Subspace Approximation for Adaptive Multichannel Radar Filtering

A.W. Bojanczyk
School of Electrical Engineering
Cornell University
Ithaca, NY
adamb@ee.cornell.edu

W.L. Melvin and E.J. Holder
Georgia Tech Research Institute/SEAL
Georgia Institute of Technology
Atlanta, GA
bill.melvin@gtri.gatech.edu or jeff.holder

Abstract

In this paper we consider subspace approximation tailored to adaptive airborne radar. Motivation for this research includes the need for reduced computational burden and approaches for practical implementation. Measured radar data only approximately satisfies the statistical assumptions intrinsic to the adaptive processor. Hence, approximate numerical methods for adaptive weight computation may successfully be used in place of exact methods. We propose a numerical procedure based on partial bi-diagonalization of the interference covariance matrix, coupled with a pre-conditioned conjugate gradient iterative method, to extract approximate basis vectors for the interference subspace. We use these basis vectors to construct an adaptive weight vector. Through example, we show the potential of this method for adaptive radar.

1. Introduction

Airborne radar must detect targets of diminishing radar cross-section, often at decreased radial velocities. The interference environment is severe and poses a significant challenge to effective target detection. Exploiting signal diversity over multiple domains offers enhanced detection performance. Space-time adaptive processing (STAP) represents a class of multi-domain adaptive techniques useful in such circumstances. The theory of adaptive radar was proposed in a series of papers by Brennan, Mallett and Reed in the early 1970's [1-2]. STAP remains an active research topic and most view advanced STAP techniques as a vital component of future airborne and spaceborne radar systems.

Recent STAP research focuses largely on mitigating computational burden and/or proposing novel processing architectures [3-8]. Eigenbased STAP, also called reduced-rank or partially adaptive STAP methods, rely on a sequence of linear transformations and selection operators to generate a set of adaptive weights [5-8]. Such methods approach optimum performance, but at the expense of considerably increased computational complexity. This occurs because the sequence of required operations amounts to a singular value decomposition (SVD) of the data matrix. Nevertheless, such approaches leverage the numerical stability of the SVD and require minimal sample support for covariance matrix estimation. These considerations are important when applying STAP in routinely encountered inhomogeneous and sample-poor signal environments.

In this paper we propose approximate numerical techniques as an alternative approach for the airborne radar STAP problem. The proposed approach capitalizes on: (1) the generally low numerical rank of the covariance matrix; (2) the realization that actual airborne radar signal environments only approximately satisfy assumed statistical models; and, (3) low-rank realizations require smaller sample populations for parameter estimation.

2. Adaptive Airborne Radar

Consider an $\boldsymbol{M}$ channel, aircraft-mounted array receiving $\boldsymbol{N}$ pulses. The space-time snapshot, $x_k \in C^{MN}$, is given by

$$X_k = [x_{(k,1,1)}, \ldots, x_{(k,M,1)}, x_{(k,1,2)}, \ldots, x_{(k,M,2)}, \ldots, x_{(k,1,N)}, \ldots, x_{(k,M,N)}]^T \quad \textbf{(1)}$$

with $x_{(k,m,n)}$ representing a complex baseband observation at the kth range, mth channel and nth pulse. Each snapshot consists of additive signal contributions from clutter, $X_{k,C}$, jamming, $X_{k,J}$, uncorrelated noise, $X_{k,N}$, and targets, $X_{k,T}$, such that

$$X_k = \begin{cases} X_{k/H_0} = X_{k,C} + X_{k,J} + X_{k,N} \\ X_{k/H_1} = X_{k,C} + X_{k,J} + X_{k,N} + X_{k,T} \end{cases} \quad \textbf{(2)}$$

for the null and alternative hypotheses. The output of the adaptive processor is given by

$$y_k = \hat{W}_k^H X_k , \quad \textbf{(3)}$$

where $\hat{W}_k \in C^{MN}$ is the adaptive weight vector. This weight vector takes the general form

$$\hat{W}_k = \hat{\alpha}_k \hat{R}_k^{-1} S_T \; ; \quad \hat{R}_k = \frac{1}{\kappa} \sum_{\substack{m=1 \\ m \neq k}}^{\kappa} X_m X_m^H \, , \qquad \textbf{(4)}$$

where $S_T \in C^{MN}$ is the target space-time steering vector, $\hat{R}_k$ is the maximum likelihood estimate (MLE) of the interference covariance matrix [2], and $\hat{\alpha}_k$ is a constant. The space-time steering vector represents the response of the array to a point source with a specific direction of arrival and Doppler frequency. Note that $\hat{R}_k$ approximates $E[X_{k/H_0} X_{k/H_0}^H]$. We subsequently subject y_k to binary hypothesis testing as a means of declaring target presence.

3. Eigenbased Adaptive Radar

The sample covariance matrix $\hat{R}_k$ is Hermitian and positive definite. For this reason, a unitary matrix Q_k exists such that

$$\hat{R}_k = Q_k \Lambda_k Q_k^H = \sum_{m=1}^{MN} \lambda_k(m) \, q_k(m) \, q_k^H(m) \, , \qquad \textbf{(5)}$$

where $\lambda_k(1) \geq \lambda_k(2) \geq \ldots \geq \lambda_k(MN)$, $q_k(m)$ is the mth column of Q_k and I_{MN} is the ***MN*** x ***MN*** identity matrix. $\lambda_k(m)$ and $q_k(m)$ are the mth eigenvalue and eigenvector of $\hat{R}_k$, respectively. Decomposing (5) into principal components (PC) and noise yields

$$Q_{PC,k} = \left[q_k(1), q_k(2), \ldots, q_k(P) \right]$$
$$\lambda_k(1) \geq \lambda_k(2) \geq \ldots \geq \lambda_k(P) \, , \qquad \textbf{(6)}$$

and

$$Q_{noise,k} = \left[q_k(P+1), q_k(P+2), \ldots, q_k(MN) \right]$$
$$\lambda_k(P) \gg \lambda_k(P+1) \approx \lambda_k(P+2) \approx \ldots \approx \lambda_k(MN) \, . \qquad \textbf{(7)}$$

Observe that $Q_{PC,k} \perp Q_{noise,k}$. The principal and noise components define the interference and noise subspaces.

An interesting result in [5] shows that explicit knowledge of the true interference subspace completely solves the optimum filtering problem. Assuming matched channels and letting $\hat{\alpha}_k \rightarrow \alpha_k = 1$, we may write (4) as

$$W_k = \frac{1}{\lambda_0} \left[S_T - \sum_{m=1}^{NM} \frac{\lambda_k(m) - \lambda_0}{\lambda_k(m)} \gamma_{k,m} \, q_k(m) \right] , \qquad \textbf{(8)}$$

where W_k denotes the optimum weight vector, λ_0 represents the noise floor, and $\gamma_{k,m} = q_k^H(m) \, S_T$ is the projection of the mth interference eigenvector onto the quiescient response. Observe that terms associated with the noise eigenvalues do not affect the optimum weight vector. This same observation is made in [6].

Other researchers provide alternative eigenbased interpretations. For example, the two methods proposed in [7] rely on constructing a weight vector lying in the noise subspace. Since the noise subspace is orthogonal to the correlated interference, selecting a weight vector $\hat{W}_k$ such that $\hat{W}_k \in span(Q_{noise,k})$ cancels correlated interference. The cross-spectral metric (CSM) method discussed in [8] applies a cost function to the problem of choosing the best low-rank eigenbasis. In this case, the weight vector generally spans the noise subspace as seen through the following alternative view of the CSM method. From (2)-(5), the signal-to-interference plus noise ratio (SINR) may be written for a normalized target signal as

$$SINR = S_T^H \hat{R}_k^{-1} S_T = \sum_{m=1}^{NM} \frac{|S_T^H q_k(m)|^2}{\lambda_k(m)} \, . \qquad \textbf{(9)}$$

Accordingly, with maximum SINR as the objective, a low-rank basis selection should choose those eigen-components which maximize the partial sum of terms in (9). The desired target response influences this selection. For example, consider the top plot in Figure 1 showing a typical eigenspectra for a simulated airborne radar clutter covariance matrix. In contrast, the bottom plot in the figure shows the individual "CSM" terms, $|S_T^H \, q_k(m)|^2 / \lambda_k(m)$, as defined in (9). It is evident from this example that terms with the largest CSM predominantly lie in the noise subspace.

4. Subspace Processing

The preceding section is central to the development of numerical approximations suited to the STAP problem. As just discussed, three choices emerge for constructing the reduced-rank weight vector: 1) $\hat{W}_k$ lies in the interference subspace; 2) $\hat{W}_k$ lies in the orthogonal noise subspace; or 3) $\hat{W}_k$ uses those basis vectors which optimize an objective function. If the basis vectors arise from the SVD, it is sensible to construct $\hat{W}_k$ from <u>all</u> principal components. Complications may arise if the interference rank is fuzzy.

An advantage of subspace processing is reduced sample support requirements for covariance matrix estimation. Generally, the airborne radar signal environment is inhomogeneous (eg., spatially varying clutter and discretes) and non-stationary (data is coherent over a limited time interval). These factors limit useful sample data. Subspace approaches appear robust in such cases and afford the possibility of localized adaptive processing schemes. To corroborate this notion, we offer Figure 2. This figure depicts the actual eigenvalues of a jammer covariance matrix (three jammers present) for a 16-element linear array. The top plot shows the principal components, whereas the bottom plot depicts the noise eigenvalues. Also included in both plots are the corresponding eigenvalues for the sample covariance matrix estimated from sample support using 1x, 2x and 10x the total degrees of freedom (DOF). The figure shows the robustness of the estimation process in representing the interference subspace, whereas the full-rank uncorrelated noise is poorly estimated. The interference eigenvectors are also usually well represented with limited sample support.

Appealing aspects of subspace methods, specifically convergence to optimum performance and mitigated sample support, is offset by computational burden. Taking into account that actual measured data only approximately satisfies the statistical assumption of homogeneous observations, we find sufficient motivation for exploring approximate subspace methods.

Several other notions influence the development of such approximation procedures. First of all, we point out that $q_k \in span\{S_I(p)\}$, where $\{S_I(p)\}$ is the set of interference space-time steering vectors. For high interference-to-noise terms, the eigenbeams generally point in a single direction. However, the discrete Fourier transform (DFT) of the eigenvectors can produce results appearing like a difference pattern when interferers appear closely spaced, as is true for ground clutter returns. Thus, a linear combination of the $S_I(p)$ comprise each eigenvector and multiple terms can dominate. Secondly, asymptotic equivalence exists between Toeplitz and circulant matrices [9]. Ideal covariance matrices are Toeplitz, whereas sample covariance matrices are non-Toeplitz. Nevertheless, this implies the columns of a discrete cosine transform (DCT) or DFT may serve as appropriate surrogates for the actual eigenvectors of the covariance matrix. (DCT and DFT vectors are eigenvectors of circulant matrices.) Let G be a unitary matrix whose columns are selected DCT or DFT vectors, D be a diagonal matrix approximating the principal eigenvalues and E represent the residual off-diagonal terms. In the case of closely spaced interferers, we generally find that

$$\| \hat{R} - G\ D\ G^H \| , \qquad \textbf{(10)}$$

where $\| \bullet \|$ is an appropriate norm, is unacceptably large. A deterministic basis may not adequately diagonalize $\hat{R}_k$. An adaptive basis, generated via efficient numerical routines, offers the potential for better performance at modest cost.

5. Numerical Subspace Approximation

We may interpret STAP as the constrained minimization of output power subject to a linear constraint,

$$\min \hat{W}_k^H\ \hat{R}_k\ \hat{W}_k \quad such\ that \quad \hat{W}_k^H\ S_T = 1\ . \qquad \textbf{(11)}$$

Define the data matrix of snapshots from (1) as

$$X^H = \left[X_1 ,\ X_2 ,\ \ldots ,\ X_L \right] . \qquad \textbf{(12)}$$

Upon substituting the MLE for $\hat{R}_k$, (11) equates to the linear least squares problem (LLSP),

$$\min \|\ X\ \hat{W}_k \|_2 \quad such\ that \quad S_T^H\ \hat{W}_k = 1\ . \qquad \textbf{(13)}$$

Assume that $\|S_T\|_2 = 1$ and let $H\ S_T = \pm e_n$, where e_n is the n[th] column of the appropriately dimensioned identity matrix. Furthermore, let $H\ \hat{W}_k = u = \left(\mu_m\right)_{m=1}^{NM}$. We then express (13) as,

$$\min \| X H^H H \hat{W}_k \|_2 \quad such\ that \quad S_T^H H^H H \hat{W}_k = 1 \qquad \textbf{(14)}$$

from which we get the unconstrained form

$$\rho = \min_{\nu} \| Z_{n-1}\,\nu + z_n \|_2\ ; \qquad \textbf{(15)}$$
$$Z = X H^H = \left[Z_{n-1}\ z_n \right]\ ;\ \nu = \left(\mu_m\right)_{m=1}^{n-1} ,$$

where $n = NM$. For a linear solver based on a direct QR-decomposition, the cost of finding ν is $O(\,N^3 M^3\,)$ with $\boldsymbol{L}$ = 2xDOF. The required computational rate can be very large.

5.1 Low-Rank Approximation of the Data Matrix

The singular values of the data matrix relate closely to the eigenvalues of $\hat{R}_k$. When Z is of low numerical rank, as is often the case in airborne radar [3, 5-8], we can approximate (15) by another minimization problem in a lower dimension subspace $span(Y) \subset span(X)$ for which minimizer p^* of

$$\min \|\ Yp\ \| \quad such\ that \quad S_T^H\,p = 1 \qquad \textbf{(16)}$$

ideally yields small residuals for $X\ p^*$ except at target locations. If Y adequately represents the dominant subspace of X, and if the dimension of Y is considerably smaller than $dim(X)$, significant savings in computational cost result without sacrificing target detection performance.

The discussion of eigenbased methods suggests selecting Y via the dominant left singular vectors of the SVD, $X = U\ \Sigma\ V^H$ [5-6]. As already mentioned, the SVD is a very costly approach to rank-reduction. However, we may obtain a suitable low-rank approximation through bi-diagonalization of X. Furthermore, we do not have to complete the bi-diagonalization to extract meaningful information regarding the dominant subspace [10]. One accomplishes bi-diagonalization via sequences of left and right Householder transformations, U and V, applied directly to X such that $B = U^H\ X\ V$ is upper (or lower) bi-diagonal. After $\boldsymbol{k}$ steps of bi-diagonalization, it is with high probability that the singular values of $B_k = B(1{:}k,1{:}k)$ tend to be very good approximations to the largest singular values of X. Thus, when X is of low numerical rank we expect

$$Y_k = U_k(:,1{:}k)\ B_k\ V_k(:,1{:}k)^H \qquad \textbf{(17)}$$

to be close to the best rank-$\boldsymbol{k}$ approximation of X.

5.2 LLSP in Lower Dimensional Space

Consider the following subspace approximation, used in effect as a rank-reducing transformation to alleviate computational complexity. First, bi-diagonalize X until an appreciable gap is found between the 2-norms of column $\boldsymbol{k}$ of $B^{(k)}$ and column $\boldsymbol{k}$+1 of $B^{(k+1)}$. Next, solve (16) in the subspace spanned by U_{k+1} alone,

$$\min \| B_{k+1}\ \tilde{p} \|_2 \quad such\ that \quad \tilde{S}_T^H \tilde{p} = 1\ , \qquad \textbf{(18)}$$

where $\tilde{S}_T = V_{k+1}\ S_T$ and $\tilde{p} = V_{k+1}^H\ p$. Note, the unitary operator U preserves measure and we select B_{k+1} after its removal. A solution for $\tilde{p}$ follows by placing the constrained minimization into an unconstrained form similar to (15). One can develop a recursive relation for $\tilde{p}$ to efficiently accomodate expanded approximate subspace dimension. Observe that $p = V\ [\tilde{p}\ ;\ \bar{0}\ ]$.

5.3 Refinement Via Conjugate Gradient

From (8) it is seen that the optimum weight vector satisfies $W_k \in span(S_T, Q_{PC,k})$. The formulation based on partial bi-diagonalization extracts information from the data matrix such that $\hat{W}_k \in span(S_T, \hat{Q}_D)$ where $\hat{Q}_D$ is an approximation to the dominant subspace. As expected, other sources of correlated interference not represented by $\hat{Q}_D$ influence target detection. When choosing a low dimension subspace, it is not necessarily true that the dominant terms most greatly influence performance. One must consider the cost function when ranking the importance of each subspace [8]. A pre-conditioned conjugate gradient (CG) iterative method applied to the partially bi-diagonalized system of equations allows us to refine the weight vector produced from the unconstrained minimization in the dominant subspace alone. This additional step effectively expands the approximate weight vector to include basis vectors representing the weaker sources of correlated interference. We use $[B_k\ ;\ \sigma_{\min}\ I\ ]$ as an (n-1)x(n-1) preconditioner, where $\sigma_{\min}$ approximates the smallest singular values of B_k. The pre-conditioner reduces singular value spread of the partially bi-diagonal system. It is known that convergence is very good for the CG method when the data matrix is well conditioned [11]. Thus, CG iterations are applicable in this instance. The starting weight vector approximation is $[\ \tilde{p}\ ;\ \bar{0}\ ]$.

5.4 Computational Cost

The computational cost of the partial bi-diagonalization (step 1) is $O(L{\bullet}NM{\bullet}k)$, with k the number of bi-diagonalization steps. The pre-conditioned CG iterations (step 2) add $O(L{\bullet}NM{\bullet}G_C)$ computations, where G_C is the number of iterations.

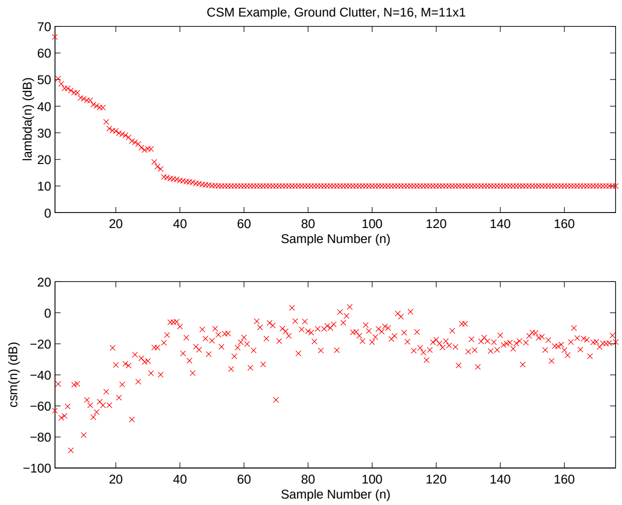

Figure 1. Eigenspectra and CSM terms.

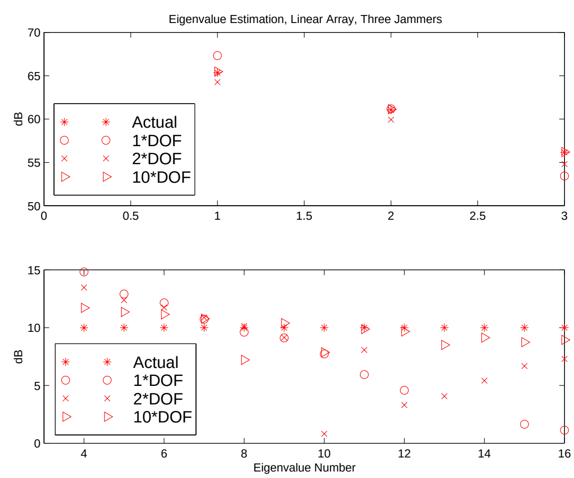

Figure 2. Robustness of dominant eigenvalues to low sample support.

6. Example

To validate the proposed method, consider the case of a 16 channel airborne linear array receiving 12 pulses. A weak target, positioned close to mainbeam clutter in Doppler space, is present in the 34th range bin (realization). Figure 3 shows the optimum filter output versus realization, whereas Figure 4 shows the adaptive filter output using the conventional sample matrix inversion (SMI) [2]. In contrast, Figure 5 shows the result using the subspace approximation of partial bi-diagonalization followed by the pre-conditioned, unconstrained CG method. The order of the partial bi-diagonalization was $\boldsymbol{k}$=16 and we used $G_C = \mathbf{18}$ CG iterations. Observe from the figures that all three methods detect the target. Interestingly enough, the approximate method gives the best results. We attribute this to improved numerical stability of the procedure in comparison to the numerical routines used to invert the covariance matrix for the optimum and SMI scenarios.

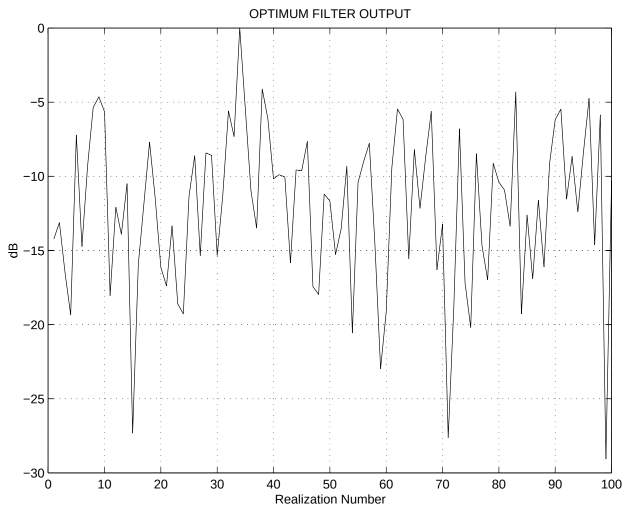

Figure 3. Optimum filter response.

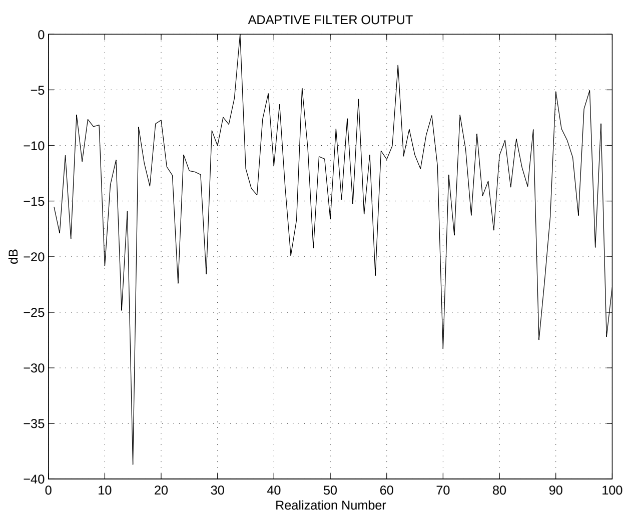

Figure 4. SMI adaptive filter response.

7. Summary

In this paper we propose an approximate subspace procedure suited to airborne radar STAP application. Recent eigenbased methods proposed by other researchers serve as motivation for our pursuit of this topic. The approximation involves partial bi-diagonalization and pre-conditioned conjugate gradient iterations to mitigate computational burden. A simple example shows that this approach has merit and warrants further consideration.

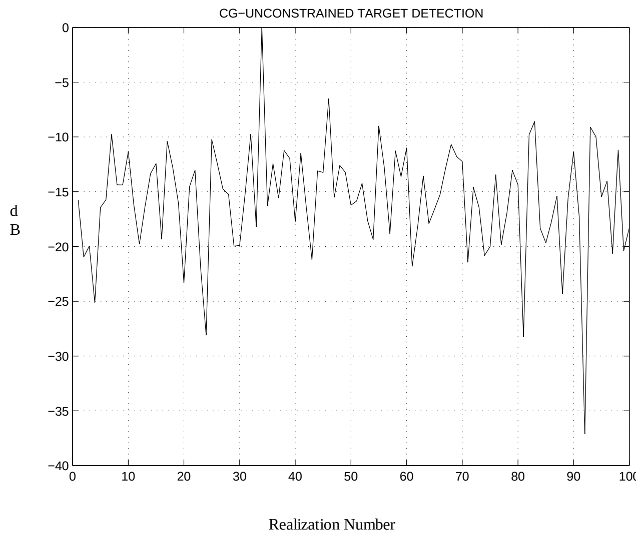

Figure 5. Unconstrained pre-conditioned CG method.

References

[1] L.E. Brennan and I.S. Reed, "Theory of Adaptive Radar", *IEEE Trans. AES*, Vol. 9, No. 2, pp. 237-252, March 1973.

[2] I.S. Reed, J.D. Mallett, and L.E. Brennan, "Rapid Convergence Rate in Adaptive Arrays", *IEEE Trans. AES*, Vol. 10, No. 6, pp. 853-863, November 1974.

[3] J. Ward, Space-Time Adaptive Processing for Airborne Radar, Lincoln Laboratory Tech. Rept., ESC-TR-94-109, December 1994.

[4] H. Wang and L. Cai, "On adaptive spatial-temporal processing for airborne surveillance radar systems," *IEEE Trans. AES*, Vol. 30, No. 3, pp. 660-670, July 1994.

[5] W.F. Gabriel, "Using Spectral Estimation Techniques in Adaptive Processing Antenna Systems," *IEEE Trans. AP*, Vol. 34, No. 3, pp. 291-300, March 1986.

[6] B.D. Van Veen, "Eigenstructure based partially adaptive array design," *IEEE Trans. AP*, Vol. 36, No. 3, pp. 357-362, March 1988.

[7] A. Haimovich, "The Eigencanceler: Adaptive Radar by Eigenanalysis Methods," *IEEE Trans. AES*, Vol. 32, No. 2, pp. 532-542, April 1996.

[8] J.S. Goldstein and I.S. Reed, "Theory of partially adaptive radar," IEEE Trans. AES, Vol. 33, No. 4, pp. 1309-1325, Oct. 1997.

[9] R. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," IEEE Trans. IT, Vol. IT-18, No. 6, pp. 725-730, Nov 1972.

[10] G.H. Golub and C.F. Van Loan, *Matrix Computations, 3rd Ed.*, Johns Hopkins Univ. Press, Baltimore, 1997.