

*SACLANT UNDERSEA
RESEARCH CENTRE
REPORT*



TOEPLITZ CONSTRAINED ADAPTIVE
BEAMFORMING WITH AUXILIARY DATA

D. A. Abraham

February 1997

The SACLANT Undersea Research Centre provides the Supreme Allied Commander Atlantic (SACLANT) with scientific and technical assistance under the terms of its NATO charter, which entered into force on 1 February 1963. Without prejudice to this main task – and under the policy direction of SACLANT – the Centre also renders scientific and technical assistance to the individual NATO nations.

This document is approved for public release.
Distribution is unlimited

SACLANT Undersea Research Centre
Viale San Bartolomeo 400
19138 San Bartolomeo (SP), Italy

tel: +39-187-540.111
fax: +39-187-524.600

e-mail: library@saclantc.nato.int

NORTH ATLANTIC TREATY ORGANIZATION

Toeplitz constrained adaptive
beamforming with auxiliary data

D. A. Abraham

The content of this document pertains to work performed under Project 041-3 of the SACLANTCEN Programme of Work. The document has been approved for release by The Director, SACLANTCEN.



Jan L. Spoelstra
Director

intentionally blank page

**Toeplitz constrained adaptive
beamforming with auxiliary data**

D. A. Abraham

Executive Summary: Active sonar detection of submarines and mines is often hindered by interfering signals such as reverberation from geological features on the sea-bottom or noises from nearby surface ships. When these interfering signals are near the target in bearing and range, adaptive beamforming may be used to form nulls in their directions, thereby increasing the beam output signal-to-interference ratio. This, however, requires estimation of the spatial covariance structure of the interfering signals at the receiving array. If the spatial covariance structure of the target signal is also estimated, it may be used to provide high angular resolution in the vicinity of the target which can improve target tracking performance in cluttered environments.

Estimation of the target and interference spatial covariance structures is a difficult problem owing to the limited amount of available data. This report proposes a method that exploits both target-plus-interference data and interference-only (auxiliary) data and enforces a particular structure on the spatial covariance. It is shown that when the actual spatial covariance has a Toeplitz structure and it is enforced by the estimation algorithm, the spatial covariance can be estimated well enough to implement adaptive beamforming.

intentionally blank page

**Toeplitz constrained adaptive
beamforming with auxiliary data**

D. A. Abraham

Abstract: Submarine and mine detection using active sonar face similar problems when adaptive beamforming is required to provide high angular resolution or to suppress angularly isolated interferences. Both situations result in a very small amount of target-plus-interference data and a moderate amount of stationary interference-only data. From this data the target-plus-interference and interference-only spatial covariance matrices (SCM's) must be estimated. When the SCM's have a Toeplitz structure, as will occur with an equi-spaced line array and plane wave propagation, beamforming performance may be improved by constraining the estimates to be Toeplitz as well.

In this report, the expectation-maximization algorithm is used to obtain the Toeplitz constrained maximum likelihood estimates of the SCM's when target-plus-interference data and interference-only (auxiliary) data are available. When the target-plus-interference SCM estimate is used in an MVDR beamformer, it is shown that the beam output SNR is non-linearly dependent on the maximum beam output SNR (S_{\max}) in such a manner that when S_{\max} is large, accurate estimation of the SCM is crucial. Simulation analysis illustrates that the proposed algorithm can provide the necessary accuracy, with further improvement as the amount of auxiliary data increases.

Keywords: adaptive beamforming ◦ Toeplitz covariance ◦ expectation maximization

Contents

1	Introduction	1
2	Algorithm	4
2.1	Problem definition	4
2.2	Algorithm derivation	5
2.3	Summary of algorithm	9
3	Initialization, convergence and estimation performance analysis	11
3.1	Algorithm initialization	11
3.2	Convergence analysis	13
3.3	Estimation performance analysis	14
4	Beamforming performance analysis	18
4.1	Beam output SNR for Gaussian target model	18
4.2	Beam output SNR with sample covariance matrix SCM estimator	21
4.3	Beam output SNR with Toeplitz constrained SCM estimator	21
4.4	Array gain improvement	24
4.5	Plane wave interference in spatially white noise	26
5	Conclusions	30
	References	31
	Annex A - Forming the diagonal elements of $\mathbf{W}_N \mathbf{R} \mathbf{W}_N^H$	33
	Annex B - Frobenius norm of error between sample and true covariance matrix	34

1

Introduction

Submarine and mine detection using active sonar face similar problems when adaptive beamforming is required to provide high angular resolution or to suppress angularly isolated interferences. In both applications the array sensor data characteristically have a time-varying interference background (i.e., reverberation and any moving interfering sources) and a target echo with a relatively short duration. Both of these characteristics lead to the requirement for accurate estimation of the spatial covariance matrix (SCM) of the array hydrophones using minimal data. To illustrate this problem, consider a continuous-wave (CW) transmit waveform with duration 0.2 seconds and a 150 m long target. The resulting target echo would have approximately 2 independent samples at each sensor which must be used to estimate the N -by- N target-plus-interference SCM when there are N hydrophones in the array. Clearly this is a formidable problem, as the sample covariance matrix in this case will not even provide a full rank estimate for $N > 2$. Estimation of the interference-only SCM is easier as the time over which one can assume stationarity is typically longer than the duration of a target echo. However, use of the interference-only SCM in adaptive beamforming will not result in high angular resolution of the target echo.

Often the full power of adaptively choosing all N of the beamforming weights is not required; that is, using fewer than all of the N spatial degrees-of-freedom (DOF) does not result in a substantial degradation in the average beamforming performance [1]. Reducing the spatial DOF also has the advantage of making the SCM estimation problem easier, resulting in improved detection performance [2]. The methods used to reduce the spatial DOF may be generally described as either nonparametric, in the sense of not assuming a particular form or structure for the SCM, or parametric, where a specific structure is assumed and exploited. The nonparametric techniques include beam-space and sub-array preprocessing where conventional beamforming (CBF) is used as a preprocessor to form $M < N$ beam outputs which are then combined using adaptive beamforming. The parametric techniques include the enhanced minimum variance distortionless response (MVDR) beamformer [3] and the exploitation of the Toeplitz structure induced by plane wave propagation for equispaced linear arrays [4, 5, 6].

Of particular interest is the maximum likelihood estimation of a Toeplitz constrained SCM using the expectation-maximization algorithm described by Miller *et al.* in [7] and applied to adaptive beamforming and detection by Fuhrmann [6]. This para-

metric adaptive beamforming technique assumes that the frequency domain array sensor data are independent multivariate complex Gaussian distributed random vectors with a Toeplitz covariance matrix,

$$\mathbf{x}_i \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{R}) \quad (1)$$

for $i = 1, \dots, L$ where $\mathbf{0}$ is a vector of zeros and the notation $\mathbf{x} \sim \mathcal{CN}_N(\mu, \Sigma)$ indicates that the dimension N random vector \mathbf{x} has a multivariate complex Gaussian distribution with mean vector μ and covariance matrix Σ . The joint probability density function (PDF) of the data $\mathbf{X} = [\mathbf{x}_1 \ \dots \ \mathbf{x}_L]$ is

$$\begin{aligned} f(\mathbf{X}; \mathbf{R}) &= \frac{1}{\pi^{NL} |\mathbf{R}|^L} \exp \left\{ - \sum_{i=1}^L \mathbf{x}_i^H \mathbf{R}^{-1} \mathbf{x}_i \right\} \\ &= \frac{1}{\pi^{NL} |\mathbf{R}|^L} \exp \left\{ -\text{tr}(\mathbf{R}^{-1} \mathbf{X} \mathbf{X}^H) \right\}, \end{aligned} \quad (2)$$

where $\text{tr}(\mathbf{A})$ is the trace (sum of the diagonal values) of the square matrix \mathbf{A} and the superscript H denotes the complex conjugate and transpose operation. Fuhrmann [6] illustrated that the Toeplitz constraint provided substantial improvement over using the unconstrained sample covariance matrix in adaptive beamforming and detection. Further, the algorithm provides full rank estimates of the covariance matrix with just one data vector — the sample covariance matrix technique requires N samples for this. These qualities imply that this type of algorithm may be successful in adaptive beamforming for active sonar.

The array data from active sonar consists of a target echo superimposed on interfering signals composed of reverberation and background noises. The interfering signals are rarely spatially white and are certainly time-varying; however, they may be assumed stationary over short periods of time. Thus, auxiliary data from times near to the target echo may be used to help estimate the the target-plus-interference and interference-only SCM's.

If the amplitude of the target echo is assumed to be a zero mean Gaussian random variable (i.e., Swerling I or II target which is also known as a Gaussian target model [8]), the resulting frequency domain array data for the target-plus-interference SCM will follow eq. (1) with

$$\mathbf{R} = s \mathbf{d} \mathbf{d}^H + \mathbf{Q} \quad (3)$$

$$= \mathbf{P} + \mathbf{Q} \quad (4)$$

where \mathbf{d} is the array direction vector pointing to the direction of the target, s is the power of the target echo and \mathbf{Q} is the interference-only SCM. The auxiliary data are assumed to be independent and identically distributed according to

$$\mathbf{y}_j \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{Q}) \quad (5)$$

for $j = 1, \dots, K$. Similar to eq. (2), the joint PDF of the auxiliary data $\mathbf{Y} = [\mathbf{y}_1 \ \dots \ \mathbf{y}_K]$ is

$$\begin{aligned} f(\mathbf{Y}; \mathbf{Q}) &= \frac{1}{\pi^{NK} |\mathbf{Q}|^K} \exp \left\{ - \sum_{j=1}^K \mathbf{y}_j^H \mathbf{Q}^{-1} \mathbf{y}_j \right\} \\ &= \frac{1}{\pi^{NK} |\mathbf{Q}|^K} \exp \left\{ - \text{tr} \left(\mathbf{Q}^{-1} \mathbf{Y} \mathbf{Y}^H \right) \right\}. \end{aligned} \quad (6)$$

If the array hydrophones are equally spaced in a straight line and the acoustic wave is assumed to propagate as a plane, the array steering vector has the form

$$\mathbf{d}(\gamma) = [1 \ e^{-j\gamma} \ e^{-j2\gamma} \ \dots \ e^{-j(N-1)\gamma}]^T \quad (7)$$

where the superscript T denotes the transpose operation and γ is the electrical angle

$$\begin{aligned} \gamma &= \pi \sin(\theta) \left[\frac{d}{\lambda} \right] \\ &= \pi \sin(\theta) \left[\frac{f}{f_d} \right]. \end{aligned} \quad (8)$$

Here, θ is the angle from broadside to the array, d is the sensor spacing, $\lambda = \frac{c}{f}$ is the wavelength at frequency f when the speed of sound is c , and $f_d = \frac{c}{2d}$ is the frequency at which the sensors are spaced every half wavelength.

The steering vector of eq. (7) leads to a Toeplitz SCM if the signals impinging on the array are either point sources uncorrelated from each other, or the result of an incoherent integration over some angular sector. Correlated point sources from different directions will not result in a Toeplitz SCM.

The intent of this report is to extend the Toeplitz constrained maximum likelihood estimator of Miller *et al.* [7] to the case where auxiliary data exists to augment the estimation of the target-plus-interference SCM for the Gaussian target model. The algorithm is derived in Section 2 and subjected to simulation analysis in Sections 3 and 4.

2

Algorithm

2.1 Problem definition

In order to implement adaptive beamforming algorithms such as MVDR or direction finding algorithms such as MUSIC, the SCM must be estimated from the observed data. In beamforming algorithms, the SCM estimate is then used to form spatial FIR filters that are applied to the array data. Assuming that the primary data (\mathbf{X}) and the auxiliary data (\mathbf{Y}) are distributed as described in eqs. (1) and (5), it is desired to form a maximum likelihood estimate (MLE) of the target-plus-interference (\mathbf{R}) and interference-only (\mathbf{Q}) SCM's subject to the following constraints:

- The matrices \mathbf{P} , \mathbf{Q} , and thus $\mathbf{R} = \mathbf{P} + \mathbf{Q}$ are Toeplitz.
- The matrix \mathbf{P} is non-negative definite (i.e., $\mathbf{P} \succeq 0$).
- The matrix \mathbf{Q} is positive definite (i.e., $\mathbf{Q} \succ 0$).

Note that the MLE of \mathbf{R} may be obtained by first forming the MLE's for \mathbf{P} and \mathbf{Q} and that by their nature the SCM's are complex-conjugate symmetric (i.e., $\mathbf{R} = \mathbf{R}^H$).

The Toeplitz matrices \mathbf{P} and \mathbf{Q} , having constant diagonal elements and being complex-conjugate symmetric are completely described by their first columns,

$$\mathbf{p} = [p_0 \ p_1 \ \cdots \ p_{N-1}]^T \quad (9)$$

and

$$\mathbf{q} = [q_0 \ q_1 \ \cdots \ q_{N-1}]^T \quad (10)$$

which may be combined into the parameter vector Θ ,

$$\Theta = \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix}. \quad (11)$$

Combining eqs. (2) and (6) leads to the log-likelihood function of Θ given the combined primary and auxiliary data

$$\begin{aligned} L(\Theta | \mathbf{X}, \mathbf{Y}) &= -L \log |\mathbf{P} + \mathbf{Q}| - \text{tr} \left((\mathbf{P} + \mathbf{Q})^{-1} \mathbf{X} \mathbf{X}^H \right) \\ &\quad - K \log |\mathbf{Q}| - \text{tr} \left(\mathbf{Q}^{-1} \mathbf{Y} \mathbf{Y}^H \right) - N(K + L) \log(\pi). \end{aligned} \quad (12)$$

The desired MLE's for \mathbf{P} and \mathbf{Q} are those matrices that maximize $L(\Theta|\mathbf{X}, \mathbf{Y})$ subject to the constraints previously mentioned.

2.2 Algorithm derivation

Miller *et al.* [7] utilized the expectation-maximization (EM) algorithm [9] to obtain the MLE of a Toeplitz constrained covariance matrix when the data are zero-mean multivariate complex Gaussian distributed. The same technique is applied here to the case when auxiliary data exists.

The EM algorithm [9] is an iterative method for obtaining a maximum likelihood estimate of Θ when the maximization of the likelihood function (eq. (12)) given the observed data $Z_o = [\mathbf{X}, \mathbf{Y}]$ is not straightforward and the data may be described as part of a larger observation with a simpler likelihood function. Following Miller *et al.* [7], assume that the observed data are part of a complete data set with the form

$$\mathbf{X}_c = \begin{bmatrix} \mathbf{X} \\ \mathbf{X}_m \end{bmatrix} \quad (13)$$

and

$$\mathbf{Y}_c = \begin{bmatrix} \mathbf{Y} \\ \mathbf{Y}_m \end{bmatrix}, \quad (14)$$

where the data $Z_m = [\mathbf{X}_m, \mathbf{Y}_m]$ are considered to be missing or not observable. The marginal PDF of the observed data must correspond to that described by eqs. (2) and (6), so the choice of distribution for the complete data is somewhat restricted. Suppose that each column of \mathbf{X}_c is zero-mean complex Gaussian distributed with covariance matrix $\tilde{\mathbf{R}}$ and that each column of \mathbf{Y}_c is similarly distributed with covariance matrix $\tilde{\mathbf{Q}}$. In this manner, the marginal distributions will be exactly eqs. (2) and (6) if \mathbf{R} and \mathbf{Q} are the primary N -by- N submatrices of $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{Q}}$, respectively.

Further suppose that $\tilde{\mathbf{P}} = \tilde{\mathbf{R}} - \tilde{\mathbf{Q}}$ and $\tilde{\mathbf{Q}}$ are M -by- M circulant matrices. A circulant matrix [10] is a Toeplitz matrix with the special form where each column (row) contains the same elements as the previous one shifted down (to the right) one space; for example,

$$\tilde{\mathbf{P}} = \begin{bmatrix} p_0 & p_{M-1} & p_{M-2} & & p_2 & p_1 \\ p_1 & p_0 & p_{M-1} & \cdots & p_3 & p_2 \\ p_2 & p_1 & p_0 & & & p_3 \\ & \vdots & & \ddots & & \vdots \\ p_{M-2} & p_{M-3} & & & p_0 & p_{M-1} \\ p_{M-1} & p_{M-2} & p_{M-3} & \cdots & p_1 & p_0 \end{bmatrix}. \quad (15)$$

If \mathbf{P} and \mathbf{Q} are to respectively form the primary N -by- N submatrices of $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{Q}}$, then clearly $M \geq 2N - 1$. The convenience of the circulant extension of \mathbf{P} and \mathbf{Q} lies in the straightforward diagonalization of $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{Q}}$. The eigenvectors of a circulant matrix are exactly the discrete Fourier transform (DFT) vectors,

$$\mathbf{w}_i = \frac{1}{\sqrt{M}} \left[1 \quad e^{j2\pi \frac{i}{M}} \quad e^{j4\pi \frac{i}{M}} \quad \dots \quad e^{j2(M-1)\pi \frac{i}{M}} \right]^T \quad (16)$$

for $i = 0, \dots, M - 1$. Thus, $\tilde{\mathbf{P}}$, $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{R}}$ are diagonalized by the matrix of DFT vectors $\mathbf{W} = [\mathbf{w}_0 \quad \dots \quad \mathbf{w}_{M-1}]$,

$$\Sigma_p = \mathbf{W}\tilde{\mathbf{P}}\mathbf{W}^H = \text{diag} \{ \sigma_p(j) \}_{j=1}^M, \quad (17)$$

$$\Sigma_q = \mathbf{W}\tilde{\mathbf{Q}}\mathbf{W}^H = \text{diag} \{ \sigma_q(j) \}_{j=1}^M \quad (18)$$

and

$$\Sigma_r = \mathbf{W}\tilde{\mathbf{R}}\mathbf{W}^H. \quad (19)$$

Restriction of the complete data to having non-negative definite circulant covariance matrices is more strongly justified by the results of Fuhrmann and Miller [11], indicating that this restriction guarantees a positive definite maximum likelihood estimate of the covariance matrix of the observed data.

The expectation step of the EM algorithm [12] requires forming the expectation of the log-likelihood function of Θ over the missing data Z_m conditioned on the observed data Z_o and an assumed set of values for the unknown parameters, say Θ_i ,

$$\begin{aligned} Q(\Theta, \Theta_i) &= E_{Z_m | \Theta_i, Z_o} [\log \{ l(\Theta | Z_o, Z_m) \}] \\ &= \int_{Z_m} L(\Theta | Z_o, Z_m) f(Z_m | \Theta_i, Z_o) dZ_m. \end{aligned} \quad (20)$$

The log-likelihood function of Θ given the complete data is

$$\begin{aligned} L(\Theta | Z_o, Z_m) &= -L \log |\tilde{\mathbf{P}} + \tilde{\mathbf{Q}}| - \text{tr} \left((\tilde{\mathbf{P}} + \tilde{\mathbf{Q}})^{-1} \mathbf{X}_c \mathbf{X}_c^H \right) \\ &\quad - K \log |\tilde{\mathbf{Q}}| - \text{tr} \left(\tilde{\mathbf{Q}}^{-1} \mathbf{Y}_c \mathbf{Y}_c^H \right) - M(K + L) \log(\pi) \end{aligned} \quad (21)$$

where the parameters from Θ (not Θ_i) are used to form $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{Q}}$. As noted in eq. (20), the expectation of $L(\Theta | Z_o, Z_m)$ is performed over the PDF of the missing data conditioned on the observed data. From eq. (21) it can be seen that this will result in the expected value of the outer products of the columns of \mathbf{X}_c and \mathbf{Y}_c . Let the k^{th} column of \mathbf{X}_c be

$$\mathbf{x}_{c,k} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{x}_{m,k} \end{bmatrix} \sim \mathcal{CN}_M(\mathbf{0}, \tilde{\mathbf{R}}_i) \quad (22)$$

where $\tilde{\mathbf{R}}_i$ is formed from the elements of Θ_i . The distribution of $\mathbf{x}_{c,k}$ conditioned on the observed value of \mathbf{x}_k is complex multivariate Gaussian [13],

$$\mathbf{x}_{c,k} | \mathbf{x}_k \sim \mathcal{CN}_M \left(\Gamma_{co} \mathbf{R}_i^{-1} \mathbf{x}_k, \tilde{\mathbf{R}}_i - \Gamma_{co} \mathbf{R}_i^{-1} \Gamma_{co}^H \right) \quad (23)$$

where Γ_{co} represents the correlation between $\mathbf{x}_{c,k}$ and \mathbf{x}_k . The observed data may be isolated from the complete data by the matrix operation

$$\mathbf{x}_k = \mathbf{U}_N \mathbf{x}_{c,k} \quad (24)$$

where premultiplying by $\mathbf{U}_N = [\mathbf{I}_N \ \mathbf{0}]$ picks off the first N elements of $\mathbf{x}_{c,k}$ (\mathbf{I}_N is the N -by- N identity matrix and $\mathbf{0}$ is a matrix of zeros). The correlation matrix Γ_{co} may then be described as

$$\begin{aligned} \Gamma_{co} &= \mathbb{E} \left[\mathbf{x}_{c,k} \mathbf{x}_k^H \right] \\ &= \mathbb{E} \left[\mathbf{x}_{c,k} \mathbf{x}_{c,k}^H \right] \mathbf{U}_N^H \\ &= \tilde{\mathbf{R}}_i \mathbf{U}_N^H. \end{aligned} \quad (25)$$

Substituting eq. (25) into eq. (23) and forming the conditional expectation of the outer product of $\mathbf{x}_{c,k}$ results in

$$\begin{aligned} \mathbb{E} \left[\mathbf{x}_{c,k} \mathbf{x}_{c,k}^H | \mathbf{x}_k \right] &= \tilde{\mathbf{R}}_i - \Gamma_{co} \mathbf{R}_i^{-1} \Gamma_{co}^H + \Gamma_{co} \mathbf{R}_i^{-1} \mathbf{x}_k \mathbf{x}_k^H \mathbf{R}_i^{-1} \Gamma_{co} \\ &= \tilde{\mathbf{R}}_i - \tilde{\mathbf{R}}_i \mathbf{U}_N^H \mathbf{R}_i^{-1} \mathbf{U}_N \tilde{\mathbf{R}}_i + \tilde{\mathbf{R}}_i \mathbf{U}_N^H \mathbf{R}_i^{-1} \mathbf{x}_k \mathbf{x}_k^H \mathbf{R}_i^{-1} \mathbf{U}_N \tilde{\mathbf{R}}_i. \end{aligned} \quad (26)$$

Summing eq. (26) over k and substituting $\tilde{\mathbf{R}}_i = \mathbf{W}^H \Sigma_{r,i} \mathbf{W}$ results in

$$\begin{aligned} \mathbb{E} \left[\mathbf{X}_c \mathbf{X}_c^H | \mathbf{X} \right] &= L \tilde{\mathbf{R}}_i - L \tilde{\mathbf{R}}_i \mathbf{U}_N^H \mathbf{R}_i^{-1} \mathbf{U}_N \tilde{\mathbf{R}}_i + L \tilde{\mathbf{R}}_i \mathbf{U}_N^H \mathbf{R}_i^{-1} \mathbf{S}_x \mathbf{R}_i^{-1} \mathbf{U}_N \tilde{\mathbf{R}}_i \\ &= L \mathbf{W}^H \mathbf{C}_i \mathbf{W} \end{aligned} \quad (27)$$

where

$$\mathbf{S}_x = \frac{1}{L} \sum_{i=1}^L \mathbf{x}_i \mathbf{x}_i^H, \quad (28)$$

$$\mathbf{C}_i = \Sigma_{r,i} - \Sigma_{r,i} \mathbf{W}_N \mathbf{R}_i^{-1} \mathbf{W}_N^H \Sigma_{r,i} + \Sigma_{r,i} \mathbf{W}_N \mathbf{R}_i^{-1} \mathbf{S}_x \mathbf{R}_i^{-1} \mathbf{W}_N^H \Sigma_{r,i} \quad (29)$$

and $\mathbf{W}_N = \mathbf{W} \mathbf{U}_N^H$ are the first N columns of the DFT matrix \mathbf{W} . Similar efforts with the auxiliary data lead to

$$\begin{aligned} \mathbb{E} \left[\mathbf{Y}_c \mathbf{Y}_c^H | \mathbf{Y} \right] &= K \tilde{\mathbf{Q}}_i - L \tilde{\mathbf{Q}}_i \mathbf{U}_N^H \mathbf{Q}_i^{-1} \mathbf{U}_N \tilde{\mathbf{Q}}_i + K \tilde{\mathbf{Q}}_i \mathbf{U}_N^H \mathbf{Q}_i^{-1} \mathbf{S}_y \mathbf{Q}_i^{-1} \mathbf{U}_N \tilde{\mathbf{Q}}_i \\ &= K \mathbf{W}^H \mathbf{D}_i \mathbf{W} \end{aligned} \quad (30)$$

where

$$\mathbf{S}_y = \frac{1}{K} \sum_{j=1}^K \mathbf{y}_j \mathbf{y}_j^H \quad (31)$$

and

$$\mathbf{D}_i = \Sigma_{q,i} - \Sigma_{q,i} \mathbf{W}_N \mathbf{Q}_i^{-1} \mathbf{W}_N^H \Sigma_{q,i} + \Sigma_{q,i} \mathbf{W}_N \mathbf{Q}_i^{-1} \mathbf{S}_y \mathbf{Q}_i^{-1} \mathbf{W}_N^H \Sigma_{q,i}. \quad (32)$$

Substitution of eqs. (27) and (30) into eqs. (21) and (20) results in

$$\begin{aligned} Q(\Theta, \Theta_i) &= -L \log |\tilde{\mathbf{P}} + \tilde{\mathbf{Q}}| - L \operatorname{tr} \left((\tilde{\mathbf{P}} + \tilde{\mathbf{Q}})^{-1} \mathbf{W}^H \mathbf{C}_i \mathbf{W} \right) \\ &\quad - K \log |\tilde{\mathbf{Q}}| - K \operatorname{tr} \left(\tilde{\mathbf{Q}}^{-1} \mathbf{W}^H \mathbf{D}_i \mathbf{W} \right) - M(K+L) \log(\pi) \end{aligned} \quad (33)$$

$$\begin{aligned} &= -L \log |\Sigma_p + \Sigma_q| - L \operatorname{tr} \left((\Sigma_p + \Sigma_q)^{-1} \mathbf{C}_i \right) \\ &\quad - K \log |\Sigma_q| - K \operatorname{tr} \left(\Sigma_q^{-1} \mathbf{D}_i \right) - M(K+L) \log(\pi). \end{aligned} \quad (34)$$

If $d_i(j)$ and $c_i(j)$ for $j = 1, \dots, M$ are respectively the diagonal elements of \mathbf{C}_i and \mathbf{D}_i then eq. (34) may be written as

$$\begin{aligned} Q(\Theta, \Theta_i) &= -L \sum_{j=1}^M \left\{ \log [\sigma_p(j) + \sigma_q(j)] + \frac{c_i(j)}{\sigma_p(j) + \sigma_q(j)} \right\} \\ &\quad - K \sum_{j=1}^M \left\{ \log [\sigma_q(j)] + \frac{d_i(j)}{\sigma_q(j)} \right\} - M(K+L) \log(\pi). \end{aligned} \quad (35)$$

The maximization step of the EM algorithm consists of choosing Θ_{i+1} as those values of Θ maximizing $Q(\Theta, \Theta_i)$. Unconstrained maximization of eq. (35) results in the estimators

$$\sigma_p(j) = c_i(j) - d_i(j) \quad (36)$$

and

$$\sigma_q(j) = d_i(j). \quad (37)$$

However, if $\sigma_p(j) \geq 0$ the nonnegative definite constraint on \mathbf{P} is satisfied and if $\sigma_q(j) > 0$ the positive definite constraint on \mathbf{Q} is satisfied. Such a constrained maximization of eq. (35) results in the estimators

$$\sigma_p(j) = [c_i(j) - d_i(j)]^+ \quad (38)$$

and

$$\sigma_q(j) = \frac{(K+L)d_i(j) - L[d_i(j) - c_i(j)]^+}{(K+L)} \quad (39)$$

where x^+ indicates taking the positive part of x ; that is,

$$x^+ = \begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases} \quad (40)$$

The new estimates of \mathbf{P} and \mathbf{Q} may then be formed by

$$\mathbf{P}_{i+1} = \mathbf{W}_N^H \Sigma_{p,i+1} \mathbf{W}_N \quad (41)$$

and

$$\mathbf{Q}_{i+1} = \mathbf{W}_N^H \Sigma_{q,i+1} \mathbf{W}_N \quad (42)$$

where $\Sigma_{p,i+1}$ and $\Sigma_{q,i+1}$ are formed from the new spectral coefficients of eqs. (38) and (39).

2.3 Summary of algorithm

The EM algorithm for obtaining maximum likelihood estimates of the Toeplitz covariance matrices is now summarized:

1. Initialization:

- Choose $M \geq 2N - 1$. Choosing M as a power of two results in some computational savings.
- Set $i = 0$.
- Form the matrix \mathbf{W}_N from the M -by- M DFT matrix \mathbf{W} .
- Obtain \mathbf{S}_x and \mathbf{S}_y from the array data.
- Form initial estimates of $\Sigma_{r,i}$ and $\Sigma_{q,i}$ according to

$$\Sigma_{r,0} = \mathbf{W} \tilde{\mathbf{R}}_0 \mathbf{W}^H \quad (43)$$

and

$$\Sigma_{q,0} = \mathbf{W} \tilde{\mathbf{Q}}_0 \mathbf{W}^H \quad (44)$$

where $\tilde{\mathbf{R}}_0$ and $\tilde{\mathbf{Q}}_0$ are formed from \mathbf{S}_x and \mathbf{S}_y as described in Section 3.1.

2. Estimation step: form \mathbf{C}_i and \mathbf{D}_i according to

$$\mathbf{C}_i = \Sigma_{r,i} - \Sigma_{r,i} \mathbf{W}_N \left(\mathbf{R}_i^{-1} - \mathbf{R}_i^{-1} \mathbf{S}_x \mathbf{R}_i^{-1} \right) \mathbf{W}_N^H \Sigma_{r,i} \quad (45)$$

$$\mathbf{D}_i = \Sigma_{q,i} - \Sigma_{q,i} \mathbf{W}_N \left(\mathbf{Q}_i^{-1} - \mathbf{Q}_i^{-1} \mathbf{S}_y \mathbf{Q}_i^{-1} \right) \mathbf{W}_N^H \Sigma_{q,i} \quad (46)$$

Note that only the elements along the main diagonal of \mathbf{C}_i and \mathbf{D}_i are required. As described by Miller *et al.* in [7], it is possible to obtain the elements of the main diagonal of the matrix $\mathbf{W}_N \mathbf{R} \mathbf{W}_N^H$ when \mathbf{R} is Toeplitz using an M -point DFT. For the convenience of the reader, this is described in Annex A.

3. **Maximization step:** update the diagonal matrices $\Sigma_{p,i+1}$ and $\Sigma_{q,i+1}$ (and thus $\Sigma_{r,i+1}$) according to the unconstrained maximizers of eqs. (36) and (37) or the constrained maximizers of eqs. (38) and (39). Note that if the unconstrained maximizers are used, the constraint should be applied after convergence and the iteration repeated. Update the covariance estimates according to

$$\mathbf{R}_{i+1} = \mathbf{W}_N^H \Sigma_{r,i+1} \mathbf{W}_N \quad (47)$$

and

$$\mathbf{Q}_{i+1} = \mathbf{W}_N^H \Sigma_{q,i+1} \mathbf{W}_N \quad (48)$$

4. Convergence:

- Form convergence criterion such as the squared Frobenius norm of the difference matrices

$$\epsilon_x = \|\mathbf{R}_{i+1} - \mathbf{R}_i\|_F^2 \quad (49)$$

or

$$\epsilon_y = \|\mathbf{Q}_{i+1} - \mathbf{Q}_i\|_F^2 \quad (50)$$

- If not converged (i.e., ϵ_x or ϵ_y is greater than some tolerable value) then set $i = i + 1$ and go to step 2.

The squared Frobenius norm of a matrix [14] is the sum of the squares of each element (magnitude squared for complex matrices) and can be related to the trace of the product

$$\begin{aligned} \|\mathbf{A}\|_F^2 &= \sum_{i=1}^N \sum_{j=1}^N |a_{i,j}|^2 \\ &= \text{tr}(\mathbf{A}\mathbf{A}^H). \end{aligned} \quad (51)$$

3

Initialization, convergence and
estimation performance analysis

There are several issues associated with evaluating the performance of this iterative technique for determining maximum likelihood estimates of the target-only, target-plus-interference and interference-only SCM's. The issues related to initialization of the algorithm are discussed in Section 3.1, convergence of the algorithm in Section 3.2 and performance of the algorithm in estimating the SCM's in Section 3.3. The performance of the SCM estimates in adaptive beamforming is examined separately in Section 4.

In the following simulation analyses, it is assumed that the target arrives broadside to the array ($\gamma_0 = 0$) and has unit power ($s = 1$), and that the interfering signals are spatially white. The array is assumed to be composed of N sensors equally spaced in a straight line, each a distance $d = \frac{\lambda}{2}$ apart (i.e., half-wavelength spacing). Thus, the interference-only SCM is

$$\mathbf{Q} = \mathbf{I}_N, \quad (52)$$

the array steering vector pointing toward the target is

$$\mathbf{d}(\gamma_0) = \mathbf{d} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (53)$$

and the target covariance matrix is

$$\mathbf{P} = \mathbf{d}\mathbf{d}^H. \quad (54)$$

3.1 Algorithm initialization

The first iteration of the EM algorithm described in Section 2.3 requires estimates of the Toeplitz constrained observed data covariances \mathbf{R}_0 and \mathbf{Q}_0 and also their circulant extensions $\tilde{\mathbf{R}}_0$ and $\tilde{\mathbf{Q}}_0$. Recall from spectral estimation the standard biased

and unbiased estimators of the autocorrelation function [15]. Extending this concept to Toeplitz covariance matrix estimation yields the unbiased estimator

$$\hat{r}_n = \frac{1}{N-n} \sum_{k=0}^{N-1-n} \mathbf{S}_x(k, k-n) \quad (55)$$

for $n = 0, 1, \dots, N-1$ where $\mathbf{S}_x(k, k-n)$ are the elements along the n^{th} subdiagonal of \mathbf{S}_x . Scharf [16] shows that this estimator, which simply averages the diagonal elements of \mathbf{S}_x , minimizes the squared error between \mathbf{S}_x and the Toeplitz matrix formed from the estimates $\{\hat{r}_n\}_{n=0}^{N-1}$.

The biased estimator forms the sum of each subdiagonal and divides by N ,

$$\hat{r}_n = \frac{1}{N} \sum_{k=0}^{N-1-n} \mathbf{S}_x(k, k-n). \quad (56)$$

The biased estimator is usually preferred in spectral estimation because the mean of the periodogram power spectral estimate (i.e., the Fourier transform of the correlation estimates) is nonnegative while the unbiased correlation estimates can result in a negative mean.

For a dimension M circulant extension of the Toeplitz SCM $\hat{\mathbf{R}}$ formed from the estimated covariances $\{\hat{r}_n\}_{n=0}^{N-1}$, $2N-1$ of the M parameters are already specified. One could reasonably choose the other parameters to be zeros, resulting in the vector

$$\begin{bmatrix} \hat{r}_0 \\ \hat{r}_1 \\ \vdots \\ \hat{r}_{N-1} \\ 0 \\ \vdots \\ 0 \\ \hat{r}_{N-1} \\ \vdots \\ \hat{r}_1 \end{bmatrix} \quad (57)$$

as the first column of $\tilde{\mathbf{R}}_0$. The algorithm described in Section 2.3 inherently enforces the matrices $\tilde{\mathbf{R}}_i$ and $\tilde{\mathbf{Q}}_i$ to be nonnegative definite. Thus, the algorithm should be initialized with nonnegative definite matrices. If $\tilde{\mathbf{R}}_0$ or $\tilde{\mathbf{Q}}_0$ contain any negative eigenvalues (i.e., the diagonal elements of $\Sigma_{r,0}$ or $\Sigma_{q,0}$), they should be set to zero. It was, however, observed that the algorithm could recover from initializations of $\tilde{\mathbf{R}}_0$ and $\tilde{\mathbf{Q}}_0$ containing negative eigenvalues. Practice has also shown that the unbiased covariance estimators of eq. (55) almost always do not satisfy the nonnegative

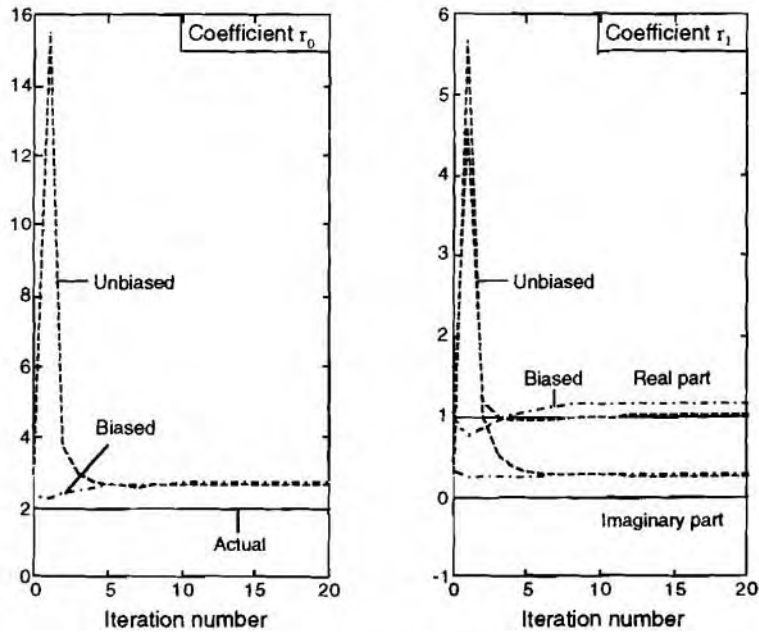


Figure 1 First four coefficients of the estimate of \mathbf{R} when initialization occurs using the unbiased and biased estimators of eqs. (55) and (56).

definite requirement when there is only one data vector (i.e., $L = 1$ or $K = 1$) and that the biased estimators of eq. (56) usually do.

As an example, consider the case where $L = 1$, $K = 4$, $N = 16$ and $M = 32$. Figure 1 depicts one realization of the convergence of the sensor correlations r_0 and r_1 when the biased and unbiased initial estimators are used. Though the unbiased initialization converges to essentially the same values as the biased initialization¹, the latter produces a more stable convergence.

3.2 Convergence analysis

Dempster *et al.* [9] indicate that the convergence of the EM algorithm is slow and related to the proportion of observed to complete data, in this case $\frac{N}{M}$. Thus, convergence is expected to be slower as M increases. To illustrate this, the mean squared Frobenius norm of the difference between consecutive estimates of \mathbf{P} and \mathbf{Q} in the EM iteration calculated from 500 trials is shown in Fig. 2 for the case of $L = 1$, $K = 4$, $N = 16$ and circulant extensions with dimension $M = 2N = 32$,

¹In this individual trial and after many iterations.

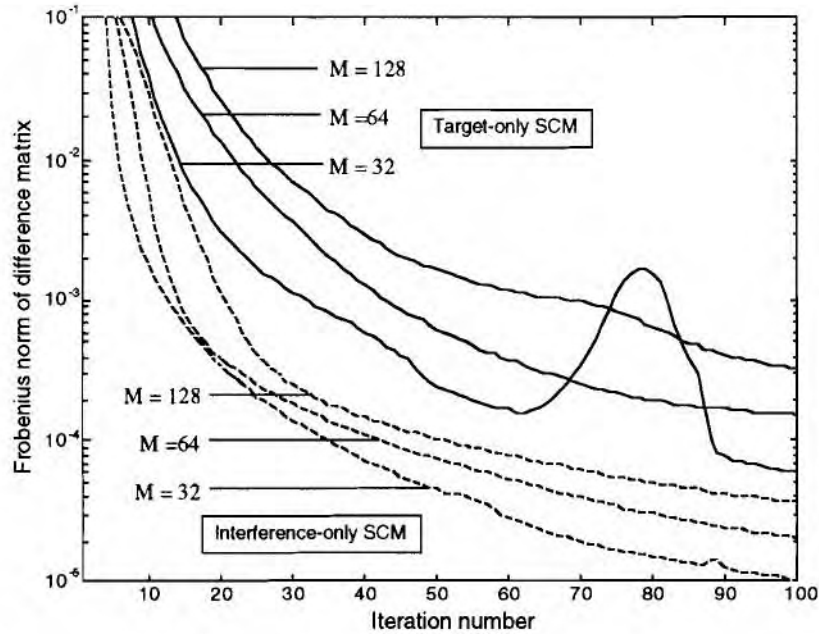


Figure 2 Mean squared Frobenius norm of difference matrix between consecutive estimates of \mathbf{P} and \mathbf{Q} for $M = 2N = 32$, $M = 4N = 64$ and $M = 8N = 128$.

$M = 4N = 64$ and $M = 8N = 128$. Though this measure does not necessarily indicate how close the iteration is to the maximum value of the likelihood function, it is used as a stopping criterion and thus worth investigating.

The results shown in Fig. 2 support the slower rate of convergence for larger M and also indicate that the convergence of the estimate of \mathbf{P} is slower than that of \mathbf{Q} . The latter effect is most likely due to the greater availability of data to estimate \mathbf{Q} (i.e., $K + L$ samples as opposed to L samples to estimate \mathbf{P} and \mathbf{R}). Although it is not shown, the estimate of \mathbf{R} converges at the slower rate between that of \mathbf{P} and \mathbf{Q} .

The faster convergence and reduced computational requirements make choosing a small M desirable. However, though it was not observed in this simulation, it is expected that choosing M too small may adversely affect estimation performance.

3.3 Estimation performance analysis

To evaluate how well the proposed algorithm estimates the matrices \mathbf{R} , \mathbf{Q} and \mathbf{P} , the squared Frobenius norm of the error between the actual and estimated matrix is considered. This is essentially a weighted mean squared error criterion where the

emphasis of the weighting decreases as the separation of the sensors being correlated increases. As shown in Annex B, the expected value of the squared Frobenius norm between the sample covariance matrix (\mathbf{S}_x) and the true covariance matrix (\mathbf{R}) is

$$E \left[\|\mathbf{R} - \mathbf{S}_x\|_F^2 \right] = \frac{[\text{tr}(\mathbf{R})]^2}{L} \quad (58)$$

when L independent samples are used to form \mathbf{S}_x . In this situation (i.e., $\mathbf{Q} = \mathbf{I}_N$, $s = 1$ and \mathbf{d} as in eq. (53)) assuming $N = 16$, $L = 1$ and $K = 4$,

$$\begin{aligned} E \left[\|\mathbf{R} - \mathbf{S}_x\|_F^2 \right] &= \frac{4N^2}{L} \\ &= 1024, \end{aligned} \quad (59)$$

$$\begin{aligned} E \left[\|\mathbf{Q} - \mathbf{S}_y\|_F^2 \right] &= \frac{N^2}{K} \\ &= 64 \end{aligned} \quad (60)$$

and

$$\begin{aligned} E \left[\|\mathbf{P} - (\mathbf{S}_x - \mathbf{S}_y)\|_F^2 \right] &= \frac{4N^2}{L} + \frac{N^2}{K} \\ &= 1088. \end{aligned} \quad (61)$$

Figures 3–5 contain histograms of the base 10 logarithm of the squared Frobenius norm of the error between the actual and estimated covariance matrices, respectively for \mathbf{R} , \mathbf{Q} , and \mathbf{P} . The estimated (and theoretical for the sample covariance matrices) mean values of the squared Frobenius norm are shown on the plots as vertical lines. The Toeplitz constrained MLE of Miller *et al.* [7] is also evaluated to observe the improvement obtained in using auxiliary data. The histograms were formed from 1000 independent trials.

Figure 3 illustrates the expected improvement when the Toeplitz constraint is applied compared to the sample covariance matrix, and minimal difference in estimating \mathbf{R} when auxiliary data is used with the Toeplitz constraint. Increasing the amount of auxiliary data from $K = 4$ to $K = 16$ did not substantially improve the estimation performance in terms of the squared Frobenius norm. This result is consistent with that of Section 3.2 where the convergence of \mathbf{R} was controlled by the dominating effect between \mathbf{P} and \mathbf{Q} .

The utility of the auxiliary data is more obvious in Fig. 4 where a visible improvement is observed over the Toeplitz constrained MLE without auxiliary data. This improvement decreases as the amount of data contributed by \mathbf{S}_x (L samples) to the estimation process becomes small as compared to that contributed by \mathbf{S}_y (K samples). Note that enforcing the Toeplitz constraint reduces the mean squared

Frobenius norm by approximately one order of magnitude compared to using the sample covariance matrix.

One of the primary advantages of using auxiliary data is the ability to estimate the target-only SCM (\mathbf{P}). This estimate may provide valuable information for target detection, localization and classification in active sonar. Figure 5 illustrates the performance of the proposed Toeplitz constrained MLE algorithm in estimating the target-only SCM. Compared to the difference of the sample covariance matrices $\mathbf{S}_x - \mathbf{S}_y$, which does not guarantee a non-negative definite estimate, the proposed estimator provides approximately a factor of 4.5 times smaller mean squared Frobenius norm in the error matrix as well as a substantially reduced variance.

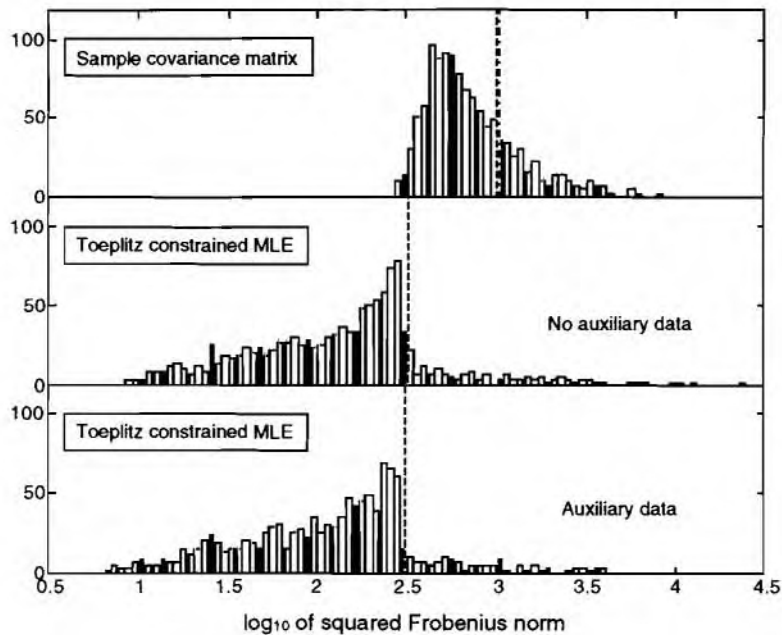


Figure 3 Histogram of squared Frobenius norm of error matrix between actual and estimated \mathbf{R} . Estimated mean values are indicated on each graph.

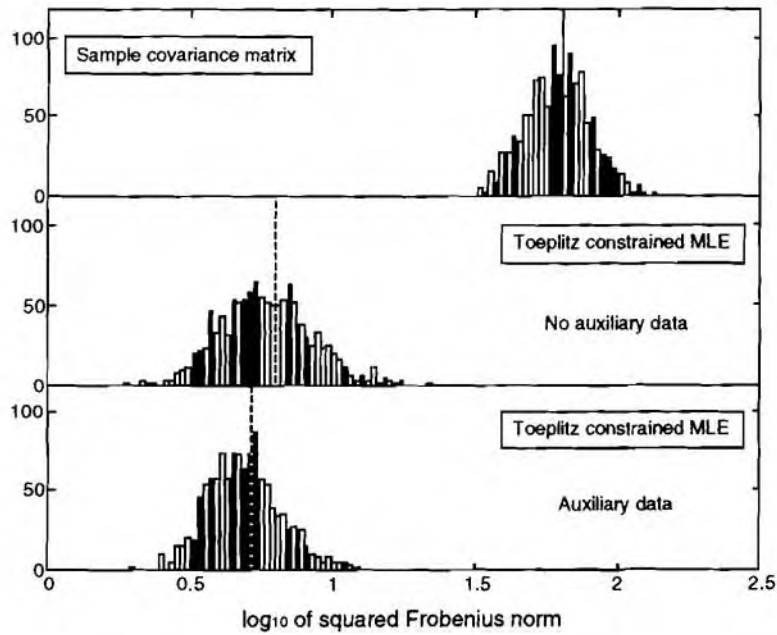


Figure 4 Histogram of squared Frobenius norm of error matrix between actual and estimated Q . Estimated mean values are indicated on each graph.

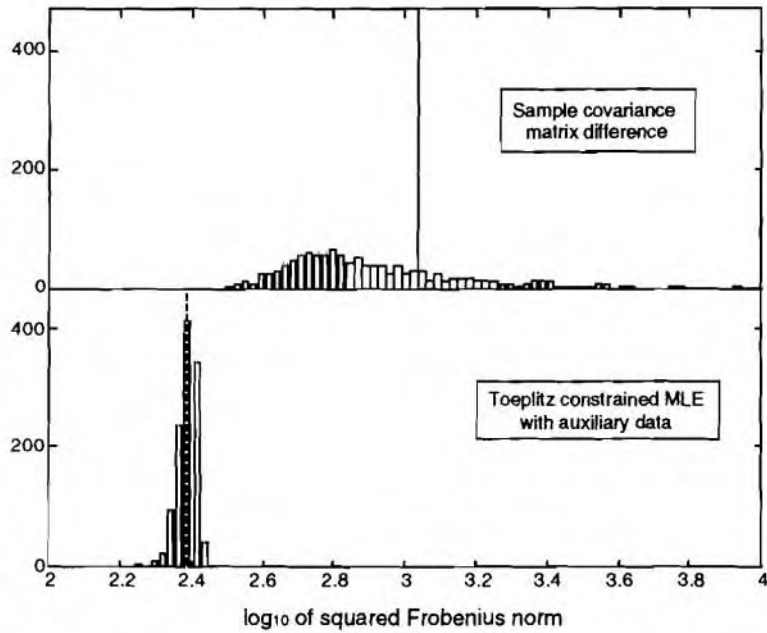


Figure 5 Histogram of squared Frobenius norm of error matrix between actual and estimated P . Estimated mean values are indicated on each graph.

4

Beamforming performance analysis

The primary purpose of adaptive beamforming is to improve the SNR at the output of the beamformer compared to that achieved with the conventional beamformer. The maximum amount of improvement obtainable depends on the structure of the interfering signals. For instance, the improvement is minimal when the interference is spatially diffuse and perhaps far from the target in angle. However, if the interference is spatially compact and near the target, the maximum improvement can be substantial.

When the beamforming filter vector is formed from an estimate of the SCM rather than the exact value, a loss in performance is incurred. Reed, Mallett and Brennan (RMB) [17] have described this loss for the deterministic target model as being a multiplicative factor reducing the beam output SNR from the maximum value. They then showed that if the sample covariance matrix is used to form the beamforming filter vector that the multiplicative factor was beta distributed. In the following sections, the beam output SNR is derived for the Gaussian target model, the array gain improvement (AGI) is described when the adaptive beamforming filter vector is estimated, and the beam output SNR and AGI are evaluated using the Toeplitz constrained SCM estimation technique proposed in this report.

4.1 Beam output SNR for Gaussian target model

Under the Gaussian target model described by eqs. (1) and (3), the beam output signal-to-noise power ratio (SNR) when \mathbf{w} is the beamforming weight vector is

$$\text{SNR} = \frac{\mathbf{w}^H \mathbf{P} \mathbf{w}}{\mathbf{w}^H \mathbf{Q} \mathbf{w}} \quad (62)$$

It is well-known that the MVDR weight vector

$$\mathbf{w}_{\text{MVDR}} = \frac{\mathbf{Q}^{-1} \mathbf{d}}{\mathbf{d}^H \mathbf{Q}^{-1} \mathbf{d}} \quad (63)$$

maximizes the beam output SNR in the direction described by the steering vector \mathbf{d} , resulting in the maximum value

$$S_{\text{max}} = s \mathbf{d}^H \mathbf{Q}^{-1} \mathbf{d}. \quad (64)$$

When $\mathbf{R} = s\mathbf{d}\mathbf{d}^H + \mathbf{Q}$ it is also known that with perfect knowledge of the SCM's, the target-plus-interference SCM may be used rather than the interference-only SCM \mathbf{Q} ; that is,

$$\frac{\mathbf{Q}^{-1}\mathbf{d}}{\mathbf{d}^H\mathbf{Q}^{-1}\mathbf{d}} = \frac{\mathbf{R}^{-1}\mathbf{d}}{\mathbf{d}^H\mathbf{R}^{-1}\mathbf{d}} \quad (65)$$

This is important when other directions than that described by \mathbf{d} are of interest or if only data containing target and interference are observable as in passive sonar. In high resolution beamforming for active sonar, the former applies as the $\mathbf{P} = s\mathbf{d}\mathbf{d}^H$ portion of the SCM is considered to be interference when the look direction is not that described by the steering vector \mathbf{d} .

Substitution of the estimate of the target-plus-interference SCM $\hat{\mathbf{R}}$ in the beamforming weight vector results in

$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{R}}^{-1}\mathbf{d}}{\mathbf{d}^H\hat{\mathbf{R}}^{-1}\mathbf{d}} \quad (66)$$

Given an observed beamforming weight vector $\hat{\mathbf{w}}$, that is assumed to be independent of the array data being beamformed, the observed beam output SNR is then

$$\begin{aligned} S &= \frac{\hat{\mathbf{w}}^H\mathbf{P}\hat{\mathbf{w}}}{\hat{\mathbf{w}}^H\mathbf{Q}\hat{\mathbf{w}}} \\ &= \frac{s}{\hat{\mathbf{w}}^H\mathbf{R}\hat{\mathbf{w}} - \hat{\mathbf{w}}^H\mathbf{P}\hat{\mathbf{w}}} \\ &= \frac{s}{\frac{1}{\rho}\mathbf{d}^H\hat{\mathbf{R}}^{-1}\mathbf{d} - s} \end{aligned} \quad (67)$$

where

$$\rho = \frac{(\mathbf{d}^H\hat{\mathbf{R}}^{-1}\mathbf{d})^2}{(\mathbf{d}^H\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{d})(\mathbf{d}^H\mathbf{R}^{-1}\mathbf{d})} \quad (68)$$

is the SNR loss factor of Reed, Mallett and Brennan [17] which was developed for a deterministic target model as opposed to the Gaussian target model considered in this report.

In simplifying eq. (67), first note that application of the matrix inversion lemma to \mathbf{R}^{-1} results in

$$\frac{\mathbf{d}^H\mathbf{Q}^{-1}\mathbf{d}}{\mathbf{d}^H\mathbf{R}^{-1}\mathbf{d}} = 1 + S_{\max} \quad (69)$$

It then follows that for the Gaussian target model with the MVDR beamforming vector formed from an estimate of the target-plus-interference SCM \mathbf{R} , the beam

output SNR is related to RMB's SNR loss factor and the maximum achievable SNR according to

$$\begin{aligned} S &= \frac{\rho s \mathbf{d}^H \mathbf{Q}^{-1} \mathbf{d}}{\frac{\mathbf{d}^H \mathbf{Q}^{-1} \mathbf{d}}{\mathbf{d}^H \mathbf{R}^{-1} \mathbf{d}} - \rho s \mathbf{d}^H \mathbf{Q}^{-1} \mathbf{d}} \\ &= \frac{\rho S_{\max}}{1 + (1 - \rho) S_{\max}}. \end{aligned} \quad (70)$$

If the MVDR beamforming vector is formed using an estimate of the interference-only SCM, the output SNR would be

$$S_Q = \rho S_{\max} \quad (71)$$

where ρ is formed using \mathbf{Q} and $\hat{\mathbf{Q}}$. Here, the relationship between the beam output SNR and S_{\max} is linear and is identical to the result of Reed, Mallett and Brennan [17] for the deterministic target model. The importance of the result of eq. (70) is that it indicates a non-linear relationship between the beam output SNR and the maximum achievable SNR for the Gaussian target model when the target-plus-interference SCM estimate is used.

For a conventional beamformer without array shading (i.e., $\mathbf{w} = \mathbf{d}$), the beam output SNR is

$$\begin{aligned} S_{\text{cbf}} &= \frac{\mathbf{d}^H \mathbf{P} \mathbf{d}}{\mathbf{d}^H \mathbf{Q} \mathbf{d}} \\ &= \frac{s N^2}{\mathbf{d}^H \mathbf{Q} \mathbf{d}} \\ &= \rho_{\text{cbf}} S_{\max} \end{aligned} \quad (72)$$

where $\mathbf{d}^H \mathbf{d} = N$ and the penalty associated with using conventional rather than optimal beamforming is quantified by

$$\rho_{\text{cbf}} = \frac{N^2}{(\mathbf{d}^H \mathbf{Q} \mathbf{d})(\mathbf{d}^H \mathbf{Q}^{-1} \mathbf{d})}. \quad (73)$$

These results provides a means for evaluating the potential improvement of adaptive beamforming over conventional beamforming, taking into consideration the estimation of the SCM's. For example, consider the case of spatially white noise and interference; that is, $\mathbf{Q} = \sigma^2 \mathbf{I}_N$. The optimality of conventional beamforming in this situation is well-known and supported by the resulting $\rho_{\text{cbf}} = 1$. This indicates that there is no possibility of increasing the beam output SNR.

4.2 Beam output SNR with sample covariance matrix SCM estimator

Reed, Mallett and Brennan [17] showed that if $\hat{\mathbf{R}}$ is complex Wishart distributed (as would be the case if the sample covariance matrix were used), ρ is beta distributed. If $L \geq N$ data vectors are used to form $\hat{\mathbf{R}} = \mathbf{S}_x$, then S has PDF

$$f(s) = \left[\frac{L!(1+S_{\max})^{L+2-N}}{(N-2)!(L+1-N)!S_{\max}} \right] \frac{\left(\frac{s}{S_{\max}}\right)^{L+1-N} \left(1 - \frac{s}{S_{\max}}\right)^{N-2}}{(1+s)^{L+1}} \quad (74)$$

for $s \in [0, S_{\max}]$. It is straightforward to show that the α^{th} moment of S is

$$\begin{aligned} E[S^\alpha] &= S_{\max}^\alpha (1+S_{\max})^{L+2-N} \left[\frac{L!(\alpha+L+1-N)!}{(\alpha+L)!(L+1-N)!} \right] \times \\ &\quad {}_2F_1(L+1, \alpha+L+2-N; \alpha+L+1; -S_{\max}) \end{aligned} \quad (75)$$

where ${}_2F_1(a, b; c; x)$ is the Gauss hypergeometric function [18] (this reference also contains an algorithm for evaluating the function).

As the estimation of \mathbf{R} improves, ρ will tend toward one. Thus, as indicated by eq. (70), the requirement for estimation accuracy increases with the maximum beam output SNR. For example, consider the PDF of S as shown in Fig. 6 for $N = 16$, $S_{\max} = 0.1, 1$ and 10 , and various values of L . As S_{\max} increases a larger amount of data is required to retain the same performance relative to the maximum.

The mean of S normalized by S_{\max} is shown in Fig. 7 for $N = 16$ and S_{\max} varying from 0.1 to 10 as a function of the amount of data used to estimate \mathbf{R} using the sample covariance matrix. As the maximum SNR decreases, the normalized mean tends toward that of the RMB SNR loss factor as would be expected from inspection of eq. (70). These curves also support the conclusion that improved estimation of \mathbf{R} is imperative for high S_{\max} scenarios if the Gaussian target model is assumed.

4.3 Beam output SNR with Toeplitz constrained SCM estimator

The previous section has illustrated the importance of improving the estimation of the target-plus-interference SCM for the Gaussian target model. This section considers the improvement obtainable by using the Toeplitz constrained maximum likelihood SCM estimators of Miller *et al.* [7] when no auxiliary data exists and the estimator described in Section 2 when auxiliary data does exist.

The PDF of the beam output SNR is estimated by histograms formed from 1000 trials in Fig. 8 for the case of $N = 16$, $L = 16$, $K = 16$, $\mathbf{Q} = \mathbf{I}_N$, and $\mathbf{P} = \mathbf{d}\mathbf{d}^H$. The theoretical PDF described in eq. (74) for the sample covariance matrix estimation of \mathbf{R} is also shown on the figure. The improvement due to application of the Toeplitz

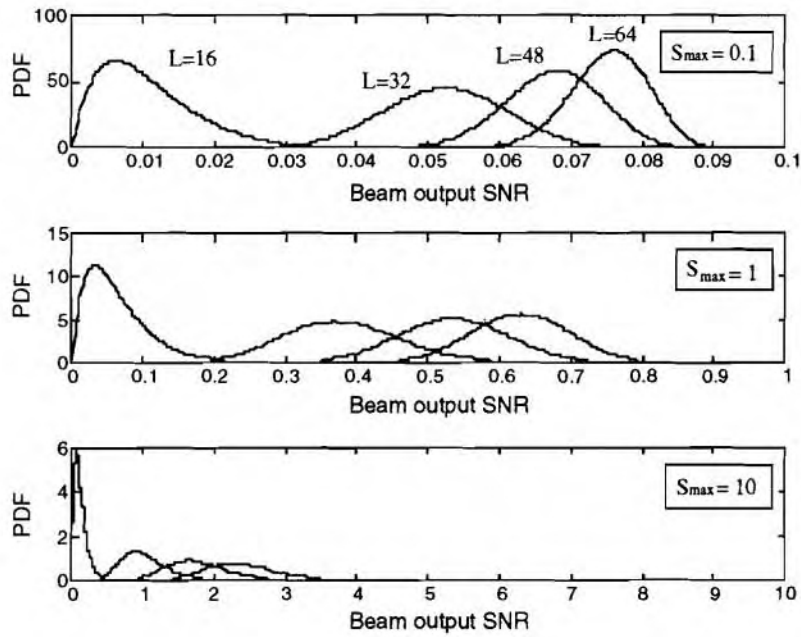


Figure 6 PDF of beam output SNR for $N = 16$, $L = 16, 32, 48$, and 64 , and $S_{\max} = 0.1, 1$, and 10 .

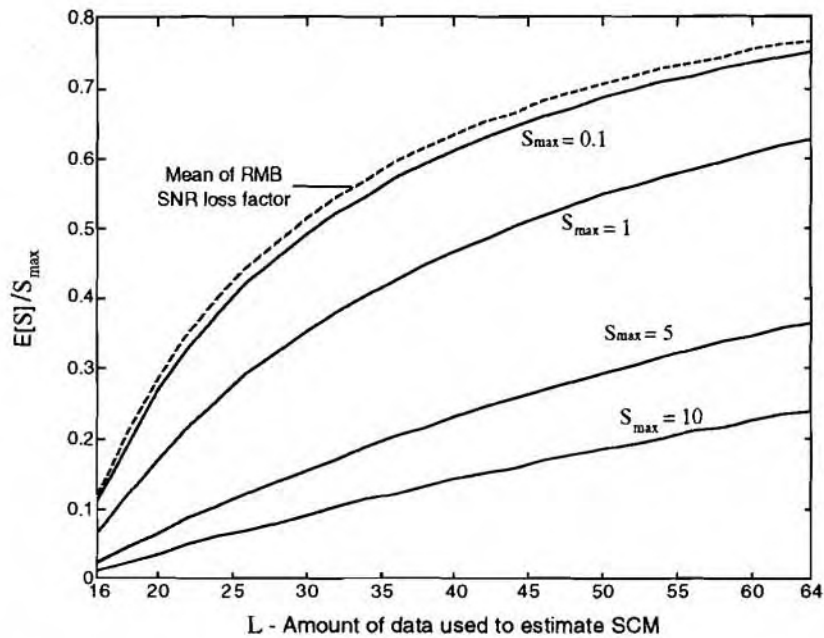


Figure 7 Mean of beam output SNR normalized by S_{\max} for $N = 16$ and $S_{\max} = 0.1, 1, 5$ and 10 as a function of L .

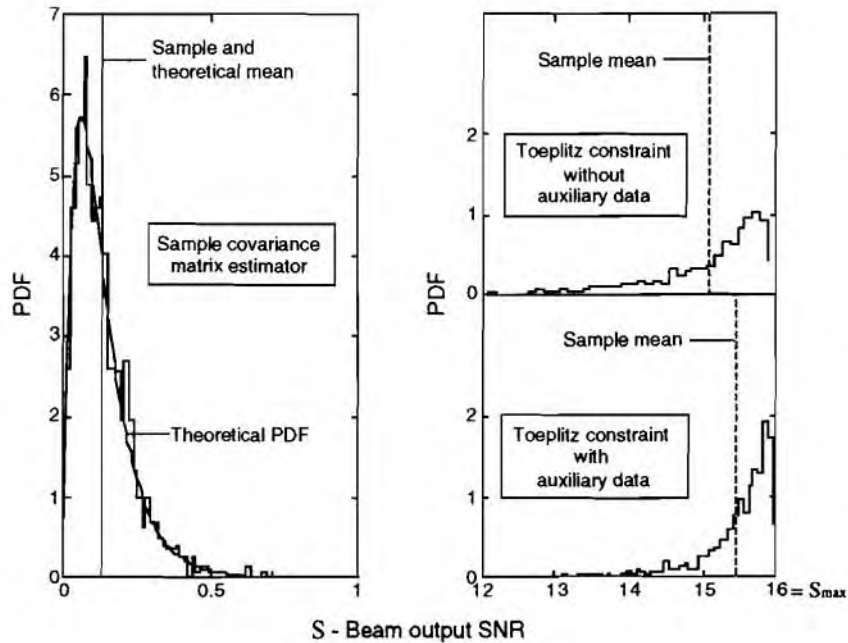


Figure 8 Histograms of beam output SNR for $N = 16$, $L = 16$, and $K = 16$.

constraint is clear in that most of the weight of the PDF's are between $S = 14$ and $S = S_{\max} = 16$ rather than in the region $S < 1$ as occurs with the sample covariance matrix. The mean values of the beam output SNR, as noted in Table 1 and in Fig. 8 as vertical lines, reflect approximately 20 dB of improvement in applying the Toeplitz constraint.

The improvement observed in the previous example when auxiliary data is used is minimal, most likely because the Toeplitz constrained estimation of \mathbf{R} using $L = 16$ samples of target-plus-interference data and no auxiliary data is quite good to begin with. In active sonar, however, there is likely to be only one or two data vectors available to estimate \mathbf{R} . Figure 9 contains histograms and boxplots of the RMB SNR loss factor, as defined in eq. (68), formed from 1000 trials when $N = 16$ and $L = 1$ as a function of K . The $K = 0$ case represents the algorithm of Miller *et al.* [7]; that is, no auxiliary data. The $K > 0$ cases represent the algorithm of Section 2. The boxplots [19] describe with lines the lower quartile, median, and upper quartile of the data. Additionally, the sample mean has been plotted with the + symbol. Observe that most of the improvement is obtained by the first 3–5 samples of auxiliary data. The expected value of the ratio of the RMB SNR loss factor with auxiliary data to without,

$$\Delta_{\text{dB}}^{\rho} = 10 \log_{10} \left\{ \mathbb{E} \left[\frac{\rho(K)}{\rho(K=0)} \right] \right\}, \quad (76)$$

Table 1 *Estimated and theoretical average beam output SNR for $N = 16$, $L = 16$, $K = 16$.*

Description	E [S]
Sample covariance matrix, theoretical value	0.1327
Sample covariance matrix, estimated value	0.13
Toeplitz constraint without auxiliary data	15.07
Toeplitz constraint with auxiliary data	15.43

increases from 0.24 dB when $K = 1$ to 0.43 dB when $K = 15$ samples of auxiliary data are used. Figure 10 contains histograms and boxplots of the beam output SNR, as described by eq. (70). Here the improvement is more notable. The expected value of the ratio of the beam output SNR with auxiliary data to without,

$$\Delta_{\text{dB}}^S = 10 \log_{10} \left\{ E \left[\frac{S(K)}{S(K=0)} \right] \right\}, \quad (77)$$

increases from 2.09 dB when $K = 1$ to 4.06 dB when $K = 15$ samples of auxiliary data are used.

4.4 Array gain improvement

The array gain improvement (AGI), as defined by Owsley in [20] or [1], is the ratio of the array gain for an adaptive beamformer to that for the conventional beamformer. As the input SNR is the same for each beamformer, it is equivalently the ratio of the beam output SNR's. Taking into account the randomness associated with the estimation of the SCM's, the AGI is

$$\begin{aligned} \text{AGI} &= \frac{S}{S_{\text{cbf}}} \\ &= \frac{\rho}{\rho_{\text{cbf}} [1 + (1 - \rho) S_{\text{max}}]}. \end{aligned} \quad (78)$$

Now consider the case of a single plane wave interfering source in spatially white noise,

$$\mathbf{Q} = s_i \mathbf{d}_i \mathbf{d}_i^H + \sigma^2 \mathbf{I}_N. \quad (79)$$

Some simple matrix algebra yields

$$\rho_{\text{cbf}} = \frac{1 + r_i}{(1 + r_i b) (1 + (1 - b) r_i)} \quad (80)$$

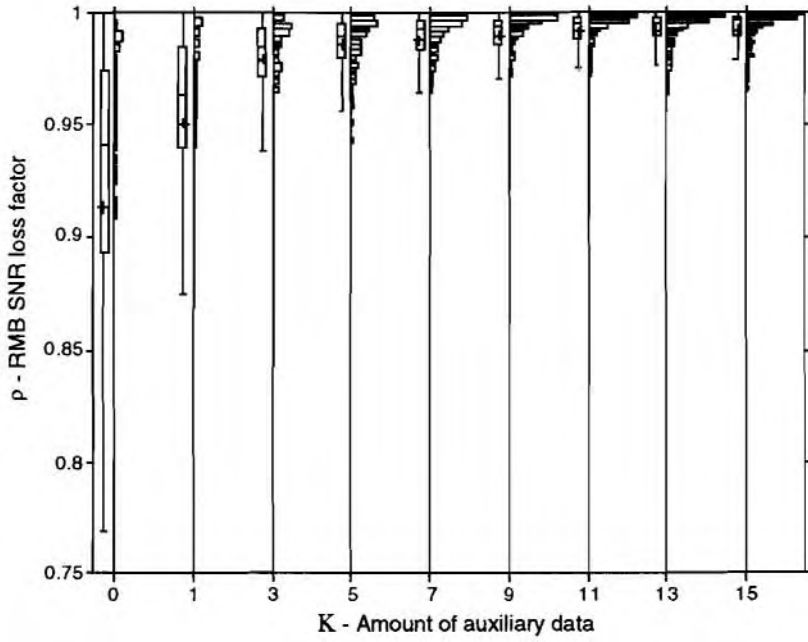


Figure 9 Histograms of RMB SNR loss factor for $N = 16$ and $L = 1$ as a function of K .

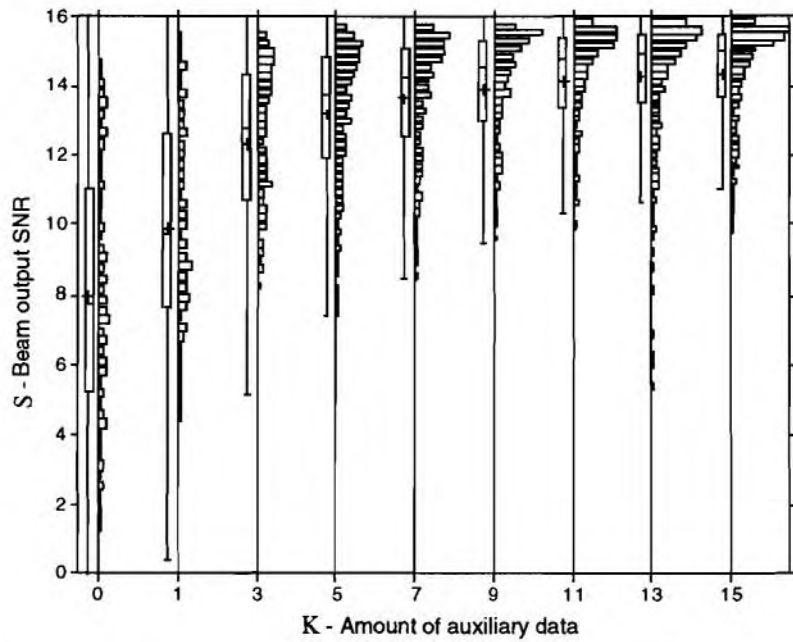


Figure 10 Histograms of beam output SNR for $N = 16$ and $L = 1$ as a function of K .

where

$$b = \frac{|\mathbf{d}^H \mathbf{d}_i|^2}{N^2} \quad (81)$$

represents the capability of the conventional beamformer to suppress the interference, and

$$r_i = \frac{N s_i}{\sigma^2} \quad (82)$$

represents the interference-to-noise ratio at the output of a conventional beamformer (i.e., gain of N). The parameter b provides some information about how far away the interference is from the direction of interest. If b is near one, the interference is in the main lobe of the beampattern of a conventional beamformer pointing to \mathbf{d} . If b is near zero, the interference is either far from the direction of interest or near a null in the beampattern of a conventional beamformer pointing to \mathbf{d} . When the SCM is known exactly (i.e., $\hat{\mathbf{R}} = \mathbf{R}$), $\rho = 1$ and the AGI of eq. (78) for the single plane wave interference simplifies to that described in [1].

The variation of ρ_{cbf} with r_i and b is illustrated in Fig. 11 for a single plane wave interference in white noise within the main lobe of the conventional beampattern pointing at the target. As the interference approaches the target ($b \rightarrow 1$) and as the interference becomes weaker compared to the noise ($r_i \rightarrow 0$), ρ_{cbf} increases toward one. This supports the common knowledge that adaptive beamforming can provide improvement when there is a strong interfering signal near the target.

The PDF for the array gain improvement when the sample covariance matrix is used to form the MVDR beamforming filter vector may be obtained by noticing from eqs. (70) and (78) that

$$\text{AGI} = \frac{S}{\rho_{\text{cbf}} S_{\text{max}}} \quad (83)$$

Thus, the PDF of the AGI for adaptive beamforming using the sample covariance matrix may be found from eq. (74) by a simple scale transformation.

4.5 Plane wave interference in spatially white noise

All of the previous simulation analyses have been for a spatially white interference-only SCM. For this interference structure the conventional beamformer is optimal. Now consider plane wave interference in spatially white noise (i.e., \mathbf{Q} as in eq. (79)) where $r_i = 10$ dB and $\sigma^2 = 1$ (i.e., $s_i = \frac{r_i \sigma^2}{N} = \frac{10}{N}$) for interferences in the main lobe so that $b = 0.95, 0.9, 0.8, \text{ and } 0.5$.

Histograms and boxplots for the AGI are found when $N = 16$, $L = 1$ and $K = 0$ (no auxiliary data) in Fig. 12 and $K = 8$ (auxiliary data) in Fig. 13. The label on the abscissa indicates the value of b and the maximum possible AGI ($\frac{1}{\rho_{cbf}}$) is indicated on each plot by an *. In all cases, the availability of auxiliary data has improved the performance. As seen in Fig. 12 and in Table 2, the AGI and its mean may be below one, in which case the losses associated with estimating \mathbf{R} are greater than the gain in SNR and conventional beamforming should be used.

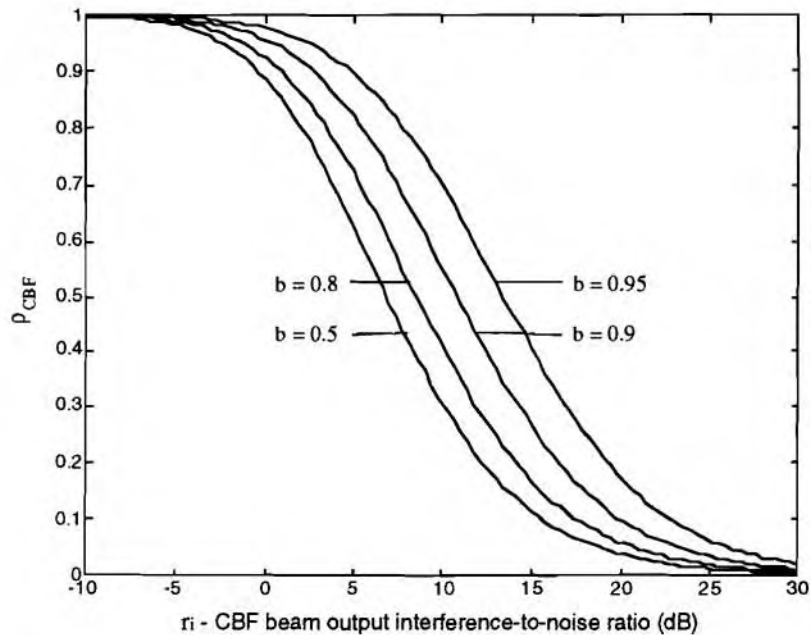


Figure 11 ρ_{cbf} for single plane wave interference in spatially white noise as a function of r_i for various values of b .

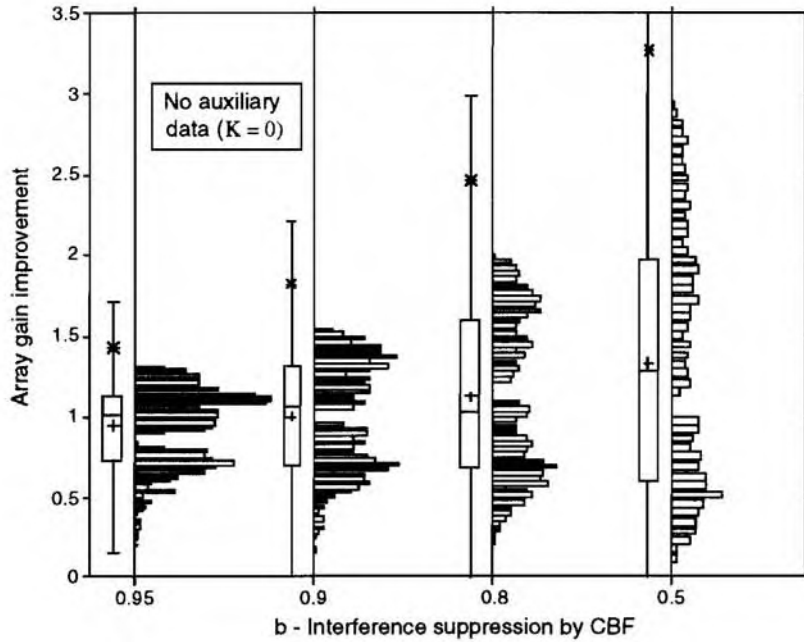


Figure 12 Histograms of AGI for $N = 16$, $L = 1$ and $K = 0$ for a single plane wave interference in spatially white noise.

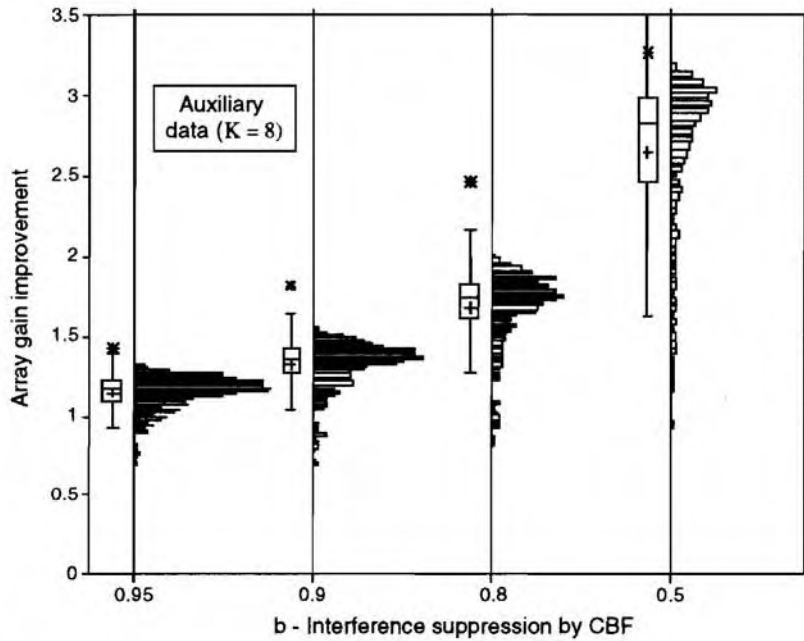


Figure 13 Histograms of AGI for $N = 16$, $L = 1$ and $K = 8$ for a single plane wave interference in spatially white noise.

Table 2 S_{\max} , maximum AGI, average AGI, ρ_{cbf} and average ρ for $N = 16$, $L = 1$ and $K = 0$ and $K = 8$ for a single plane wave interference in spatially white noise.

b	Angle (deg)	S_{\max}	AGI _{max}	E[AGI]		ρ_{cbf}	E[ρ]	
				$K = 0$	$K = 8$		$K = 0$	$K = 8$
0.95	0.894	2.18	1.432	0.95	1.15	0.698	0.840	0.925
0.9	1.278	2.91	1.818	1.01	1.33	0.550	0.800	0.906
0.8	1.848	4.36	2.455	1.13	1.68	0.407	0.777	0.913
0.5	3.179	8.73	3.273	1.34	2.65	0.306	0.799	0.971

5

Conclusions

A method for estimating the target-only, target-plus-interference and interference-only SCM's was proposed when both target-plus-interference data and interference-only (auxiliary) data are available. The algorithm constrains the estimation to matrices that are Toeplitz, substantially improving estimation and beamforming performance over the unconstrained sample covariance matrix estimator. The beam output SNR for the Gaussian target model when MVDR beamforming is performed using an estimate of the target-plus-interference SCM was shown to be non-linearly dependent on the maximum beam output SNR in such a manner that when S_{\max} is large, accurate estimation of the SCM is crucial. Simulation analysis illustrated that the proposed algorithm can provide the necessary accuracy, with further improvement as the amount of auxiliary data increases.

Adaptive beamforming for submarine or mine detection presents the specific situation of a very small amount of target-plus-interference data and a moderate amount of stationary interference-only data. If the array configuration and propagation conditions result in a Toeplitz structure for the SCM's, the algorithm presented in this report may be utilized to perform adaptive beamforming to provide high angular resolution or to suppress angularly isolated interferences. However, this has yet to be shown using real data and the performance of the algorithm when the actual covariance matrix is not exactly Toeplitz is unknown and needs to be investigated.

References

-
- [1] D. A. Abraham and N. L. Owsley, "Preprocessing for high resolution beamforming," in *Proceedings of Twenty-Third Asilomar Conference on Signals, Systems, and Computers*, November 1989.
 - [2] K. A. Burgess and B. D. VanVeen, "Improved adaptive detection performance via subspace processing," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 1992, pp. V-353-356.
 - [3] N. L. Owsley, "Enhanced minimum variance beamforming," Tech. Rep. 8305, Naval Underwater Systems Center, Nov. 1988.
 - [4] J. Burg, D. Luenberger, and D. Wenger, "Estimation of structured covariance matrices," *Proceedings of IEEE*, vol. 70, pp. 963-974, September 1982.
 - [5] D. Gray, B. Anderson, and P. Sim, "Estimation of structured covariances with application to array beamforming," *Circuits, Systems and Signal Processing*, vol. 6, no. 4, pp. 421-447, 1987.
 - [6] D. R. Fuhrmann, "Application of Toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Transactions on Signal Processing*, vol. 39, no. 10, pp. 2194-2198, October 1991.
 - [7] S. Haykin, Ed., *Advances in Spectrum Analysis and Array Processing*, vol. II, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1991.
 - [8] J. V. DiFranco and W. L. Rubin, *Radar Detection*, Artech House, Inc., Dedham, Massachusetts, 1980.
 - [9] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 2, pp. 1-38, 1977.
 - [10] G. Strang, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, Massachusetts, 1986.
 - [11] D. R. Fuhrmann and M. I. Miller, "On the existence of positive-definite maximum-likelihood estimates of structured covariance matrices," *IEEE Transactions on Information Theory*, vol. 34, no. 4, pp. 722-729, July 1988.
 - [12] M. A. Tanner, *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, Springer-Verlag, New York, second edition, 1993.

- [13] S. R. Searle, *Linear Models*, John Wiley & Sons, Inc., 1971.
- [14] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1992.
- [15] S. M. Kay, *Modern Spectral Estimation Theory and Application*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1988.
- [16] L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley Pub. Co., Reading, MA, 1991.
- [17] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-10, no. 6, pp. 853-863, November 1974.
- [18] J. Spanier and K. B. Oldham, *An Atlas of Functions*, Hemisphere Pub. Corp., Washington, 1987.
- [19] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey, *Graphical Methods for Data Analysis*, Wadsworth & Brooks/Cole Pub. Co., Pacific Grove, California, 1983.
- [20] N. L. Owsley, "Enhanced minimum variance beamforming," in *NATO Advanced Study Institute on Underwater Acoustic Data Processing*, Y. T. Chan, Ed., Kingston, Ontario, July 1988, pp. 285-291, Kluwer Academic Publishers.

Annex A

Forming the diagonal elements of $\mathbf{W}_N \mathbf{R} \mathbf{W}_N^H$

When \mathbf{W}_N is a matrix containing the first N columns of an M -point DFT matrix and \mathbf{R} is an N -by- N Toeplitz matrix, the diagonal elements of $\mathbf{W}_N \mathbf{R} \mathbf{W}_N^H$ may be efficiently determined using an M -point DFT. This relationship is described by Miller *et al.* in [7] and repeated here for convenience. Let $\mathbf{T} = \mathbf{W}_N \mathbf{R} \mathbf{W}_N^H$, $\{t_m\}_{m=0}^{M-1}$ be the diagonal elements of \mathbf{T} , and $\{r[l]\}_{l=-(N-1)}^{(N-1)}$ be the l^{th} diagonal element of \mathbf{R} where $l > 0$ indicates the subdiagonals and $l < 0$ indicates the superdiagonals. Then,

$$t_m = \sum_{i=0}^{N-1} W_{m,i} \sum_{k=0}^{N-1} R_{i,k} W_{m,k}^* \quad (84)$$

$$= \frac{1}{M} \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} R_{i,k} e^{j2\pi \frac{m(i-k)}{M}} \quad (85)$$

$$= \frac{1}{M} \sum_{l=-(N-1)}^{N-1} (N - |l|) r[l] e^{j2\pi \frac{ml}{M}} \quad (86)$$

$$= \frac{1}{M} \sum_{l=0}^{N-1} (N - l) r[-l] e^{-j2\pi \frac{ml}{M}} \quad (87)$$

$$+ \frac{1}{M} \sum_{l=M-N+1}^{M-1} (N - M + l) r[M - l] e^{-j2\pi \frac{ml}{M}} \quad (88)$$

$$= \sum_{l=0}^{N-1} v[l] e^{-j2\pi \frac{ml}{M}} \quad (89)$$

where

$$v[l] = \begin{cases} \frac{(N-l)}{M} r[l]^* & 0 \leq l \leq N-1 \\ \frac{(N-M+l)}{M} r[M-l] & M-N+1 \leq l \leq M-1 \\ 0 & \text{otherwise} \end{cases} \quad (90)$$

Thus, the diagonal elements of $\mathbf{W}_N \mathbf{R} \mathbf{W}_N^H$ may be computed using an M -point DFT rather than requiring $2MN^2$ multiplications of a brute force method.

Annex B

Frobenius norm of error between
sample and true covariance matrix

Suppose the data $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_L]$ are independent, zero-mean, complex normally distributed with covariance matrix \mathbf{R} as described by eqs. (1) and (2). Let the statistic T be the squared Frobenius norm between the sample covariance matrix

$$\begin{aligned} \mathbf{S}_x &= \frac{1}{L} \sum_{i=1}^L \mathbf{x}_i \mathbf{x}_i^H \\ &= \frac{1}{L} \mathbf{X} \mathbf{X}^H \end{aligned} \quad (91)$$

and the true covariance matrix,

$$\begin{aligned} T &= \|\mathbf{R} - \mathbf{S}_x\|_F^2 \\ &= \text{tr} \left[\left(\mathbf{R} - \frac{1}{L} \mathbf{X} \mathbf{X}^H \right) \left(\mathbf{R} - \frac{1}{L} \mathbf{X} \mathbf{X}^H \right) \right] \\ &= \text{tr}(\mathbf{R}^2) - \frac{2}{L} \text{tr}(\mathbf{R} \mathbf{X} \mathbf{X}^H) + \frac{1}{L^2} \text{tr}(\mathbf{X} \mathbf{X}^H \mathbf{X} \mathbf{X}^H) \end{aligned} \quad (92)$$

The expected value of T may be further simplified to

$$\begin{aligned} E[T] &= \text{tr}(\mathbf{R}^2) - 2 \text{tr}(\mathbf{R}^2) + \frac{1}{L^2} E \left[\text{tr}(\mathbf{X} \mathbf{X}^H \mathbf{X} \mathbf{X}^H) \right] \\ &= \frac{1}{L^2} E \left[\text{tr}(\mathbf{X}^H \mathbf{X} \mathbf{X}^H \mathbf{X}) \right] - \text{tr}(\mathbf{R}^2) \\ &= \frac{1}{L^2} E \left[\text{tr}(\mathbf{A}^2) \right] - \text{tr}(\mathbf{R}^2) \end{aligned} \quad (93)$$

where $\mathbf{A} = \mathbf{X}^H \mathbf{X}$. The first term in eq. (93) may be described as

$$\begin{aligned} E \left[\text{tr}(\mathbf{A}^2) \right] &= E \left[\|\mathbf{A}\|_F^2 \right] \\ &= \sum_{i=1}^L \sum_{j=1}^L E \left[|a_{i,j}|^2 \right] \end{aligned} \quad (94)$$

where $a_{i,j} = \mathbf{x}_i^H \mathbf{x}_j$ is the (i,j) element of $\mathbf{A} = \mathbf{X}^H \mathbf{X}$.

If $i \neq j$ then

$$\begin{aligned}
 E[|a_{i,j}|^2] &= E\left[|\mathbf{x}_i^H \mathbf{x}_j|^2\right] \\
 &= E\left[\mathbf{x}_i^H \mathbf{x}_j \mathbf{x}_j^H \mathbf{x}_i\right] \\
 &= E\left[\text{tr}\left(\mathbf{x}_j \mathbf{x}_j^H \mathbf{x}_i \mathbf{x}_i^H\right)\right] \\
 &= \text{tr}\left(E\left[\mathbf{x}_j \mathbf{x}_j^H\right] E\left[\mathbf{x}_i \mathbf{x}_i^H\right]\right) \\
 &= \text{tr}\left(\mathbf{R}^2\right).
 \end{aligned} \tag{95}$$

If $i = j$ then

$$E[|a_{i,i}|^2] = E\left[\left(\mathbf{x}_i^H \mathbf{x}_i\right)^2\right] \tag{96}$$

$$= [\text{tr}(\mathbf{R})]^2 + \text{tr}(\mathbf{R}^2) \tag{97}$$

where the latter step is obtained by noting that $\mathbf{x}_i^H \mathbf{x}_i$ is a quadratic form and exploiting the results of Searle [13]. Note that Searle only considers real Gaussian random vectors; the extension to the complex case is straightforward.

Substituting eqs. (95) and (97) in eq. (94) and then eq. (93), the expected value of the squared Frobenius norm between the sample covariance matrix and the true covariance matrix is seen to be

$$E[T] = \frac{1}{L^2} \left\{ L^2 \text{tr}(\mathbf{R}^2) + L [\text{tr}(\mathbf{R})]^2 \right\}. \tag{98}$$

Thus,

$$E[T] = \frac{[\text{tr}(\mathbf{R})]^2}{L}. \tag{99}$$

Using similar techniques, it can be shown that

$$E\left[\|(\mathbf{R} - \mathbf{Q}) - (\mathbf{S}_x - \mathbf{S}_y)\|_F^2\right] = \frac{[\text{tr}(\mathbf{R})]^2}{L} + \frac{[\text{tr}(\mathbf{Q})]^2}{K} \tag{100}$$

when $\mathbf{S}_y = \frac{1}{K} \mathbf{Y} \mathbf{Y}^H$ and $\mathbf{Y} = [\mathbf{y}_1 \ \cdots \ \mathbf{y}_K]$ is distributed as described in eqs. (5) and (6).

Security Classification NATO UNCLASSIFIED		Project No. 041-3
Document Serial No. SR-261	Date of Issue March 1997	Total Pages 41 pp.
Author(s) Abraham, D.A.		
Title Toeplitz constrained adaptive beamforming with auxiliary data		
<p>Abstract</p> <p>Submarine and mine detection using active sonar face similar problems when adaptive beamforming is required to provide high angular resolution or to suppress angularly isolated interferences. Both situations result in a very small amount of target-plus-interference data and a moderate amount of stationary interference-only data. From this data the target-plus-interference and interference-only spatial covariance matrices (SCM's) must be estimated. When the SCM's have a Toeplitz structure, as will occur with an equi-spaced line array and plane wave propagation, beamforming performance may be improved by constraining the estimates to be Toeplitz as well.</p> <p>In this report, the expectation-maximization algorithm is used to obtain the Toeplitz constrained maximum likelihood estimates of the SCM's when target-plus-interference data and interference-only (auxiliary) data are available. When the target-plus-interference SCM estimate is used in an MVDR beamformer, it is shown that the beam output SNR is non-linearly dependent on the maximum beam output SNR S_{\max} in such a manner that when S_{\max} is large, accurate estimation of the SCM is crucial. Simulation analysis illustrates that the proposed algorithm can provide the necessary accuracy, with further improvement as the amount of auxiliary data increases.</p>		
<p>Keywords</p> <p>Adaptive beamforming – Toeplitz covariance – expectation maximization</p>		
<p>Issuing Organization</p> <p>North Atlantic Treaty Organization SACLANT Undersea Research Centre Viale San Bartolomeo 400, 19138 La Spezia, Italy</p> <p>[From N. America: SACLANTCEN (New York) APO AE 09613]</p>		<p>Tel: +39 (0)187 540 111 Fax: +39 (0)187 524 600 E-mail: library@saclantc.nato.int</p>

Report no. changed (Mar 2006): SR-261-UU

Initial Distribution for SR-261

Ministries of Defence

DND Canada	10
CHOD Denmark	8
DGA France	8
MOD Germany	15
HNDGS Greece	12
MARISTAT Italy	10
MOD (Navy) Netherlands	12
NDRE Norway	10
MOD Portugal	5
MDN Spain	2
TDKK and DNHO Turkey	5
MOD UK	20
ONR USA	42

Scientific Committee of National Representatives

SCNR Belgium	1
SCNR Canada	1
SCNR Denmark	1
SCNR Germany	1
SCNR Greece	1
SCNR Italy	1
SCNR Netherlands	2
SCNR Norway	1
SCNR Portugal	1
SCNR Spain	1
SCNR Turkey	1
SCNR UK	1
SCNR USA	2
French Delegate	1
SECGEN Rep. SCNR	1
NAMILCOM Rep. SCNR	1

NATO Commands and Agencies

NAMILCOM	2
SACLANT	3
CINCEASTLANT/	
COMNAV/NORTHWEST	1
CINCIBERLANT	1
CINCWESTLANT	1
COMASWSTRIKFOR	1
COMMAIREASTLANT	1
COMSTRIKFLTANT	1
COMSUBACLANT	1
SACLANTREPEUR	1
SACEUR	2
CINC/NORTHWEST	1
CINCSOUTH	1
COMEDCENT	1
COMMARAIRMED	1
COMNAVSOUTH	1
COMSTRIKFOR/SOUTH	1
COMSUBMED	1
NC3A	1
PAT	1

National Liaison Officers

NLO Canada	1
NLO Denmark	1
NLO Germany	1
NLO Italy	1
NLO Netherlands	1
NLO UK	1
NLO USA	1

Sub-total	208
SACLANTCEN	30
Total	238