13}$$

with Jacobian as shown. The latter is a special case of Equation (A1-66) of Appendix 1. Substituting, and using Equation (A3-7) again, we obtain

$$\int \mathcal{F}(S)\, d_0(S) = \frac{\Gamma_J(J)}{\pi^{J^2}} |GG^H|^J \int \mathcal{F}(G\tilde{Z}\tilde{Z}^H G^H)\, d(\tilde{Z})$$

$$= |GG^H|^J \int \mathcal{F}(G\tilde{S}G^H)\, d_0(\tilde{S}) .$$

It follows that the change of the volume element of integration associated with transformation (A3-12) is given by

$$d_0(S) = |GG^H|^J\, d_0(\tilde{S}) . \tag{A3-14}$$

The validity of Equation (A3-7) depends on the postulated condition $J \leq M$. If, however, Z is a $J \times M$ array with $J \geq M$, then Z^H satisfies the requirements of the theorem. We also have $d(Z^H) = d(Z)$, as a direct consequence of the definition (A1-63) of Appendix 1. We therefore obtain the identity

$$\int \mathcal{F}(Z^H Z)\, d(Z) = \frac{\pi^{JM}}{\Gamma_M(J)} \int \mathcal{F}(S)\, |S|^{J-M}\, d_0(S) . \tag{A3-15}$$

In this case, of course, S is of order M.

To prove the integration theorem [formula (A3-7)], we first verify its validity for the special case in which $J=1$. A general proof will then be established by induction. When $J=1$, we write z instead of Z, where z is a row vector of M elements. Putting $z_m = x_m + iy_m$, we have

$$zz^H = \sum_{m=1}^{M} |z_n|^2 = \sum_{m=1}^{M} (x_n^2 + y_n^2) \equiv r^2 . \tag{A3-16}$$

The volume element of integration is, of course,

$$d(z) = dx_1 \ldots dx_M dy_1 \ldots dy_M ,$$

and we now change to spherical coordinates in the real space of 2M dimensions. The radial coordinate is r, defined in Equation (A3-16), and we write Ω_{2M} for the solid angle in this space. We also write $d\Omega_{2M}$ for the differential of this solid angle. Then, we get

$$\int \mathcal{F}(zz^H)\, d(z) = \iint \mathcal{F}(r^2)\, r^{2M-1} dr\, d\Omega_{2M}$$

for the integral of an arbitrary function of zz^H. The integrand depends only on r, and we can therefore integrate over the solid angle, using the well-known formula

$$\Omega_{2M} = 2\frac{\pi^M}{(M-1)!} .$$

Changing variables again, we let $x = r^2$, and then we have

$$\int \mathcal{F}(zz^H)\, d(z) = \frac{\pi^M}{(M-1)!} \int_0^\infty \mathcal{F}(x)\, x^{M-1} dx . \tag{A3-17}$$

From definition (A3-8), we see that $(M-1)! = \Gamma_1(M)$, and we also note that x corresponds to S, which is a scalar in this case. Thus, Equation (A3-17) agrees with Equation (A3-7), including the restriction on the range of integration to non-negative values, for the special case under consideration.

To prove the general case, we assume the validity of the integration theorem for $J < M$, and show that it also holds when J is replaced by $J+1$. We begin by writing

$$Z = \begin{bmatrix} v \\ W \end{bmatrix} ,$$

where W is a complex array of dimension $J \times M$ and v is a row vector of M complex components, and we study the integral

$$\mathcal{I} \equiv \int \mathcal{F}(ZZ^H)\, d(Z) \tag{A3-18}$$

We exclude from this integral all points for which the Z matrix is not of full rank. It may be shown that the measure of the set of points so excluded is zero; hence, the integral itself is not affected. Similarly, all integrals over the space of non-negative definite Hermitian matrices may be replaced by the corresponding integrals over the subset of positive-definite Hermitian matrices, again with no effect on the results. The latter matrices form an open, dense subset of the non-negative definites, and this subset carries full measure, which is an equivalent statement of our assertion. Since the full-rank restriction on Z implies that ZZ^H is always positive definite, it is sufficient to prove the integration theorem under these two restrictions on the respective ranges of integration.

The volume element of integration in Equation (A3-18) is simply $d(Z) = d(v)d(W)$, and we also have

$$ZZ^H = \begin{bmatrix} vv^H & vW^H \\ Wv^H & WW^H \end{bmatrix} . \tag{A3-19}$$

The key to the proof is provided by the form of the determinant:

$$|ZZ^H| = |WW^H| \left[vv^H - vW^H (WW^H)^{-1} Wv^H \right] , \tag{A3-20}$$

which is evaluated by an application of Equation (A1-2) of Appendix 1. The second factor on the right may be written

$$v \left[I_M - W^H (WW^H)^{-1} W \right] v^H ,$$

which shows that only the component of v which is orthogonal to the row space of W enters the expression for this determinant. The fact that WW^H is non-singular follows directly from the non-singularity of ZZ^H itself.

Following the procedure first used in Section 2, we introduce the $J \times M$ array

$$\alpha \equiv (WW^H)^{-1/2} W , \tag{A3-21}$$

which serves as a basis array in the J-dimensional row space of W. The properties

$$\alpha\alpha^H = I_J$$

$$\alpha^H\alpha = W^H(WW^H)^{-1} W$$

$$W = (WW^H)^{1/2} \alpha$$

follow directly. Continuing as in Section 2, we let β be an arbitrary basis array, of dimension $(M-J)\times M$, in the orthogonal complement of the row space of W, so that

$$\beta\beta^H = I_{M-J}$$

$$\beta\alpha^H = 0$$

$$\alpha^H\alpha + \beta^H\beta = I_M .$$

Then, α and β together form a unitary matrix of order M:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \equiv U_M .$$

We apply this matrix to v and partition the result:

$$v U_M^H \equiv [v_1 \ v_2] . \tag{A3-22}$$

The new components are given by

$$v_1 = v\alpha^H$$

$$v_2 = v\beta^H . \tag{A3-23}$$

Note that v_1 and v_2 are row vectors, of dimension J and $M - J$, respectively.

With these conventions established, we have

$$vW^H = v\alpha^H(WW^H)^{1/2} = v_1(WW^H)^{1/2} ,$$

and the determinant of ZZ^H becomes

$$|ZZ^H| = |WW^H|\, v[I_M - W^H(WW^H)^{-1}W]v^H$$

$$= |WW^H|\, v(I_M - \alpha^H\alpha)v^H = |WW^H|\, v_2v_2^H . \tag{A3-24}$$

The argument of $\mathcal{F}$ can now be written

$$ZZ^H = \begin{bmatrix} y & v_1(WW^H)^{1/2} \\ (WW^H)^{1/2}v_1^H & WW^H \end{bmatrix} , \tag{A3-25}$$

where

$$y \equiv v_1v_1^H + v_2v_2^H . \tag{A3-26}$$

In the integral itself, the volume element involves $d(v) = d(v_1)d(v_2)$, since the Jacobian associated with transformation (A3-22) is unity. Our integral is now expressed in a form which depends on W only through the product WW^H, and we can therefore invoke Equation (A3-7) to transform the W integral.

This allows us to write Equation (A3-18) as

$$\mathcal{J} = \frac{\pi^{JM}}{\Gamma_J(M)} \int \mathcal{F}(ZZ^H)|S|^{M-J} d_0(S)d(v_1)d(v_2) , \tag{A3-27}$$

where it is understood that

$$ZZ^H = \begin{bmatrix} y & v_1 S^{1/2} \\ S^{1/2} v_1^H & S \end{bmatrix} .$$

The integrations over v_1 and v_2 are unrestricted in Equation (A3-27), but the integration over S is limited to positive-definite matrices. The determinant of ZZ^H is, of course, given by

$$|ZZ^H| = |S| \, v_2 v_2^H . \tag{A3-28}$$

We now introduce a change of variables by the linear transformation

$$v_1 \equiv u S^{-1/2} \quad , \qquad d(v_1) = |S|^{-1} d(u) , \tag{A3-29}$$

with Jacobian as shown. This Jacobian is a special case of Equation (A1-66). The matrix ZZ^H now assumes the form

$$ZZ^H = \begin{bmatrix} y & u \\ u^H & S \end{bmatrix} , \tag{A3-30}$$

and y is given by

$$y = v_2 v_2^H + u S^{-1} u^H .$$

Next, we define

$$x \equiv v_2 v_2^H , \tag{A3-31}$$

and note that our integral depends on v_2 only through x. Since v_2 is a row vector, of $M-J$ components, we can apply Equation (A3-7) to the integration over v_2, which is of the same kind as the special case first evaluated as Equation (A3-17). When this is carried out, together with the change of variable from v_1 to u, we obtain

$$\mathcal{J} = \frac{\pi^{JM+M-J}}{(M-J-1)!\,\Gamma_J(M)} \int \mathcal{F}(ZZ^H)(|S|x)^{M-J-1}\, dx\, d(u)\, d_0(S) \,. \qquad \text{(A3-32)}$$

The integration over u is unrestricted, as was the integration over v_1, but x is limited to positive values by our application of Equation (A3-7). The matrix ZZ^H is still given by Equation (A3-30), with the understanding that

$$y = x + uS^{-1}u^H \,, \qquad \text{(A3-33)}$$

and its determinant [according to Equations (A3-28) and (A3-31)] is simply

$$|ZZ^H| = |S|\,x \,.$$

It may be verified directly that

$$\pi^J (M-J-1)!\,\Gamma_J(M) = \Gamma_{J+1}(M) \,,$$

and we therefore have

$$\mathcal{J} = \frac{\pi^{(J+1)M}}{\Gamma_{J+1}(M)} \int \mathcal{F}(ZZ^H)|ZZ^H|^{M-J-1}\, dx\, d(u)\, d_0(S) \,. \qquad \text{(A3-34)}$$

We make a final change of variable, replacing x by y, which is defined in Equation (A3-33). The only change in Equation (A3-34) is the replacement of dx by dy, together with the restriction

$$y > uS^{-1}u^H$$

on the range of integration over y. But it is easily shown that this condition, together with $S > 0$ (positive definiteness), is necessary and sufficient to ensure the positivity of ZZ^H, as defined by Equation (A3-30). This claim can, in fact, be verified by an application of Equation (A1-9) to an arbitrary quadratic form in the matrix ZZ^H.

According to definition (A3-6), the volume element of integration in Equation (A3-34) can be expressed as

$$dy\, d(u) d_0(S) = d_0(T) ,$$

where T is a $(J+1) \times M$-dimensional array of integration variables. Thus, we obtain the final result

$$\mathcal{I} \equiv \int \mathcal{F}(ZZ^H)\, d(Z) = \frac{\pi^{(J+1)M}}{\Gamma_{J+1}(M)} \int \mathcal{F}(T) |T|^{M-J-1} d_0(T) , \tag{A3-35}$$

and this completes the proof.

Next we consider the multivariate generalization of the complex central F distribution. Let V and W be independent Gaussian arrays, both of which have mean values of zero. Their dimensions are implied by the covariance matrices

$$\mathrm{Cov}(V) = I_J \otimes I_M$$

$$\mathrm{Cov}(W) = I_J \otimes I_{J+K} .$$

We wish to study the random array

$$\mathcal{A}(J,M,K) \equiv V^H T^{-1} V , \tag{A3-36}$$

where

$$T \equiv WW^H . \tag{A3-37}$$

The notation is analogous to that used for the $\mathcal{B}$ array in Section 4, which is obviously given by

$$\mathcal{B}(J,M,K) = I_M + \mathcal{A}(J,M,K) .$$

As before, we assume that $K \geq 0$, so that T obeys the complex Wishart pdf, expressed by Equation (A3-10).

Again we consider the expected value of an arbitrary function of $\mathscr{A}$, which may be written

$$E\mathscr{F}[\mathscr{A}(J,M,K)] = \frac{1}{\pi^{JM}} \iint \mathscr{F}(V^H T^{-1} V)\, f_W(T;J,K|I)\, e^{-\mathrm{Tr}(VV^H)}\, d_0(T) d(V) . \qquad (A3\text{-}38)$$

The double integral signifies integration over the complex Wishart pdf of T and the complex Gaussian pdf of V, the latter having been explicitly introduced in Equation (A3-38). Holding the integration over T in abeyance, we make the change of variable

$$V \equiv T^{1/2} Z \quad , \quad d(V) = |T|^M d(Z) ,$$

with Jacobian as shown. This change of variable is exactly like the one given in Equation (A3-13), and the notation is meant to signify the positive-definite square root of T. Thus, substituting for the complex Wishart pdf, we obtain

$$E\mathscr{F}[\mathscr{A}(J,M,K)] = \frac{1}{\pi^{JM} \Gamma_J(J+K)} \iint \mathscr{F}(Z^H Z)\, |T|^{M+K}\, e^{-\mathrm{Tr}[(I_J + ZZ^H)T]}\, d_0(T) d(Z) .$$

We now reverse the order of integration, and also make the change of variable

$$T \equiv (I_J + ZZ^H)^{-1/2} S (I_J + ZZ^H)^{-1/2} .$$

The Jacobian of this transformation, according to Equation (A3-14), is

$$d_0(T) = |I_J + ZZ^H|^{-J} d_0(S) ,$$

and, therefore,

$$\int |T|^{M+K} e^{-\mathrm{Tr}[(I_J + ZZ^H)T]} d_0(T)$$

$$= |I_J + ZZ^H|^{-(J+M+K)} \int |S|^{M+K} e^{-\mathrm{Tr}(S)} d_0(S) \; .$$

The above integral is evaluated as the Wishart normalization factor [see Equation (A3-11)], and we obtain

$$\int |T|^{M+K} e^{-\mathrm{Tr}[(I_J + ZZ^H)T]} d_0(T) = \frac{\Gamma_J(J+M+K)}{|I_J + ZZ^H|^{J+M+K}} \; .$$

According to Equation (A1-3) of Appendix 1,

$$|I_J + ZZ^H| = |I_M + Z^H Z| \, , \tag{A3-39}$$

and, hence, we have

$$E\,\mathcal{F}[\mathcal{A}(J,M,K)] = \frac{\Gamma_J(J+M+K)}{\pi^{JM}\,\Gamma_J(J+K)} \int \mathcal{F}(Z^H Z) \frac{d(Z)}{|I_M + Z^H Z|^{J+M+K}} \; . \tag{A3-40}$$

At this point, we postulate that $J \geq M$. Without this assumption, $\mathcal{A}$ is always rank deficient and a discussion of its pdf, although possible, is more complicated. With this assumption, Equation (A3-15) may be applied to integral (A3-40), and we obtain

$$E\,\mathcal{F}[\mathcal{A}(J,M,K)] = \frac{\Gamma_J(J+M+K)}{\Gamma_J(J+K)\,\Gamma_M(J)} \int \mathcal{F}(A) \frac{|A|^{J-M}}{|I_M + A|^{J+M+K}} d_0(A) \, , \tag{A3-41}$$

where A is a matrix of integration variables. It may be verified directly from definition (A3-8) that

$$\frac{\Gamma_J(J+M+K)}{\Gamma_J(J+K)} = \frac{\Gamma_M(J+M+K)}{\Gamma_M(M+K)} ,$$

and we may therefore write Equation (A3-41) in the form

$$E\,\mathcal{F}[\mathcal{A}(J,M,K)] = \frac{\Gamma_M(J+M+K)}{\Gamma_M(M+K)\,\Gamma_M(J)} \int \mathcal{F}(A)\,\frac{|A|^{J-M}}{|I_M + A|^{J+M+K}}\, d_0(A) . \qquad \text{(A3-42)}$$

We introduce the definition

$$B_n(b,c) \equiv \frac{\Gamma_n(b)\,\Gamma_n(c)}{\Gamma_n(b+c)} , \qquad \text{(A3-43)}$$

where n, b, and c are all positive integers. This quantity is a generalization of the Beta function, and we note that

$$B_1(b,c) = \frac{\Gamma(b)\,\Gamma(c)}{\Gamma(b+c)} = B(b,c) , \qquad \text{(A3-44)}$$

which is analogous to the reduction of the generalized Gamma function when its subscript equals unity. The multivariate complex Beta pdf is now introduced with the definition

$$f_\phi(A;M,J,K) \equiv \frac{1}{B_M(J,K)}\,\frac{|A|^{J-M}}{|I_M + A|^{J+K}} . \qquad \text{(A3-45)}$$

The parameter M specifies the matrix dimension of the complex Beta variable in this distribution. When $M=1$, the pdf reduces to the scalar complex Beta pdf, already defined by Equation (A2-9) of Appendix 2:

$$f_\phi(A;1,J,K) = \frac{1}{B(J,K)}\,\frac{A^{J-1}}{(1+A)^{J+K}} = f_F(A;J,K) . \qquad \text{(A3-46)}$$

In terms of the multivariate complex Beta pdf, integral (A3-42) can be expressed in the form

$$E\,\mathcal{F}[\mathcal{A}(J,M,K)] = \int \mathcal{F}(A)\, f_\phi(A;M,J,K+M)\, d_0(A)\ .$$

A complex multivariate analog of the complex Beta random variable can be defined in terms of $\mathcal{A}$, still under the assumption that $J \geq M$. It is simply the inverse of the matrix $\mathcal{C}$:

$$\mathcal{R}(J,M,K) \equiv \mathcal{C}(J,M,K)^{-1} = (I_M + V^H T^{-1} V)^{-1}\ . \tag{A3-47}$$

If A is a positive semi-definite matrix of order M, then R, defined by

$$R \equiv (I_M + A)^{-1}\ , \tag{A3-48}$$

is clearly positive definite. In addition, the eigenvalues of R will lie in the range zero to unity, hence $I_M - R$ will be positive semi-definite.

We solve Equation (A3-48) for A:

$$A = R^{-1} - I_M\ , \tag{A3-49}$$

and consider the elements of A to be functions of the elements of R. Using the well-known formula for the differential of the inverse of a matrix, we get

$$dA = d(R^{-1}) = -R^{-1}\,dR\,R^{-1}\ , \tag{A3-50}$$

where dA and dR are matrices of differentials. We view Equation (A3-49) as a change of variables, and note that the relation expressed by Equation (A3-50) is of the same form as the linear transformation (A3-12), but applied now to the differential arrays dA and dR. Then, Equation (A3-14) provides the Jacobian for the change of variable, and we can write

$$d_0(A) = |R|^{-2M}\, d_0(R)\ . \tag{A3-51}$$

Finally, by expressing A in the form

$$A = R^{-1}(I_M - R) ,$$

we can easily evaluate the required determinants in formula (A3-45) and supply the Jacobian from Equation (A3-51). As a result, we can write the expected value of any well-behaved function of $\mathcal{R}$ in the form

$$E\,\mathcal{F}[\mathcal{R}(J,M,K)] = E\,\mathcal{F}\{[I_M + \mathcal{A}(J,M,K)]^{-1}\} = \int \mathcal{F}(R)\, f_B(R;M,K+M,J)\, d_0(R) , \quad \text{(A3-52)}$$

where

$$f_B(R;M,K,J) \equiv \frac{1}{B_M(K,J)} |R|^{K-M} |I_M - R|^{J-M} \quad \text{(A3-53)}$$

is the complex multivariate Beta probability density function. The similarity to the scalar pdf is apparent, and when $M=1$ it is complete:

$$f_B(R;1,K,J) = \frac{1}{B(K,J)} R^{K-1}(1-R)^{J-1} = f_\beta(R;K,J) . \quad \text{(A3-54)}$$

An identity is used in Section 5 which follows directly from the definition of the complex multivariate Beta pdf. Multiplying both sides of Equation (A3-53) by the n^{th} power of the determinant of R, we have

$$|R|^n f_B(R;M,K,J) = \frac{1}{B_M(K,J)} |R|^{K+n-M} |I_M - R|^{J-M}$$

$$= \frac{B_M(K+n,J)}{B_M(K,J)} f_B(R;M,K+n,J) , \quad \text{(A3-55)}$$

and by direct evaluation we obtain

$$\frac{B_M(K+n,J)}{B_M(K,J)} = \frac{\Gamma_M(K+n)\Gamma_M(K+J)}{\Gamma_M(K+n+J)\Gamma_M(K)} = \prod_{m=0}^{M-1}\prod_{j=0}^{J-1} \frac{K+j-m}{K+j+n-m}$$

$$= \prod_{j=0}^{J-1} \frac{(K+j)!\,(K-M+n+j)!}{(K-M+j)!\,(K+n+j)!} \,. \qquad \text{(A3-56)}$$

Combining these results, we obtain the desired identity:

$$|R|^n \; f_B(R;M,K,J) = \prod_{j=0}^{J-1} \frac{(K+j)!\,(K-M+n+j)!}{(K-M+j)!\,(K+n+j)!} \; f_B(R;M,K+n,J) \,. \qquad \text{(A3-57)}$$

In Section 4, the GLR test statistic was defined as

$$\ell(J,M,K) \equiv |I_M + V^H T^{-1} V| \,,$$

where V and T have the same meanings as defined here, assuming the absence of signal components in the original data array. No restriction on the relative magnitudes of J and M is imposed at this point. If $\mathcal{F}$ is now an arbitrary function of a scalar argument, we can write

$$E\mathcal{F}[\ell(J,M,K)] = \frac{1}{\pi^{JM}} \iint \mathcal{F}(|I_M + V^H T^{-1} V|)\, f_W(T;J,K|I)\, e^{-\mathrm{Tr}(VV^H)}\, d_0(T)\, d(V) \,,$$

which is a particular case of Equation (A3-38) above. By following the same analysis we used to derive Equation (A3-40), and recalling also Equation (A3-39), we obtain the two equivalent forms

$$E\mathcal{F}[\ell(J,M,K)] = \frac{\Gamma_J(J+M+K)}{\pi^{JM}\,\Gamma_J(J+K)} \int \mathcal{F}(|I_M + Z^H Z|)\, \frac{d(Z)}{|I_M + Z^H Z|^{J+M+K}}$$

$$= \frac{\Gamma_J(J+M+K)}{\pi^{JM}\,\Gamma_J(J+K)} \int \mathcal{F}(|I_J + Z Z^H|)\, \frac{d(Z)}{|I_J + Z Z^H|^{J+M+K}} \,. \qquad \text{(A3-58)}$$

If $J \geq M$, we can continue as before and apply Equation (A3-15), with the result

$$E\mathcal{F}[\ell(J,M,K)] = \int \mathcal{F}(|I_M + A|)\, f_\Phi(A;M,J,K+M)\, d_0(A) \,. \qquad \text{(A3-59)}$$

If, on the other hand, $J \leq M$, we continue with the second line of Equation (A3-58) and apply the original integration identity [Equation (A3-7)] to obtain

$$E\mathcal{F}[\ell(J,M,K)] = \frac{\Gamma_J(J+M+K)}{\Gamma_J(J+K)\,\Gamma_J(M)} \int \mathcal{F}(|I_J + S|)\, \frac{|S|^{M-J}}{|I_J + S|^{J+M+K}}\, d_0(S) \,.$$

Since

$$\frac{\Gamma_J(J+K)\,\Gamma_J(M)}{\Gamma_J(J+M+K)} = B_J(M,J+K) \,,$$

we obtain the analogous formula:

$$E\mathcal{F}[\ell(J,M,K)] = \int \mathcal{F}(|I_J + A|)\, f_\Phi(A;J,M,J+K)\, d_0(A) \,. \qquad \text{(A3-60)}$$

Equations (A3-59) and (A3-60) represent formal statements of the statistical character of the signal-free GLR test statistic, expressed in terms of the complex multivariate F distribution. Later in this Appendix, this formal representation will be developed to produce the explicit characterization of the test statistic as a product of scalar complex Beta random variables, in agreement with the results obtained in Section 4. This exercise will also illustrate some useful techniques for carrying out explicit integration in the space of Hermitian matrices.

The integral representation, Equation (A3-58), will now be used to prove an important identity concerning members of the family of GLR test statistics, again in the signal-free case. Suppose that V is partitioned as follows:

$$V = [V_1 \;\; V_2] \,,$$

where the dimension of V_1 is $J \times M_1$, V_2 is $J \times M_2$ in dimension, and $M_1 + M_2 = M$. Then, we can write

$$\ell(J,M,K) = |I_M + V^H T^{-1} V| = \frac{|T + VV^H|}{|T|} ,$$

and we also have

$$VV^H = V_1 V_1^H + V_2 V_2^H .$$

These expressions permit us to make the factorization

$$\ell(J,M,K) = \frac{|T + V_1 V_1^H + V_2 V_2^H|}{|T + V_1 V_1^H|} \frac{|T + V_1 V_1^H|}{|T|} .$$

Recalling the definition of T [Equation (A3-37)], we note that

$$T + V_1 V_1^H = [\,W \;\; V_1\,][\,W \;\; V_1\,]^H ,$$

which is another complex Wishart matrix, of the same dimension J, but with $J + K + M_1$ degrees of freedom. Thus, we can write

$$\frac{|T + V_1 V_1^H + V_2 V_2^H|}{|T + V_1 V_1^H|} = |I_{M_2} + V_2^H (T + V_1 V_1^H)^{-1} V_2| = \ell(J,M_2,K+M_1) , \qquad \text{(A3-61)}$$

and also

$$\frac{|T + V_1 V_1^H|}{|T|} = |I_{M_1} + V_1^H T^{-1} V_1| = \ell(J,M_1,K) . \qquad \text{(A3-62)}$$

The notation on the right sides of these equations has been introduced as a way of indicating the statistical character of the quantities involved.

We have shown that

$$\ell(J,M,K) = \ell(J,M_2,K+M_1)\,\ell(J,M_1,K) . \qquad \text{(A3-63)}$$

and we now wish to prove that the factors on the right are statistically independent. Equation (A3-63) is therefore analogous to the representation of the GLR statistic as a product of independent factors, as given by Equation (4-31) of Section 4. By choosing $M_1 = 1$, and then iterating this identity, we can obtain from Equation (A3-63) the representation

$$\ell(J,M,K) = \prod_{m=0}^{M-1} \ell(J,1,K+m) \ , \qquad \text{(A3-64)}$$

which is directly analogous to Equation (4-32). The factors in this product are independent, and from this representation we can again obtain the double-product form, Equation (4-36).

To prove the independence of the factors in Equation (A3-63), we let $\mathcal{F}$ be an arbitrary function of two scalar arguments and consider the expectation value

$$E\,\mathcal{F}(\ell_a,\ell_b) \ ,$$

where

$$\ell_a \equiv \ell(J,M_2,K+M_1)$$
$$\ell_b \equiv \ell(J,M_1,K) \ . \qquad \text{(A3-65)}$$

This notation is adopted for brevity, and the variables on the right sides of these definitions are given explicitly in Equations (A3-61) and (A3-62), respectively. Since V_1 and V_2 are independent complex Gaussian arrays, we can write

$$E\,\mathcal{F}(\ell_a,\ell_b) = \frac{1}{\pi^{JM}} \iiint \mathcal{F}(\ell_a,\ell_b)\, f_W(T;J,K|1)\, e^{-\mathrm{Tr}(V_1V_1^H + V_2V_2^H)}\, d_0(T)d(V_1)d(V_2) \ . \qquad \text{(A3-66)}$$

The proof is carried out by means of a sequence of linear transformations, applied to the variables of integration. The first transformation, together with its Jacobian, is given by

$$V_2 = (T+V_1V_1^H)^{1/2} Z_a \quad , \qquad d(V_2) = |T+V_1V_1^H|^{M_2}\, d(Z_a) \ .$$

In terms of Z_a, we have [recalling Equation (A3-61)]

$$\ell_a = |I_{M_2} + Z_a^H Z_a| = |I_J + Z_a Z_a^H| ,$$

while ℓ_b is unaffected. We carry out this transformation, and also substitute for the complex Wishart pdf in Equation (A3-66), with the result

$$E\,\mathcal{F}(\ell_a, \ell_b) = \frac{1}{\pi^{JM}\Gamma_J(J+K)} \iiint \mathcal{F}(\ell_a, \ell_b)\,|T|^K\,|T + V_1V_1^H|^{M_2}$$

$$\times\; e^{-\mathrm{Tr}(T + V_1V_1^H)(I_J + Z_aZ_a^H)}\, d_0(T)\,d(V_1)\,d(Z_a) \,. \qquad \text{(A3-67)}$$

Next, we carry out the simultaneous transformations

$$V_1 \equiv (I_J + Z_aZ_a^H)^{-1/2}\,\tilde{V}_1$$

$$T \equiv (I_J + Z_aZ_a^H)^{-1/2}\,\tilde{T}\,(I_J + Z_aZ_a^H)^{-1/2} .$$

The corresponding Jacobians are expressed by the equations

$$d(V_1) = |I_J + Z_aZ_a^H|^{-M_1}\, d(\tilde{V}_1)$$

$$d_0(T) = |I_J + Z_aZ_a^H|^{-J}\, d_0(\tilde{T}) ,$$

and we note that ℓ_b is unchanged in form:

$$\ell_b = |I_{M_1} + \tilde{V}_1^H\,\tilde{T}^{-1}\,\tilde{V}_1| .$$

We make the evaluations

$$\mathrm{Tr}(T + V_1V_1^H)(I_J + Z_aZ_a^H) = \mathrm{Tr}(I_J + Z_aZ_a^H)^{1/2}\,(T + V_1V_1^H)\,(I_J + Z_aZ_a^H)^{1/2}$$

$$= \mathrm{Tr}(\tilde{T} + \tilde{V}_1\tilde{V}_1^H) ,$$

and

$$|T|^K |T + V_1 V_1^H|^{M_2} = |I_J + Z_a Z_a^H|^{-K-M_2} |\tilde{T}|^K |\tilde{T} + \tilde{V}_1 \tilde{V}_1^H|^{M_2} ,$$

and substitute in the integrand of Equation (A3-67), with the result

$$E\,\mathcal{F}(\ell_a, \ell_b) = \frac{1}{\pi^{JM} \Gamma_J(J+K)} \iiint \mathcal{F}(\ell_a, \ell_b) |\tilde{T}|^K |\tilde{T} + \tilde{V}_1 \tilde{V}_1^H|^{M_2}$$

$$\times\ e^{-\mathrm{Tr}(\tilde{T} + \tilde{V}_1 \tilde{V}_1^H)} \frac{d_0(\tilde{T}) d(\tilde{V}_1) d(Z_a)}{|I_J + Z_a Z_a^H|^{J+M+K}} . \tag{A3-68}$$

The last step of the proof is similar to our previous analysis of the pdf of the $\mathcal{A}$ matrix. We let

$$\tilde{V}_1 = \tilde{T}^{1/2} Z_b \quad , \quad d(\tilde{V}_1) = |\tilde{T}|^{M_1} d(Z_b) ,$$

and note that now

$$\ell_b = |I_{M_1} + Z_b^H Z_b| = |I_J + Z_b Z_b^H|$$

and

$$|\tilde{T} + \tilde{V}_1 \tilde{V}_1^H| = |\tilde{T}|\, |I_J + Z_b Z_b^H| .$$

With these changes of variable, integral (A3-68) becomes

$$E\,\mathcal{F}(\ell_a, \ell_b) = \frac{1}{\pi^{JM} \Gamma_J(J+K)} \iiint \mathcal{F}(\ell_a, \ell_b) |T|^{K+M} e^{-\mathrm{Tr}[(I_J + Z_b Z_b^H) T]}$$

$$\times \frac{|I_J + Z_b Z_b^H|^{M_2}}{|I_J + Z_a Z_a^H|^{J+M+K}} d_0(T) d(Z_b) d(Z_a) . \tag{A3-69}$$

In order to simplify the notation, we have dropped the tilde from the matrix of integration variables. The integration over T is carried out as before:

$$\int |T|^{K+M} e^{-\mathrm{Tr}[(I_J + Z_b Z_b^H)T]} d_0(T) = \frac{\Gamma_J(J+M+K)}{|I_J + Z_b Z_b^H|^{J+M+K}} .$$

We are left with the double integral

$$E\,\mathcal{F}(\ell_a,\ell_b) = \frac{\Gamma_J(J+M+K)}{\pi^{JM}\,\Gamma_J(J+K)} \iint \mathcal{F}(\ell_a,\ell_b) \frac{d(Z_a)}{|I_J + Z_a Z_a^H|^{J+M+K}} \frac{d(Z_b)}{|I_J + Z_b Z_b^H|^{J+M_1+K}} .$$

By an obvious factoring of the expression which precedes the integral in this formula, we can write it in the form

$$E\,\mathcal{F}(\ell_a,\ell_b) = \iint \mathcal{F}(\ell_a,\ell_b) \frac{\Gamma_J(J+M+K)}{\pi^{JM_2}\,\Gamma_J(J+M_1+K)} \frac{d(Z_a)}{|I_J + Z_a Z_a^H|^{J+M+K}}$$
$$\times \frac{\Gamma_J(J+M_1+K)}{\pi^{JM_1}\,\Gamma_J(J+K)} \frac{d(Z_b)}{|I_J + Z_b Z_b^H|^{J+M_1+K}} . \qquad \text{(A3-70)}$$

Comparison with Equation (A3-58) shows that the proof is complete, and that ℓ_a and ℓ_b are indeed independent random variables.

In Section 4, under the assumption that no signal components are present in the data array, it was shown that the inverse of the GLR test statistic can be expressed as a product of independent random variables, each of which obeys a Beta distribution. This result will now be obtained independently, using the methods of this Appendix, starting with one of the formal integral representations derived above. We assume that $J \le M$, in which case Equation (A3-60) will be our starting point. A similar derivation, proceeding from Equation (A3-59), would apply in the case where $J \ge M$.

We begin by partitioning the A matrix, as follows:

$$A \equiv \begin{bmatrix} y & u \\ u^H & B \end{bmatrix} . \qquad \text{(A3-71)}$$

where y is a scalar, and B is a square matrix of order $J-1$. Since A is positive definite over the range of integration, the new variables are subject to the restrictions

$$B > 0$$

$$y > uB^{-1}u^H \ .$$

We have noted these conditions before, and we make the change of variable

$$y \equiv x + uB^{-1}u^H \tag{A3-72}$$

to facilitate the application of this constraint. It is only necessary to require that $x > 0$, and the integration over u is completely unconstrained. It is permissible, therefore, to put

$$d_0(A) = dx\,d(u)\,d_0(B)$$

in Equation (A3-60). We also compute

$$|A| = |B|\,x$$

and, dropping the subscripts on the identity matrices now, we have

$$|I + A| = |I + B|\left[1 + x + uB^{-1}u^H - u(I+B)^{-1}u^H\right] .$$

We define the matrix Q by means of the equation

$$Q^{-1} \equiv B^{-1} - (I+B)^{-1} = B^{-1}(I+B)^{-1} \ ,$$

from which it follows that

$$Q = (I+B)\,B \ .$$

We introduce the new variable v by means of the definition

$$u \equiv vQ^{1/2}x^{1/2} \ ,$$

and recall that u is a row vector of $J-1$ elements. It follows that

$$d(u) = x^{J-1}|Q|d(v) = x^{J-1}|B||I + B|d(v) ,$$

and also that

$$|I + A| = |I + B|\left[1 + x(1 + vv^H)\right] . \tag{A3-73}$$

Making the appropriate substitutions in Equation (A3-60), we obtain

$$E\,\mathcal{F}[\ell(J,M,K)] = \frac{1}{B_J(M,J+K)}\iiint \mathcal{F}(|I + A|)\,\frac{|B|^{M-J+1}}{|I + B|^{J+M+K-1}}$$

$$\times \frac{x^{M-1}}{\left[1 + x(1 + vv^H)\right]^{J+M+K}}\,d_0(B)dxd(v) . \tag{A3-74}$$

The integration over x in this multiple integral is confined to positive values and, in the argument of $\mathcal{F}$, it is understood that Equation (A3-73) is to be applied.

Next we replace x by a new variable ρ, by means of the definition

$$1 + x(1 + vv^H) \equiv \rho^{-1} .$$

Obviously, we will have $0 \le \rho \le 1$, and also

$$|dx| = (1 + vv^H)^{-1}|d\rho/\rho^2| .$$

We can therefore write

$$|I + A| = |I + B|\rho^{-1} ,$$

and make the evaluation

$$\frac{x^{M-1}}{\left[1 + x(1 + vv^H)\right]^{J+M+K}}\,dxd(v) = \frac{\rho^{J+K-1}(1-\rho)^{M-1}}{(1 + vv^H)^M}\,d\rho\, d(v) .$$

An application of Equation (A3•17), together with the normalization integral of the scalar complex F distribution [Equation (A2-9)], yields the evaluation

$$\int \frac{d(v)}{(1 + vv^H)^M} = \frac{\pi^{J-1}}{(J-2)!} \int_0^\infty \frac{\xi^{J-2}}{(1+\xi)^M} d\xi = \pi^{J-1} \frac{(M-J)!}{(M-1)!} ,$$

and, therefore, we have

$$E\,\mathcal{F}[\ell(J,M,K)] = \frac{\pi^{J-1}(M-J)!}{(M-1)!\,B_J(M,J+K)} \iint \mathcal{F}(|I + B|\rho^{-1}) \frac{|B|^{M-J+1}}{|I + B|^{J+M+K-1}}$$

$$\times \rho^{J+K-1}(1-\rho)^{M-1}\, d\rho\, d_0(B) . \qquad \text{(A3-75)}$$

Recalling the scalar complex Beta pdf [Equation (A2•12)], we can write

$$\rho^{J+K-1}(1-\rho)^{M-1} = \frac{(J+K-1)!(M-1)!}{(J+M+K-1)!} f_\beta(\rho;J+K,M) .$$

and it is easily verified that

$$\frac{(J+M+K-1)!\,B_J(M,J+K)}{\pi^{J-1}(M-J)!(J+K-1)!} = B_{J-1}(M,J+K-1) .$$

We therefore find that Equation (A3-75) can be written

$$E\,\mathcal{F}[\ell(J,M,K)] = \frac{1}{B_{J-1}(M,J+K-1)} \int_0^1 \left[\int \mathcal{F}(|I + B|\rho^{-1}) \frac{|B|^{M-J+1}}{|I + B|^{J+M+K-1}} d_0(B) \right]$$

$$\times f_\beta(\rho;J+K,M)\, d\rho ,$$

and, by iteration, it follows that

$$E\,\mathcal{F}\left[\ell(J,M,K)\right] = \int \cdots \int \mathcal{F}\left[(\rho_1 \cdots \rho_J)^{-1}\right] f(\rho_1 \cdots \rho_J)\, d\rho_1 \cdots d\rho_J \; , \tag{A3-76}$$

where

$$f(\rho_1 \cdots \rho_J) = \prod_{j=1}^{J} f_\beta(\rho_j; K+j, M) \; . \tag{A3-77}$$

In the final step in this iteration, the scalar complex F distribution appears and is easily transformed to an integral over the scalar complex Beta density, with the result as stated above. Equation (A3-76) is equivalent to Equation (4-33) of Section 4, and with this observation the proof is complete.

APPENDIX 4

AN ALTERNATIVE DERIVATION OF THE GLR TEST

In this Appendix we provide an alternate derivation of the GLR test, which is particularly appropriate for the signal model described by Equation (1-4). We return to Equation (2-25), as a starting point, and write it in the form

$$l = \frac{|Z Z^H|}{\operatorname*{Min}_b F(b)} ,$$

where as in Section 2,

$$F(b) = (Z - ebp)(Z - ebp)^H .$$

Recall the arrays e and f, introduced in that section, and also the unitary matrix

$$U_N = [e \quad f] .$$

We now introduce a decomposition of the data array Z, by means of the definition

$$U_N^H Z \equiv \begin{bmatrix} X_A \\ X_B \end{bmatrix} \tag{A4-1}$$

or, equivalently,

$$X_A \equiv e^H Z$$

$$X_B \equiv f^H Z . \tag{A4-2}$$

In terms of the components defined in Section 2, we note that

$$X_A U_L^H = [Z_A \ W_A]$$

$$X_B U_L^H = [Z_B \ W_B] . \tag{A4-3}$$

The new components are brought into the analysis by means of the definition

$$\tilde{F}(b) \equiv U_N^H F(b) U_N ,$$

and the observation that

$$\operatorname*{Min}_b |F(b)| = \operatorname*{Min}_b |\tilde{F}(b)| .$$

Substituting for U_N and $F(b)$, we obtain

$$\tilde{F}(b) = \begin{bmatrix} (X_A - bp)(X_A - bp)^H & (X_A - bp)X_B^H \\ X_B(X_A - bp)^H & X_B X_B^H \end{bmatrix} . \tag{A4-4}$$

The required determinant is evaluated using identity (A1-2):

$$|\tilde{F}(b)| \equiv |X_B X_B^H| \, |J(b)| ,$$

where

$$J(b) \equiv (X_A - bp)(X_A - bp)^H - (X_A - bp) X_B^H (X_B X_B^H)^{-1} X_B (X_A - bp)^H .$$

Since $X_B X_B^H$ satisfies a complex Wishart distribution of dimension $N-J$, with L complex degrees of freedom, its inverse exists with probability one as a result of our assumption in Equation (1-9). We restrict the present analysis to the case $J < N$.

In terms of the matrix

$$R \equiv I_L - X_B^H (X_B X_B^H)^{-1} X_B , \tag{A4-5}$$

we have

$$\begin{aligned} J(b) &= (X_A - bp) R (X_A - bp)^H \\ &= b p R p^H b^H - b p R X_A^H - X_A R p^H b^H + X_A R X_A^H \end{aligned} \tag{A4-6}$$

Since the dimension of X_B is $(N-J) \times L$, we can compute the trace of R as follows:

$$\mathrm{Tr}(R) = L - \mathrm{Tr}[X_B^H (X_B X_B^H)^{-1} X_B]$$

$$= L - \mathrm{Tr}[(X_B X_B^H)^{-1} (X_B X_B^H)]$$

$$= L + J - N .$$

Since R is obviously idempotent, its eigenvalues are either zero or unity, and the trace evaluation shows that $N-J$ of them must vanish. Thus, R is a projection matrix and singular, except in the special case $J = N$. However, the matrix pRp^H is positive definite (with probability one), as will now be shown.

From Equations (A4-3) and the definition of U_L, we have

$$X_B p^H = Z_B$$

and also

$$X_B X_B^H = Z_B Z_B^H + W_B W_B^H .$$

It follows that

$$pRp^H = I_M - Z_B^H (Z_B Z_B^H + W_B W_B^H)^{-1} Z_B , \tag{A4-7}$$

and the existence of the inverse in this formula has already been noted. But the right side of Equation (A4-7) is itself just the inverse of the matrix C_M, defined by Equation (3-14), which we know to be positive definite, and this completes the proof.

We can therefore define

$$\hat{b} \equiv X_A Rp^H (pRp^H)^{-1} = X_A Rp^H C_M , \tag{A4-8}$$

and complete the square in Equation (A4-6). The result is

$$J(b) = (b - \hat{b}) pRp^H (b - \hat{b})^H + X_A R X_A^H - \hat{b} pRp^H \hat{b}^H .$$

The use of identity (2-30) then yields

$$\underset{b}{\mathrm{Min}}\,|J(b)| = |J(\hat{b})| ,$$

provided only that

$$J(\hat{b}) = X_A [R - Rp^H(pRp^H)^{-1} pR] X_A^H$$

is a positive-definite matrix, as will be shown below. Since the numerator of the test statistic is the determinant of F(0), we easily obtain the desired result

$$\ell = \frac{|J(0)|}{|J(\hat{b})|} = \frac{|X_A R X_A^H|}{|X_A Q X_A^H|} , \tag{A4-9}$$

where

$$Q \equiv R - Rp^H(pRp^H)^{-1} pR . \tag{A4-10}$$

An efficient algorithm for carrying out this computation with actual data can be devised, using the techniques which are described in Appendix 6.

It interesting to evaluate the performance of the test, as expressed in the form just derived. We assume that the true signal parameter array is B and that the covariance matrix of the columns of the data array is Σ. Taking the expected values of both sides of Equations (A4-2), we obtain

$$E\, X_A = b p$$

$$E\, X_B = 0 .$$

The array b which appears in the first of these formulas, is given in terms of B by Equation (2-23). These component arrays have independent columns, but they are, of course, correlated with one another.

It is expedient to carry out the whitening operation at this point, rather than at a later stage, as was the case in the analysis of Section 3. First, however, we eliminate the correlation between X_A and X_B by writing the former array as the sum of its conditional expectation given X_B (i.e., the linear predictor) and a "remainder" term (the prediction error):

$$X_A = \Sigma_{AB} \Sigma_{BB}^{-1} X_B + \widetilde{X}_A .$$

The remainder term is Gaussian and independent of X_B, and it is characterized by the relations

$$E \widetilde{X}_A = b p$$

$$Cov(\widetilde{X}_A) = (\Sigma^{AA})^{-1} \otimes I_L .$$

The R matrix is a projection onto the subspace orthogonal to the span of the the columns of X_B, and it is obvious from its definition that $X_B R = 0$. Therefore, we can simply replace X_A by the remainder term in the numerator of the test statistic. Because of the form of Q, the same is true of the denominator; hence, we have

$$l = \frac{|\widetilde{X}_A R \widetilde{X}_A^H|}{|\widetilde{X}_A Q \widetilde{X}_A^H|} . \tag{A4-11}$$

Whitened arrays can now be introduced, as follows:

$$\widetilde{X}_{A0} \equiv (\Sigma^{AA})^{1/2} X_A$$

$$X_{B0} \equiv (\Sigma_{BB})^{-1/2} X_B . \tag{A4-12}$$

These Gaussian arrays are independent and are characterized by the equations

$$E \widetilde{X}_{A0} = (\Sigma^{AA})^{1/2} b p$$

$$E X_{B0} = 0$$

$$Cov(\widetilde{X}_{A0}) = I_J \otimes I_L$$

$$Cov(X_{B0}) = I_{N-J} \otimes I_L . \tag{A4-13}$$

In terms of these quantities, we have

$$\ell = \frac{|\tilde{X}_{A0} R \tilde{X}_{A0}^H|}{|\tilde{X}_{A0} Q \tilde{X}_{A0}^H|} , \tag{A4-14}$$

since the determinants of the whitening matrices will cancel out. Moreover, the R matrix is unchanged in form as a result of this transformation:

$$R = I_L - X_{B0}^H (X_{B0} X_{B0}^H)^{-1} X_{B0} . \tag{A4-15}$$

It still remains to be proved that the matrix whose determinant forms the denominator of the test statistic is positive definite.

At this point, we simplify the notation by dropping the tilde and the subscript 0. Then, the test statistic is again given by the right side of Equation (A4-9), but the X_A and X_B arrays now have the properties given by Equations (A4-13). Turning to the R matrix, we follow the pattern established in Section 3 by introducing the array

$$\eta \equiv (X_B X_B^H)^{-1/2} X_B , \tag{A4-16}$$

assuming that the positive-definite square root of the matrix $X_B X_B^H$ has been chosen. The basic properties

$$\eta \eta^H = I_{N-J}$$

$$\eta^H \eta = X_B^H (X_B X_B^H)^{-1} X_B$$

$$X_B = (X_B X_B^H)^{1/2} \eta$$

then follow as before. The η array forms a basis for the row space of X_B. A basis array θ is chosen in the orthogonal complement of this subspace which, together with η, forms a basis for $\mathbb{C}^L$ itself. We then have

$$\theta \theta^H = I_{L+J-N}$$

$$\theta \eta^H = 0$$

$$\eta^H \eta + \theta^H \theta = I_L .$$

A unitary matrix is formed from these basis arrays, as follows

$$\mathcal{U}_L \equiv \begin{bmatrix} \theta \\ \eta \end{bmatrix} ,$$

and it is used to perform a rotation and partitioning of the array X_A:

$$X_A \mathcal{U}_L^H \equiv [\Psi_1 \ \Psi_2] . \tag{A4-17}$$

The new components are, of course, also given by the equations

$$\Psi_1 = X_A \theta^H$$

$$\Psi_2 = X_A \eta^H .$$

The R matrix finds a simple expression in terms of these arrays, namely,

$$R = I_L - \eta^H \eta = \theta^H \theta , \tag{A4-18}$$

and we also have

$$Q = \theta^H P \theta ,$$

where

$$P \equiv I_{L+J-N} - \theta p^H (p\theta^H \theta p^H)^{-1} p\theta^H . \tag{A4-19}$$

The GLR test statistic can now be written

$$\ell = \frac{|\Psi_1 \Psi_1^H|}{|\Psi_1 P \Psi_1^H|} . \tag{A4-20}$$

Note that only the first of the two components of X_A, introduced in Equation (A4-17), has survived in this formula. Next, we define the array

$$\mu \equiv p\theta^H . \tag{A4-21}$$

since this combination appears in the matrix P, which describes another projection:

$$P = I_{L+J-N} - \mu^H (\mu\mu^H)^{-1} \mu .$$

The expected value of Ψ_1 can also be expressed in terms of μ:

$$E\,\Psi_1 = (\Sigma^{AA})^{1/2} b \mu .$$

To deal with the decomposition imposed by P, we define the basis array

$$\gamma \equiv (\mu\mu^H)^{-1/2} \mu , \qquad (A4\text{-}22)$$

in direct analogy to previous derivations Then, we have

$$\gamma\gamma^H = I_M$$

$$\gamma^H\gamma = \mu^H (\mu\mu^H)^{-1} \mu$$

$$\mu = (\mu\mu^H)^{1/2} \gamma .$$

The γ array forms a basis of the $(L+J-N)$-dimensional row space of μ, and the orthogonal complement of this space is given a basis array which we will call δ. Then,

$$\delta\delta^H = I_{L+J-N-M}$$

$$\delta\gamma^H = 0$$

$$\gamma^H\gamma + \delta^H\delta = I_{L+J-N} . \qquad (A4\text{-}23)$$

Continuing in the usual way, we form the unitary matrix

$$\mathcal{U}_{L+J-N} \equiv \begin{bmatrix} \gamma \\ \delta \end{bmatrix} . \qquad (A4\text{-}24)$$

and then decompose Ψ_1:

$$\Psi_1 \mathcal{U}^H_{L+J-N} \equiv [\psi_1 \; \psi_2] . \qquad (A4\text{-}25)$$

Individually, these components are given by the equations

$$\psi_1 = \Psi_1 \gamma^H$$

$$\psi_2 = \Psi_1 \delta^H . \qquad \text{(A4-26)}$$

Since the expected value of Ψ_1 can be written

$$E\Psi_1 = (\Sigma^{AA})^{1/2} b (\mu\mu^H)^{1/2} \gamma ,$$

we compute

$$E\psi_1 = (\Sigma^{AA})^{1/2} b (\mu\mu^H)^{1/2}$$

$$E\psi_2 = 0 .$$

Working back through the definitions, we find that

$$\mu\mu^H = pRp^H = C_M^{-1} ,$$

and, consequently,

$$E\psi_1 = (\Sigma^{AA})^{1/2} b C_M^{-1/2} = V_{0s} . \qquad \text{(A4-27)}$$

The "signal array" V_{0s} was defined in Equation (3-37) of Section 3.

From definition (A4-26) and the last of Equations (A4-23), we obtain

$$\Psi_1 \Psi_1^H = \psi_1 \psi_1^H + \psi_2 \psi_2^H$$

and

$$P = \delta^H \delta .$$

These results, in turn, lead to the simple form

$$\ell = \frac{|\psi_1 \psi_1^H + \psi_2 \psi_2^H|}{|\psi_2 \psi_2^H|} . \tag{A4-28}$$

Since ψ_2 is a zero-mean Gaussian array, with covariance equal to the identity and dimension $J \times (L + J - N - M)$, the matrix in the denominator obeys a Wishart distribution with sufficient degrees of freedom to assure its positivity, hence this property is finally established.

The dimension of the array ψ_1 is $J \times M$ and its covariance matrix is the identity. Since its mean is V_{0s}, it is statistically identical to V_0 introduced in Section 3. In addition, the matrix $\psi_2 \psi_2^H$ is statistically identical to T_0 of that section, hence we write

$$\psi_1 = V_0$$

$$\psi_2 \psi_2^H = T_0 .$$

and obtain, for the GLR test statistic:

$$\ell = \frac{|V_0 V_0^H + T_0|}{|T_0|} . \tag{A4-29}$$

From the determinant identity (A1-2), we see that this expression is identical to formula (3-41) for the GLR test, hence the two approaches are entirely equivalent.

APPENDIX 5
THE CONSTRAINT ON THE DIMENSIONAL PARAMETERS

The general decision problem discussed in the main body of this report is characterized, in part, by the four dimensional parameters N, L, J, and M. The original data array is $N \times L$ and the signal parameter array is $J \times M$ in dimension. We pointed out in Section 1 that these parameters are constrained by the condition $L \geq N + M$, if we are to have a meaningful GLR test. The condition was used at several points in the analysis, always to ensure that some matrix was positive definite, and its sufficiency has therefore been established. We claimed that the constraint is also necessary, and that property is proved here This fact is of importance only because it affects the applicability of the model itself.

As shown in Section 2, the GLR test statistic is

$$\ell = \frac{|ZZ^H|}{\underset{b}{\text{Min}}\, F(b)} ,$$

where

$$F(b) \equiv (eb - Z_p)(eb - Z_p)^H + S$$

$$S \equiv Z_q Z_q^H .$$

The notation is that of Sections 2 and 3. We now assume that $L < N + M$ and show that b can be chosen to make F(b) singular, in which case the GLR test statistic will not exist. The proof will be probabilistic, and it will actually be shown that an array b can be found with probablity one. We introduce the "whitened" arrays

$$Z_{q0} \equiv \Sigma^{-1/2} Z_q$$

$$Z_{p0} \equiv \Sigma^{-1/2} Z_p$$

$$e_0 \equiv \Sigma^{-1/2} e , \qquad \text{(A5-1)}$$

and consider the matrix

$$F_0(b) \equiv (e_0 b - Z_{p0})(e_0 b - Z_{p0})^H + S_0 . \tag{A5-2}$$

where

$$S_0 \equiv Z_{q0} Z_{q0}^H . \tag{A5-3}$$

Since

$$F_0(b) = \Sigma^{1/2} F(b) \Sigma^{1/2} ,$$

it will suffice to show that $F_0(b)$ can be made singular by an appropriate choice of b.

Since the $N \times N$ matrix S_0 is composed of $L - M$ dyads, formed from the columns of Z_{q0}, it will be rank deficient under our assumption. For a given data array Z, let v be a vector in the null space of S_0, so that

$$v^H S_0 v = 0 . \tag{A5-4}$$

We must now find an array b, for which

$$v^H F_0(b) v = 0 , \tag{A5-5}$$

in order to show that $F_0(b)$ is singular. Obviously, Equations (A5-4) and (A5-5) together imply that

$$v^H (e_0 b - Z_{p0}) = 0 ,$$

or,

$$v^H e_0 b = v^H Z_{p0} . \tag{A5-6}$$

We must show that these equations can be solved for the b array, with probability one.

We can express the $J \times M$ array b in the form

$$b = [b_1, \ldots, b_M] .$$

where each b_m is a J vector. We can also write

$$v^H Z_{p0} = [\xi_1, \ldots, \xi_M] ,$$

where the ξ_m are simply scalars. Finally, we note that $v^H e_0$ is a row vector of J elements. Equations (A5-6) can therefore be written

$$v^H e_0 b_m = \xi_m \quad , \quad 1 \le m \le M . \tag{A5-7}$$

If $v^H e_0$ has at least one non-vanishing component, then the b array can be chosen (in many ways) to make $F_0(b)$ singular. The procedure fails only if v is orthogonal to every column of e_0, and this must be true for every v in the null space of S_0. Equivalently, each column of e_0 must be orthogonal to the null space of S_0. There is nothing special about the columns of e_0, and we now propose to show that for any fixed unit vector in $\mathbb{C}^N$, say A, the probability that A is orthogonal to the null space of S_0 is zero, and with this our proof will be completed.

Let P_0 be a projection matrix which projects onto the column space of Z_{q0}. Then, $I_N - P_0$ projects onto the null space of S_0, and for A to be orthogonal to this null space we must have

$$A^H(I_N - P_0)A = 0 ,$$

or,

$$u \equiv 1 - A^H P_0 A = 0 .$$

The projector P_0 is constructed directly from Z_{q0}, as follows:

$$P_0 = Z_{q0}(Z_{q0}^H Z_{q0})^{-1} Z_{q0}^H . \tag{A5-8}$$

As a result of our whitening, the array Z_{q0} is Gaussian, with zero mean, and with independent elements. It is also circular, which in this case means that the real and imaginary parts of its elements are all independent. Then,

$$V \equiv Z_{q0}^H$$

is a Gaussian array, with identical properties; hence,

$$V V^H = Z_{q0}^H Z_{q0}$$

obeys a complex Wishart distribution, of dimension $L - M$, with N complex degrees of freedom. Therefore, the inverse indicated in Equation (A5-8) exists with probability one.

In terms of V, we have

$$u = 1 - A^H V^H (V V^H)^{-1} V A \ . \tag{A5-9}$$

The unit vector A defines a subspace of $\mathbb{C}^N$ and we introduce a basis array, say D, in its orthogonal complement. Then,

$$U_0 \equiv [\, A \quad D \,]$$

is a unitary matrix, and we write

$$V U_0 = [\, V_1 \ V_2 \,] \ ,$$

where

$$V_1 \equiv V A$$

$$V_2 \equiv V D \ .$$

We also have

$$V V^H = V_1 V_1^H + V_2 V_2^H \ .$$

The array V_2 is just like V, except that its dimension is $(L - M) \times (N - 1)$. Since, by our assumption,

$$N - 1 \geq L - M \ ,$$

the Wishart matrix $V_2 V_2^H$ is also positive definite with probability one. We can now express u in the form

$$u = 1 - V_1^H (V_1 V_1^H + V_2 V_2^H)^{-1} V_1$$

$$= \left[1 + V_1^H (V_2 V_2^H)^{-1} V_1\right]^{-1} , \qquad (A5\text{-}10)$$

where the Woodbury identity [Equation (A1-5)] has been utilized.

The form found above for the random variable u is exactly like the inverse of the test statistic in the absence of signals, for the special case $M=1$ discussed in Section 4. It was shown there that this random variable is subject to a Beta distribution; hence, u assumes the value zero (or any other discrete value) with probability zero, and this completes our proof.

APPENDIX 6

NUMERICAL COMPUTATION OF THE FALSE ALARM PROBABILITY

In Section 4 it was shown that the GLR test statistic can be expressed as a product of independent random variables, in the case when no signal components are present. The probability distribution function of this product provides the PFA of the test as a function of the threshold. The product representation derived in Section 4 is

$$1/\ell = \prod_{j=1}^{J} \prod_{m=1}^{M} x_\beta(K+j+m-1,1) \, . \tag{A6-1}$$

where $x_\beta(n,1)$ is subject to the Beta distribution:

$$f_\beta(x;n,1) = n\,x^{n-1} \, .$$

It is understood that the factors in Equation (A6-1) are all independent, and the notation signifies the statistical character of each factor.

We introduce the logarithm of the GLR test statistic:

$$\lambda \equiv \log \ell \, ,$$

and the generating function:

$$\Phi(z) \equiv E\ell^z = Ee^{z\lambda} \, , \tag{A6-2}$$

which will be evaluated later. $\Phi(iu)$ is the characteristic function of the random variable λ, and the pdf of λ is therefore

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iu\lambda}\, \Phi(iu)\, du$$

$$= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{-z\lambda}\, \Phi(z)\, dz \, . \tag{A6-3}$$

If $\lambda_0 = \log \ell_0$, the PFA of the test will be

$$\text{PFA} = \text{Prob}(\ell \geq \ell_0) = \int_{\lambda_0}^{\infty} f(\lambda)\, d\lambda \ . \tag{A6-4}$$

We substitute Equation (A6-3) into Equation (A6-4) and shift the contour of integration over z to the right of the imaginary axis by a small amount μ. This permits an interchange of the order of integration and the evaluation

$$\text{PFA} = \frac{1}{2\pi i} \int_{-i\infty+\mu}^{i\infty+\mu} e^{-\lambda_0 z}\, \Phi(z)\, \frac{dz}{z}$$

$$= \frac{1}{2\pi i} \int_{-i\infty+\mu}^{i\infty+\mu} \ell_0^{-z}\, \Phi(z)\, \frac{dz}{z} \ . \tag{A6-5}$$

To evaluate $\Phi(z)$, we first compute

$$\int_0^1 x^{-z} f_\beta(x;n,1)\, dx = \frac{n}{n-z} \ ,$$

and then, from definition (A6-2), we obtain

$$\Phi(z) = \prod_{j=1}^{J} \prod_{m=1}^{M} \frac{K+j+m-1}{K+j+m-1-z} \ . \tag{A6-6}$$

The poles of this function are all on the real axis between $x = K+1$ and $x = K+J+M-1$. The extreme poles are simple, but the others have varying multiplicities, and this makes an evaluation by means of the residue series quite awkward. We note that $\Phi(z)$ is analytic over the entire z-plane, with the exception of the poles, and, in particular, it is analytic in the strip

$$0 < x < K+1 \ ,$$

(where $z = x + iy$), hence μ may have any value in this range. Since

$$[\Phi(z)]^* = \Phi(z^*) ,$$

the integral in Equation (A6-5) over the portion of the contour in the lower half plane is the negative of the complex conjugate of the integral over the upper portion and, therefore,

$$\mathrm{PFA} = \frac{1}{\pi} \mathrm{Im} \int_{\mu}^{i\infty+\mu} \ell_0^{-z} \Phi(z) \frac{dz}{z} . \tag{A6-7}$$

The contour for this integral may be deformed so that it passes to infinity anywhere in the first quadrant, as long as the poles are avoided.

We now show that μ and the contour can be chosen in a way which makes the integral converge rapidly, while the integrand remains positive and monotonically decreasing. By following this contour, the integral can be efficiently evaluated by numerical integration. Our procedure follows closely the work of Helstrom, especially the technique used in Reference 22.

We observe that the function $x^{-1}a^x$ is convex, for real positive values of a and x:

$$\frac{d^2}{dx^2}(x^{-1}a^x) = \left[(\log a - \frac{1}{x})^2 + \frac{1}{x^2}\right](x^{-1}a^x) \geq 0 .$$

Putting $a = \ell/\ell_0$ and taking the expected value of both sides of this equation, we obtain

$$\frac{d^2}{dx^2} E[x^{-1}(\ell/\ell_0)^x] = \frac{d^2}{dx^2}[x^{-1}\ell_0^{-x}\Phi(x)] \geq 0 . \tag{A6-8}$$

For values of z on the real axis between zero and $K+1$, the integrand of Equation (A6-7) is real and positive, and Equation (A6-8) shows that it is also convex.

The integrand has poles at the ends of this interval, and it must therefore have a single minimum at some interior point. We choose this point for μ, and discuss later the procedure for finding it. We also define the function

$$\Psi(z) \equiv \log[z^{-1}\ell_0^{-z}\Phi(z)] , \tag{A6-9}$$

so that Equation (A6-7) may be written

$$\mathrm{PFA} = \frac{1}{\pi} \operatorname{Im} \int_{\mu}^{i\infty+\mu} \exp[\Psi(z)]\, dz \,. \tag{A6-10}$$

The derivative $d\Psi(x)/dx$ obviously vanishes at $x=\mu$, hence $d\Psi(z)/dz=0$ at $z=\mu$ since Ψ is an analytic function of z. Therefore, the real and imaginary parts of Ψ, being solutions of Laplace's equation, both exhibit saddle points at $z=\mu$. The imaginary part of Ψ is zero on the real axis; hence, another contour on which $\operatorname{Im}(\Psi)=0$ must cross the real axis at $x=\mu$, in a direction parallel to the imaginary axis. These contours, on which the imaginary part of Ψ is zero, are contours of steepest descent or ascent of the real part of Ψ which pass through its saddle point. We know that the real part increases away from $x=\mu$ on the real axis; therefore, the other contour, crossing the axis of reals at right angles, is the one along which the real part of Ψ descends most rapidly from its value at $z=\mu$.

By choosing the portion of this contour which lies in the upper half-plane for our integral, we are assured of rapid convergence. Since the integrand is real and monotonically decreasing on the contour, we are also assured of numerical stability when the integral is carried out numerically. For large values of $|z|$, $\Psi(z)$ is dominated by the term

$$\Psi(z) \underset{|z|\to\infty}{\longrightarrow} -z \log \ell_0 \,.$$

In consequence, the contour $\operatorname{Im}(\Psi)=0$ will eventually level off with zero slope. It will therefore pass to infinity in the first quadrant of the complex plane and there is no difficulty in deforming the path of the integral of Equation (A6-7) to follow it. In order to show how an algorithm may be constructed along these lines, the remainder of this Appendix is given over to a discussion of the following topics: (1) a procedure for finding the saddle point, (2) the behavior of the contour in its vicinity, (3) a procedure for locating points on the contour for numerical integration, and (4) a stopping rule, or truncation bound, for the integration.

We have shown that the integrand in Equation (A6-10) has a unique minimum on the real axis between the origin and the first pole at $x=K+1$. It follows that the first derivative of Ψ has a unique zero in this range, and it may be located by Newton's method using the iteration

$$x_{n+1} = x_n - \frac{\Psi'(x_n)}{\Psi''(x_n)} .$$

Substituting Equation (A6-6) into (A6-9), we obtain the explicit formula

$$\Psi(x) = -\lambda_0 x - \log x + \sum_{j=1}^{J} \sum_{m=1}^{M} \log\left(\frac{K+j+m-1}{K+j+m-1-x} \right) . \qquad \text{(A6-11)}$$

and the required derivatives are then given by

$$\Psi'(x) = -\lambda_0 - \frac{1}{x} + \sum_{j=1}^{J} \sum_{m=1}^{M} \frac{1}{K+j+m-1-x} , \qquad \text{(A6-12)}$$

and

$$\Psi''(x) = \frac{1}{x^2} + \sum_{j=1}^{J} \sum_{m=1}^{M} \frac{1}{(K+j+m-1-x)^2} . \qquad \text{(A6-13)}$$

The technique works well in the present case, provided a good initial value is used for x. One approach is to approximate the derivative [Equation (A6-12)] and equate it to zero, as follows:

$$-\lambda_0 - \frac{1}{x} + \frac{JM}{b-x} = 0 .$$

In this approximation,

$$b \equiv K + (J+M)/2$$

is the "average" value of $K + j + m - 1$. The appropriate solution of this quadratic is

$$x = \frac{b}{2} - \frac{JM+1}{2\lambda_0} + \sqrt{\left[\frac{b}{2} - \frac{JM+1}{2\lambda_0} \right]^2 + \frac{b}{\lambda_0}} , \qquad \text{(A6-14)}$$

and this value has been successfully used as a starting value for x in the Newton iteration. When λ_0 is zero, or when it is small compared with b, the limiting value

$$x = \frac{b}{JM+1}$$

should be used instead. If the PFA is to be computed for a series of values of λ_0, it is a good idea to save the final value of x obtained in each case, and use it as a starting point for the next value of λ_0.

As a function of x, $\Psi(x)$ and its derivatives are real, and the first derivative vanishes at the saddle point $x=\mu$. Since $\Psi(z)$ is an analytic function of z, its derivatives at the saddle point are the same as those of $\Psi(x)$, and the expansion

$$\mathrm{Im}\,\Psi(z) = \Psi''(\mu)\,\mathrm{Im}\,(z-\mu)^2/2 + \Psi'''(\mu)\,\mathrm{Im}\,(z-\mu)^3/6 + \ldots$$

is valid. From this expansion, we find the equation of the contour: $\mathrm{Im}\Psi(z)=0$, in the immediate vicinity of the point $z=\mu$:

$$y\,\{\Psi''(\mu)(x-\mu) + \frac{1}{6}\Psi'''(\mu)[3(x-\mu)^2 - y^2] + \ldots\} = 0 .$$

The solution $y=0$ falls on the real axis through the saddle point, and the other solution is described by

$$x = \mu + \frac{\Psi'''(\mu)}{6\Psi''(\mu)}\,y^2 + \ldots ,$$

which approximates the equation of a parabola.

Equation (A6-13) shows that the second derivative of Ψ is positive, but the third derivative (evaluated at $x=\mu$) may have either sign. For large values of λ_0, which correspond to small values of the PFA, the saddle point moves toward the pole at $x=K+1$, and the third derivative will be positive. Then, the contour curves to the right as it leaves the saddle point, and (in the examples studied) it has a simple shape, leveling off as x increases. For sufficiently small values of λ_0, the contour curves initially to the left and then swings around to the right, leveling off again as it passes to infinity in the first quadrant of the complex plane.

The second derivative of Ψ, evaluated at the saddle point, also controls the behavior of the real part of Ψ on the contour in the vicinity of $z=\mu$. The shape of this variation will also be parabolic, and its curvature can be used to establish an initial step size Δ for the numerical evaluation of our integral, using a formula such as

$$\Delta = \text{constant} \times [\Psi''(\mu)]^{-1/2} , \qquad \text{(A6-15)}$$

with a suitable value for the constant. When the second derivative of Ψ is small, the value

$$\Delta = \text{constant} \times (K+1)$$

may be used instead, again with a suitable value for the constant. In the latter case, we are attempting to gauge the scale of the variation of the integrand by the distance from the origin to the first pole. When the final algorithm is applied, the step size can be adjusted until the desired accuracy is attained.

With the saddle point located and a step size chosen, we can begin to find points on the desired contour. The first point is obviously the saddle point itself, and the starting value of a search for the second point is chosen at a distance Δ, in the positive Y direction. A search for the contour is carried out in a direction parallel to the real axis. In general, given two successive points z_{N-1} and z_N on the contour, we compute the angle Θ_N according to

$$\tan \theta_N = \frac{\mathrm{Im}\,(z_N - z_{N-1})}{\mathrm{Re}\,(z_N - z_{N-1})} . \qquad \text{(A6-16)}$$

This angle is the slope of the line joining these two points, and we project ahead a distance Δ along this line to obtain the starting value, say w_0, of a search for z_{N+1}:

$$w_0 \equiv z_N + \Delta e^{i\theta_N} . \qquad \text{(A6-17)}$$

Using Newton's method again, we drive the imaginary part of $\Psi(z)$ to zero along a line at right angles to the first line, in other words along the line

$$w = w_0 - i\alpha e^{i\theta_N} , \qquad \text{(A6-18)}$$

where α is a real variable. The iteration begins with $\alpha = 0$ and is terminated when the change in α is sufficiently small.

To carry out this iteration, we require the derivative of the imaginary part of $\Psi(z)$ along this new line, and to obtain it we use the fact that Ψ is analytic. Thus, we have

$$\frac{d}{d\alpha} \operatorname{Im} \Psi(w) = \operatorname{Im} \frac{d}{d\alpha} \Psi(w) = \operatorname{Im}\left[-ie^{i\theta_N} \Psi'(w)\right] .$$

We define the real and imaginary parts of this derivative as follows:

$$\Psi'(w) \equiv X(w) + iY(w) ,$$

and the iteration can then be written

$$\alpha_{n+1} = \alpha_n - \frac{\operatorname{Im} \Psi(w_n)}{-X(w_n)\cos\theta_N + Y(w_n)\sin\theta_N} . \qquad \text{(A6-19)}$$

In this formula w_n is given by the right side of Equation (A6-18), with α replaced by α_n. Finally, if we write $w_n = \xi_n + i\eta_n$, we obtain the pair of iteration equations:

$$\xi_{n+1} = \xi_n + (\alpha_{n+1} - \alpha_n)\sin\theta_N$$

$$\eta_{n+1} = \eta_n - (\alpha_{n+1} - \alpha_n)\cos\theta_N . \qquad \text{(A6-20)}$$

When the iteration is terminated, the final value of w becomes the next point on the contour: z_{N+1}.

If the contour is followed exactly, the integrand will remain real by definition. If the contour is followed only approximately, a valid numerical approximation to the integral can still be obtained but the imaginary part of the integrand must also be taken into account, as in Helstrom's procedure. It is feasible, however, to continue the iteration far enough to locate the contour with such precision that we can ignore the imaginary part of the integrand, and this method has been chosen for our algorithm. As a check, the correction terms due to the imaginary part of the integrand were carried along in some examples, and they were found to contribute negligibly to the result, being many orders of magnitude lower than the contributions of the real part. In these examples, the iteration was stopped when the change in the imaginary part of Ψ fell below 10^{-4} in magnitude. We also found that very few iterations were needed to locate the contour in this way. A further advantage of this approach lies in the fact that the real part of Ψ changes only slightly during the search, hence the accuracy of the resulting value is enhanced.

It remains to derive a truncation bound, assuming that the integral (evaluated by a simple rectangular or trapezoidal rule) is terminated at the point z' on the contour. Let R be the remainder after truncation. Instead of following the steepest descent contour, we express the remainder as an integral along a path parallel to the real axis, beginning at the point z':

$$R = \frac{1}{\pi} \operatorname{Im} \int_{z'}^{\infty} G(z)\, dz \, , \qquad \text{(A6-21)}$$

where G(z) is the original integrand:

$$G(z) \equiv \exp[\Psi(z)] = z^{-1} \ell_0^{-z} \Phi(z) \, . \qquad \text{(A6-22)}$$

Along the steepest descent contour this integrand is real, and only the differential dz is complex. Since that contour tends to level off for large z, the effect of the imaginary part, applied to dz, is to improve the convergence of the integral. On the remainder portion, the differential is real and the integrand becomes complex. We put $z = z' + \xi$ on the remainder contour, and write R in the form

$$R = \frac{1}{\pi} G(z') \int_0^{\infty} \operatorname{Im}\left[\frac{G(z'+\xi)}{G(z')} \right] d\xi \, .$$

Note that G(z') is real and that this is an exact expression for the remainder.

Substituting from Equation (A6-22), we can write

$$\operatorname{Im}\left[\frac{G(z'+\xi)}{G(z')} \right] = e^{-\lambda_0 \xi} H(\xi) \, ,$$

where

$$H(\xi) \equiv \text{Im}\left[\frac{z'}{z'+\xi}\,\frac{\Phi(z'+\xi)}{\Phi(z')}\right] . \tag{A6-23}$$

The remainder can now be bounded as follows:

$$R \le \frac{1}{\pi} G(z') \int_0^{\infty} e^{-\lambda_0 \xi} \, |H(\xi)| \, d\xi , \tag{A6-24}$$

and we also have

$$|H(\xi)| \le \prod_{j=1}^{J} \prod_{m=1}^{M} \left| \frac{K+j+m-1-z'}{K+j+m-1-z'-\xi} \right| .$$

If we define

$$x_{j,m} \equiv K+j+m-1 ,$$

and also put

$$z' \equiv x' + iy' ,$$

then we can write

$$|H(\xi)| \le \prod_{j=1}^{J} \prod_{m=1}^{M} \left| \frac{x'-x_{j,m}+iy'}{x'-x_{j,m}+\xi+iy'} \right| . \tag{A6-25}$$

In those factors for which

$$x'-x_{j,m} \ge 0 ,$$

we have

$$\left| \frac{x'-x_{j,m}+iy'}{x'-x_{j,m}+\xi+iy'} \right| \le 1 .$$

This situation occurs for those factors corresponding to poles to the left of the stopping point for truncation. On the other hand, when

$$x' - x_{j,m} < 0 \ ,$$

we obtain the bound

$$\left| \frac{x' - x_{j,m} + i\,y'}{x' - x_{j,m} + \xi + i\,y'} \right| \le \left| \frac{x' - x_{j,m} + i\,y'}{y'} \right| \le 1 + \frac{|x' - x_{j,m}|}{y'} \ . \qquad \text{(A6-26)}$$

In this way, we compute the bound

$$|H(\xi)| \le H_0 \ , \qquad \text{(A6-27)}$$

where H_0 is a product of factors like those of Equation (A6-26). This bound is now independent of ξ and, when it is substituted in Equation (A6-24), the final result

$$R \le \frac{1}{\pi} G(z') \frac{H_0}{\lambda_0} \qquad \text{(A6-28)}$$

is obtained. The bound is easily computed as the numerical integration progresses, and the latter is terminated when the bound falls below a preset value.

APPENDIX 7
COMPUTATIONAL ALGORITHMS

In Section 2, a Generalized Likelihood Ratio (GLR) test was derived in which detection is based on the comparison of a test statistic to a fixed threshold. The quantity to be evaluated is reproduced here:

$$\ell = \frac{|I_M + Z_p^H S^{-1} Z_p|}{|I_M + Z_p^H P Z_p|} , \tag{A7-1}$$

where

$$Z_p = Z\tau^H(\tau\tau^H)^{-1/2}$$

$$S = Z[I_L - \tau^H(\tau\tau^H)^{-1}\tau]Z^H \tag{A7-2}$$

and

$$P \equiv S^{-1} - S^{-1}\sigma(\sigma^H S^{-1}\sigma)^{-1}\sigma^H S^{-1} . \tag{A7-3}$$

The data array Z and the known signal arrays σ and τ were introduced in Section 1. Here, we present an algorithm for the computation of the right side of Equation (A7-1). This algorithm utilizes a standard technique of signal processing, namely the construction of a unitary matrix which, multiplying a known array, converts it into "triangular" form. More precisely, when the unitary matrix pre-multiplies the (generally rectangular) known array, the resulting array has all zeros below the main diagonal. When post-multiplying, the result has zeros above the main diagonal. Several techniques are available for constructing these unitary matrices.[7] They are iterative in nature, building the unitary matrix as a product of factors, each of which is, for example, a Householder reflection matrix or a Givens rotation. We take this construction for granted, without further discussion here.

We begin with the known array τ and assume that U_τ is any unitary matrix with the property that

$$\tau U_\tau = [\,\rho \quad 0\,] ,$$

where ρ is an $M \times M$ matrix. Since τ has rank M, ρ will be non-singular. The procedure described above will suffice for the construction of U_τ, but in this particular case it is unimportant that ρ be triangular in form. Notice that the construction of U_τ cannot be charged to the cost of computing the test statistic, since τ is known and U_τ can be developed once and for all. We multiply the data array by U_τ on the right, and then partition the result, as follows:

$$Z U_\tau = [Z_1 \ \ Z_2] \ , \tag{A7-4}$$

where Z_1 is $N \times M$ and Z_2 is $N \times (L-M)$ in dimension. From these definitions it follows that

$$\tau\tau^H = \tau U_\tau U_\tau^H \tau^H = [\rho \quad 0] \begin{bmatrix} \rho^H \\ 0 \end{bmatrix} = \rho\rho^H \ , \tag{A7-5}$$

and

$$\begin{aligned} Z_p &= Z U_\tau U_\tau^H \tau^H (\tau\tau^H)^{-1/2} \\ &= [Z_1 \ \ Z_2] \begin{bmatrix} \rho^H \\ 0 \end{bmatrix} (\rho\rho^H)^{-1/2} \\ &= Z_1 \rho^H (\rho\rho^H)^{-1/2} \ . \end{aligned} \tag{A7-6}$$

Since the matrix $\rho^H(\rho\rho^H)^{-1/2}$ is unitary, it is easily shown that the GLR test is the same as

$$\ell = \frac{|I_M + Z_1^H S^{-1} Z_1|}{|I_M + Z_1^H P Z_1|} \ . \tag{A7-7}$$

From Equation (A7-4), we obtain

$$Z Z^H = Z_1 Z_1^H + Z_2 Z_2^H \ .$$

and from Equation (A7-6) we have

$$Z_p Z_p^H = Z_1 Z_1^H .$$

These facts give us the result

$$S = ZZ^H - Z_p Z_p^H = Z_2 Z_2^H . \qquad (A7\text{-}8)$$

The component arrays Z_1 and Z_2 are directly analogous to Z_p and Z_q, and, in the special case described by Equation (1-3), the former are identical to the latter.

Having found Z_2, we now generate a unitary matrix U_2 which converts it to the form

$$Z_2 U_2 = [\, L_2 \;\; 0 \,] , \qquad (A7\text{-}9)$$

where L_2 itself is lower triangular. Then,

$$S = L_2 L_2^H ,$$

in analogy to the derivation of Equation (A7-5). Since S is non-singular, the same is true of L_2, and the numerator of Equation (A7-7) can therefore be written in the form

$$| I_M + (L_2^{-1} Z_1)^H (L_2^{-1} Z_1) | . \qquad (A7\text{-}10)$$

Using the definition (A7-3), we have

$$P = (L_2^{-1})^H (I_N - P_N) L_2^{-1} ,$$

where

$$P_N = (L_2^{-1}\sigma) [(L_2^{-1}\sigma)^H (L_2^{-1}\sigma)]^{-1} (L_2^{-1}\sigma)^H .$$

These results give us the expression

$$\ell = \frac{|I_M + (L_2^{-1}Z_1)^H (L_2^{-1}Z_1)|}{|I_M + (L_2^{-1}Z_1)^H (I_N - P_N)(L_2^{-1}Z_1)|} \tag{A7-11}$$

for the GLR test statistic.

Next, we introduce the arrays V and μ as solutions of the sets of equations

$$L_2 V = Z_1$$
$$L_2 \mu = \sigma \; . \tag{A7-12}$$

These equations are easily solved because of the triangular form of L_2; they are just like the "back solutions" which arise in conventional adaptive nulling algorithms. In terms of the new quantities, we have

$$P_N = \mu(\mu^H \mu)^{-1} \mu^H \tag{A7-13}$$

and

$$\ell = \frac{|I_M + V^H V|}{|I_M + V^H (I_N - P_N) V|} \; . \tag{A7-14}$$

Now we find a unitary matrix U_μ which converts μ to the form

$$U_\mu \mu = \begin{bmatrix} \nu \\ 0 \end{bmatrix} ,$$

where ν is an upper triangular matrix. Since μ, like σ, is $N \times J$ in dimension and of rank J, the new array ν will be $J \times J$ and non-singular. A simple calculation now shows that

$$U_\mu P_N U_\mu^H = \begin{bmatrix} I_J & 0 \\ 0 & 0 \end{bmatrix} .$$

Hence we find

$$U_\mu (I_N - P_N) U_\mu^H = \begin{bmatrix} 0 & 0 \\ 0 & I_{N-J} \end{bmatrix} . \qquad \text{(A7-15)}$$

This treatment of projection matrices, such as P_N, has been used[32] as a means of deriving an architecture for their implementation in hardware. Note that the right side of Equation (A7-15) is simply zero in the special case $J = N$.

In the algorithm itself we find U_μ and apply it to V, calling the result W:

$$U_\mu V \equiv W . \qquad \text{(A7-16)}$$

The matrix μ is discarded when the development of U_μ is complete. Obviously,

$$V^H V = W^H W ,$$

and also

$$V^H (I_N - P_N) V = W^H \begin{bmatrix} 0 & 0 \\ 0 & I_{N-J} \end{bmatrix} W .$$

The array W is then partitioned:

$$W \equiv \begin{bmatrix} W_A \\ W_B \end{bmatrix} , \qquad \text{(A7-17)}$$

where W_A is $J \times M$ and W_B is $(N-J) \times M$ in dimension. Arrays W, V, and Z_1 all have the dimension of Z_p.

We substitute now, and obtain the form

$$\ell = \frac{|I_M + W^H W|}{|I_M + W_B^H W_B|} , \qquad \text{(A7-18)}$$

for the test. But we can write

$$I_M + W^H W = [W^H \; I_M] \begin{bmatrix} W \\ I_M \end{bmatrix}$$

and then find a unitary matrix U_n, which has the property

$$U_n \begin{bmatrix} W \\ I_M \end{bmatrix} = \begin{bmatrix} Y_n \\ 0 \end{bmatrix} .$$

where Y_n is upper triangular. Similarly, we choose U_d to make

$$U_d \begin{bmatrix} W_B \\ I_M \end{bmatrix} = \begin{bmatrix} Y_d \\ 0 \end{bmatrix} .$$

where Y_d is also upper triangular. Arrays Y_n and Y_d are both of dimension $M \times M$. With these transformations, we obtain

$$\ell = \frac{|Y_n^H Y_n|}{|Y_d^H Y_d|} . \tag{A7-19}$$

The determinants are now easy to evaluate and, for simplicity, we assume that U_n and U_d have been chosen so that the diagonal elements of Y_n and Y_d are real (this is easily accomplished). If the diagonal elements of Y_n are $(a_1, \ldots, a_M)$ and those of Y_d are $(b_1, \ldots, b_M)$, then

$$\ell = \prod_{m=1}^{M} \left(\frac{a_m}{b_m} \right)^2 , \tag{A7-20}$$

and the test can actually be carried out in the form

$$\sum_{m=1}^{M} \log\left(\frac{a_m}{b_m} \right) \geq \frac{1}{2} \log \ell_0 .$$

Note that all operations except the last involve linear operations on the data and signal arrays.

The same technique can be applied to the alternative form of the GLR test statistic, expressed by Equation (2-57). The components Z_1 and Z_2 are formed, as described above, and the matrix S is then evaluated using Equation (A7-8). Equation (2-57) is written in the form

$$\ell = \frac{|\sigma^H (Z_2 Z_2^H)^{-1} \sigma|}{|\sigma^H (Z Z^H)^{-1} \sigma|} , \tag{A7-21}$$

and the unitary matrix U_2 [defined by Equation (A7-9)] is found as before. Another unitary matrix, say U_1, is generated which will convert Z itself to lower triangular form, according to

$$Z U_1 = [\, L_1 \quad 0 \,] .$$

Then, we have

$$\ell = \frac{|(L_2^{-1}\sigma)^H (L_2^{-1}\sigma)|}{|(L_1^{-1}\sigma)^H (L_1^{-1}\sigma)|} , \tag{A7-22}$$

and the next step is the introduction of new arrays μ_1 and μ_2 as the solutions of the equations

$$L_1 \mu_1 = \sigma$$

$$L_2 \mu_2 = \sigma .$$

These arrays are of dimension $N \times J$, and μ_2 is identical to μ, defined in Equation (A7-12).

The test statistic takes the simple form

$$\ell = \frac{|\mu_2^H \mu_2|}{|\mu_1^H \mu_1|} \tag{A7-23}$$

in terms of these arrays. Finally, we form two $J \times J$ unitary matrices, which will again be called U_n and U_d, and which convert the μ arrays into upper triangular form by premultiplication:

$$U_n \mu_2 = \begin{bmatrix} Y_n \\ 0 \end{bmatrix}$$

$$U_d \mu_1 = \begin{bmatrix} Y_d \\ 0 \end{bmatrix} .$$

With these transformations, the test statistic assumes the same form as Equation (A7-19), and the remainder of the analysis is unchanged.

REFERENCES

1. E.L. Lehmann, *Testing Statistical Hypotheses* (Wiley, New York, 1959).

2. S.S. Wilks, Biometrika **24**, 471 (1932).

3. E.J. Kelly, IEEE Trans. Aerosp. Electron. Syst. **AES-22**, 115 (1986), DTIC AD-A174799.

4. E.J. Kelly, "Adaptive Detection in Non-Stationary Interference, Part I and Part II," Technical Report 724, Lincoln Laboratory, MIT (25 June 1985), DTIC AD-A158810.

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

6. J.Y. Chen and I.S. Reed, IEEE Trans. Aerosp. Electron. Syst. **AES-23**, 46 (1987).

7. G.H. Golub and C.F. Van Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, Maryland, 1983).

8. T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd edition (Wiley, New York, 1985).

9. K.V. Mardia, J.T. Kent, and J.M. Bibby, *Multivariate Analysis* (Academic Press, New York, 1979).

10. R.J. Muirhead, *Aspects of Multivariate Statistical Theory* (Wiley, New York, 1982).

11. C.G. Khatri, Ann. Inst. Statist. Math. **18**, 75 (1966).

12. R.F. Potthoff and S.N. Roy, Biometrika **51**, 315 (1964).

13. C.G. Khatri and C.R. Rao, "Test for a Specified Signal when the Noise Covariance Matrix is Unknown," Technical Report 85-47, Center for Multivariate Analysis, University of Pittsburgh (November 1985).

14. C.G. Khatri and C.R. Rao, IEEE Trans. Acoust., Speech, Signal Process. **ASSP-35**, 671 (1987).

15. N.R. Goodman, Ann. Math. Stat. **34**, 152 (1963).

16. G. Strang, *Linear Algebra and Its Applications* (Harcourt Brace Jovanovich, San Diego, California, 1988).

17. D.R. Brillinger, IEEE Trans. Acoust., Speech, Signal Process. **ASSP-33**, 1076 (1985).

18. J. Capon, Proc. IEEE **57**, 1408 (1969), DDC AD-696880.

19. D.N. Lawley, Biometrika **30**, 180 (1938).

20. C.G. Khatri, Ann. Math. Stat. **36**, 98 (1965).

21. M. Schatzoff, Biometrika **53**, 347 (1966).

22. C.W. Helstrom, IEEE Trans. Aerosp. Electron. Syst. **AES-19**, 428 (1983).

23. S.O. Rice, Bell Syst. Tech. J. **52**, 707 (1973).

24. J.I. Marcum, IRE Trans. Inf. Theory **IT-6**, 59 (1960).

25. E.T. Copson, *An Introduction to the Theory of Functions of a Complex Variable* (Oxford University Press, London, 1935).

26. M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards Applied Mathematics Series 55 (US Government Printing Office, Washington, DC, 1964).

27. I.S. Reed, J.D. Mallett, and L.E. Brennan, IEEE Trans. Aerosp. Electron. Syst. **AES-10**, 853 (1974).

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

29. D.A. Shnidman, IEEE Trans. Inf. Theory **IT-22**, 746 (1976).

30. R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, England, 1985).

31. H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1951).

32. C.P. Rialan and L.L. Scharf, "Vector and Cellular Pipelines for Implementing Projection Operators," Conference Record, Twentieth Asilomar Conference on Signals, Systems and Components, November 1986.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report 848	5. MONITORING ORGANIZATION REPORT NUMBER(S) ESD-TR-88-333
6a. NAME OF PERFORMING ORGANIZATION Lincoln Laboratory, MIT 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Electronic Systems Division
6c. ADDRESS (City, State, and Zip Code) P.O. Box 73 Lexington, MA 02173-0073	7b. ADDRESS (City, State, and Zip Code) Hanscom AFB, MA 01731
8a. NAME OF FUNDING/SPONSORING ORGANIZATION HQ AF Systems Command 8b. OFFICE SYMBOL (If applicable) AFSC/XTKT	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F19628-85-C-0002
8c. ADDRESS (City, State, and Zip Code) Andrews AFB Washington, DC 20334-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: 62301E, 63424F PROJECT NO.: 280 TASK NO.: WORK UNIT ACCESSION NO.:

11. TITLE (Include Security Classification)
Adaptive Detection and Parameter Estimation for Multidimensional Signal Models

12. PERSONAL AUTHOR(S)
Edward J. Kelly and Keith M. Forsythe

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Report	FROM ____ TO ____	1989, April, 19	252

16. SUPPLEMENTARY NOTATION
None

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
adaptive detection; maximum-likelihood estimation; adaptive nulling; statistical hypothesis testing

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The problem of target detection and signal parameter estimation in a background of unknown interference is studied, using a multidimensional generalization of the signal models usually employed for radar, sonar, and similar applications. The required techniques of multivariate statistical analysis are developed and extensively used throughout the study, and the necessary mathematical background is provided in Appendices. Target detection performance is shown to be governed by a form of the Wilks' Lambda statistic, and a new method for its numerical evaluation is given which applies to the probability of false alarm of the detector. Signal parameter estimation is shown to be directly related to known techniques of adaptive nulling, and several new results relevant to adaptive nulling performance are obtained.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Lt. Col. Hugh L. Southall, USAF	22b. TELEPHONE (Include Area Code) (617) 981-2330 22c. OFFICE SYMBOL ESD/TML