

STRUCTURED COVARIANCE ESTIMATION FOR SPACE-TIME ADAPTIVE PROCESSING

Timothy A. Barton and Steven T. Smith

Lincoln Laboratory
Massachusetts Institute of Technology
Lexington, Massachusetts 02173
email: barton@ll.mit.edu and stsmith@ll.mit.edu

ABSTRACT

Adaptive algorithms require a good estimate of the interference covariance matrix. In situations with limited sample support such an estimate is not available unless there is structure to be exploited. In applications such as space-time adaptive processing (STAP) the underlying covariance matrix is structured (e.g., block Toeplitz), and it is possible to exploit this structure to arrive at improved covariance estimates. Several structured covariance estimators have been proposed for this purpose. The efficacy of several of these are analyzed in this paper in the context of a variety of STAP algorithms. The SINR losses resulting from the different methods are compared. An example illustrating the superior performance resulting from a new maximum likelihood algorithm (based upon the expectation-maximization algorithm) is demonstrated using simulation and experimental data.

1. INTRODUCTION

The presence of ground clutter can seriously degrade the detection performance of airborne surveillance radars because the clutter energy may exist in all angular directions and doppler bins. Because the precise structure and location of clutter interference is not known a priori, adaptive nulling methods, especially space-time adaptive processing (STAP), are used to mitigate its effects.

Such adaptive processing methods require knowledge of the interference covariance matrix. If this covariance matrix is not known, it must be estimated in some way. Typically, the sample covariance matrix is used as the estimate of the covariance. However, in situations with insufficient sample support, this estimate yields poor detection performance [8]. This problem motivates the consideration of structured covariance estimation procedures in which improved adaptive processing performance may be achieved [6, 5, 9, 4]. This is the situation for STAP, where the clutter covariance matrix is modeled as having block Toeplitz or Toeplitz block Toeplitz structure [12, 10], depending on prior knowledge of the array geometry. The subject of this paper is to analyze the efficacy of a variety of structured covariance estimation techniques for STAP, including a new maximum likelihood (ML) algorithm.

This work was sponsored by DARPA under Air Force contract F19628-95-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Air Force.

2. COVARIANCE MATRIX ESTIMATORS

2.1. Sample Covariance Matrix Estimator

The sample covariance matrix is usually used as an estimate of the true interference covariance matrix in STAP algorithms. This approach is reasonable because the sample covariance matrix is easy to compute and is the unconstrained ML estimate under a Gaussian data model. However, this method yields a poor estimate with insufficient sample support. Diagonal loading is a common technique used to improve performance in this scenario; however, this method is not effective against lower power yet significant interference sources. Furthermore, neither method incorporates constraints that arise from consideration of the physical phenomena that generate clutter data.

2.2. ML Structured Covariance Estimator

In most ML structured covariance estimation problems no closed-form solution is known; this is true for covariance matrices with generic block Toeplitz or Toeplitz block Toeplitz structure. It is this difficulty that leads one to an iterative numeric technique for ML estimation.

Structured covariance estimation is considered in its full generality by Burg et al. [2], who propose an iterative algorithm for ML structured covariance matrix estimation based upon the idea of driving the derivative of the log-likelihood function to zero. Alternatively, the iterative expectation-maximization (EM) algorithm of Dempster et al. [3] has been used to generate ML structured covariance matrix estimates [6, 5, 9, 4]. The attractive properties of the EM-based ML algorithm include: (1) the sequence of covariance estimates are non-decreasing in likelihood; (2) the constraints imposed on the covariance estimates are easily incorporated into the estimation procedure; (3) the procedure reduces to a sequence of simple matrix operations. Unfortunately, because of its iterative nature, the EM algorithm for structured covariance estimation is computationally intensive.

An EM algorithm for generating block Toeplitz and Toeplitz block Toeplitz covariance matrix estimates derived by Fuhrmann and Barton [5] is summarized by the following sequence:

$$\Sigma_{k+1} = B(\Sigma_k + \Sigma_k \mathbf{A}^H \mathbf{R}_k^{-1} (\mathbf{S} - \mathbf{R}_k) \mathbf{R}_k^{-1} \mathbf{A} \Sigma_k), \quad (1)$$

$$\mathbf{R}_{k+1} = \mathbf{A} \Sigma_{k+1} \mathbf{A}^H, \quad (2)$$

where \mathbf{R}_k is the structured covariance estimate, \mathbf{S} is the sample covariance matrix, and $B(\cdot)$ denotes the block di-

agonal part. The sequence is initiated by providing a positive definite guess for Σ_0 (say $\Sigma_0 = \mathbf{I}$), then taking $\mathbf{R}_0 = \mathbf{A}\Sigma_0\mathbf{A}^H$. For the case of block Toeplitz covariance matrices of order MN -by- MN (M^2 blocks each of order N),

$$\mathbf{A} = ([\mathbf{I}_M, \mathbf{0}_{M \times (P-M)}]\mathbf{W}_P) \otimes \mathbf{I}_N, \quad (3)$$

where \mathbf{W}_P is the P -by- P ($P > M$) normalized discrete Fourier transform matrix. The use of this EM algorithm imposes the additional constraint that the estimated MN -by- MN block Toeplitz covariance matrix has nonnegative definite block circulant extensions of size P .

For the case of Toeplitz block Toeplitz covariance matrices of order MN -by- MN (M^2 blocks each Toeplitz of order N)

$$\mathbf{A} = ([\mathbf{I}_M, \mathbf{0}_{M \times (P-M)}]\mathbf{W}_P) \otimes ([\mathbf{I}_N, \mathbf{0}_{N \times (Q-N)}]\mathbf{W}_Q), \quad (4)$$

where $P > M$ and $Q > N$.

Given the EM iteration sequence of Eqs. (1) and (2), the number of computations incurred when estimating an ML block Toeplitz structured covariance matrix is significantly increased over that incurred in generating the unstructured sample covariance matrix. However, the ML structured covariance matrix estimates result in improved performance of STAP algorithms.

For the case of block Toeplitz and Toeplitz block Toeplitz covariance structure, these issues lead one to consider fast non-ML structured covariance estimation procedures, two of which are discussed in the following sections.

2.3. Projected Covariance Estimators

Projection offers one simple method of estimating structured covariance matrices. This idea comes from the fact that for any vector space with an inner product, there exists a unique orthogonal projection onto a given subspace. Here the covariance matrices are viewed as the vector space of Hermitian matrices, and the subspace is viewed as the subspace of structured Hermitian matrices. Looking ahead, we forecast difficulties with this approach because its assumptions are invalid: the covariance matrices do not have a vector space structure. We shall nevertheless describe and analyze this method because it is an obvious procedure and because its deficiencies will be improved somewhat in section 2.4.

Given two Hermitian matrices \mathbf{A} and \mathbf{B} , there is a natural inner product between \mathbf{A} and \mathbf{B} defined by:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr } \mathbf{A}\mathbf{B}. \quad (5)$$

The square length of \mathbf{A} is written $\|\mathbf{A}\|^2 = \langle \mathbf{A}, \mathbf{A} \rangle$ (i.e., the Frobenius norm). Let S be an n -dimensional subspace of the Hermitian matrices, and \mathbf{E}_i ($i = 1, \dots, n$) be an orthonormal basis of S , i.e., $\langle \mathbf{E}_i, \mathbf{E}_j \rangle = \delta_{ij}$ (Kronecker delta). For example, if S is the vector space of Hermitian Toeplitz matrices, then we may take $\mathbf{E}_1 = \text{toeplitz}(1, 0, \dots, 0)/\sqrt{n}$, $\mathbf{E}_2 = \text{toeplitz}(0, 1, 0, \dots, 0)/\sqrt{2(n-1)}$, and so on (using Moler's Matlab notation). The unique orthogonal projection of an arbitrary Hermitian matrix $\mathbf{A} = (\mathbf{A}_{ij})$ onto S is defined to be the vector $\sum_i \alpha_i \mathbf{E}_i$ such that the square length $\|\sum_i \alpha_i \mathbf{E}_i - \mathbf{A}\|^2$ is minimized over all α_i . This results in the projection

$$\Pi(\mathbf{A}) = \sum_i \langle \mathbf{E}_i, \mathbf{A} \rangle \mathbf{E}_i. \quad (6)$$

For the Hermitian Toeplitz example, this projection results in the formulae

$$\text{diagonal of } \Pi(\mathbf{A}) = \frac{\mathbf{A}_{11} + \mathbf{A}_{22} + \dots + \mathbf{A}_{nn}}{n}, \quad (7)$$

$$\text{off-diagonal of } \Pi(\mathbf{A}) = \frac{\mathbf{A}_{12} + \mathbf{A}_{23} + \dots + \mathbf{A}_{n-1,n}}{n-1}, \quad (8)$$

and so on. That is, take the mean down diagonals. The projection for block Toeplitz and Toeplitz block Toeplitz matrices is similar: simply take the mean down diagonals and block diagonals.

The problem with this method is that there is no guarantee that $\Pi(\mathbf{A})$ will be positive definite; in fact, it is very easy to generate examples where $\Pi(\mathbf{A})$ fails to be a covariance matrix. Furthermore, our purpose for estimating structured covariance matrices is to construct adaptive weight vectors that null interference. There is no guarantee that $\Pi(\mathbf{A})$ will result in adaptive weights that sufficiently null the interference, especially in sample limited scenarios.

2.4. Weighted Projected Covariance Estimators

By a simple modification in the case of block Toeplitz and Toeplitz block Toeplitz matrices, the projected covariance estimator of the previous section almost surely yields positive definite Hermitian matrices. Instead of taking sample means down (block) diagonals as in Eqs. (7) and (8), take the weighted means

$$\mathbf{P}_0 = \frac{\mathbf{A}_{11} + \mathbf{A}_{22} + \dots + \mathbf{A}_{nn}}{n}, \quad (9)$$

$$\mathbf{P}_1 = \frac{\mathbf{A}_{12} + \mathbf{A}_{23} + \dots + \mathbf{A}_{n-1,n}}{n}, \quad (10)$$

$$\dots \dots \dots$$

$$\mathbf{P}_{n-1} = \frac{\mathbf{A}_{1,n}}{n}, \quad (11)$$

then set the covariance estimate to be $\mathbf{R} = \text{toeplitz}(\mathbf{P}_0, \mathbf{P}_1, \dots, \mathbf{P}_{n-1})$. This weighted projection almost surely yields a positive definite block Toeplitz or Toeplitz block Toeplitz covariance estimate. To see this, note that if $\mathbf{A} = (\mathbf{A}_{ij})$ is a sample covariance matrix, then \mathbf{P}_k ($k = -n+1, \dots, n-1$) represents a (matrix) autocorrelation sequence [7], which has a positive power spectrum. This property is very attractive when there are fewer samples than degrees of freedom.

In addition to producing a positive definite estimate, this method also has the advantages of speed and ease of computation. However, this estimator is not based upon a statistical model, and for the application of adaptive nulling, there is no reason to believe that it will yield improved performance. In fact, as will be seen in the next section, this method may not yield better performance than the sample covariance matrix estimate.

3. RESULTS AND DISCUSSION

We will compare different covariance matrix estimation methods by examining the the SINR loss achieved by the fully optimum and several suboptimum STAP algorithms [12]. The SINR loss measures the loss in target power resulting from adaptive nulling. In symbols,

$$\text{SINR loss} = \frac{|\mathbf{w}^H \mathbf{v}|^2}{(\mathbf{w}^H \mathbf{R} \mathbf{w})(\mathbf{v}^H \mathbf{v})}, \quad (12)$$

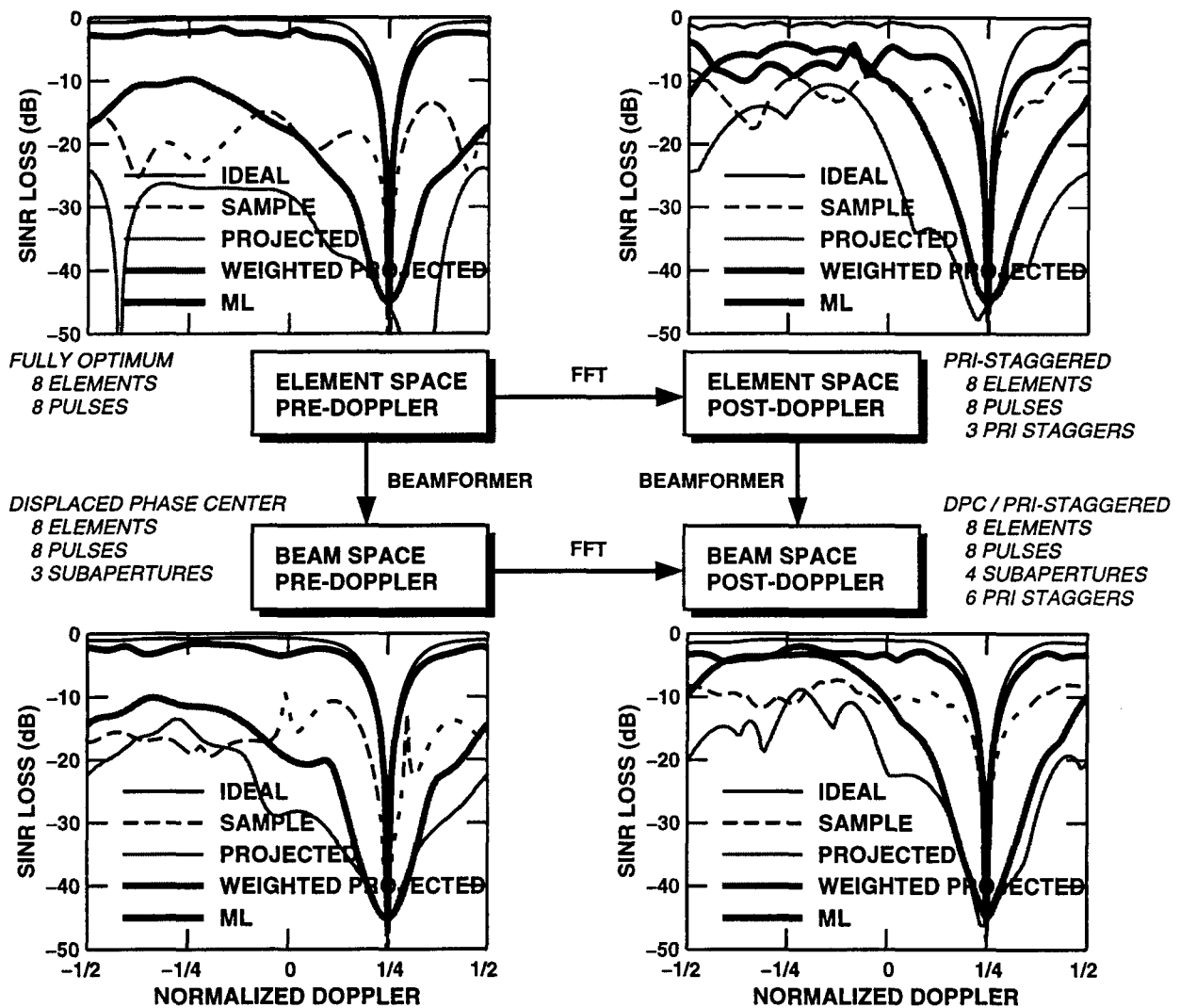


Figure 1: SINR Loss versus doppler (sample support = $1 \times$ degrees of freedom) for four classes of STAP algorithms. The ideal performance is shown by the thin black curve. Note that the ML's performance (thick black curve) exceeds all other methods both in the width of the clutter notch and the SINR loss away from mainlobe clutter (at $1/4$ doppler), except for the PRI staggered post-doppler algorithm. In this case, we have 3^2 blocks each of order 8, which has less structure than the other algorithms—fully optimum has 8^2 blocks each of order 8, displaced phase center (DPC) pre-doppler has 8^2 blocks each of order 3, and DPC PRI-staggered has 6^2 blocks each of order 4.

where \mathbf{v} is the steering vector, \mathbf{w} is the adaptive weight vector, and \mathbf{R} is the covariance of the interference-plus-noise. Of course all adaptive algorithms will suffer SINR loss near interference; however good nulling algorithms will not have significant losses away from the interference.

The scenario considered is a airborne radar with $N = 8$ elements and $M = 8$ pulses per CPI, whose velocity is such that 1 half-interelement spacing is traversed per PRI, and a uniform clutter environment with 40 dB clutter-to-noise ratio. A crab angle of 30° is used, resulting in mainlobe clutter at a normalized doppler frequency of $1/4$. There is no back lobe clutter. The ideal clutter covariance matrix for this scenario is Toeplitz block Toeplitz; however, we shall only assume a block Toeplitz structure, i.e., no pulse-to-pulse mismatches.

STAP algorithms may be classified by the domain in

which adaptivity occurs—element or beamspace, and pre- or post-doppler processing [12]. In any domain, the optimum (w.r.t. SINR) weight vector is given by the well-known equation

$$\mathbf{w} = \mathbf{R}^{-1}\mathbf{v}, \quad (13)$$

whose degrees of freedom is MN . Because this product can be quite large for moderately sized systems, reduced dimension suboptimum STAP algorithms are attractive. In this case, the suboptimum weight vector is given by the equation

$$\mathbf{w} = \mathbf{T}(\mathbf{T}^H\mathbf{R}\mathbf{T})^{-1}\mathbf{T}^H\mathbf{v}, \quad (14)$$

whose degrees of freedom is the column dimension of \mathbf{T} , where \mathbf{T} is a transformation matrix into one of the STAP domains shown in Figure 1. If \mathbf{T} is block Toeplitz, then block Toeplitz structure is preserved in the reduced dimension space, and structured covariance matrix estimators

may be considered.

A set of complex Gaussian data with the ideal clutter covariance matrix for this scenario [10] is generated. The number of independent samples available for covariance estimation is chosen to be exactly the adaptive degrees of freedom—a low sample support scenario. This equals 64 for the fully optimum STAP algorithm, and 24 for the reduced dimension STAP algorithms. The SINR loss of these STAP algorithms using the covariance estimation techniques described in section 2 are shown in Figure 1. For comparison, the loss using the known clutter covariance matrix is plotted using a thin black curve. Because of the small sample support, the sample covariance matrix yields a poor estimate of the clutter covariance matrix, resulting in severe SINR loss (8–25 dB) away from mainlobe clutter. The projected covariance estimator (thin gray curve) yields abysmal performance for all algorithms because its estimate is not necessarily positive definite. The weighted projected covariance estimator (thick gray curve) yields similar performance to the sample covariance matrix, except that the notch at mainlobe clutter is significantly wider, resulting in decreased doppler coverage for all algorithms. In contrast, for all but the PRI-staggered post-doppler algorithm, the ML covariance estimator (thick black curve) [for $P = 4M$ in Eq. (3)] suffers only minor losses (2–4 dB) away from mainlobe clutter, and, for all algorithms, has a comparable notch width at mainlobe clutter as the ideal (known covariance) STAP algorithm.

We observe that the ML's performance improves as the amount of structure increases, a trend exhibited in Figure 1. PRI-staggered post-doppler has 3^2 blocks each of order 8 and yields 4–10 dB of loss away from mainlobe clutter, whereas the other algorithms have more structure and less loss.

Structured covariance matrix estimation algorithms applied to Mountaintop data [11] are shown in Figure 2. This plot shows residual clutter-plus-noise to signal ratio versus range for a DPC pre-doppler STAP algorithm (14 elements, 16 pulses, 3 subapertures) for an adaptive weight vector (unit gain on target) steered near mainlobe clutter.

4. CONCLUSIONS

In this paper several structured covariance matrix estimators were considered in conjunction with the performance of STAP algorithms that use them. It is shown using simulation and experimental data that with limited sample support for a block Toeplitz covariance matrix, an ML estimator based upon the EM algorithm can yield superior adaptive nulling performance to other estimators. Based upon the results in this paper, we make the following recommendations to those who wish to estimate block Toeplitz covariance matrices for adaptive nulling: (1) make sure that the assumption of block Toeplitz structure is valid (e.g., array errors and multipath reflections ruin spatial and temporal stationarity, respectively); (2) in situations (such as data analysis) where greater computational intensity can be afforded, use an ML estimator; (3) if the computational intensity of ML cannot be afforded, use a diagonally loaded sample covariance matrix; (4) if the weighted projected estimator is used, verify that it yields improved performance in an idealized setting.

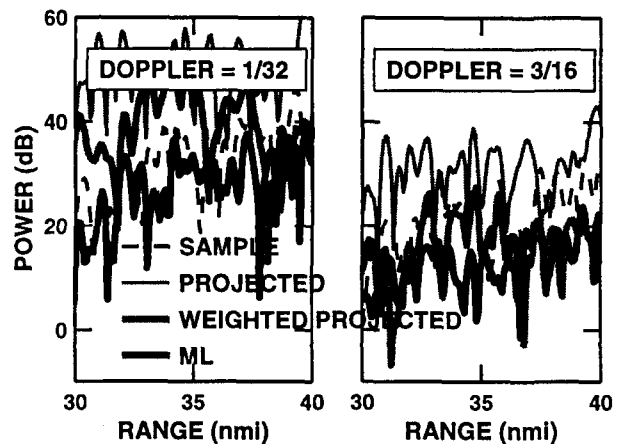


Figure 2: Residual clutter-plus-noise to signal ratio versus range using the Mountaintop database [11] with a DPC pre-doppler algorithm (data collected 9 March 1994 at White Sands, New Mexico). An adaptive weight vector steered near and away from mainlobe clutter ($\text{DOF} = 3 \cdot 16$, unit gain on target) generated using 61 samples of data between 25 and 30 nmi is applied to clutter between 30 and 40 nmi. The algorithms' performance agrees with the simulation results shown in Figure 1.

REFERENCES

- [1] T. A. BARTON. *Covariance Estimation for Multidimensional Data*. D.Sc. diss., Washington University, 1993.
- [2] J. P. BURG, D. G. LUENBERGER, and D. L. WENGER. "Estimation of structured covariance matrices," *Proc. IEEE* 70(9):963–974, 1982.
- [3] A. P. DEMPSTER, N. M. LAIRD, and D. B. RUBIN. "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statistical Society B-36*:1–38, 1977.
- [4] D. R. FUHRMANN. "Application of Toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Trans. Signal Processing* 39(10), 1991.
- [5] D. R. FUHRMANN and T. A. BARTON. "Estimation of block Toeplitz covariance matrices," in *Proc. 24th Asilomar Conf. Signals, Systems, Computers* 2:779–783, 1990.
- [6] M. I. MILLER and D. L. SNYDER. "The role of likelihood and entropy in incomplete-data problems: Applications to estimating point-process intensities and Toeplitz constrained covariances," *Proc. IEEE* 75(7):892–907, 1987.
- [7] A. V. OPPENHEIM and R. W. SCHAFER. *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975.
- [8] I. S. REED, J. D. MALLETT, and L. E. BRENNAN. "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace and Electronic Systems* AES-10(6):853–863, 1974.
- [9] F. C. ROBEY. *A Covariance Modeling Approach to Adaptive Beamforming and Detection*. D.Sc. diss., Washington University, 1990.
- [10] S. T. SMITH. "Space-Time clutter covariance matrix computation and interference subspace tracking," in *Proc. 29th Asilomar Conf. Signals, Systems, Computers*, 1995.
- [11] G. W. TITI. "An overview of the ARPA/Navy Mountaintop Program," in *Proc. IEEE Adaptive Antenna Systems Symposium*, November 1994. See also www.ll.mit.edu/11_darpa-sto/mtn_top.
- [12] J. WARD. "Space-Time adaptive processing for airborne radar," MIT Lincoln Laboratory Technical Report 1015, DTIC No. ESC-TR-94-109, 13 December 1994.