DTIC FILE COPY

(4)

# STATISTICAL ANALYSIS OF ADAPTIVE BEAM-FORMING METHODS

AD-A196 725

DTIC
SELECTED
JUN 1 4 1988
D

### ESL Final Report

by
R. C. McCarty

Senior Staff Scientist

ESL-88-0099

88 6 003

ESL INCORPORATED
A Subsidiary of TRW Incorporated
Sunnyvale, California


STATISTICAL ANALYSIS OF ADAPTIVE BEAM-FORMING METHODS


ESL Final Report


May 1, 1988


by


R.C. McCarty


Senior Staff Scientist

# CONTENTS (U)

## ACKNOWLEDGEMENTS

# ABSTRACT

A statistical measure of the performance of an adaptive beam-forming technique for a multisensor linear array has been developed. The sample complex vector signal process, $Z(t_k)$ as observed at sample times $(t_k) \in (O,T)$ is used to calculate a positive, definite, maximum likelihood estimate (M.L.E.) $\hat{R}$ of the signal process covariance matrix $R$. Then, an estimated measure of the performance of the array in terms of its ability to suppress an interfering signal is developed from an *a priori* specified process matrix $E(\hat{R})$, its sample M.L.E. estimate $\hat{R}$ and a form of the Wishart distribution. In particular, such an estimated measure of performance in terms of suppressing an interfering signal with an angle of arrival (AOA) of $\theta_2$ relative to a beam AOA of $\theta_1$ is the statistic represented by the absolute difference of the $L_2$ norms, $\{ |\, \|\hat{\rho}^{-1}\|_2$ $- \|\rho^{-1}\|_2 | \}$. The estimate $\hat{\rho}^{-1} = V^*(\theta_1) \; \hat{R}^{-1} \; V(\theta_1)$ , it's exact value is $\rho^{-1} = V^*(\theta_1) \; R^{-1} \; V(\theta_1)$ and $R, \hat{R}$ are functions of $V(\theta_2)$ and $V^*(\theta_2)$.

# 1. BACKGROUND.

The research undertaken here is a statistical study of the performance properties of an adaptive beam-forming algorithm when utilized to process stochastic signals as received by a multi-element linear array.

A performance measure is defined which represents the degree to which the adaption process minimizes off-boresight signal interference from a single source.

It is well known, particularly in acoustics, that multipath signal interference in a dispersive medium such as salt water presents many difficulties to the detection and acquisition of wanted signals when in the presence of unwanted or interfering signals. In fact, even in the case of a single interferer, it is often very difficult to supress such signals at low signal-to-noise (SNR) ratios without the processing enhancement as offered by adaptive beam-forming techniques.

Most of the existing adaptive signal processing techniques deal with the deterministic, expected or steady state properties of stochastic signal processes which enable one to defer or even ignore the distributional aspects of the signal processes being considered. Therefore, since most interest and concern has been given heretofore to beam-forming techniques using assumed expected values of signal processes parameter estimates rather than to the distributional aspects thereof, it was decided that improvements in the analysis of beam-forming techniques could best be achieved by a rigorous statistical analysis of the problem. Thus, major emphasis was placed on development of the distribution of the estimated measure of supression of off boresight signals.

It has long been recognized that the distributional aspects of statistical sampling theory could be adapted to provide quantitative measures of performance for physical processes such as the beam forming of signals from an array of sensors. In some cases, such as for multi-variate, (vector) Gaussian stochastic processes, many of the essential sampling distributional properties and characteristics were developed and published by John Wishart[2] of Rothamsted Experimental station in England, circa 1928. Then in those lean, worldwide depression years of 1932 and 1933, Wishart and Bartlett[3,4], while at Cambridge University, utilized a characteristic function approach, made possible by the work of Ingram[5], to rigorously established the existence of a class of distributions which became known later as Wishart distributions. In 1938, the noted Chinese mathematician, P. L. Hsu[6], generated an elegant proof for the existence of a general form of the Wishart distribution by mathematical induction.

In 1956 Cramer[7] and in 1963, Fisz[8] exemplified this form of the Wishart distribution, citing the earlier works of Wishart and Bartlett and the existence proof of Hsu, respectively. With the advent of later works in the vector calculus, linear algebra, matrix theory and modern abstract analysis, multivariate statistical analysis became a very powerful tool in the mathematical

modeling of multidimensional (vector) stochastic physical processes such as encountered in the adaptive beam forming of passive signals, received from an array of sensors consisting of a large number of elements.

In 1984, Anderson[9] of Stanford, published the second edition of his "Encyclopedic" text on multi-variate analysis in which he rigorously developed all of the properties of the several classes of Wishart distributions both in the real and complex domains. Thus, it is the Wishart distribution and the various transformed versions thereof which provide the statistical foundations for the work to be presented in this report.

## 2. INTRODUCTION.

In the research results presented here, a statistical measure of performance for an adaptive beam-forming technique and its associated distribution function has been developed for the case of a complex Gaussian signal process. The statistical measure is a sample estimate of the degree to which and adaptive beam-forming process for a linear array can minimize the off-axis interference of a plane wave propagated from a single interferer. Most of the existing signal processing references are concerned only with the expected value of requisite statistical estimators. That is, the existing work has assumed only that the sample estimators of interest exist and are unbiased. What is not well understood by many signal processing practitioners, however, is how system parameters and the number of sensors effect the distribution of the estimators. When complete, this research can be utilized to provide answers to the following types of questions:

- Given a noise model, interference model and sensor array size, what is the optimal sampling method in order to obtain a specified probability of detection while at the same time utilizing a minimum amount of computing resources?

- What are the tradeoffs being made when a system design selects *block averaging* over *exponential averaging*? Will system performance improve and if so by how much?

- It is well known that the noise sampled at adjacent sensors is not statistically independent. What are the effects of this anisotropy on the distribution of the beam-former output as a function of sampling technique?

- For arrays with a very large number of sensors ( $> 10^3$ ) a long sampling time would be required to form a good estimate of the covariance matrix. In this case, how does one develop a strategy to combine sensors to produce the best improvement in the estimation of the covariance matrix while keeping within the desired integration times?

The distribution of the performance estimate (in terms of the inverse covariance matrix estimate) enables the specification of confidence bounds for the process performance measure as a function of the number of samples and an *a priori* selected level of significance. These results are valid for any signal processing system operating in a low SNR environment and where statements of precision for requisite estimates are most desirable.

# 3. THE SIGNAL SAMPLE SPACE.

Consider an $n$ element $(n \geq 2)$, multi-sensor linear array with a spacing between sensors of one half the wave length $(\lambda/2)$. The sample signal process $Z(t_k) \in C^n, k = 1,2, \dots, M$ is considered to be a complex, non-zero mean, wide-sense stationary, Gaussian vector process with sample components, $z_1(t_k), z_2(t_k), \dots, z_n(t_k)$, for each $\{t_k\} \in [O,T], T < \infty, k = 1,2, \dots, M$. Thus, the sample vector

$$Z(t_k) = \begin{bmatrix} z_1(t_k) \\ z_2(t_k) \\ \vdots \\ z_n(t_k) \end{bmatrix} \text{ and } Z^*(t_k) = [\bar{z}_1(t_k) \ \bar{z}_2(t_k) \dots \bar{z}_n(t_k)] \tag{3.1}$$

Then, for each $t_k \in [O,T]$; $k = 1,2, \dots, M$, the complex sample vectors $Z(t_1), Z(t_2), \dots, Z(t_M)$ are assumed to be independent identically distributed (I.I.D.) over a temporal sample interval $[O,T]$ for $T < \infty$. Perhaps a less restrictive condition than general independence is to consider the sample vectors as random variables from a quasi-stationary, Gaussian sample process which is (strict sense) stationary over the finite sample interval $[O,T]$, but temporally uncorrelated through at lease the second degree. In general, a second degree, uncorrelated sample process is a process in which the sample values about the mean are uncorrelated such that $E[(\overline{Z_k - \Theta_j})^2]_{k,y} = 0$ for $(k,y) = 1,2, \dots, M$.

A principle reason for meeting such a sampling condition is that the sample estimates of the process parameters which define the covariance matrix are statistically consistent and hence so is the covariance matrix $\hat{R}$.

# 4. COMPLEX MULTIVARIATE GAUSSIAN CASE.

Consider a sample set of size $M$ of second degree, uncorrelated, Gaussian random vectors $\{Z_1, Z_2 ..., Z_M\} \in C^n$ where $Z_i = Z(t_i)$, $i \in [1,M]$, as previously defined. Then M (n element) I.I.D. complex vectors have the joint probability density function,

$$f(Z_1, Z_2, ..., Z_M) = \prod_{k=1}^{M} (\pi^{\frac{-n}{2}}) |R|^{-1/2} \; e^{-1/2(Z-\Theta_s)^* R^{-1}(Z-\Theta_s)}$$

$$= \pi^{\frac{-Mn}{2}} |R|^{\frac{-M}{2}} \; e^{\frac{-M}{2}[(Z-\Theta_s)^* R^{-1}(Z-\Theta_s)]}$$

(4.1)

where the complex process means $\Theta_1 = E(Z)$ and the covariance matrix $R \in M_n$ ($n \times n$ complex matrix) is a nonsingular Hermitian matrix such that $(Z - \Theta_s)^* R(Z - \Theta_s) > 0$ for all non-zero $(Z - \Theta_s) \in C^n$, and hence $R$ is positive definite. Therefore, the matrix $R^{-1} \in M_n$ exists and is also a positive Hermitian matrix.

# 5. MAXIMUM LIKELIHOOD ESTIMATES OF $\Theta$ R.

It can be demonstrated that if $Z_1, Z_2, \ldots, Z_M$ constitutes a sample of size $M$ of an $n$ element I.I.D. complex Gaussian vector process where each complex vector is $N(\Theta, R)$ with $M > n$, then the maximum likelihood estimators of $\Theta$ and $R$ are:

$$\hat{\Theta}_z = \frac{1}{M} \sum_{k=1}^{M} Z_k \hat{R} = \frac{1}{M} \sum_{k=1}^{M} (Z_k - \hat{\Theta}_z)(Z_k - \hat{\Theta}_z)^* \tag{5.1}$$

For computational purposes we write

$$\hat{\Theta}_z = \frac{1}{M} \sum_{k=1}^{M} Z_{ik} \; ; \; \hat{\Theta}_z^* = \frac{1}{M} \sum_{k=1}^{M} \overline{Z}_{jk} \tag{5.2}$$

$$\hat{R}_z = \frac{1}{M} \sum_{k=1}^{M} Z_k Z_k^* - M \hat{\Theta}_z \hat{\Theta}_z^* \tag{5.3}$$

$$Z(t_k) = Z_k = \begin{bmatrix} Z_{1k} \\ Z_{2k} \\ \cdot \\ \cdot \\ \cdot \\ Z_{nk} \end{bmatrix} \text{ and } Z^*(t_k) = Z_k^* = [\overline{z}_{1k} \ \overline{z}_{2k} \cdots \overline{z}_{nk}] \tag{5.4}$$

$$\frac{1}{M} \sum_{k=1}^{M} Z_k Z_k^* = \frac{1}{M} \sum_{k=1}^{M} Z_{ik} \overline{Z}_{jk} \tag{5.5}$$

$$M \hat{\Theta}_z \hat{\Theta}_z^* = \frac{1}{M} \sum_{k=1}^{M} Z_{ik} \sum_{k=1}^{M} \overline{Z}_{jk} \tag{5.6}$$

and hence, the sample covariance matrix is

$$\hat{R}_z = [f_{ij}] \; ; \; f_{ij} = \frac{1}{M} \left[ \sum_{k=1}^{M} Z_{ik} \overline{Z}_{jk} - \sum_{k=1}^{M} Z_{ik} \sum_{k=1}^{M} \overline{Z}_{jk} \right] \tag{5.7}$$

# 6. THE WISHART DISTRIBUTION.

In 1928 John Wishart prublished the first of several papers [2,3,4] in multivariate statistical analysis. Anderson[9] shows that if the sample covariance matrix, $\hat{R} = (M-1)S = \sum_{k=1}^{M}(X_k - \bar{X})'$, where $X_1, \ldots, X_M(M > n)$ are $n$ component, real valued, statistically independent vectors, each with distribution $N(\mu, R)$ then $\hat{R}$ has the density

$$g(\hat{R}\,|\,R,M,n) = \frac{|\hat{R}|^{1/2(M-n-2)}e^{-1/2tr(R^{-1}\hat{R})}}{2^{1/2n(M-1)}\frac{\pi^{\frac{n(n-1)}{4}}}{}|R|^{1/2(M-1)}\prod_{i=1}^{n}\Gamma[1/2(M-i)]} \tag{6.1}$$

for $\hat{R}$ positive definite and 0 otherwise. The expression (5.1) is the original Wishart result in real vector form. For $Z_1, Z_2, \ldots, Z_N$ normally distributed complex vectors, each with mean $\Theta$ and Hermitian covariance matrix, $R$, the sample estimate $\hat{R}$ of the covariance matrix $R$ has the complex Wishart density

$$g(\hat{R} \in M_n)\,|\,R,M,n) = \frac{|\hat{R}|^{(M-n)}e^{1/2tr(\hat{R}R^{-1})}}{\pi^{1/2n(n-1)}|R|^{M}\prod_{i=1}^{n}\Gamma(M-i+1)} \tag{6.2}$$

where $\hat{R} = \sum_{k=1}^{N}(Z_k - \hat{\Theta}_s)(Z_k - \hat{\Theta}_Z)^*$, $\hat{\Theta}_s = \frac{1}{N}\sum_{k=1}^{N}Z_k$, $M = N-1$, and $(R$ and $\hat{R}) \in M_n$ are Hermitian positive definite matrices, since $Z^* RZ > 0$ and $Z^* RZ > 0$ for all non-zero $Z \in C^n$. The matrices $R$ and $\hat{R}$ have the additional property, that if the number of sample vectors $M$ are greater than $n$ the number of components of each vector, then the probability is unity of drawing a sample so that $\hat{R}$ is positive definite $(Z^* \hat{R}Z) > 0$ and on the other hand, if $M < n$ (i.e., the number of sample vectors is less than the number $n$ of the components of which each vector is composed) then it can be demonstrated that $\hat{R}$ does not have a density, but is does have a well defined cumulative distribution function (c.d.f.) (monotone non-decreasing).

$$G(\hat{R}\,|\,R,M,n) = \begin{cases} 0 & |\hat{R}| \le 0 \\ 0 \le \alpha \le 1 & |\hat{R}| > 0 \end{cases} \tag{6.3}$$

For $R$ and $\hat{R}$ Hermitian positive definite it follows that $R^{-1}$ and $\hat{R}^{-1}$ are also Hermitian positive definite. In addition by[10]

Theorem 1

If $H$ is a Hermitian matrix, then $x^*Hx$ is real for all $x \in C^n$.

**Proof:**

Write $\overline{(x^*Hx)} = (x^*Hx)^* = (x)^*(H)^*(x^*)^* = x^*H^*x) = (x^*Hx)$, i.e., $(x^*Hx)$ equals its complex conjugate and is therefore real by definition.

# 7. BEAM-FORMING MEASURE OF PERFORMANCE.

In adaptive beam-forming problems, an appropriate measure of performance of the beam-froming technique is the degree to which the beam-former output approaches an ideal supression response to unwanted signals in an otherwise "noise free" environment. By noise free it is meant in the absence of external and/or internal system noise, deterministic or random.

Thus, an ideal measure of performance in the sense of suppressing the unwanted signals of a single interferer at an AOA of $\theta_2$ by the beam-forming processing at an AOA of $\theta_1$, is

$$\rho^{-1} = V^*(\theta_1)R^{-1}(\theta_2)V(\theta_1) \qquad (7.05)$$

Then the sample estimator $\hat{\rho}^{-1}$ of $\rho^{-1}$ and its destributional properties with resepct to $\rho^{-1}$ are defined as follows.

Consider an interfering signal with an AOA of $\theta_2$ relative to a linear array with a beam formed at an AOA of $\theta_1$. For $\hat{R}_M^{-1}(\theta_2) = [\hat{r}_{ij}(\theta_2)]$ Hermitian and positive definite, it follows $\hat{\rho}_M^{-1}(\theta_1, \theta_2) = V^*(\theta_1)\hat{R}^{-1}(\theta_2)V(\theta_1) > 0$ (postive definite) and real by Theorem 1. $V^*(\theta)$, $V(\theta)$ are the usual phase vectors used in conventional beam forming. Because $\hat{\rho}_M^{-1}(\theta_1, \theta_2)$ is a function of the number of time samples $M$ for any AOA's $\theta_1, \theta_2$ it can be demonstrated that

$$\lim_{M \to \infty} E(\hat{\rho}_M^{-1}) = \rho^{-1} \qquad (7.1)$$

and

$$\lim_{M \to \infty} E(\hat{\rho}_M^{-1} - \rho^{-1})^2 = 0 \qquad (7.2)$$

Hence, $\hat{\rho}_M^{-1}$ converges in expectation to $\rho^{-1}$ and is also a consistent estimate by (7.2).

For notational sake, we will use $(\hat{\eta}_{ij}) = \hat{R}^{-1}$ and $(\eta_{ij}) = R^{-1}$. The sample statistic $\hat{\rho}_M^{-1}$ can be written as a matrix

$$\hat{\rho}_M^{-1} = \sum_{ij}^{n} [\hat{\eta}_{ij}\bar{v}_i(\theta_1)v_j(\theta_1)] \qquad (7.3)$$

which is positive definite and thus real by Theorem 1.

Deriving the distribution of $\hat{R}^{-1}$ from the complex Wishart density given by (5.2) and computing the Jacobian of the transformation $(\hat{B} = \hat{R}^{-1})$, which is the determinant of the $n^2 \times n^2$ positive definite matrix

$$J(\hat{B}) = \left| \frac{\delta \eta_{ij}}{\delta \eta_{kj}} \right| = |\hat{R}|^{-2n} \tag{7.4}$$

Consequently,

$$g(\hat{B}) = |(R,M,n)| \hat{R} |^{-2n} = g(\hat{R}^{-1} | R^{-1}, M, n) \tag{7.5}$$

if $\hat{R}^{-1}$ is positive definite and 0 otherwise.

The density distribution for the sample statistic, $\hat{\rho}^{-1}$ can then be obtained from (7.5) for $\hat{B} = \hat{\rho}^{-1} = V^* (\theta_1) \hat{R}^{-1} V(\theta_1)$ and calculating $J(\hat{B}) = J[V^*(\theta_1) \hat{R}^{-1} V(\theta_1)] = J[\hat{b}_{ij}]$ where $\hat{b}_{ij} = [\eta_{ij} \bar{v}_i(\theta_1) v_j(\theta_1)]$. These calculations yield

$$g(\hat{b} | P, M, n) J(\hat{B}) = g(V^*(\theta_1) \hat{R}^{-1} V(\theta_1) | P, M, n) \tag{7.6}$$

for $\hat{\rho}^{-1} \geq 0$ and 0 otherwise.

Using the Wishart density function (7.6) it is found that the sample statistic,

$$\|\hat{B}\|_2 = \left[ \sum_{ij}^{n} |\eta_{ij}|^2 v_i(\theta_1) v_j(\theta_1) \right]^{1/2} = \|\hat{\rho}^{-1}\|_2 \tag{7.7}$$

with

$$\|P\|_2 = \left[ \sum_{ij}^{n} |\eta_{ij}|^2 v_i(\theta_1) v_j(\theta_1) \right]^{1/2} = \|\rho^{-1}\|_2 \tag{7.8}$$

has density

$$g(\hat{b} | P, M, n) J(\hat{B}) = \begin{Bmatrix} g(\|\hat{\rho}^{-1}\|_2 | \|\rho^{-1}\|_2, M, n) & \text{for } \|\hat{\rho}^{-1}\|_2 \geq 0 \\ 0 & \text{otherwise} \end{Bmatrix} \tag{7.9}$$

then it follows quickly that,

$$P\left[ | \|\hat{\rho}^{-1}\|_2 - \|\rho^{-1}\|_2 | \leq \epsilon(M, \alpha) \right] = \int_0^\epsilon dG(\|\hat{\rho}^{-1}\|_2 | \|\rho^{-1}\|_{2, M, n}) = 1 - \alpha \tag{7.10}$$

or

$$P\left( \left[ \|\hat{\rho}^{-1}\|_2 - \|\rho^{-1}\|_2 \right]^2 \leq \epsilon^2(M, \alpha) \right) = \int_0^{\epsilon_2(M, \alpha)} dG(\|\hat{\rho}^{-1}\|_2 | \|\rho^{-1}\|_{2, M, n}) = 1 - \alpha \tag{7.11}$$

Where $M$ is the number of time samples and $0 < \alpha < 0.10$ is an *a priori* selected level of significance. The cumulative distribution function represented by (7.11) yields the probability $1 - \alpha$, that $\|\hat{\rho}^{-1}\|_2$ will fall into an interval (confidence interval) of width $2c(\alpha, M)$ with midpoint $\|\rho^{-1}\|_2$.

## 8. A QUICK LOWER BOUND IN PROBABILITY.

Many times a useful lower bound in probability can be obtained by the use of Chebyshev's Inequality [11]. Consider the sample statistic $\|\hat{\rho}^{-1}\|_2$ with $E\|\hat{\rho}^{-1}\|_2 < \infty$ and $\sigma^2(\|\hat{\rho}^{-1}\|_2) < \infty$ as before. Then by Chebyshev's inequality and the sample estimate $\hat{\sigma}_M^2$,

$$P\left(\left[\|\hat{\rho}^{-1}\|_2 - E\|\hat{\rho}^{-1}\|_2\right]^2 \le \epsilon^2\right) = 1 - \hat{\sigma}_M^2 \frac{(\|\hat{\rho}^{-1}\|_2}{\epsilon^2} \tag{8.1}$$

This bound can be quite sharp in this case, since $\sigma_M^2(\|\hat{\rho}^{-1}\|_2) = O(1/M)$. Another very useful inequality for a quick estimate of the statistical magnitude of the random variable $\|\hat{\rho}^{-1}\|_2$ is given in terms of the Markov inequality

$$P[\|\hat{\rho}^{-1}\|_2 \ge \alpha] = \frac{E\|\hat{\rho}^{-1}\|_2}{\alpha} \tag{8.2}$$

where $\alpha > 0$ and $\|\hat{\rho}^{-1}\|_2 \ge 0$, regardless of the actual distribution of $\|\hat{\rho}^{-1}\|_2$, which of course has been demonstrated to be Wishart by expressions (7.9) and (7.10), respectively. For large numbers of temporal samples $M < \infty$, both (8.1) and (8.2) can often give very useful results quite rapidly. The Markov inequality has the additional advantage that if the sample vector components used to calculate $\hat{R}^{-1}$ are temporally correlated as well as spatially correlated across an array, $E\|\hat{\rho}^{-1}\|_2 < \infty$ can be demonstrated to be $O(1/M)$ and is thus asymptotically unbiassed in expectation.

If it can be demonstrated [15] that the estimated variances and corvariances have temporal component correlations of less than $M^{n+1.5}$ for the number of temporal samples $M$, where $1 \ll M < \infty$ for each of the sensors, $2 \le n < \infty$, in the array, then the $L_2$ sample estimates such as $\|\hat{\rho}_M^{-1}\|_2$ are $\|\hat{\rho}_M^{-1} - \hat{\rho}^{-1}\|_2$ are stoclastically convergent, and asymptotically unbiassed, $O(1/M)$.

## 9. CONCLUSIONS.

- Optimal sampling occurs in the adaptive beam-forming estimation process when the number of temporal samples $M \geq n$, the number sensors, $n \gg 2$. If $M < n$, see e.g. (6.3), the probability density in the form of (7.9) does not exist.

- Block averaging as required in the sample estimation of $\hat{R}^{-1}$ and hence $\hat{\rho}_M^{-1}$ provides statistically consistent estimates of the process parameters necessary to the efficient interference supression of unwanted signals by the beamformer process.

- Exponential averaging [12,13,14,15] does not provide numerically or statistically consistent estimates of the requisite process parameters either for deterministic or stochastic sample process.

- Spatially correlated signal process samples between adjacent sensor elements of an array are accounted for in the estimation of $\hat{R}^{-1}$, $\hat{\rho}^{-1}$ and hence the subsequent density distribution given by 7.9.

- In the case of a large number of sensors $n$, $(n > 10^\nu)$ and the requirement that the number of temporal samples $M \geq n$, for the obtainment of statistically consistent supression performance estimates, integration times consistent with specific improvements in SNR ratios should pose no unusual requirements for computational agility.

- A derived form of the complex Wishart distribution for the sample estimator of an array's performance, $G\{\|\hat{\rho}^{-1}\|_2\}$, provides a probability measure for the precision of the estimated performance magnitude $\|\hat{\rho}^{-1}\|_2$ as a function of the number of temporal samples $M$, the number of sensor elements $n$ and a significance level $\alpha$.

- For $N$ temporally correlated sample values on a finite sample interval $[O, T < \infty]$ which are such that the sample process covariances estimates $\{Cov(X_i, X_j); i,j = 1,2 \ldots, N \text{ and } i \neq j\}$ $\leq N^{\nu \geq 1.5}$, the derived form of the Wishart distribution $G_N\{\|\hat{\rho}^{-1}\|_2\} \rightarrow G\{\|\hat{\rho}^{-1}\|_2 | \rho^{-1}, M, n\}$ asymptotically, even for processes which do not strictly satisfy the initial Gaussian assumptions.

# REFERENCES

1.  ESL Interim Report, Kandle, D. A. and McCarty, R.C., *Statistical Analysis of Adaptive Beamforming* (ONR Contract N00014-85-C-0891), ESL Inc., Sunnyvale, CA, September 1986.

2.  Wishart, John, "The Generalized Product Moment Distribution in Samples from a Normal Multivariate Population" *Biometrica*, vol. 20A, 1928.

3.  Wishart, John and Bartlett, M. S., "The Distribution of Second Order Moment Statistics in a Normal System," *Proceeding Cambridge Philosophical Society* vol. 28, pp 455-459, 1932.

4.  Wishart, John and Bartlett, M. S., "The Generalized Product Moment Distribution in a Normal System," *Proceedings Cambridge Philosophical Society*, vol. 29, pp 260-270, 1933.

5.  Ingram, A. E., "An Integral Which Occurs in Statistics," *Proceedings Cambridge Philosophical Society*, vol. 29 pp 271-276, 1933.

6.  Hsu, P. L., "A New Proof on the Joint Product Moment Distribution," *Proceedings Cambridge Philosophical Society*, vol. 35, pp 336-338, 1939.

7.  Cramer, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, New Jersey 1951.

8.  Fisz, M., *Probability Theory and Mathematical Statists (3rd Ed.)*, John Wiley & Sons, New York, 1963.

9.  Anderson, T. W., *An Introduction to Multivariate Statistical Analysis (2nd Ed.)*, John Wiley & Sons, New York, 1984.

10. Horn, R. A. and Johnson, C. A. *Matrix Analysis*, Cambridge University Press, New York, 1985.

11. Birnbaum, Z. W., *Introduction to Probability and Mathematical Statistics*, Harper and Brothers, New York, 1962.

12. Hildebrand, F. B., *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956.

13. Lanczos, C., *Introduction to Numerical Analysis*, Prentice Hall, Englewood Cliffs, New Jersey, 1956.

14. McCarty, R. C. and Jolley, D. K. *Tutorial On Prony's Method*, ESL TM-1989, Unclassified, ESL Incorporated, Sunnyvale, CA, May 1986.

15. McCarty, R. C., "Sample Estimators for Entropic Measures of Mutual Information," *SPIE Proceedings*, San Diego, CA, August 1988. (to be published as above)

# ADDITIONAL REFERENCES. (NOT CITED IN TEXT)

Bienvenu, G., Kopp, L., "High-Resolution Array Processing Using Eigensystem: Estimation Aspect," *High-Resolution Spatial Processing in Underwater Acoustics*, R. A. Wagstaff and A. B. Baggeroer (eds.) 1984.

Bienvenu, G., Kopp, L., "Principal de la Goniometrie Passive Adaptive," *Proc. 7'ieme Collopue GRESTI*, Nice France, 1979, pp. 106/1-106/10.

Bienvenu, G., Kopp, L., "Adaptivity to Background Noise Spatial Coherence for High-Resolution Passive Methods," *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Processing*, Denver, CO, 1980, pp. 307-310.

Bienvenu, G., Kopp, L., "Source Power Estimation Method Associated With High-Resolution Bearing Estimator," *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Processing*, Atlanta, GA, 1981, pp. 302-305.

Cron, B. F., Sherman, C. H., "Spatial Correlation Functions for Various Noise Models," *J. Acoust. Soc. Am.*, vol. 11, 1962, pp. 1732-1736.

Golub, G. H., Van Loan, C. V., *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 1983.

Ligget, W. S., "Passive Sonar Fitting Models to Multiple Time Series," *NATO A.S.I on Signal Processing*, ed. G. W. R. Griffiths et al., New York Academic Press, 1973, pp. 327-345.

Owsley, N. L., "Sonar Array Processing," *Array Signal Processing*, S. Haykin, ed., 1985.

Stewart, G. W., "Error and Perturbation Bounds for Subspaces Associated With Certain Eigenvalue Problems," *SIAM Review*, vol. 15, 1973.

Wax, M., Shan, T-J., Kailath, T., "Spatio-temporal Spectral Analysis by Eigenstructure Methods," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 32, pp. 817-827 and 727-764.

Horibe, Y., "Entropy and Correlation," *IEEE Trans. Systems, Man and Cybernetics*, vol. SMC-15, No. Sept/Oct 1985.

Blachman, N., *Noise and Its Effect On Communication*, McGraw-Hill, New York.

Blum, J., and Rosenblatt, M. *Probability and Statistics*, Chapter 5, W. B. Sanders, Philadelphia, PA 1972.

Parzen, E., *Stochastic Processes*, Chapter 1, Holden Day Inc., San Francisco, CA 1962.

Feller, W., *An Introduction to Probability Theory and Its Applications*, vol. 2, John Wiley & Sons, New York 1966.

Grenander, U. and Rosenblatt, M., *Statistical Analysis of Stationary Time Series*, John Wiley & Sons, New York 1957

Parzen, E., *Modern Probability Theory and Its Applications* , John Wiley & Sons, New York 1960.

Parzen, E., *Stochastic Processes*, Holden-Day, San Francisco, CA 1965.

Monzingo, R. A. and Miller, T. W., *Introduction to Adaptive Arrays*, John Wiley & Sons, New York 1980.

McCarty, R. C., *A Sample Estimator of Beam-Forming Performance and Its Distribution Function*, ONR Report ESL-87-0317, ESL Inc., Sunnyvale, CA 1988.