SOURCE LOCALIZATION USING ADAPTIVE SUBSPACE BEAMFORMER OUTPUTS

Edward J. Baranoski *James Ward*

MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02173
ejb@ll.mit.edu, jward@ll.mit.edu

ABSTRACT

Maximum likelihood (ML) parameter estimation for multi-dimensional adaptive problems is addressed. Multiple adaptive outputs are ordinarily combined by utilizing the full dimension data. However, many adaptive problems utilize subspace processing for each adaptive beam which can increase the difficulty of many super-resolution techniques. This paper shows that steering vector structure can be utilized to allow ML techniques for a fixed grid of hypothesis vectors to be computationally feasible for many scenarios.

1. INTRODUCTION

Many adaptive array applications now require adaptivity in multiple dimensions, such as planar arrays or space-time adaptive processing (STAP) algorithms for airborne nulling of clutter [1]. Because of the increasing size of these problems, subspace array processing is often used to reduce the degrees of freedom with each adaptive weight vector often using a different subspace. This reduces a large dimension problem into many smaller problems, but it can make parameter estimation more difficult. This paper addresses multi-dimensional parameter estimation using adaptive outputs where each adaptive output may have been formed from a different subspace.

When parameter estimation of source locations is done independently in each dimension of a multi-dimensional system, the estimates can be heavily biased. Multi-dimensional parameter estimation has been shown to provide much more accurate source locations [2, 3]. Maximum likelihood (ML) source localization [4] is well understood but is often deemed too computationally expensive for practical use since the computational load goes up exponentially with multiple dimensions. Several alternative techniques have been developed to offer multi-dimensional parameter estimation, such as adaptive monopulse radar techniques [2] and a rooting algorithm [5]. However, these algorithms generally generally require access to the full dimension data and are not able to utilize the subspace adaptive beamformed outputs. In addition, monopulse techniques rely on a first-order Taylor expansion at the beam center and are biased away from this point. This paper re-examines maximum likelihood estimation using these varied subspace adaptive beamformer outputs. Section 2 details the theory behind adaptive subspace ML multi-dimensional parameter estimation, including simplifications when the steering vectors can be decomposed as the Kronecker product between vectors in each dimension. In Section 3, this technique is demonstrated for an example STAP application to provide efficient two-dimensional azimuth-Doppler target localization in the presence jammer and clutter for an airborne radar system.

This work was supported in part by Air Force contract F19628-95-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Air Force.

2. BEAMSPLITTING IN ADAPTIVE BEAMSPACE

A common model in adaptive processing for radar consists of a data snapshot

$$\mathbf{x} = \alpha \mathbf{v}(\Theta) + \mathbf{x}_u \ , \tag{1}$$

where α is the target amplitude and $\mathbf{v}$ is the steering vector for the unknown target location parameter Θ. The interference and noise component of the snapshot, $\mathbf{x}_u$, is assumed to have covariance matrix $\mathbf{R} = E\{\mathbf{x}_u \mathbf{x}_u^H\}$. Typically there are available target free snapshots from which the sample covariance matrix $\widehat{\mathbf{R}}$ is estimated. For this data model the ML estimator for Θ is [4]

$$\widehat{\Theta} = \arg\max_{\Theta} \frac{\left|\mathbf{v}^H(\Theta)\widehat{\mathbf{R}}^{-1}\mathbf{x}\right|^2}{\mathbf{v}^H(\Theta)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\Theta)} = \arg\max_{\Theta} \left|\mathbf{w}^H(\Theta)\mathbf{x}\right| \tag{2}$$

where $\mathbf{w}(\Theta)$ is the AMF weight vector [6]

$$\mathbf{w}(\Theta) = \frac{\widehat{\mathbf{R}}^{-1}\mathbf{v}(\Theta)}{\sqrt{\mathbf{v}^H(\Theta)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\Theta)}}. \tag{3}$$

The ML expression of (2) applies equally well when beamspace processing is used. If the set of beams is denoted by the matrix $\mathbf{T}$, the ML estimator for $\widehat{\theta}$ becomes

$$\widehat{\Theta}_T = \arg\max_{\Theta} \frac{\left|\mathbf{v}_T^H(\Theta)\widehat{\mathbf{R}}_T^{-1}\mathbf{x}_T\right|^2}{\mathbf{v}_T^H(\Theta)\widehat{\mathbf{R}}_T^{-1}\mathbf{v}_T(\Theta)} \tag{4}$$

where the T subscript denotes projection onto $\mathbf{T}$ (i.e., $\mathbf{x}_T = \mathbf{T}^H\mathbf{x}$ and $\widehat{\mathbf{R}}_T = \mathbf{T}^H\widehat{\mathbf{R}}\mathbf{T}$).

In general, we will not have the luxury of computing $\mathbf{w}(\Theta)$ for all Θ hypotheses and must instead utilize a small set of p weight vectors

$$\mathbf{W} = \widehat{\mathbf{R}}^{-1}\mathbf{V} \tag{5}$$

where $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_p]$. Multiple constraint vectors are often used in monopulse radar [7], space-time adaptive processing for nulling clutter [1], or when using derivative constraints on beamshape [8]. These weight vectors can be applied as an adaptive beamspace in (4) to produce

$$\widehat{\Theta} = \arg\max_{\Theta} \frac{\left|\mathbf{v}^H(\Theta)\mathbf{W}\widehat{\mathbf{R}}_W^{-1}\mathbf{W}^H\mathbf{x}\right|^2}{\mathbf{v}^H(\Theta)\mathbf{W}\widehat{\mathbf{R}}_W^{-1}\mathbf{W}^H\mathbf{v}(\Theta)}. \tag{6}$$

$\widehat{\mathbf{R}}_W = \mathbf{W}^H\widehat{\mathbf{R}}\mathbf{W}$ is the $p \times p$ sample covariance matrix of the adaptive beamformer outputs. When the target component of $\mathbf{x}$ is contained within the space of $\mathbf{V}$, (6) will still provide an asymptotically unbiased estimate of target bearing. However, evaluation of (6) requires that the full weight vector be applied to all steering vector hypotheses $\mathbf{v}(\Theta)$. Depending upon the desired resolution of the bearing estimates, this can be a large computational burden particularly when the search space is multi-dimensional such as with a 2-D array or STAP algorithms.

Rooting estimation techniques can experience difficulties when using multiple outputs for estimation since the output filters do not have identical patterns, as is generally the case when the weight vectors are formed from different subspaces. For example, the i-th vector in (5) may be formed using the $\mathbf{T}_i$ subspace, so the subspace weight vector becomes

$$\mathbf{w}_i = \left(\mathbf{T}_i^H\widehat{\mathbf{R}}\mathbf{T}_i\right)^{-1}\mathbf{T}_i^H\mathbf{v}_i. \tag{7}$$

However, the ML estimation of (6) accounts for pattern distortions by explicitly forming the pattern responses $\mathbf{W}^H\mathbf{v}(\Theta)$. Each of the terms in $\mathbf{W}^H\mathbf{v}(\Theta)$ is computed in the beamspace corresponding to the appropriate vector, thus saving computations. Similarly, the outputs $\mathbf{W}^H\mathbf{x}$ are also computed within each beamspace. Changes in the correlation between subspace adaptive weight vectors are adjusted with the $\mathbf{R}_W$ term in (6). Therefore, the benefits of subspace adaptive processing are still preserved with the full ML estimation procedure.

The cost of the direct ML calculation can be prohibitive, but many adaptive radar problems can exploit two simplifications on the steering vector structure to reduce the computation of (6) to something more manageable. These simplifications include representing the steering vector as a Kronecker product and approximating each Kronecker component as a linear combination of a smaller number of component vectors.

First, multi-dimensional arrays on a regular grid layout (such as a rectilinear or rectangular grid) and STAP algorithms (whose space-time sampling forms a rectangular grid) can partition the steering vector as a Kronecker product of two smaller vectors which can dramatically reduce the cost of forming the ML test for all candidate hypotheses. Let the snapshot length be NM, where N and M are the lengths in each of the two dimensions. The steering vector can be written as

$$\mathbf{v}(\Theta) = \mathbf{b}(\omega) \otimes \mathbf{a}(\psi) \tag{8}$$

where $\Theta = (\psi, \omega)$ are the two location parameters, $\mathbf{a}(\psi)$ is the $N \times 1$ steering vector for the ψ dimension, and $\mathbf{b}(\omega)$ is the $M \times 1$ steering vector for the second dimension.

Once the steering vector is presented in a Kronecker form of (8), the $(NM \times p)$ weight matrix $\mathbf{W}$ should be partitioned in a similar way. Let the number of hypothesis vectors $\mathbf{v}(\Theta)$ be mn where n and m are the number of $\mathbf{a}$ and $\mathbf{b}$ vector hypotheses, respectively. Typically radar systems require accuracies to be a small fraction of a beamwidth in each dimension. Thus, ten or twenty hypotheses per beamwidth in each dimension, or a total of several hundred per resolution cell, is typical. This forces all hypothesis weight vectors to lie on a grid in Θ space composed of lines having constant cone angles relative to the the Kronecker component vectors making up the array (a specific implementation of the Kronecker form for hexagonal arrays is given in Appendix A). The weight matrix $\mathbf{W}$ can be broken up into M submatrices $\mathbf{W}_k$ each being of size $N \times p$ so that

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 \\ \vdots \\ \mathbf{W}_M \end{bmatrix}. \tag{9}$$

Each matrix-vector product $\mathbf{W}^H\mathbf{v}$ can then be computed as

$$\mathbf{W}^H\mathbf{v}(\Theta) = \begin{bmatrix} \mathbf{W}_1^H\mathbf{a}(\psi), & \cdots, & \mathbf{W}_M^H\mathbf{a}(\psi) \end{bmatrix}\mathbf{b}(\omega). \tag{10}$$

The bracketed term of (10) need only be computed for each of the $\mathbf{a}$ hypothesis vectors, requiring a total of $pnMN$ operations (where an operation is a complex multiply and addition). Once formed, post-multiplying by each $\mathbf{b}(\omega)$ requires an additional pM operations which must be performed for each of the nm hypothesis combinations of (ψ, ω). The total cost to compute $\mathbf{W}^H\mathbf{v}(\Theta)$ for all hypotheses is $pnM(N+m)$ operations. This is less than the $pmnMN$ operations required to compute all $\mathbf{W}^H\mathbf{v}(\Theta)$ without using the Kronecker simplification by a factor of $mN/(m+N)$ which is bounded by $\min(m,N)/2$ and $\min(m,N)$.

Further reduction in computations can be achieved when the hypothesis vectors in each dimension are approximated as the linear combination of a much smaller number of basis vectors. Let $\mathbf{a}(\psi)$ and $\mathbf{b}(\omega)$ be represented as linear combinations of basis vectors from matrices $\mathbf{A}$ and $\mathbf{B}$, respectively,

$$\hat{\mathbf{a}}(\psi) = \mathbf{A}\mathbf{c}_a(\psi) \quad \hat{\mathbf{b}}(\omega) = \mathbf{B}\mathbf{c}_b(\psi) \tag{11}$$

where

$$\mathbf{c}_a(\psi) = \left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{a}(\psi) \tag{12}$$

$$\mathbf{c}_b(\omega) = \left(\mathbf{B}^H\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{b}(\omega) \tag{13}$$

are the expansion coefficients. Let n_b and m_b be the number of basis vectors in $\mathbf{A}$ and $\mathbf{B}$, respectively. The accuracy of this representation increases with the number of beams in $\mathbf{A}$ and $\mathbf{B}$, but generally three to five beams is adequate for modeling any $\mathbf{v}(\Theta)$ within a beamwidth of some desired location. The $\mathbf{c}_a(\psi)$ and $\mathbf{c}_b(\omega)$ coefficients can also be pre-computed and stored provided the appropriate grid of hypothesis vectors is used around each origin for Θ. The full steering vector $\mathbf{v}(\Theta)$ can then be approximated by

$$\hat{\mathbf{v}}(\Theta) = \mathbf{B}\mathbf{c}_b(\psi) \otimes \mathbf{A}\mathbf{c}_a(\omega) = (\mathbf{B} \otimes \mathbf{A})\left(\mathbf{c}_b(\psi) \otimes \mathbf{c}_a(\omega)\right) \tag{14}$$

For each ψ basis vector $\mathbf{a}_i$, form the $p \times m_b$ matrix

$$\mathbf{D}_i = \widehat{\mathbf{R}}_W^{-1/2}\begin{bmatrix} \mathbf{W}_1^H\mathbf{a}_i, & \cdots, & \mathbf{W}_M^H\mathbf{a}_i \end{bmatrix}\mathbf{B}. \tag{15}$$

Table 1. Operation counts of ML estimation using Kronecker and steering subspace operations

Quantity	Operations
Non-recurring:	
$\widehat{\mathbf{R}}_W^{-1/2}$	$p^2MN + p^3$
$\mathbf{D}_i$	$pn_bMN + pm_bn_bM + p^2m_bn_b$
$\mathbf{E}(\phi)$	pmm_bn_b
$\mathbf{g}(\psi,\omega)$	$pmnn_b$
$\|\mathbf{g}(\psi,\omega)\|^2$	pmn
Recurring:	
$\dfrac{\|\mathbf{g}^H(\psi,\omega)\hat{\mathbf{x}}\|^2}{\|\mathbf{g}(\psi,\omega)\|^2}$	$(p+2)mn$

These matrices may then be combined for each ω hypothesis to provide the $p \times n_b$ matrix

$$\mathbf{E}(\omega) = \left[\ \mathbf{D}_1^H\mathbf{c}_b(\omega), \quad \cdots, \quad \mathbf{D}_{n_b}^H\mathbf{c}_b(\omega)\ \right] \tag{16}$$

Finally, for a specific Θ, this then yields the p-length vector

$$\mathbf{g}(\psi,\omega) = \mathbf{E}(\omega)\mathbf{c}_a(\psi) \tag{17}$$

such that the ML estimator (6) becomes

$$\widehat{\Theta} = \arg\max_{\Theta} \frac{\left|\mathbf{g}^H(\psi,\omega)\hat{\mathbf{x}}\right|^2}{\|\mathbf{g}(\psi,\omega)\|^2} \tag{18}$$

where $\hat{\mathbf{x}} = \widehat{\mathbf{R}}_W^{-1/2}\mathbf{W}^H\mathbf{x}$. The total computations required to compute all terms in (15)-(17), and $\|\mathbf{g}(\psi,\omega)\|^2$ only need to be computed a single time for a specific weight vector matrix $\mathbf{W}$. This cost is often comparable to the cost of computing the weights, which can be quite small compared to the cost of actually applying the weights to the data.

The computation cost of (18) is recurring for each detected data sample for which estimation is required. This requires $(p+2)mn$ computations. Thus, the overall cost of direct ML beamsplitting a detected data sample using this approach has been reduced to about $(p+2)$ operations per hypothesis vector, which is significantly less than the MN operations per hypothesis to do full-dimension ML beamsplitting. This result can be thought of as comparing the whitened data vector $\hat{\mathbf{x}}$ to a bank of mn matched filters. The number of samples with a declared detection is generally quite small compared to the overall number of range gates, so this cost does not stress the computation level of the processor. The overall costs of this approach are parameterized in Table 1.

3. STAP NULLING APPLICATION EXAMPLES

Airborne early warning radars are required to detect targets in the presence of heavy ground clutter and/or jamming. The platform motion induces an angle-dependent Doppler shift on the clutter; for low PRF radars the clutter may in fact fill the available Doppler space. Clutter cancellation for such a system requires space-time adaptive processing (STAP) over the antenna element (space) and the radar pulse repetition interval (time) dimensions. STAP in this context is joint angle-Doppler adaptive processing. This problem fits naturally into the multidimensional framework described above. For the STAP problem the data snapshot is comprised of the samples from N antenna elements and M pulses that

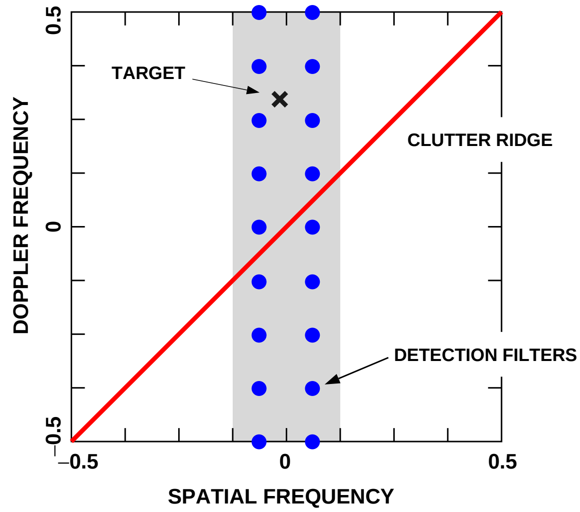

Figure 1. STAP signal and beam grid scenario

correspond to a single radar range gate. The target component of the snapshot has the space-time steering vector given by (8). For the STAP problem ψ represents spatial frequency or azimuth and ω represents Doppler frequency (normalized to the radar PRF):

$$\psi = \frac{d}{\lambda}\sin\phi\ , \ \omega = \frac{2v_pT_r}{\lambda}\ . \tag{19}$$

Here d is the array interelement spacing, T_r is the time interval between successive pulses, and v_p is the radar platform velocity. Thus, $\mathbf{b}(\omega)$ is a temporal steering vector that contains the interpulse phase shift commensurate with the target Doppler, and $\mathbf{a}(\psi)$ is the usual spatial steering vector that contains the interelement phases for the target angle of arrival.

A surveillance radar devotes a waveform dwell to a specific angle, and for that period of time can reasonably expect targets to only come from the direction of the transmit beam. Therefore the number of receive beams to be formed is typically small (two or three). The target velocity is completely unknown, so the STAP radar must form a bank of adaptive filters that cover the whole Doppler space. Figure 1 shows a scenario for an $N = 8$ element, $M = 8$ pulse STAP radar. The radar PRF, frequency, and platform velocity are such that the clutter spread equals the radar PRF. A two beam by eight Doppler bin cluster of adaptive filters is processed and fed to the beamsplitting algorithm. Each of these adaptive filters will have a deep adaptive null on the clutter line. After an initial adaptive matched filter detection with this coarsely sampled grid of adaptive filters, the approximate ML form from (18) was implemented. A finely sampled 20×20 grid of angle-Doppler hypotheses was searched to provide accuracy commensurate with 20:1 beamsplitting in both angle and Doppler. The basis set for the hypotheses consisted of five steering vectors in each dimension, equally spaced across a nominal beamwidth centered at the filter of the initial detection.

The performance of the estimator is evaluated by Monte-Carlo simulation and shown in Figure 2. For each target Doppler 200 trials were conducted. The target SNR is 0 dB per element per pulse. This SNR corresponds to 18 dB SNR at the output of a matched filter in the absence of any interference, which is just enough

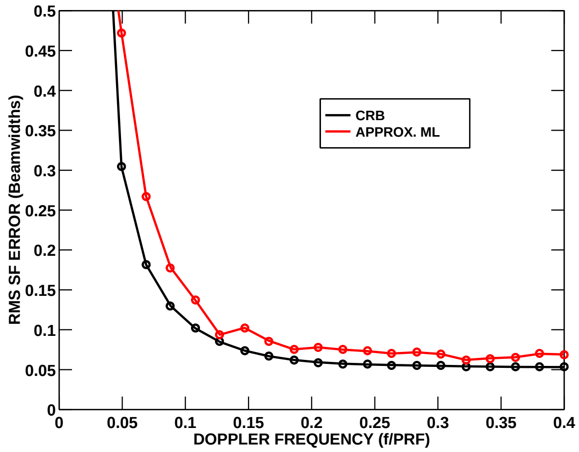

Figure 2. Monte Carlo simulation of approximate ML approach

to support 20:1 beamsplitting in each dimension. Each trial consisted of a sample covariance matrix with $5NM$ clutter samples. The RMS spatial frequency (azimuth) error is plotted versus target Doppler; the Cramér-Rao bound [9] is shown for reference. The error increases as target Doppler gets smaller because the target is approaching the clutter ridge; i.e. the target starts to fall in the STAP clutter null. The results show that the proposed approach provides performance very close to the Cramér-Rao bound, while providing substantial computational savings over a full 2-D search.

4. SUMMARY

This paper presented a multi-dimensional maximum-likelihood (ML) estimation procedure using adaptive subspace beamformed outputs. A search over hypothesis vectors lying on a fixed grid in the multi-dimensional estimation space allows a Kronecker representation of the hypothesis steering vectors. In addition, a least-squares fit of the hypothesis vectors to the constraint beams employed by the adaptive weight vectors further reduce the required computations. These two effects show that the outputs from p adaptive beams can be used as a matched filter for the grid of hypothesis estimation vectors to find the ML parameter estimate. A simulation result demonstrated this procedure on a space-time adaptive processing example and showed very close matching with the Cramér-Rao bound.

A KRONECKER STEERING VECTOR FOR A HEXAGONAL GRID

This appendix details the Kronecker steering vector formulations for a two-dimensional arrays on a hexagonal grid as shown in Figure 3. A steering vector can be factored as a Kronecker product for azimuth θ and elevation ϕ as $\mathbf{v}(\theta,\phi) = \mathbf{v}_{\mathrm{hex}}(\theta,\phi) \otimes \mathbf{v}_{\mathrm{az}}(\theta,\phi)$ where $\mathbf{v}_{\mathrm{hex}}(\theta,\phi)$ is the Vandermonde steering vector as seen by the diagonal column of elements marked by '$*$' in Figure 3 and $\mathbf{v}_{\mathrm{az}}(\theta,\phi)$ is the steering vector as seen by the horizontal column of elements denoted by the 'o' marks. The Kronecker form of these two vectors fills out the trapezoidal grid shown by the dots in Figure 3. Often the array will be constrained to only a portion of this grid as indicated by the elements within the circle of Figure 3. When the

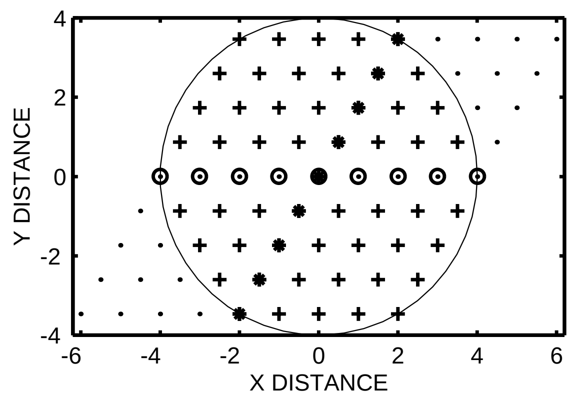

Figure 3. 2-D array on a hexagonal grid

vector product $\mathbf{w}^H \mathbf{v}(\theta,\phi)$ is computed, $\mathbf{w}$ is expanded to include zeros in the trapezoidal grid which does not correspond to an actual antenna element. The steering vectors to be used for the ML hypotheses should then consist of a fixed set of points on a grid defined by lines having a constant cone angle with respect to the each of the diagonal and horizontal axes marked by '$*$' and 'o' respectively.

REFERENCES

[1] J. Ward, "Space-time adaptive processing for airborne radar clutter suppression," Tech. Rep. TR-1015, M.I.T. Lincoln Laboratory, Lexington, MA, November 1994.

[2] U. Nickel, "Monopulse estimation with adaptive arrays," *IEE Proc.-F*, vol. 140, pp. 303–308, October 1993.

[3] J. Ward, "Maximum likelihood angle and velocity estimation with space-time adaptive processing radar," in *30th Asilomar Conf. on Signals, Sys., and Comp.*, November 1996.

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[5] G. F. Hatke and K. W. Forsythe, "A class of polynomial rooting algorithms for joint azimuth/elevation estimation using multidimensional arrays," in *Proc. 28th Asliomar Conf. on Sig., Sys., and Computers*, Oct. 1994.

[6] F. C. Robey, F. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. AES*, vol. 28, pp. 208–216, January 1992.

[7] M. I. Skolnick, *Introduction to Radar Systems*. New York: McGraw-Hill, 1980.

[8] S. Haykin and A. Steinhardt, eds., *Adaptive Radar and Detection*. New York: J. Wiley & Sons, 1992.

[9] J. Ward, "Cramér-Rao bounds for target angle and doppler estimation with stap radar," in *29th Asilomar Conf. on Signals, Sys., and Comp.*, November 1995.