ROBUSTNESS OF THE SUBSPACE GLRT TO SIGNAL MISMATCH

Keith A. Burgess
Barry D. Van Veen

Department of Electrical and Computer Engineering
1415 Johnson Drive; University of Wisconsin
Madison, WI 53706; U.S.A.

Abstract

The robustness of a subspace generalized likelihood ratio test (GLRT) detector to signal mismatch is addressed for data conforming to the generalized multivariate analysis of variance model. This model assumes a deterministic signal of known form in the presence of unknown, colored, Gaussian noise. The subspace GLRT compresses data into a lower-dimensional subspace prior to detection. It is shown in this paper that a subspace GLRT reduces the performance loss due to mismatch relative to that of a non-subspace detector.

I. Introduction

A common signal processing problem is the detection of a signal having a known form in the presence of noise. This problem occurs in active radar, active sonar, communication systems, biomedical stimulus response, and many other multi-sensor or time-series applications. For a radar array, a signal form is typically derived from known sensor locations, transmitter location, and frequency. In time series applications, a signal form is derived from the relative amplitudes as a function of time. We will consider as unknowns the noise covariance matrix and the signal magnitude. The signal is absent when this magnitude is zero. Note that this detection problem differs from a direction of arrival estimation problem where the task is to ascertain the existence of a signal of unknown form.

Subspace generalized likelihood ratio test (GLRT) detectors have been shown to yield strong performance gains over full data space detectors (e.g. [1] and [2]) when the number of data vectors is limited. These subspace detectors are constructed by passing data through a linear 'subspace' (matrix) transformation prior to performing signal detection. *See figure 1.* Performance in this case is measured by the detector's probability of detection (PD) for a fixed probability of false alarm (PFA). A subspace detector gains this performance in three primary ways. First, *a priori* information about the interference scenario often can be folded indirectly into the detector through a judicious choice of subspace transformation. Direct inclusion of additional information into the full data space model usually leads to an intractable problem. Second, a subspace detector has significantly fewer model parameters. This means that there is effectively more data to estimate each parameter. Hence, the estimates tend to be more stable which improves subspace detector performance. Third, because a subspace detector has fewer parameters to estimate it requires less data to exist. In many cases, a subspace detector can exist with good performance when insufficient data exists to construct the full data space detector. The disadvantage of subspace processing is that most non-adaptive transformations will tend to diminish the signal-to-noise ratio (SNR). A loss in SNR corresponds to a loss in detection performance.

The relative advantage of subspace detection over full data space detection (when $\mathbf{T} = \mathbf{I}_N$) depends almost entirely upon the amount of available data. It is shown in [1] that with short data records, statistical stability dominates detector performance and subspace detection is preferable to full data space detection in spite of

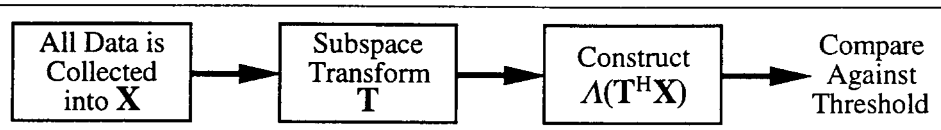

Figure 1: Block Diagram of a Non-Adaptive Subspace Detector.

typical SNR losses. With large data records, SNR dominates detector performance and the full data space detector has better performance.

This paper discusses a fourth advantage of subspace detectors. It is shown that subspace GLRT detectors are more robust than full data space GLRT detectors to signal model mismatch. Kelly and Forsythe [4] define signal mismatch as the effect of imperfect knowledge of the form of signals in the GLRT model. Robustness is demonstrated in this paper by showing that subspace processing reduces the effect of mismatch between the assumed signal model forms and true signal model forms. These reductions occur for two reasons. Primarily, mismatch is reduced when data is transformed into a subspace because linear compression tends to diminish the differences between data forms. Secondarily, subspace detectors can be less sensitive to changes in SNR that result from erroneous signal model assumptions.

For notation, bold lower and upper case roman letters designate column vectors and matrices, respectively. Script $\mathbb{C}^{J \times K}$ refers to the $J \times K$ complex product space. Superscript '$^{\mathrm{H}}$' denotes the complex conjugate transpose. Let vec(**Z**) be the operator which constructs a column vector by stacking columns of the matrix **Z**. Finally, '$\otimes$' denotes the Kronecker product (*See* [4]).

II. A Subspace GLRT Detection Model

Let $\mathbf{X}$ represent an N by L matrix of data formed by concatenating L snapshots of the N-dimensional data observations. The snapshots are assumed to be stationary, statistically independent, colored, zero-mean, multivariate Gaussian distributed noise with an additive, deterministic signal. The covariance, $\mathbf{R}_\mathbf{x}$, of each snapshot is assumed unknown, and the form of the signal to be detected is given by $\mathrm{E}[\mathbf{X}] = \mathbf{a}\mathbf{b}^{\mathrm{H}}\mathbf{C}$ with $\mathbf{a} \in \mathbb{C}^N$, $\mathbf{C} \in \mathbb{C}^{M \times L}$ known, and $\mathbf{b} \in \mathbb{C}^M$ is an unknown, non-stochastic, complex matrix of regression parameters. Hence the measured data array is modeled as $\mathrm{vec}(\mathbf{X}) \sim \mathrm{N}(\mathrm{vec}(\mathbf{a}\mathbf{b}^{\mathrm{H}}\mathbf{C}), \mathbf{R}_\mathbf{x} \otimes \mathbf{I}_L)$. This is a generalized multivariate analysis of variance model (GMANOVA) [2].

A subspace transformation $\mathbf{T} \in \mathbb{C}^{N \times P}$, is a fixed linear mapping from an N-dimensional space into a smaller, P-dimension space. The columns of $\mathbf{T}$ are assumed to be linearly independent. Methods for choosing $\mathbf{T}$ are given in [1] and [2]. The transformed data $\mathbf{Z} \equiv \mathbf{T}^{\mathrm{H}}\mathbf{X} \in \mathbb{C}^{P \times L}$ is in the same parametric family:

$$\mathrm{vec}(\mathbf{Z}) \sim \mathrm{N}(\mathrm{vec}(\mathbf{T}^{\mathrm{H}}\mathbf{a}\mathbf{b}^{\mathrm{H}}\mathbf{C}), (\mathbf{T}^{\mathrm{H}}\mathbf{R}_\mathbf{x}\mathbf{T}) \otimes \mathbf{I}_L). \qquad (1)$$

To yield a meaningful test, we require that $\mathbf{T}$ be chosen such that $\mathbf{T}^{\mathrm{H}}\mathbf{a} \neq \mathbf{0}$. The GLRT under this distribution can be expressed as [4; p.23]

$$\Lambda(\mathbf{Z}) = \frac{\mathbf{a}^{\mathrm{H}}\mathbf{T}\left(\mathbf{Z}\mathbf{P}_{\mathbf{C}\perp}\mathbf{Z}^{\mathrm{H}}\right)^{-1}\mathbf{T}^{\mathrm{H}}\mathbf{a}}{\mathbf{a}^{\mathrm{H}}\mathbf{T}\left(\mathbf{Z}\mathbf{Z}^{\mathrm{H}}\right)^{-1}\mathbf{T}^{\mathrm{H}}\mathbf{a}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \ell_0 \qquad (2)$$

where $\mathbf{P}_{\mathbf{C}\perp} \equiv \mathbf{I}_L - \mathbf{C}^{\mathrm{H}}(\mathbf{C}\mathbf{C}^{\mathrm{H}})^{-1}\mathbf{C}$, and where we require $L \geq M+N$ in the full data space (when $\mathbf{T} = \mathbf{I}_N$) and $L \geq M+P$ in a P-dimensional subspace. The PD and PFA of (2) are known in closed form and are given in [1] and [2]. The PFA is a function only of the test threshold ℓ_0 and dimensionality parameters L, M, and P. The PD is a function of these parameters as well as the non-adaptive subspace SNR

$$SNR_\mathbf{T} = \left(\mathbf{b}^{\mathrm{H}}\mathbf{C}\mathbf{C}^{\mathrm{H}}\mathbf{b}\right)\left(\mathbf{a}^{\mathrm{H}}\mathbf{T}\left(\mathbf{T}^{\mathrm{H}}\mathbf{R}_\mathbf{x}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{H}}\mathbf{a}\right). \qquad (3)$$

III. Signal Mismatch

Signal mismatch results when there is imperfect knowledge of the vector $\mathbf{a}$. Hence signal mismatch is an issue of primary concern to a detector's robustness. This section will show that subspace processing reduces the effect of mismatch between the assumed signal vector $\mathbf{a}$ and a true signal vector $\tilde{\mathbf{a}}$.

The hypothesis test's performance depends upon the signal vector only through the non-adaptive SNR parameter (3) used in the PD, and then only through the term $\mathbf{T}^{\mathrm{H}}\mathbf{a}$. The PD is a monotonically increasing function of SNR. The actual performance reduction resulting from using the assumed subspace signal $\mathbf{T}^{\mathrm{H}}\mathbf{a}$ instead of the true quantity $\mathbf{T}^{\mathrm{H}}\tilde{\mathbf{a}}$ depends upon the angle between these two vectors. Kelly [3] shows that the true, full space ($\mathbf{T} = \mathbf{I}_N$), non-adaptive SNR parameter $SNR_{\mathbf{I}_N} = (\mathbf{b}^{\mathrm{H}}\mathbf{C}\mathbf{C}^{\mathrm{H}}\mathbf{b})(\tilde{\mathbf{a}}^{\mathrm{H}}\mathbf{R}_\mathbf{x}^{-1}\tilde{\mathbf{a}})$ is attenuated by an amount

$$\cos^2(\theta_N) = \frac{\left|\tilde{\mathbf{a}}^{\mathrm{H}}\mathbf{R}_\mathbf{x}^{-1}\mathbf{a}\right|^2}{\left(\tilde{\mathbf{a}}^{\mathrm{H}}\mathbf{R}_\mathbf{x}^{-1}\tilde{\mathbf{a}}\right)\left(\mathbf{a}^{\mathrm{H}}\mathbf{R}_\mathbf{x}^{-1}\mathbf{a}\right)} \qquad (4)$$

due to mismatch between the assumed signal model $\mathbf{a}$ and true signal vector $\tilde{\mathbf{a}}$. Here, $\cos^2(\theta_N)$ represents the cosine squared of the angle between $\mathbf{a}$ and $\tilde{\mathbf{a}}$ in an appropriate inner product space. This quantity is less than one whenever $\mathbf{a} \neq \tilde{\mathbf{a}}$. It is easily shown that the subspace non-adaptive SNR (3) is attenuated by a similar amount

$$\cos^2(\theta_P) = \frac{\left|\tilde{\mathbf{a}}^H\mathbf{T}(\mathbf{T}^H\mathbf{R_x}\mathbf{T})^{-1}\mathbf{T}^H\mathbf{a}\right|^2}{\left(\tilde{\mathbf{a}}^H\mathbf{T}(\mathbf{T}^H\mathbf{R_x}\mathbf{T})^{-1}\mathbf{T}^H\tilde{\mathbf{a}}\right)\left(\mathbf{a}^H\mathbf{T}(\mathbf{T}^H\mathbf{R_x}\mathbf{T})^{-1}\mathbf{T}^H\mathbf{a}\right)}. \tag{5}$$

The relative mismatch performance of the subspace detector is assessed by comparing $\cos^2(\theta_P)$ to $\cos^2(\theta_N)$. For instance, if $\cos^2(\theta_N)$ is smaller than $\cos^2(\theta_P)$, then the full-space SNR loss is greater than the subspace SNR loss.

To compare (4) with (5), begin by using the singular value decomposition (SVD) of $\mathbf{R_x^{H/2}T} = \mathbf{U\Sigma V^H}$. The first P columns of $\mathbf{U}$, denoted by $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \cdots, \mathbf{u}_P\}$, form an orthonormal basis for the range space of $\mathbf{R_x^{H/2}T}$. The remaining $N-P$ columns, denoted by $\{\mathbf{u}_{P+1}, \mathbf{u}_{P+2}, \cdots, \mathbf{u}_N\}$, form an orthonormal basis for its orthogonal complement. Let $\mathbf{d} \equiv \mathbf{U}^H(\mathbf{R_x^{-1/2}a})$ be the coordinate representation of the mismatched quantity $\mathbf{R_x^{-1/2}a}$ in this new basis. The coordinates for $\mathbf{R_x^{-1/2}a}$ can be separated into two sets $\mathbf{g}$ and $\mathbf{h}$ such that $\mathbf{d}^H = [\mathbf{g}^H, \mathbf{h}^H]$. In particular, the P coordinates of $\mathbf{R_x^{-1/2}a}$ in the range space of $\mathbf{R_x^{H/2}T}$ are given by $\mathbf{g} \equiv [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_P]^H(\mathbf{R_x^{-1/2}a})$, and the $(N-P)$ coordinates of $\mathbf{R_x^{-1/2}a}$ in the null space of $\mathbf{R_x^{H/2}T}$ are given by $\mathbf{h} \equiv [\mathbf{u}_{P+1}, \mathbf{u}_{P+2}, \cdots, \mathbf{u}_N]^H(\mathbf{R_x^{-1/2}a})$. A similar, true quantity $\mathbf{R_x^{-1/2}\tilde{a}}$ has a coordinate representation defined by $\tilde{\mathbf{d}} \equiv \mathbf{U}^H(\mathbf{R_x^{-1/2}\tilde{a}})$ in the same new basis. As was true above, $\tilde{\mathbf{d}}$ can be separated into two sets $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{h}}$ such that $\tilde{\mathbf{d}}^H = [\tilde{\mathbf{g}}^H, \tilde{\mathbf{h}}^H]$.

Using these definitions, it is easily shown that:

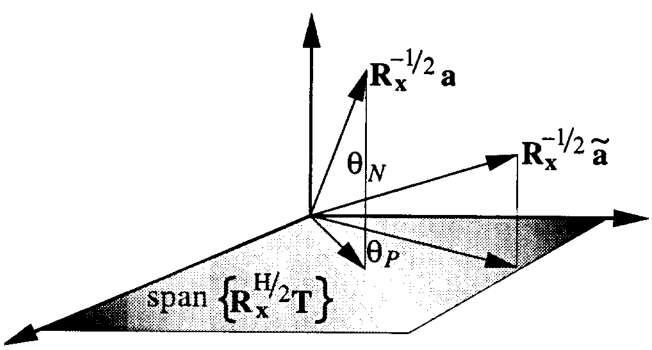

Figure 2: The Subspace Angle θ_P is Smaller.

$$\cos^2(\theta_N) = \frac{|\tilde{\mathbf{d}}^H\mathbf{d}|^2}{(\tilde{\mathbf{d}}^H\tilde{\mathbf{d}})(\mathbf{d}^H\mathbf{d})} \text{ and}$$

$$\cos^2(\theta_P) = \frac{|\tilde{\mathbf{g}}^H\mathbf{g}|^2}{(\tilde{\mathbf{g}}^H\tilde{\mathbf{g}})(\mathbf{g}^H\mathbf{g})}. \tag{6}$$

With these definitions, the full space angle cosine can be rewritten as follows:

$$\begin{aligned}\cos^2(\theta_N) &= \frac{|\tilde{\mathbf{g}}^H\mathbf{g}+\tilde{\mathbf{h}}^H\mathbf{h}|^2}{(\tilde{\mathbf{g}}^H\tilde{\mathbf{g}}+\tilde{\mathbf{h}}^H\tilde{\mathbf{h}})(\mathbf{g}^H\mathbf{g}+\mathbf{h}^H\mathbf{h})} \\ &= \frac{|\tilde{\mathbf{g}}^H\mathbf{g}|^2 + |\tilde{\mathbf{h}}^H\mathbf{h}|^2 + 2\mathrm{Re}\left((\tilde{\mathbf{g}}^H\mathbf{g})(\tilde{\mathbf{h}}^H\mathbf{h})\right)}{(|\tilde{\mathbf{g}}|^2+|\tilde{\mathbf{h}}|^2)(|\mathbf{g}|^2+|\mathbf{h}|^2)} \\ &\le \frac{\left(|\tilde{\mathbf{g}}^H\mathbf{g}| + |\tilde{\mathbf{h}}^H\mathbf{h}|\right)^2}{(|\tilde{\mathbf{g}}|^2+|\tilde{\mathbf{h}}|^2)(|\mathbf{g}|^2+|\mathbf{h}|^2)}\end{aligned} \tag{7}$$

where we have used

$$\mathrm{Re}\left((\tilde{\mathbf{g}}^H\mathbf{g})(\tilde{\mathbf{h}}^H\mathbf{h})\right) \le \left|\mathrm{Re}\left((\tilde{\mathbf{g}}^H\mathbf{g})(\tilde{\mathbf{h}}^H\mathbf{h})\right)\right|$$
$$\le \mathrm{Re}\left(\left|(\tilde{\mathbf{g}}^H\mathbf{g})(\tilde{\mathbf{h}}^H\mathbf{h})\right|\right) = \left|(\tilde{\mathbf{g}}^H\mathbf{g})(\tilde{\mathbf{h}}^H\mathbf{h})\right| \le |\tilde{\mathbf{g}}^H\mathbf{g}|\,|\tilde{\mathbf{h}}^H\mathbf{h}|.$$

The transformation $\mathbf{T}$ is constrained to span the assumed signal $\mathbf{a}$. Therefore, the term $|\mathbf{g}|^2$ is greater than zero. If the assumed $\mathbf{a}$ is not perpendicular to the true $\tilde{\mathbf{a}}$, then the term $|\tilde{\mathbf{g}}|^2$ is also greater than zero. The goal in designing a subspace transformation $\mathbf{T}$ is to make the non-adaptive SNR, (3), approach the upper bound SNR_{I_N} (*see* [2]). This upper bound is reached when $\mathbf{R_x^{-1/2}a}$ lies in the range of columns of $\mathbf{R_x^{H/2}T}$. For such a well designed $\mathbf{T}$, $\mathbf{h} = \mathbf{0}$. In this case,

$$\begin{aligned}\cos^2(\theta_N) &\le \frac{|\tilde{\mathbf{g}}^H\mathbf{g}|^2}{(|\tilde{\mathbf{g}}|^2+|\tilde{\mathbf{h}}|^2)|\mathbf{g}|^2} \\ &\le \frac{|\tilde{\mathbf{g}}^H\mathbf{g}|^2}{|\tilde{\mathbf{g}}|^2|\mathbf{g}|^2} = \cos^2(\theta_P)\end{aligned} \tag{8}$$

with equality when $\mathbf{R_x^{-1/2}\tilde{a}}$ also lies in the range of $\mathbf{R_x^{H/2}T}$. By symmetry, if $\mathbf{R_x^{-1/2}\tilde{a}}$ lies in the range of $\mathbf{R_x^{H/2}T}$, then $\tilde{\mathbf{h}} = \mathbf{0}$, and again $\cos^2(\theta_N) \le \cos^2(\theta_P)$. A perturbation argument shows that subspace detection remains more robust than full space detection under mismatch when either $\mathbf{R_x^{-1/2}a}$ or $\mathbf{R_x^{-1/2}\tilde{a}}$ lies close to the range space of $\mathbf{R_x^{H/2}T}$. Thus if $\mathbf{T}$ and $\mathbf{a}$ are chosen reasonably well, then $\cos^2(\theta_N) \le \cos^2(\theta_P)$, and the subspace SNR loss is less than or equal to the full space

SNR loss resulting from a mismatched signal. Figure 2 illustrates how the angle between vectors is reduced by the subspace mapping.

Not only does a subspace detector loose less SNR to signal mismatch, but the loss of SNR can affect detector performance to a smaller degree. This effect is illustrated in the following section.

IV. Examples

To demonstrate the effect of mismatch loss on SNR, consider a ten-dimensional system ($N = 10$) with an assumed steering vector $\mathbf{a}=[1,0,0,\ldots,\alpha]$. For simplicity, assume that the noise is white ($\mathbf{R_x} = \mathbf{I}_{10}$) and that the subspace transformation $\mathbf{T}$ compresses data by throwing away data from the last six sensors. This simple transformation is non-optimal because it knowingly looses SNR by throwing away signal energy (α) from the last sensor. However, it offers a simple demonstration of a situation where $\mathbf{R_x}^{-1/2}\mathbf{a}$ does not lie in the range of columns of $\mathbf{R_x}^{H/2}\mathbf{T}$. Let the true steering vector $\tilde{\mathbf{a}}$ be perturbed from $\mathbf{a}$ by a Gaussian distributed random vector with covariance $\frac{1}{\sqrt{10}}\mathbf{I}_{10}$. Figure 3 plots the mean of the difference in mismatch loss ($\cos^2(\theta_P) - \cos^2(\theta_N)$) over 30,000 trials as the amount of signal amplitude, α, outside of the range of $\mathbf{R_x}^{H/2}\mathbf{T}$ increases. Also plotted in figure 3 are the percentage of trials where subspace SNR loss was greater than the full space SNR loss. Note that even when one half of the signal model lies outside of the range of $\mathbf{R_x}^{H/2}\mathbf{T}$, i.e. $\alpha = 1$, the SNR loss due to mismatch in the subspace detector is less than that for the full space detector in more than 70% of the random trials. If the amount of signal amplitude outside $\mathbf{R_x}^{H/2}\mathbf{T}$ is small, then subspace

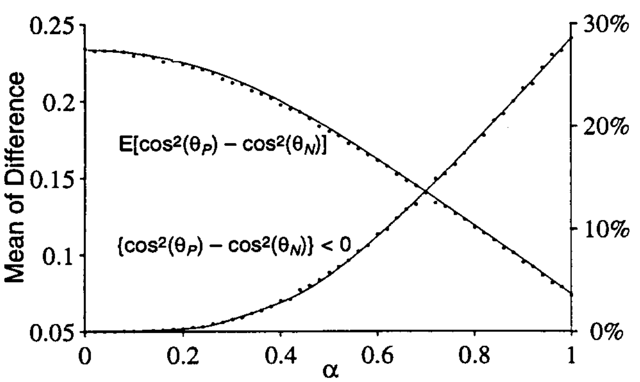

Figure 3: Mean of mismatch loss differences and cases (%) with inferior subspace performance.

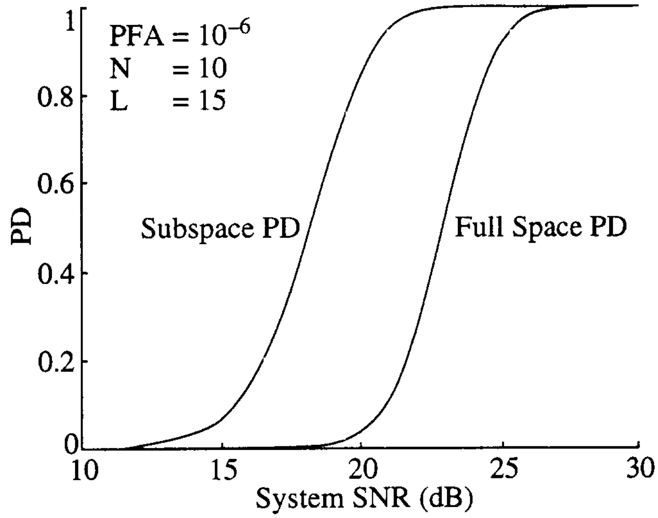

Figure 4: Full and Subspace Performance.

SNR loss is nearly always less than the full space loss, as predicted by the analysis in the previous section.

Now consider figure 4. This figure depicts the PD as a function of non-adaptive SNR in the absence of signal mismatch. Fifteen snapshots of data ($L = 15$) are assumed with $N = 10$ and a PFA $= 10^{-6}$. The subspace results assume that a transformation from ten to four dimensions leads to an SNR loss of 3 dB; i.e. the subspace SNR is 3 dB less than the values depicted on the x-axis as "System SNR". If the system SNR were 27 dB and the full space mismatch resulted in a 4 dB loss, then the full space PD would decrease from 0.997 to 0.508. Because the subspace mismatch is less than 4 dB, the corresponding subspace PD decreases from 1.000 to an amount still greater than 0.998.

V. References

[1] Keith A. Burgess and Barry D. Van Veen, "Improved Adaptive Detection Performance via Subspace Processing," *1992 ICASSP, San Francisco, CA*, vol. 5, pp. 353-356, Mar. 1992.

[2] Keith A. Burgess, "Subspace Based Adaptive Detection Using a Generalized Likelihood Ratio Test," *Ph.D. Thesis, The University of Wisconsin–Madison*, Aug. 1994.

[3] Edward J. Kelly, "Adaptive Detection in Non-Stationary Interference, Part III," *Lincoln Laboratory Technical Report*, no. 761, August 24, 1987.

[4] Edward J. Kelly and Keith M. Forsythe, "Adaptive Detection and Parameter Estimation for Multidimensional Signal Models," *Lincoln Laboratory Technical Report*, no. 848, April 19, 1989.