

# Sample Support Requirements for Nonlinear Adaptive Beamforming

Peter Vouras

Naval Research Laboratory, Washington, DC 20375

**Abstract**—In theory, nonlinear adaptive beamforming techniques which form an extended virtual array from a sparsely populated physical array offer significantly enhanced adaptive degrees of freedom towards mitigating environmental interference as opposed to conventional linear beamformers. A neglected practical issue however, central to the implementation of nonlinear beamformers in radar, is an accurate determination of the data support necessary for computing the adaptive weights and the adapted array output. This paper investigates the data support requirements for a new nonlinear beamformer operating on the streaming output of a multichannel array. The array exists within a quasi-stationary or slowly varying environment where more interference sources are present than array elements.

## I. ANALYTICAL BEAMFORMER: KNOWN COVARIANCE MATRIX AND INTERFERENCE DIRECTIONS

This paper considers the effects of finite sample support on nonlinear adaptive beamforming as applied to a nested linear array architecture [1]. Nested linear arrays are sparse arrays formed by concatenating two or more uniform linear arrays (ULAs) to create a single array of  $N$  elements with increasing inter-sensor spacing. One example of a two-level nested linear array is an array of length 12 with 6 elements located at the positions  $\{0, 1, 2, 3, 7, 11\}$ . The smallest spacing between the elements of a nested array corresponds to one-half the wavelength of the highest received frequency to avoid spatial aliasing.

To describe the adapted output of the ideal nonlinear beamformer assume that all signal sources are mutually uncorrelated and arriving from distinct directions. Consider  $K$  interference sources impinging on the array and denote by  $x_n[p]$  the output of the  $n$ th array element at time instant  $p$ . Let  $s_k[p]$  be the signal emitted by the  $k$ th source from direction  $\theta_k$ . The noise  $n_n[p]$  is assumed wide-sense stationary (WSS) with an unknown spatial covariance matrix and is statistically independent from the interference sources. Define  $\mathbf{x}[p] = [x_1[p], \dots, x_N[p]]^T$ ,  $\mathbf{s}[p] = [s_1[p], \dots, s_K[p]]^T$  and  $\mathbf{n}[p] = [n_1[p], \dots, n_N[p]]^T$ . The received signal can be written as

$$\mathbf{x}[p] = \mathbf{A}\mathbf{s}[p] + \mathbf{n}[p] \quad (1)$$

where  $\mathbf{A} = [\mathbf{a}(\theta_1) \ \mathbf{a}(\theta_2) \ \dots \ \mathbf{a}(\theta_K)]$  is the array manifold matrix. The  $N$ -by-1 steering vector  $\mathbf{a}(\theta)$  is

$$\mathbf{a}(\theta) = \left[ 1 \quad e^{j(2\pi/\lambda)r_1 \sin \theta} \quad \dots \quad e^{j(2\pi/\lambda)r_{N-1} \sin \theta} \right]^T \quad (2)$$

where  $r_i$  denotes the position of the  $i$ th sensor which is assumed to be an integer multiple of  $\lambda/2$ .

In general, the covariance matrix for a nonstationary process is  $\mathbf{R}(p, l) = E[\mathbf{x}[p]\mathbf{x}[p-l]^H]$ . In this paper it is assumed that the radar environment is WSS on a sufficiently short time scale and the data covariance matrix for one such interval is defined as

$$\mathbf{R} = E[\mathbf{x}[p]\mathbf{x}[p]^H] = \mathbf{A}\mathbf{S}\mathbf{A}^H + \mathbf{C} \quad (3)$$

where  $\mathbf{C} = E[\mathbf{n}[p]\mathbf{n}[p]^H]$  and  $\mathbf{S}$  is a diagonal matrix of signal powers  $\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2$ . Vectorizing the covariance matrix  $\mathbf{R}$  yields

$$\begin{aligned} \mathbf{z} &= \text{vec}(\mathbf{R}) = \text{vec}(\mathbf{A}\mathbf{S}\mathbf{A}^H) + \text{vec}(\mathbf{C}) \\ &= (\mathbf{A}^* \odot \mathbf{A}) \mathbf{d} + \text{vec}(\mathbf{C}) \end{aligned} \quad (4)$$

where

$$\mathbf{A}^* \odot \mathbf{A} = [\mathbf{a}(\theta_1)^* \otimes \mathbf{a}(\theta_1) \ \dots \ \mathbf{a}(\theta_K)^* \otimes \mathbf{a}(\theta_K)] \quad (5)$$

is the Khatri-Rao product of the matrices  $\mathbf{A}^*$  and  $\mathbf{A}$  and  $\otimes$  represents the Kronecker product. The vector  $\mathbf{d} = [\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2]^T$ . The distinct entries in the columns of  $\mathbf{A}^* \odot \mathbf{A}$  correspond to steering vector components of a virtual array whose element locations are given by the distinct values in the set  $\{r_i - r_j : 1 \leq i, j \leq N\}$ . The virtual array is known as the difference co-array and has length equal to  $(N^2 - 2)/2 + N$  for  $N$  even. For a two-level nested array with six elements at positions  $\{0, 1, 2, 3, 7, 11\}$  the difference co-array is a ULA with 23 elements spaced  $\lambda/2$  apart. The adapted output of the nonlinear beamformer is formed by determining a length  $N^2$  weight vector  $\mathbf{w}$  and computing the inner product  $\mathbf{w}^H \mathbf{z}$ .

### A. Derivation of Analytical Beamformer Solution

The desired beamformer satisfies specified null constraints in the directions  $\theta_k$  of interferences and a mainbeam constraint in the desired pointing direction  $\theta_0$  while simultaneously minimizing the integrated sidelobe level of the adapted antenna pattern. The objective function to be optimized is to minimize

$$J(\mathbf{w}) = \sum_{j=0}^{L-1} |\mathbf{w}^H (\mathbf{a}(\theta_j)^* \otimes \mathbf{a}(\theta_j))|^2 \quad (6)$$

such that

$$\mathbf{G}^H \mathbf{w} = \begin{bmatrix} (\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0))^H \\ (\mathbf{a}(\theta_1)^* \otimes \mathbf{a}(\theta_1))^H \\ \vdots \\ (\mathbf{a}(\theta_K)^* \otimes \mathbf{a}(\theta_K))^H \end{bmatrix} \mathbf{w} = \mathbf{b} \quad (7)$$

where the  $L$  angles  $\theta_j$  in the summation are sampled on a discrete grid in the sidelobe region of the antenna pattern. Define the Hermitian matrix  $\mathbf{Q}$  to be

$$\mathbf{Q} \equiv \sum_{j=0}^{L-1} [\mathbf{a}(\theta_j)^* \otimes \mathbf{a}(\theta_j)] [\mathbf{a}(\theta_j)^* \otimes \mathbf{a}(\theta_j)]^H \quad (8)$$

and the optimization program becomes

$$\begin{aligned} & \text{minimize} && J(\mathbf{w}) = \mathbf{w}^H \mathbf{Q} \mathbf{w} \\ & \text{such that} && \mathbf{G}^H \mathbf{w} = \mathbf{b}. \end{aligned} \quad (9)$$

The constraint matrix  $\mathbf{G}$  has linearly independent columns. Note that the standard analytic solution  $\mathbf{w}_{OPT} = \mathbf{Q}^{-1} \mathbf{G} [\mathbf{G}^H \mathbf{Q}^{-1} \mathbf{G}]^{-1} \mathbf{b}$  for quadratic functions with linear constraints is not applicable here because the matrix  $\mathbf{Q}$  is not full rank since each vector  $\mathbf{a}(\theta_j)^* \otimes \mathbf{a}(\theta_j)$  contains redundant components. In this section an analytic solution for the program in (9) will be derived using the method of Lagrange multipliers.

Adjoining the constraints in (9) to the objective function yields the real-valued Lagrangian,

$$H(\mathbf{w}) = \frac{1}{2} \mathbf{w}^H \mathbf{Q} \mathbf{w} + \lambda_R^T \Re\{\mathbf{G}^H \mathbf{w} - \mathbf{b}\} + \lambda_I^T \Im\{\mathbf{G}^H \mathbf{w} - \mathbf{b}\}. \quad (10)$$

The factor of 1/2 is included to simplify the arithmetic. Minimizing the Lagrangian transforms the original optimization program with constraints into an unconstrained minimization. The optimum solution is obtained by taking the gradient of  $H(\mathbf{w})$  with respect to the real and imaginary parts of  $\mathbf{w} = \mathbf{w}_R + j\mathbf{w}_I$  and then setting the resulting quantities equal to zero. This step yields,

$$\frac{\partial H(\mathbf{w})}{\partial \mathbf{w}_R} = \frac{1}{2} \mathbf{w}^H \mathbf{Q} + \lambda_R^T \Re\{\mathbf{G}^H\} + \lambda_I^T \Im\{\mathbf{G}^H\} = \mathbf{0} \quad (11)$$

$$\frac{\partial H(\mathbf{w})}{\partial \mathbf{w}_I} = \frac{1}{2} j \mathbf{w}^H \mathbf{Q} - \lambda_R^T \Im\{\mathbf{G}^H\} + \lambda_I^T \Re\{\mathbf{G}^H\} = \mathbf{0}$$

Defining the derivative of  $H(\mathbf{w})$  with respect to  $\mathbf{w}$  as [2]

$$\frac{\partial H(\mathbf{w})}{\partial \mathbf{w}} = \frac{\partial H(\mathbf{w})}{\partial \mathbf{w}_R} - j \frac{\partial H(\mathbf{w})}{\partial \mathbf{w}_I} \quad (12)$$

and the complex  $(K+1) \times 1$  Lagrange multiplier vector  $\lambda = \lambda_R + j\lambda_I$  yields

$$\mathbf{Q} \mathbf{w} = -\mathbf{G} \lambda. \quad (13)$$

Next the singular value decomposition (SVD) of  $\mathbf{Q}$  is computed. The rank of  $\mathbf{Q}$ , denoted  $r$ , is equal to the length of the difference co-array associated with the nested array of physical elements. For example, if the nested array consists of 6 elements, then  $r = 23$ .

$$\mathbf{Q} = [\mathbf{U}_1 \quad \mathbf{U}_2] \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^H \\ \mathbf{V}_2^H \end{bmatrix}. \quad (14)$$

The matrix sizes are;  $\mathbf{U}_1$  is  $N^2 \times r$ ,  $\mathbf{U}_2$  is  $N^2 \times (N^2 - r)$ ,  $\mathbf{V}_1$  is  $N^2 \times r$ ,  $\mathbf{V}_2$  is  $N^2 \times (N^2 - r)$ , and  $\Sigma$  is  $r \times r$ . The matrix  $\mathbf{V}_1 \mathbf{V}_1^H$  is the projection onto  $N(\mathbf{Q})^\perp = R(\mathbf{Q}^H)$  and

the matrix  $\mathbf{V}_2 \mathbf{V}_2^H$  is the projection onto  $N(\mathbf{Q})$  [3]. Here,  $R$  denotes the range space of a matrix,  $N$  the null space, and  $^\perp$  the orthogonal complement. Decomposing the weight vector  $\mathbf{w}$  into

$$\mathbf{w} = \mathbf{V}_1 \mathbf{V}_1^H \mathbf{w} + \mathbf{V}_2 \mathbf{V}_2^H \mathbf{w} \quad (15)$$

and substituting into (13) produces

$$\mathbf{V}_1 \mathbf{V}_1^H \mathbf{w} = -\mathbf{V}_1 \Sigma^{-1} \mathbf{U}_1^H \mathbf{G} \lambda. \quad (16)$$

Since  $\mathbf{Q} \mathbf{V}_2 \mathbf{V}_2^H \mathbf{w} = \mathbf{0}$ , the orthogonal component of  $\mathbf{w}$ ,  $\mathbf{V}_2 \mathbf{V}_2^H \mathbf{w}$ , does not contribute to the solution. Consequently, the optimal weight vector can now be set equal to

$$\mathbf{w}_{OPT} = -\mathbf{V}_1 \Sigma^{-1} \mathbf{U}_1^H \mathbf{G} \lambda_{OPT} \quad (17)$$

where  $\lambda_{OPT}$  is the corresponding optimal Lagrange multiplier vector. The astute reader will recognize the matrix  $\mathbf{V}_1 \Sigma^{-1} \mathbf{U}_1^H$  as the pseudoinverse of  $\mathbf{Q}$ . The vector  $\lambda_{OPT}$  is determined by applying the constraint equation  $\mathbf{G}^H \mathbf{w} = \mathbf{b}$ . In most practical applications, the  $(K+1) \times (K+1)$  matrix  $\mathbf{G}^H \mathbf{V}_1 \Sigma^{-1} \mathbf{U}_1^H \mathbf{G}$  is invertible so the final result becomes

$$\mathbf{w}_{OPT} = \mathbf{V}_1 \Sigma^{-1} \mathbf{U}_1^H \mathbf{G} [\mathbf{G}^H \mathbf{V}_1 \Sigma^{-1} \mathbf{U}_1^H \mathbf{G}]^{-1} \mathbf{b}. \quad (18)$$

## II. PROPOSED CLOSED-LOOP BEAMFORMER STEP 1; DOA ESTIMATION

This section summarizes an approach described in [4] for estimating the directions of arrival (DOAs) for environmental interference sources. Once the directions of the interference sources are estimated, the adaptive beamformer weights can be computed using an algorithm proposed in Section III. The interference sources are assumed to be mutually uncorrelated and arriving from distinct directions. A target signal is assumed not present in the received data.

Define a local covariance matrix valid for a block of  $L$  vector samples as

$$\begin{aligned} \mathbf{R}_m &= E[\mathbf{x}[p]\mathbf{x}[p]^H] \text{ for } (m-1)L \leq p \leq mL-1 \\ &= \mathbf{A} \mathbf{S}_m \mathbf{A}^H + \mathbf{C} \end{aligned} \quad (19)$$

where  $\mathbf{S}_m$  is a diagonal matrix of signal powers  $\sigma_{m1}^2, \sigma_{m2}^2, \dots, \sigma_{mK}^2$  at frame  $m$ . The first step in the proposed beamformer is to estimate the signal directions  $\theta_1, \dots, \theta_K$  from a sequence of  $M$  local covariance matrices  $\mathbf{R}_1, \dots, \mathbf{R}_M$  without knowledge of the spatial noise covariance matrix  $\mathbf{C}$  or the signal covariances  $\mathbf{S}_m$ . Vectorizing the local covariance matrix  $\mathbf{R}_m$  yields

$$\mathbf{z}_m = \text{vec}(\mathbf{R}_m) = (\mathbf{A}^* \odot \mathbf{A}) \mathbf{d}_m + \text{vec}(\mathbf{C}) \quad (20)$$

where the vector  $\mathbf{d}_m = [\sigma_{m1}^2, \sigma_{m2}^2, \dots, \sigma_{mK}^2]^T$ . Stacking  $M$  vectors  $\mathbf{z}_m$  into a data matrix  $\mathbf{Y} = [\mathbf{z}_1, \dots, \mathbf{z}_M]$  yields

$$\mathbf{Y} = (\mathbf{A}^* \odot \mathbf{A}) \Psi^T + \text{vec}(\mathbf{C}) \mathbf{1}_M^T \quad (21)$$

where  $\Psi = [\mathbf{d}_1, \dots, \mathbf{d}_M]^T$  and the  $M$ -by-1 vector  $\mathbf{1}_M$  consists of  $M$  ones. It is assumed that the rank of  $\Psi$  is  $K$ . This condition is satisfied in practice if the temporal powers of the signal sources are slightly varying due to platform motion or changes in aspect angle. Changes in signal angles of arrival

are assumed to occur over a longer time scale than temporal variations in the signal powers primarily due to the distance between the signal sources and the radar.

Next the deterministic noise term  $\text{vec}(\mathbf{C})$  is eliminated from  $\mathbf{Y}$  by forming the projection matrix

$$\mathbf{P}_{\mathbf{1}_M}^\perp = \mathbf{I}_M - (1/M)\mathbf{1}_M\mathbf{1}_M^T. \quad (22)$$

Multiplying  $\mathbf{P}_{\mathbf{1}_M}^\perp$  and  $\mathbf{Y}$  produces the noise-free data matrix

$$\mathbf{Y}\mathbf{P}_{\mathbf{1}_M}^\perp = (\mathbf{A}^* \odot \mathbf{A})(\mathbf{P}_{\mathbf{1}_M}^\perp \boldsymbol{\Psi})^T. \quad (23)$$

Since the angles  $\theta_i$  are assumed distinct,  $(\mathbf{A}^* \odot \mathbf{A})$  has full column rank. Computing the SVD yields

$$\mathbf{Y}\mathbf{P}_{\mathbf{1}_M}^\perp = [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^H \\ \mathbf{V}_n^H \end{bmatrix}. \quad (24)$$

The dimensions of  $\mathbf{U}_s$  are  $N^2 \times K$ ,  $\mathbf{U}_n$  is  $N^2 \times (N^2 - K)$ ,  $\mathbf{V}_s$  is  $M \times K$ ,  $\mathbf{V}_n$  is  $M \times (N^2 - K)$ , and  $\boldsymbol{\Sigma}$  is a  $K \times K$  diagonal matrix. Note that by the properties of the SVD,

$$R^\perp(\mathbf{A}^* \odot \mathbf{A}) = R^\perp(\mathbf{U}_s) = R(\mathbf{U}_n) \quad (25)$$

and the matrix  $\mathbf{P}_n = \mathbf{U}_n\mathbf{U}_n^H$  is the projection matrix onto  $R^\perp(\mathbf{A}^* \odot \mathbf{A})$ . Since the source and noise subspaces are orthogonal, the  $k$ th source DOA satisfies

$$\mathbf{r}(\theta_k) = \mathbf{U}_n^H [\mathbf{a}(\theta_k)^* \otimes \mathbf{a}(\theta_k)] = \mathbf{0}, \quad k = 1, \dots, K \quad (26)$$

and the proposed criterion for estimating each DOA  $\theta_k$  becomes

$$\begin{aligned} &\text{find } \theta \\ &\text{such that } \|\mathbf{r}(\theta)\|^2 = 0. \end{aligned} \quad (27)$$

#### A. Finite Sample Effects

The effects of finite sample support for linear beamformers acting on the streaming output of a radar receiver have been well studied in the literature [5]–[7]. These analyses however do not adequately address the data support requirements for nonlinear beamformers. This paper provides some preliminary results illustrating the effects of finite data support on nonlinear adaptive processing by considering the performance of the beamformer proposed in the next section. This closed-loop beamformer operates in two steps. The first step uses the estimated DOAs  $\hat{\theta}_k$  of interference to compute an adaptive weight vector  $\mathbf{w}$  and the second step forms the adapted array output  $\mathbf{w}^H \mathbf{z}$ .

To begin the analysis, assume that only estimated local covariance matrices  $\hat{\mathbf{R}}_m$  are available and that the second order statistics of the received radar signal are slowly varying. Under this scenario, the available data support becomes very important since the signal covariance matrix  $\hat{\mathbf{S}}_m$  implicit in the estimate  $\hat{\mathbf{R}}_m$  may no longer be strictly diagonal even if the signal sources are perfectly uncorrelated since the second order statistics of the received signals are computed from a finite number of samples.

Define the estimated local covariance matrix  $\hat{\mathbf{R}}_m$  to be

$$\hat{\mathbf{R}}_m = \frac{1}{L} \sum_{p=(m-1)L}^{mL-1} \mathbf{x}[p]\mathbf{x}[p]^H \quad (28)$$

with  $\hat{\mathbf{z}}_m = \text{vec}(\hat{\mathbf{R}}_m)$ . Form a data matrix  $\hat{\mathbf{Y}} = [\hat{\mathbf{z}}_1, \dots, \hat{\mathbf{z}}_M]$  consisting of  $M$  vectorized local covariance matrix estimates and compute the SVD of the noise-free data matrix as in

$$\hat{\mathbf{Y}}\mathbf{P}_{\mathbf{1}_M}^\perp = [\hat{\mathbf{U}}_s \quad \hat{\mathbf{U}}_n] \begin{bmatrix} \hat{\boldsymbol{\Sigma}} & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{\Delta}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{V}}_s^H \\ \hat{\mathbf{V}}_n^H \end{bmatrix} \quad (29)$$

where  $\hat{\boldsymbol{\Delta}}$  corresponds to a diagonal submatrix of small scalar values. Now the criterion for estimating the angles  $\theta_k$  becomes

$$\begin{aligned} &\text{find } \theta \\ &\text{such that } \|\hat{\mathbf{U}}_n^H [\mathbf{a}(\theta)^* \otimes \mathbf{a}(\theta)]\|^2 \leq \epsilon \end{aligned} \quad (30)$$

where  $\epsilon$  represents a detection threshold.

### III. PROPOSED CLOSED-LOOP BEAMFORMER STEP 2; PROJECTED GRADIENT ITERATIONS

To avoid the matrix inversion computations required in the analytical solution (18) an iterative algorithm is sometimes preferred. Given an initial weight vector  $\mathbf{w}_0$  that satisfies the desired null and mainbeam constraints, the optimization program in (9) can be solved iteratively by projecting the gradient of  $J(\mathbf{w})$  onto the orthogonal complement of the constraint subspace [8]. In other words,

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu \nabla^\perp J(\mathbf{w}_k) \quad (31)$$

where  $\mu$  is a small step-size parameter that ensures convergence and

$$\nabla^\perp J(\mathbf{w}_k) = \mathbf{P}^\perp \nabla J(\mathbf{w}_k). \quad (32)$$

The matrix  $\mathbf{P}^\perp = \mathbf{I} - \mathbf{G}^H(\mathbf{G}\mathbf{G}^H)^{-1}\mathbf{G}$  is the projection matrix onto the orthogonal complement of the constraint subspace. Note that for

$$\begin{aligned} J(\mathbf{w}_k) &= \frac{1}{2} \mathbf{w}_k^H \mathbf{Q} \mathbf{w}_k, \\ \nabla J(\mathbf{w}_k) &= \mathbf{Q} \mathbf{w}_k. \end{aligned} \quad (33)$$

#### A. Initial Condition A

One method for computing an initial weight vector  $\mathbf{w}_0$  after the interference directions have been estimated is to solve a system of equations. For example, if the mainbeam is pointed towards  $\theta_0$  and  $K$  jammers arrive from the directions  $\hat{\theta}_k$  for  $k = 1, \dots, K$ , then  $\mathbf{w}$  should be chosen as the solution to

$$\mathbf{L} \mathbf{w} = \begin{bmatrix} (\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0))^H \\ (\mathbf{a}(\hat{\theta}_1)^* \otimes \mathbf{a}(\hat{\theta}_1))^H \\ \vdots \\ (\mathbf{a}(\hat{\theta}_K)^* \otimes \mathbf{a}(\hat{\theta}_K))^H \end{bmatrix} \mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{b}. \quad (34)$$

The system of equations in (34) is under-determined if  $K < N^2$ , in which case infinitely many solutions are possible.

The unique minimum norm solution is given by

$$\mathbf{w}_{min} = \mathbf{L}^H(\mathbf{L}\mathbf{L}^H)^{-1}\mathbf{b}. \quad (35)$$

### B. Initial Condition B

This section describes another procedure for initializing the iterated beamformer with a vector  $\mathbf{w}_0$  that satisfies the desired null and mainbeam constraints. Note that the  $K$  null constraints at  $\hat{\theta}_1, \dots, \hat{\theta}_K$  are satisfied if and only if  $\mathbf{w} \in R^\perp(\mathbf{A}^* \odot \mathbf{A})$ , the orthogonal complement of the interference subspace. Recall that  $\mathbf{A}^* \odot \mathbf{A}$  does not contain  $\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0)$ . Thus, there exists a vector  $\mathbf{v}$  such that

$$\mathbf{w} = \hat{\mathbf{U}}_n \mathbf{v} \quad (36)$$

and  $(\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0))^H \mathbf{w} \neq 0$ . This implies that

$$\mathbf{L} \hat{\mathbf{U}}_n \mathbf{v} = \mathbf{b} \quad (37)$$

for some  $\mathbf{v}$ . Let  $\mathbf{f}^H = (\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0))^H \hat{\mathbf{U}}_n$ . Then,

$$\begin{bmatrix} \mathbf{f}^H \\ \mathbf{0}^H \\ \vdots \\ \mathbf{0}^H \end{bmatrix} \mathbf{v} = \mathbf{b} \quad (38)$$

which implies that

$$\mathbf{w} = \frac{\hat{\mathbf{P}}_n(\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0))}{(\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0))^H \hat{\mathbf{P}}_n(\mathbf{a}(\theta_0)^* \otimes \mathbf{a}(\theta_0))} \quad (39)$$

where  $\hat{\mathbf{P}}_n = \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H$ . A similar linear beamformer solution is described in [9] but derived using a very different approach based on polynomials.

### C. Polak-Ribiere Conjugate Gradient Algorithm

An alternative iterative algorithm which forms Polak-Ribiere conjugate directions from the projected gradient may increase the convergence rate. This algorithm is listed below. Figure 1 illustrates adapted beampatterns created using the conjugate gradient algorithm with initial condition (B) for a two-level nested array. For this case, the data matrix  $\Psi$  was constructed as

$$\Psi = [\mathbf{d}_1, \dots, \mathbf{d}_1]^T + \nu \mathbf{D} \quad (40)$$

where  $\nu = 10^{-2}$  and  $\mathbf{D}$  consists of random entries uniformly distributed in the interval  $(0, 1)$ . The random matrix  $\nu \mathbf{D}$  is meant to model a slowly changing radar environment and guarantees that the rank of  $\Psi$  equals  $K$ . The vector  $\mathbf{d}_1 = [3 \ 10 \ 20 \ 15 \ 5 \ 30 \ 25 \ 5 \ 12]^T$  contains the interference powers (shown in dB) for jammers located at  $\{-53^\circ, -40^\circ, -26^\circ, -20^\circ, -10^\circ, 10^\circ, 15^\circ, 33^\circ, 47^\circ\}$ . The step size constant  $\lambda_j = 10^{-4}$  for all iterations. The mainbeam is steered to  $0^\circ$ .

Observe that Fig. 1 illustrates the progression of the adapted beampattern over 100 iterations of the algorithm. The curve labeled  $\mathbf{w}_0$  is the adapted beampattern using the initial weight vector defined in (39). The curve labeled  $\mathbf{w}_{100}$  is the adapted

beampattern after 100 iterations and the red curve corresponds to the analytical beamformer computed using (18). The circle at  $0^\circ$  represents the gain of the uniformly weighted nested array (equal to 7.78 dBi). As intended, the  $\mathbf{w}_{100}$  curve shows a significant reduction in the ambient sidelobe level (at the expense of a wider mainbeam). In this case, an interesting result is that the initial weight vector  $\mathbf{w}_0$  is the same with either initializing procedure (A) or (B). In general however, for larger values of  $\nu$ ,  $\mathbf{w}_0$  will be different for each initial condition.

---

Algorithm 1. Conjugate Gradient Algorithm to Compute Weight Vector

---

**Require:** Initial weight vector  $\mathbf{w}_0$  that satisfies constraints.

- 1:  $\mathbf{h}_0 = -\nabla^\perp J(\mathbf{w}_0)$
- 2:  $\mathbf{g}_0 = \mathbf{h}_0$
- 3: For each iteration  $j = 0, 1, \dots$  compute a step size  $\lambda_j$  using any line optimization routine. Alternatively, set  $\lambda_j$  to a small value that ensures convergence.
- 4:  $\mathbf{w}_{j+1} = \mathbf{w}_j + \lambda_j \mathbf{h}_j$
- 5:  $\mathbf{g}_{j+1} = -\nabla^\perp J(\mathbf{w}_{j+1})$
- 6:  $\mu_j = \frac{(\mathbf{g}_{j+1} - \mathbf{g}_j)^T \mathbf{g}_{j+1}}{\|\mathbf{g}_j\|^2}$
- 7:  $\mathbf{h}_{j+1} = \mathbf{g}_{j+1} + \mu_j \mathbf{h}_j$
- 8:  $j = j + 1$
- 9: **Go to 3**

---

## IV. SIMULATED RESULTS

### A. Simulation Framework

Simulated results are based on a radar that utilizes a two-level nested linear array antenna with an independent receiver channel behind each array element. The processing for each array element channel assumes the received signal is perfectly downconverted to baseband. It is then sampled by an analog-to-digital converter (ADC) and convolved with a matched filter. The matched filter output from each array element sampled at the target range bin is used to form an array steering vector corresponding to each received pulse. The received waveform is a linear chirp with a pulse width of 3.2  $\mu\text{sec}$  and a swept bandwidth of 10 MHz. The ADC sampling rate is 40 MHz and the initial phase of the received chirp at each array element is set equal to the interelement phase shift given by (2). The energy of each radar pulse is normalized such that the waveform ambiguity function evaluated at the origin (corresponding to zero target delay and zero target Doppler shift) is equal to unity.

### B. Effect of Sample Support on DOA Estimation

Figure 2 illustrates the effect of finite sample support on DOA estimation performance. DOA estimation was performed using data samples before pulse compression. The plot shows the DOA search function defined in (30) for different values of  $L$  in (28) and the number  $M$  of vectorized local covariance matrix estimates. Figure 2 demonstrates that increasing the total number of data snapshots improves the accuracy of DOA estimates by making the nulls in the DOA search function deeper and more precise. The jammer locations were

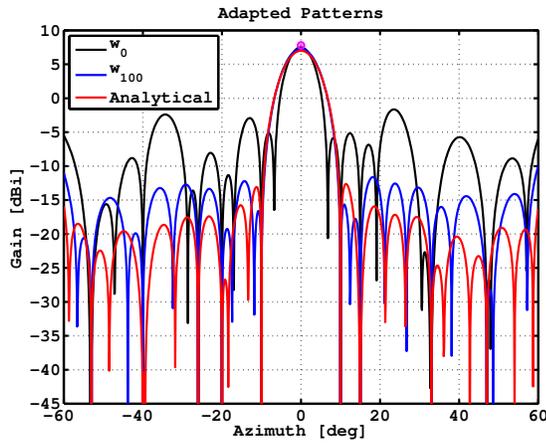


Fig. 1. Adapted Patterns; Initial Condition A or B

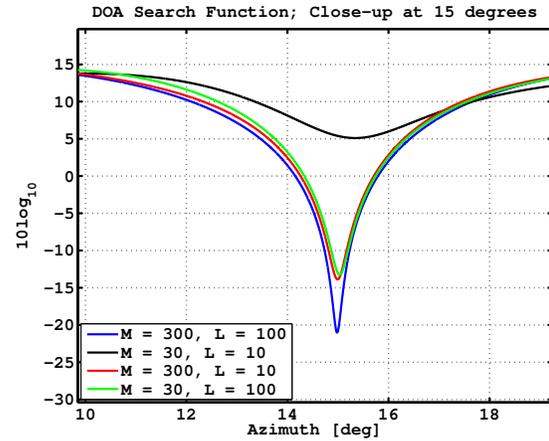


Fig. 3. DOA Search Function; Close-Up

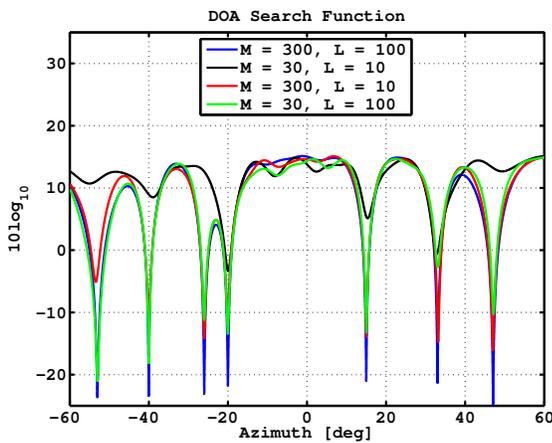


Fig. 2. DOA Search Function

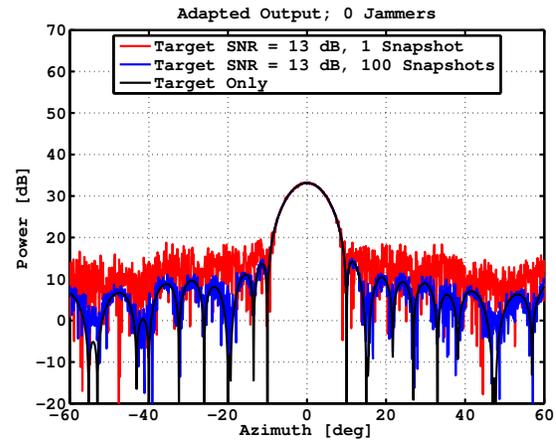


Fig. 4. Adapted Output; SNR 13 dB, 0 Jammers

specified at  $\{-53^\circ, -40^\circ, -26^\circ, -20^\circ, 15^\circ, 33^\circ, 47^\circ\}$  and the jamming-to-noise ratio (JNR) for each jammer was randomly chosen between 0 and 30 dB after every interval of 40 pulses. Figure 3 is a close-up of the DOA search function in the vicinity of  $15^\circ$ .

### C. Effect of Sample Support on Adapted Array Output

For these results, the adaptive weight vector  $\mathbf{w}$  was assumed to be known apriori and the vectorized estimated covariance matrix  $\hat{\mathbf{z}}$  includes a target signal. Figure 4 illustrates  $10\log_{10}|\mathbf{w}^H \hat{\mathbf{z}}|$  for the case where the single pulse, target signal-to-noise ratio (SNR) at a single array element is 13 dB. The black curve corresponds to the adapted output when only the target signal is present without noise. The plot shows that with just a single data snapshot the adapted output is very noisy. However, by using  $L = 100$  pulses to form  $\hat{\mathbf{z}}$ , the adapted output more closely follows the desired beampattern. Figure 5 illustrates similar results for a single snapshot ( $L = 1$ ) with the target SNR equal to 30 dB.

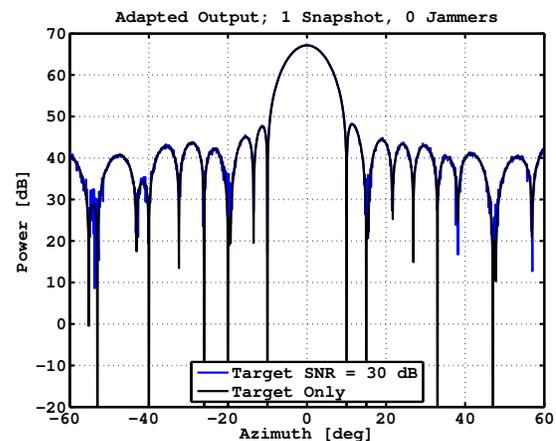


Fig. 5. Adapted Output; SNR 30 dB, 0 Jammers

Figures 6 and 7 illustrate the case where a single jammer is present at  $-53^\circ$ . The total input JNR is 15 dB. Figures

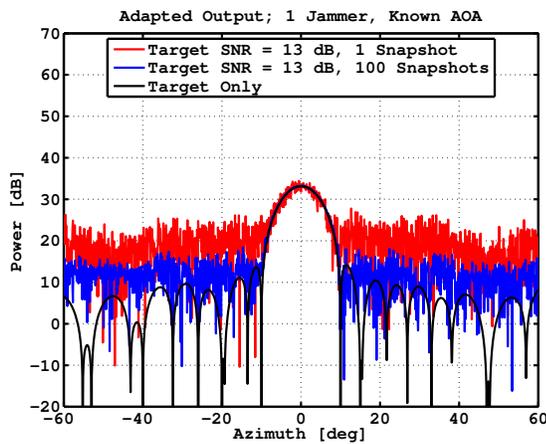


Fig. 6. Adapted Output; SNR 13 dB, 1 Jammer

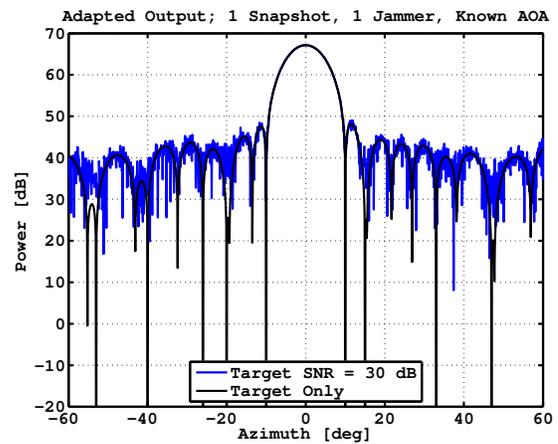


Fig. 7. Adapted Output; SNR 30 dB, 1 Jammer

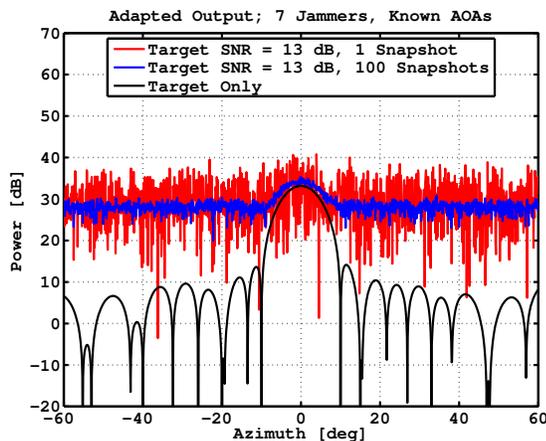


Fig. 8. Adapted Output; SNR 13 dB, 7 Jammers

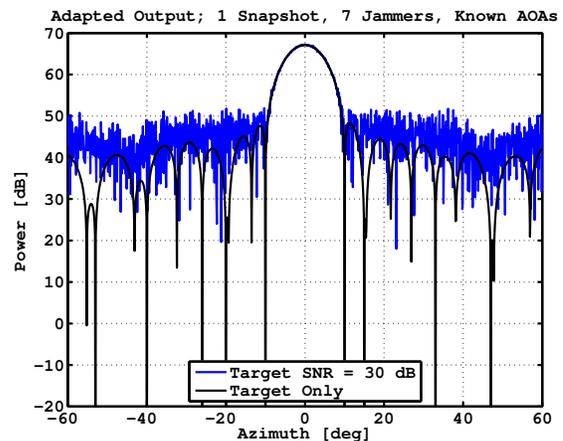


Fig. 9. Adapted Output; SNR 30 dB, 7 Jammers

8 and 9 illustrate the case where 7 jammers are located at  $\{-53^\circ, -40^\circ, -26^\circ, -20^\circ, 15^\circ, 33^\circ, 47^\circ\}$  and the JNR of each jammer is  $\{20, 15, 5, 30, 25, 5, 12\}$  dB respectively. The total JNR at each array element is now 31.75 dB.

## V. CONCLUSION

This paper considers the impact of sample support size on nonlinear adaptive beamforming in radar. A new nonlinear beamformer is proposed which operates on the streaming output of a multichannel array. Simulation results suggest that the nulling performance of a closed-loop nonlinear beamformer is especially sensitive to sample size at low SNRs.

## VI. ACKNOWLEDGEMENTS

This work was carried out under the Naval Research Laboratory Base Program sponsored by the Office of Naval Research.

## REFERENCES

[1] P. Pal and P. P. Vaidyanathan, "Nested Arrays: A Novel Approach to Array Processing with Enhanced Degrees of Freedom," *IEEE Transactions on Signal Processing*, Vol. 58, No. 8, August 2010.

[2] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1996.

[3] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: The Johns Hopkins University Press, 1996.

[4] W. Ma, T. Hsieh, and C. Chi, "DOA Estimation of Quasi-Stationary Signals via Khatri-Rao Subspace," Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Taipei, Taiwan, April 19-24, 2009.

[5] 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.

[6] D. M. Boroson, "Sample Size Considerations for Adaptive Arrays," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-16, No. 4, pp. 446-451, July 1980.

[7] B. D. Carlson, "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 24, No. 4, pp. 397-401, July 1988.

[8] D. Day, "Robust Phase-Only Nulling for Adaptive and Non-Adaptive Phased Arrays," Proceedings 2007 Asilomar Conference on Signal, Systems and Computers, Monterey, CA, November 4-7, 2007.

[9] Y. Bresler, V. U. Reddy, T. Kailath, "Optimum Beamforming for Coherent Signal and Interferences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 36, No. 6, June 1988.