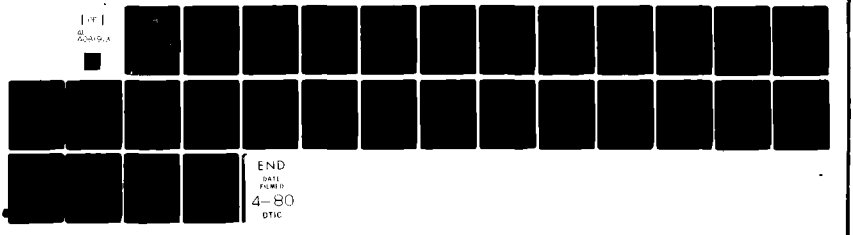
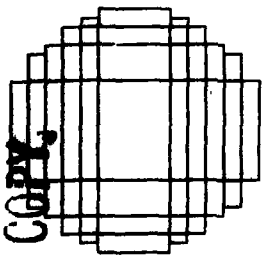
AD-A081 913 TECHNOLOGY SERVICE CORP SANTA MONICA CALIF F/G 9/3
ADAPTIVE ARRAYS IN AN ADVANCED INTERFERENCE ENVIRONMENT.(U)
OCT 79 L E BRENNAN, I S REED N00019-78-C-0497
UNCLASSIFIED TSC-PD-A211-4 NL



END
DATE FILMED
4-80
DTIC

LEVEL II

12



Technology Service Corporation

Technology Service Corporation

2811 WILSHIRE BOULEVARD • SANTA MONICA, CALIFORNIA 90403 • PH. (213) 829-7411

12

6 ADAPTIVE ARRAYS IN AN ADVANCED INTERFERENCE ENVIRONMENT

14 TSC-PD-A211-4

10 L. E. Brennan
I. S. Reed

11 24 October 1979

12 Sec

9 A Final Progress Report,
Submitted to
The Naval Air Systems Command
on Contract N00019-78-C-0497

15

404453

FE

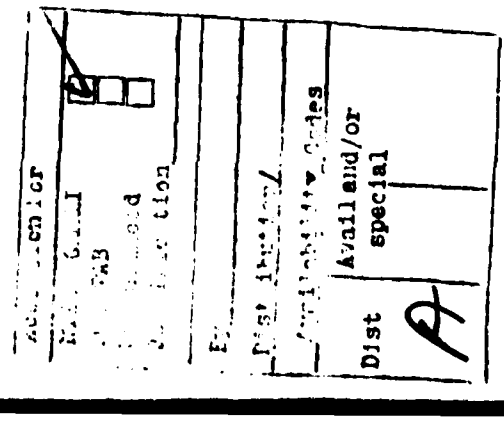
1. INTRODUCTION

When an adaptive array is implemented digitally, the sample covariance matrix algorithm provides a direct method of computing the adaptive weights and rapid convergence independent of the eigenvalues of the covariance matrix. Previous analyses of this algorithm[1] have assumed that the weights are computed using one set of array element outputs and these weights are applied to later array outputs. This report considers the case where the adaptive weights are tested against the same set of data used in the weight computation.

For many applications, the multiple channel sidelobe canceller is the preferred adaptive configuration. It can be shown that the sidelobe canceller is a special case of the more general adaptive array.[2] In the next section it is shown that the general adaptive array problem can be transformed to an equivalent sidelobe canceller problem, a form which is more convenient for some analyses. It is also shown that the array performance is independent of this transformation, and that the effective weights, output S/N, etc., can be computed in any convenient coordinate system provided the transformation of coordinates is non-singular.

2. COORDINATE TRANSFORMATIONS

Let X denote the column vector of array element outputs, $X_T = (x_1, x_2, \ldots, x_N)$, and S_x denote the corresponding signal vector. The noise covariance matrix of the array outputs is



$$M_x = E\ X\ X^{\dagger}, \tag{1}$$

where M_x is a NxN Hermitian matrix for an N element array, E denotes the expectation, † the complex transpose, and all noise components (but no signal) are included in M_x.

The weights which maximize the S/N ratio are

$$W_x = M_x^{-1}\ S_x \tag{2}$$

and the corresponding array output is

$$Z = W_x^{\dagger}\ X = S_x^{\dagger}\ M_x^{-1}\ X \tag{3}$$

With optimum weights, the output S/N ratio is

$$r = \frac{(W_x^{\dagger}\ S_x)^2}{W_x^{\dagger}\ M_x\ W_x} = S_x^{\dagger}\ M_x^{-1}\ S_x \tag{4}$$

Let T denote any non-singular transformation, and

$$Y = T\ X \tag{5}$$

In the new coordinate system,

$$M_y = E\ Y\ Y^{\dagger} = T\ M_x\ T^{\dagger} \tag{6}$$

$$S_y = T\ S_x \tag{7}$$

$$W_y = M_y^{-1}\ S_y \tag{8}$$

Combining 5 through 8,

$$W_y = (T^\dagger)^{-1} W_x \tag{9}$$

and

$$Z = W_y^\dagger Y = W_x^\dagger X \tag{10}$$

i.e., the output of the array is unchanged by the transformation.

The sample covariance matrix in X coordinates is

$$\hat{M}_x = \frac{1}{K} \sum_{k=1}^{K} X_k X_k^\dagger \tag{11}$$

where X_k denotes the k^{th} independent sample of array element outputs and K is the number of samples in the estimator of M_x. The weights based on a sample covariance matrix are

$$\hat{W}_x = \hat{M}_x^{-1} S_x \tag{12}$$

Replacing M_x and M_y with the corresponding esimators, $\hat{M}_x$ and $\hat{M}_y$, and following the analysis of the preceding paragraph, it can easily be shown the array output with weights based on a sample covariance matrix is also independent of the transformation T. Hence, the analysis of adaptive array performance, including the sample covariance matrix algorithm, can be performed in any convenient coordinate system provided the required transformation is non-singular.

3. PROBABILITY DISTRIBUTION OF S/N

For any arbitrary adaptive array, the input vector X can be transformed to a new set of coordinates in which the signal is present in only one component and the noise covariance matrix is diagonal. First, let $V = M_x^{-1/2} X$ to diagonalize the noise covariance matrix. Then,

$$M_v = E\, M_x^{-1/2} \overline{X\, X^{\dagger}}\, M_x^{-1/2} = I$$
$$S_v = M_x^{-1/2} S_x \tag{13}$$

Next, rotate the coordinates by a unitary transformation U so that the S vector is non-zero only in the first component, and normalize its amplitude to unity

$$Y = (S_x^{\dagger} M_x^{-1} S_x)^{-1/2} U V$$
$$= r_o^{-1/2} U M_x^{-1/2} X \tag{14}$$

where r_o is the S/N ratio with optimum weights, given by (4). Then,

$$M_y = r_o^{-1} U M_x^{-1/2} M_x M_x^{-1/2} U^{\dagger} = \frac{1}{r_o} I \tag{15}$$

and

$$S_y = r_o^{-1/2} U M_x^{-1/2} S_x = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{16}$$

Note that the output S/N ratio in the new coordinate system is

$$S_y^{\dagger} M_y^{-1} S_y = (M_y)_{11}^{-1} \tag{17}$$

i.e., the (1,1) element of the matrix M_y^{-1}.

The sample covariance matrix algorithm can be analyzed conveniently in the new coordinate system. The subscript y will be dropped in the following equations. Again, the sample covariance matrix is

$$\hat{M} = \frac{1}{K} \sum_{k=1}^{K} Y_k Y_k^{\dagger} \tag{18}$$

The weights based on this estimator are

$$\hat{W} = \hat{M}^{-1} S \tag{19}$$

When these weights are tested on a different set of samples than those used in estimating $\hat{W}$, the S/N ratio r_1 is

$$r_1 = \frac{|\hat{W}^{\dagger} S|^2}{\hat{W}^{\dagger} M \hat{W}} \tag{20}$$

The ratio of r_1 to the S/N with optimum weights is

$$\rho_1 = \frac{r_1}{r_o} = \frac{(S^{\dagger} \hat{M}^{-1} S)^2}{(S^{\dagger} M^{-1} S)(S^{\dagger} \hat{M}^{-1} M \hat{M}^{-1} S)} \tag{21}$$

The probability density of this variable ρ_1 was derived in [1].

In some cases of interest, the weights may be applied to the same set of samples used in computing $\hat{W}$. In this case, the output S/N ratio is

$$r = S^{+} \hat{M}^{-1} S = (\hat{M}^{-1})_{11} \tag{22}$$

The analysis of [1] can be extended to obtain an expression for the probability density of r.

The sample covariance matrix has a complex Wishart distribution[1,3], i.e.,

$$P(A) = \frac{|A|^{K-N}}{I(M)} \exp\left\{-\mathrm{tr}(M^{-1} A)\right\} \tag{23}$$

where $|A|$ denotes the determinant of A, N is the number of elements in the array, tr denotes the trace of the matrix, and $A = K \hat{M}$. The constant I(M) is a function of K, N, and the covariance matrix M. In (23), P(A) is the joint probability density of the elements of A, and is restricted to those matrices A which are positive definite. It assumes that the underlying noise process is complex Gaussian.

Consider the following representation of the matrix A.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \tag{24}$$

where A_{11} is a scalar, A_{21} is a (N-1)x1 column vector equal to A_{12}^{+}, and A_{22} is a (N-1)x(N-1) matrix. As in [1], A can be factored as follows

$$A = \begin{pmatrix} 1 & A_{21}^{\dagger} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}-A_{12}\ A_{22}^{-1}\ A_{21}, & 0 \\ A_{22}^{-1}\ A_{21} & , I \end{pmatrix}, \tag{25}$$

and

$$|A| = |A_{11}-A_{12}\ A_{22}^{-1}\ A_{21}|\ \ |A_{22}| \tag{26}$$

Let

$$\begin{aligned} D_{11} &= A_{11}-A_{12}\ A_{22}^{-1}\ A_{12}^{\dagger} \\ D_{12} &= A_{12} = A_{21}^{\dagger} \\ D_{22} &= A_{22} \end{aligned} \tag{27}$$

The Jacobian of the transformation from (A_{11}, A_{12}, A_{22}) to (D_{11}, D_{12}, D_{22}) is one, so

$$\begin{aligned} P(D_{11},D_{12},D_{22}) &= D_{11}^{K-N}\ |D_{22}|^{K-N}\ \frac{1}{I(M)}\ \exp\left\{-(D_{11}+D_{12}\ D_{22}^{-1}\ D_{12}^{\dagger} + \text{tr}\ D_{22})r_o\right\} \\ &= P(D_{11})\ P(D_{12},D_{22}) \quad , \end{aligned} \tag{28}$$

where

$$P(D_{11}) = C_1\ D_{11}^{K-N}\ \exp\left\{-r_o D_{11}\right\} \tag{29}$$

and C_1 is a constant.

Representing A^{-1} in the same form as (24),

$$A^{-1} = \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} , \tag{30}$$

it can easily be shown that $A^{11} = \frac{1}{D_{11}}$. Since $A = K \hat{M}$ and, from (22), the output S/N ratio is

$$r = (\hat{M}^{-1})_{11} = K A^{11} = \frac{K}{D_{11}} \tag{31}$$

Let $\rho = r/r_o = \frac{K}{r_o D_{11}}$. From (29)

$$P(\rho) = \frac{C_2 K^{K-N+1}}{\rho^{K-N+2}} \exp\left\{-\frac{K}{\rho}\right\} , \tag{32}$$

Normalizing this distribution, $C_2 = \frac{1}{(K-N)!}$, and

$$P(\rho) = \frac{K^{K-N+1}}{(K-N)!\ \rho^{K-N+2}} \exp\left[-\frac{K}{\rho}\right] , \rho > 0 \tag{33}$$

This is the probability density function for the normalized output S/N ratio, $\rho = r/r_o$, when the same samples are used in $\hat{M}$ for computing the weights and for testing the weights.

From (33),

$$\bar{\rho} = \frac{K}{K-N} \tag{34}$$

Note that $\bar{\rho} > 1$. When the same set of samples is used for computing the adaptive weights and for testing these weights, the output S/N ratio is greater than it would be with "optimum" weights, i.e., $W_o = M^{-1}S$. (This statement assumes that a signal is not present in the sample set, and signal is defined by the steering vector S. It will be shown in Section 5 that this algorithm using the same samples leads to a signal cancellation effect when a signal is present.)

4. PROBABILITY DENSITY OF MEAN OUTPUT NOISE POWER

In the analysis of the preceding section, the steering signal had the form $S_T = (1,0,0,0,...0)$. This represents the usual sidelobe canceller problem where the signal is present in the main channel and is either absent or negligible in the auxiliary channels. For example, the main channel may be fed by a high gain antenna directed toward the desired signal source and the auxiliaries fed by near-omnidirectional auxiliary elements. As formulated above, the magnitudes of the weights applied to all channels vary with the noise field.

A more convertional sidelobe canceller implementation, illustrated in Figure 1, constrains the weight on the main channel to a constant value, e.g., unity. While the performance of the two sidelobe canceller algorithms is equivalent, the configuration of Figure 1 with constant main channel gain is usually more convenient to implement. With a constant unit weight on the main channel, and no signal components in the auxiliary channels, the output signal power is independent of the adaptive weights.

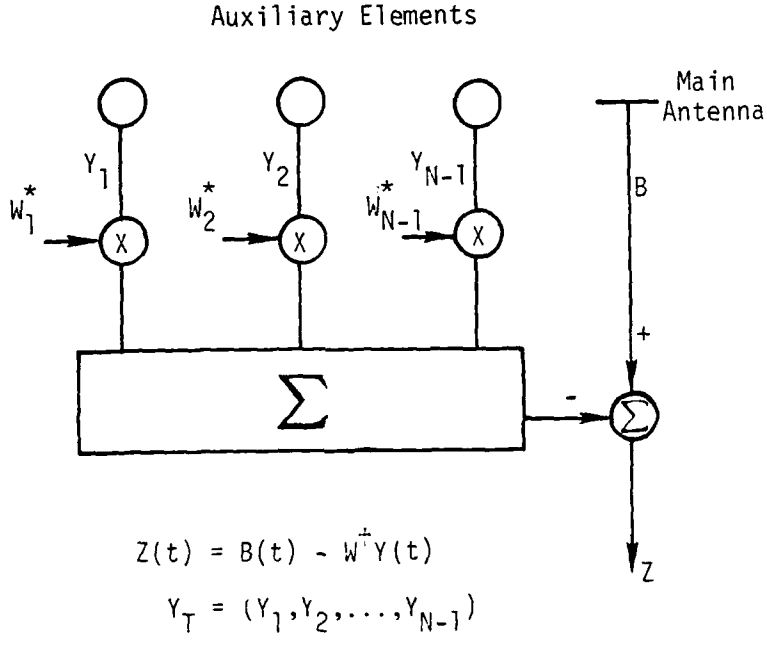


Figure 1. Sidelobe Canceller with Unit Weight on Main Channel

In the configuration of Figure 1, the output of the main channel is B and the (N-1) auxiliary element outputs are represented by the column vector Y, where $Y_T = (Y_1, Y_2, \ldots, Y_{N-1})$. The canceller output is

$$Z = B - W^{\dagger}Y \tag{35}$$

The output noise power,

$$\overline{|Z|^2} = \overline{BB^*} - W^{\dagger}\overline{YB^*} - \overline{BY^{\dagger}}W + W^{\dagger}M_y W \tag{36}$$

$$M_y = \overline{YY^{\dagger}}$$

is minimized when

$$W = M_y^{-1}\,\overline{YB^*} \tag{37}$$

Since the output signal power is independent of the weights, the probability distribution of the mean output noise power can be derived from the distribution of output S/N. First consider the case of [1] where the adaptive weights are based on a sample covariance matrix and the resulting weights are tested on a different set of samples. As derived in [1] the normalized output S/N ratio, given by ρ_1 of (21) has the probability density

$$P(\rho_1) = \frac{K!}{(N-2)!(K-N+1)!}(1-\rho_1)^{N-2}\rho_1^{K-N+1}, \quad 0 \le \rho_1 \le 1 \tag{38}$$

Let η_1, denote the normalized mean output noise power, i.e.,

$$\eta_1 = \frac{\overline{|Z|^2}_{\hat{W}}}{\overline{|Z|^2}_{W_o}} = \frac{1}{\rho_1}, \tag{39}$$

where $\hat{W}$ are weights based on a sample covariance matrix and W_o are optimum weights of (37). From (38) and (39)

$$P(\eta_1) = \frac{K!}{(N-2)!(K-N+1)!} \frac{(\eta_1-1)^{N-2}}{\eta_1^{K+1}} \quad 1 \leq \eta_1 \leq \infty \tag{40}$$

The mean output noise power exceeds that with optimum weights by the factor

$$\bar{\eta}_1 = \frac{K}{K-N} \tag{41}$$

As before, K denotes the number of samples in the sample covariance matrix and N is the total number of channels. The number of auxiliary channels in the sidelobe canceller configuration of Figure 1 is (N-1).

Similar expressions for the distribution and mean of the normalized output noise power, when the weights are computed and tested on the same sample set, can be derived from the results of the preceding section. The density function for the output S/N ratio, ρ, is given by (33). Again

$$\eta = \frac{\overline{|Z|^2}_{\hat{W}}}{\overline{|Z|^2}_{W_o}} = \frac{1}{\rho} \tag{42}$$

From (33) and (42),

$$P(\eta) = \frac{K^{K-N+1}}{(K-N)!}\,\eta^{K-N}\,e^{-K\eta}\,, \qquad 0 < \eta \tag{43}$$

and

$$\bar{\eta} = \frac{K-N+1}{K} = \left(1 - \frac{N-1}{K}\right) \tag{44}$$

Note that the mean value of η is less than unity. The weights based on a set of K noise samples reduce the noise in that sample set to a lower mean power level than weights based on the true covariance matrix of the noise process.

5. PROBABILITY DENSITY OF OUTPUT SIGNAL

As observed in the preceding section, the sample covariance matrix algorithm yields weights which are more effective in cancelling the noise in the same sample set than optimum weights. These weights are optimized for the particular set of samples and would not perform as well as the theoretically optimum weights (based on the true covariance matrix) against other input noise samples. The weights are optimized

to minimize the output noise power in the same specific set of samples. This suggests that these weights will also tend to cancel any signal present in the sample set when the weights are computed from and applied to the same samples. The following analysis shows that this is indeed the case.

Again, consider the sidelobe canceller configuration of Figure 1, where the weight in the main channel is constrained to unity and a signal, if present, occurs only in the main channel. Let

$$B = Y_o + S \tag{45}$$

where Y_o is the noise component and S is the signal component of B. The weights based on a set of K input samples, $\{Y_k\}$ in the auxiliaries and $\{B_k\}$ in the main channel, are

$$\hat{W} = \hat{M}_Y^{-1} \, \widehat{\overline{YB}^*} \tag{46}$$

where

$$\begin{aligned} \hat{M}_Y &= \frac{1}{K} \sum_{k=1}^{K} Y_k Y_k^\dagger \\ \widehat{\overline{YB}^*} &= \frac{1}{K} \sum_{k=1}^{K} Y_k \, B_k^* \quad . \end{aligned} \tag{47}$$

The sidelobe canceller output is then

$$Z = B - \hat{W}^{\dagger}Y \tag{48}$$

We are interested in obtaining the probability density and mean value of the output signal when a signal is present in B and the same samples are used for computing and testing the weights. The mean output signal voltage in Z is the correlation of Z with S, i.e.,

$$\begin{aligned} \overline{SZ^*} &= \overline{SB^*} - \overline{SY^{\dagger}\hat{W}} \\ &= \overline{SY_o^*} + \overline{SS^*} - \overline{SY^{\dagger}}\hat{M}_y^{-1}\,\overline{Y(Y_o^*+S^*)} \\ &= \overline{S(Y_o^* - Y^{\dagger}\hat{M}^{-1}YY_o^*)} + (\overline{SS^*} - \overline{SY^{\dagger}}M_y^{-1}\overline{YS^*}) \end{aligned} \tag{49}$$

In (49), the first term contains the product of S and a noise output voltage, so its mean value is zero. The noise component of this first term is the output noise voltage that would be present in the absence of a signal. When the signal is constant, the second term reduces to

$$\begin{aligned} \overline{SZ^*} &= \overline{SS^*}(1 - \overline{Y^{\dagger}}\,\hat{M}_y^{-1}\,\overline{Y}) \\ &= \overline{SS^*}\,(1 - q) \end{aligned} \tag{50}$$

where q is the fractional loss in signal. The same expression of (50) is obtained when the signal is any function of time, not necessarily a constant, provided its amplitude remains constant. In particular, (50) gives the mean output signal amplitude for a constant amplitude phase coded signal.

The probability density function for $q = \bar{Y}^{\dagger}\hat{M}_y^{-1}\bar{Y}$ is derived in the appendix. The same set of samples, $\{Y_k\}$, is used in forming the sample covariance matrix $\hat{M}_y$ and in obtaining the mean, $\bar{Y}$. The probability density of q, assuming K samples of Y_k are contained in both $\bar{Y}$ and $\hat{M}_y$ and that Y_k is a zero mean complex Gaussian process, is

$$P(q) = \frac{\Gamma(K)}{\Gamma(K-N+1)\,\Gamma(N-1)}\, q^{N-2}(1-q)^{K-N} \qquad 0 < q < 1 \qquad (51)$$

where N-1 is the number of auxiliary elements. The mean value of q is

$$\bar{q} = \frac{N-1}{K}\,, \qquad (52)$$

so the mean signal voltage is reduced by the factor $(1 - \frac{N-1}{K})$. This is the same factor obtained earlier in (44) for the reduction in mean output noise power.

6. CONCLUSIONS

In a digital adaptive array or sidelobe canceller, the sample covariance matrix algorithm can be used to obtain adaptive weights based on a set of array output samples. Applying these weights to the same set of samples to obtain the corresponding set of array

outputs is an effective method of optimizing the system performance in a rapidly changing noise field. The weights are always optimized for the specific noise field which is present during the sampling period. The probability density functions for the output noise power and output signal amplitude have been obtained for this same sample algorithm. It was shown that the application of same sample weights reduces the mean output noise power (Eq. 44) and mean signal amplitude (Eqs. 50 and 52) by the same factor.

REFERENCES

1. Reed, I. S., J. D. Mallett, and L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," IEEE Trans. AES, Nov. 1974, pp. 853-863.

2. Applebaum, S. P., "Adaptive Arrays," IEEE Trans. AP, Sept. 1976, pp. 585-598.

3. Goodman, N. R., "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution," Ann. Math. Stat., March 1963, pp. 152-177.

4. Cramér, "Mathematical Methods of Statistics," Princeton Univ. Press, 1951.

APPENDIX

Distribution of $\overline{X}^{\dagger}\hat{M}^{-1}\overline{X}$

To obtain the probability distribution of this quadratic form consisting of complex variables, a corresponding quadratic form will first be defined in terms of real variables and real matrix elements. Let

$$X = U + i\,V$$

$$Z = \begin{pmatrix} U \\ V \end{pmatrix} \tag{A1}$$

where U and V are column vectors of real variables. The corresponding covariance matrices are

$$M = \overline{XX^{\dagger}}$$

$$H' = \overline{ZZ^{T}} \tag{A2}$$

Expanding these matrices in terms of U and V

$$M = \overline{(U+iV)(U-iV)^{T}} = \overline{UU^{T}} + \overline{VV^{T}} + i\,(\overline{VU^{T}} - \overline{UV^{T}})$$

$$= (A + iB) \tag{A3}$$

$$H' = \overline{\begin{pmatrix} U \\ V \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix}^{T}} = \begin{pmatrix} \overline{UU^{T}} & \overline{UV^{T}} \\ \overline{VU^{T}} & \overline{VV^{T}} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

where $\overline{UU^T} = \overline{VV^T} = A/2$

$\overline{VU^T} = \overline{UV^T} = B/2$

These relationships between the second moments of U and V follow from the usual assumptions of zero mean normal distributions for the components of U and V, with $\overline{U_m U_n} = \overline{V_m V_n}$ and $\overline{U_n V_m} = -\overline{U_m V_n}$.

We will be interested in sample covariance matrix estimates of M and H'. From the equalities of (A3), better estimates of A and B for the real covariance matrix can be obtained by using

$$A = \overline{UU^T} + \overline{VV^T}$$

$$B = \overline{VU^T} - \overline{UV^T} \qquad \text{(A4)}$$

$$H = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} ,$$

where K complex samples in the estimates of M correspond to 2K real samples in the estimate of H.

Next consider the inverses of the covariance matrices M and H. Because of the form of H, H^{-1} can be expressed as

$$H^{-1} = \begin{pmatrix} C & -D \\ D & C \end{pmatrix} \qquad \text{(A5)}$$

where

$$H^{-1}H = \begin{pmatrix} C & -D \\ D & C \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = I$$

and

$$\begin{aligned} CA - DB &= I \\ DA + CB &= 0 \end{aligned} \tag{A6}$$

In terms of the real matrices C and D, the inverse of the complex matrix M is

$$M^{-1} = (C + iD) \tag{A7}$$

$$\text{since } M^{-1}M = (C+iD)(A+iB) = (CA-DB) + i(DA+CB) = I \tag{A8}$$

from (A6)

Next, the quadratic form

$$\overline{Z}^T H^{-1} \overline{Z} = \begin{pmatrix} \overline{U} \\ \overline{V} \end{pmatrix}^T \begin{pmatrix} C & -D \\ D & C \end{pmatrix} \begin{pmatrix} \overline{U} \\ \overline{V} \end{pmatrix} =$$

$$= \overline{U}^T C \overline{U} + \overline{V}^T C \overline{V} - \overline{U}^T D \overline{V} + \overline{V}^T D \overline{U} \quad . \tag{A9}$$

The corresponding quadratic form of complex variables is

$$\begin{aligned} \overline{X}^{\dagger} M^{-1} \overline{X} &= (\overline{U} - i\overline{V})^T (C+iD)(\overline{U}+i\overline{V}) \\ &= (\overline{U}^T C \overline{U} + \overline{V}^T C \overline{V} + \overline{V}^T D \overline{U} - \overline{U}^T D \overline{V}) \\ &= \overline{Z}^T H^{-1} \overline{Z} \end{aligned} \tag{A10}$$

Since the covariance matrix M and its inverse are Hermitian , the imaginary part of (A10) is zero.

Having derived the relationship (A10) between the real quadratic form $\overline{Z^T}H^{-1}\overline{Z}$ and the corresponding complex quadratic form $\overline{X^\dagger}M^{-1}\overline{X}$, results for the real variables case can be used to obtain the probability density and mean of $\overline{\hat{X}^\dagger}M^{-1}\hat{\overline{X}}$ in (50). Consider a set of K independent samples of the variables $X = U + iV$, where U and V are column vectors of N real variables, normally distributed with zero mean. Let

$$Z_k = \begin{pmatrix} U_k \\ V_k \end{pmatrix},$$

with the sample mean

$$\hat{\overline{Z}} = \frac{1}{K} \sum_{k=1}^{K} Z_k , \tag{A11}$$

and moment matrix

$$\hat{H} = \frac{1}{K} \sum_{k=1}^{K} Z_k Z_k^T \tag{A12}$$

Also, let $\hat{H}_1 = \frac{1}{K} \sum_{k=1}^{K} (Z_k - \hat{\overline{Z}})(Z_k - \hat{\overline{Z}})^T$

$$= \hat{H} - \hat{\overline{Z}}\,\overline{\hat{Z}^T} \tag{A13}$$

The elements of $\hat{H}_1$ are Wishart distributed [Eq. 29.9.6, p. 405 of Ref. 4], and independent of $\hat{\overline{Z}}$. The sample means are normally distributed with zero mean and a moment matrix H' of (A2).

The quadratic form of interest, $\hat{\bar{X}}^T M^{-1} \hat{\bar{X}}$, can now be related to the Hotelling distribution. Let

$$T^2 = (K-1)\hat{\bar{Z}}^T \hat{H}_1^{-1} \hat{\bar{Z}} \tag{A14}$$

The probability density function of T [Eq. 29.11.4, page 409 of Ref. 4] is

$$P(T) = \frac{2\Gamma(\frac{K}{2})}{(K-1)^{N/2}\Gamma(\frac{K-N}{2})\Gamma(\frac{N}{2})} T^{N-1}(1 + \frac{T^2}{K-1})^{-K/2}, \quad T>0 \tag{A15}$$

where K is the number of samples in $\hat{H}_1$ and $\hat{\bar{Z}}$, $\hat{H}_1$ is a N x N sample dispersion matrix, and $\hat{\bar{Z}}$ is a column vector of N sample means.

Next, from the matrix inversion identity and (A13),

$$\hat{H}^{-1} = (\hat{H}_1 + \hat{\bar{Z}}\hat{\bar{Z}}^T)^{-1} = \hat{H}_1^{-1} - \frac{\hat{H}_1^{-1}\hat{\bar{Z}}\hat{\bar{Z}}^T\hat{H}_1^{-1}}{1 + \hat{\bar{Z}}^T\hat{H}_1^{-1}\hat{\bar{Z}}} \tag{A16}$$

The quadratic form of interest is

$$Q = \hat{\bar{X}}^T \hat{M}^{-1} \hat{\bar{X}} = \hat{\bar{Z}}^T \hat{H}^{-1} \hat{\bar{Z}} \tag{A17}$$

In terms of the sample dispersion matrix $\hat{H}_1$,

$$Q = \hat{\bar{Z}}^T\left(\hat{H}_1^{-1} - \frac{\hat{H}_1^{-1}\hat{\bar{Z}}\hat{\bar{Z}}^T\hat{H}_1^{-1}}{1+\hat{\bar{Z}}^T\hat{H}_1^{-1}\hat{\bar{Z}}}\right)\hat{\bar{Z}}$$

$$= \frac{\hat{\bar{Z}}^T\hat{H}_1^{-1}\hat{\bar{Z}}}{(1+\hat{\bar{Z}}^T\hat{H}_1^{-1}\hat{\bar{Z}})} \tag{A18}$$

From (A14),

$$Q = -\frac{1}{\left(1 + \frac{T^2}{K-1}\right)}\left(\frac{T^2}{K-1}\right) \tag{A19}$$

where the probability density function of T is given in (A15).

To obtain the corresponding probability density function for Q, note that

$$\frac{dT}{dQ} = -\frac{(K-1)^{1/2}}{2Q^{1/2}(1-Q)^{3/2}} \tag{A20}$$

From (A15), (A19), and (A20)

$$P(Q) = \frac{2\Gamma(\frac{K}{2})(K-1)^{\frac{N-1}{2}}(1-Q)^{K/2}}{(K-1)^{N/2}\,\Gamma(\frac{K-N}{2})\Gamma(\frac{N}{2})}\left(\frac{Q}{1-Q}\right)^{\frac{N-1}{2}}\frac{\sqrt{K-1}}{2Q^{1/2}(1-Q)^{3/2}}$$

$$= \frac{\Gamma(\frac{K}{2})(Q)^{N/2-1}(1-Q)^{K/2-N/2-1}}{\Gamma(N/2)\Gamma(\frac{K-N}{2})}, \quad 0<Q<1 \tag{A21}$$

where the variables N and K represent the size of the real matrix and number of real samples in $\hat{H}$, respectively.

As noted in (A4), the effective number of samples in the sample covariance matrix H is 2K. Also, when (N-1) represents the size of the complex covariance matrix and sample mean vector, the corresponding dimension for the real variable case is 2(N-1). Replacing K by 2K and N by 2(N-1) in (A21), the corresponding probability density function for the q of (50) is

$$P(q) = \frac{\Gamma(K)\, q^{N-2}(1-q)^{K-N}}{\Gamma(N-1)\Gamma(K-N+1)} \qquad 0<q<1 \tag{A22}$$

This density function can be derived in a different way which does not require the analogy to the real variable problem. As before, let

$$\hat{M} = \frac{1}{K} \sum_{k=1}^{K} X_k X_k^{\dagger} \tag{A23}$$

and

$$\hat{M}_1 = \frac{1}{K} \sum_{k=1}^{K} (X_k - \hat{\bar{X}})(X_k - \hat{\bar{X}})^{\dagger} \tag{A24}$$

where

$$\hat{\bar{X}} = \frac{1}{K} \sum_{k=1}^{K} X_k \ .$$

Again, the X_k are column vectors of N zero mean complex Gaussian variables. The variables $(X_k - \hat{\bar{X}})$ are also zero mean Gaussian variables,

and the sample matrix $\hat{M}_1$ has the complex Wishart distribution of (23) with K replaced by K-1.

Next, consider the quantity

$$R = K \frac{\underline{\hat{X}}^{\dagger} M^{-1} \underline{\hat{X}}}{\underline{\hat{X}}^{\dagger} \hat{M}_1^{-1} \underline{\hat{X}}} \tag{A25}$$

For any column vector $\underline{\hat{X}}$, there is a unitary transformation on X which leaves the value of each quadratic form in (A25) unchanged, but reduces $\underline{\hat{X}}$ to the form $(a,o,o,o\ldots,o)_T$. Without loss of generality, it can be assumed that M=I. Then

$$R = \frac{K}{(\hat{M}_1^{-1})_{11}}, \tag{A26}$$

and from the analysis of Section 3 and (29),

$$P(R) = \frac{R^{K-1-N}}{(K-1-N)!} e^{-R}, \qquad R>0 \tag{A27}$$

The numerator of (A25),

$$a = K \underline{\hat{X}}^{\dagger} M^{-1} \underline{\hat{X}} = K \underline{\hat{X}}^{\dagger} \underline{\hat{X}} \tag{A28}$$

is the sum of 2N real Gaussian variables and has the probability density

$$P(a) = \frac{a^{N-1} e^{-a}}{(N-1)!}, \qquad a>0 \tag{A29}$$

The sample means are independent of the sample matrix $\hat{M}_1$, so R and a are independent. Let

$$Q = \underline{\hat{X}}^{\dagger}\hat{M}_1^{-1}\underline{\hat{X}} = \frac{a}{R} \tag{A30}$$

From (A27) and (A30)

$$\begin{aligned} P(Q) &= \int_0^{\infty} P(R)P(Q|R)dR \\ &= \int_0^{\infty} \frac{R^{K-N-1}}{(K-N-1)!} e^{-R} \frac{(RQ)^{N-1} e^{-RQ}}{(N-1)!} R\, dR \qquad \text{(A31)} \\ &= \frac{(K-1)!}{(K-N-1)!(N-1)!} \frac{Q^{N-1}}{(1+Q)^K} , \qquad Q>0 \end{aligned}$$

We are interested in obtaining the distribution of

$q = \underline{\hat{X}}^{\dagger}\hat{M}^{-1}\underline{\hat{X}}$, where

$$\hat{M} = \hat{M}_1 + \underline{\hat{X}}\,\underline{\hat{X}}^{\dagger} \tag{A32}$$

Since

$$\hat{M}^{-1} = \hat{M}_1^{-1} - \frac{\hat{M}_1^{-1}\underline{\hat{X}}\,\underline{\hat{X}}^{\dagger}\hat{M}_1^{-1}}{1+\underline{\hat{X}}^{\dagger}\hat{M}_1^{-1}\underline{\hat{X}}} \tag{A33}$$

The quadratic forms q and Q satisfy

$$q = \frac{Q}{1+Q} \tag{A34}$$

From (A31) and (A34)

$$P(q) = \frac{(K-1)!}{(K-N-1)!(N-1)!} q^{N-1}(1-q)^{K-N-1}, \quad 0<q<1 \qquad \text{(A35)}$$

Replacing N with N-1 in (A35), where (N-1) is the number of auxiliary elements as in (A22), the two results for P(q) are the same.