Parametric GLRT for Multichannel Adaptive Signal Detection

Kwang June Sohn, *Student Member, IEEE*, Hongbin Li, *Member, IEEE*, and Braham Himed, *Fellow, IEEE*

***Abstract*—This paper considers the problem of detecting a multichannel signal in the presence of spatially and temporally colored disturbance. A parametric generalized likelihood ratio test (GLRT) is developed by modeling the disturbance as a multichannel autoregressive (AR) process. Maximum likelihood (ML) parameter estimation underlying the parametric GLRT is examined. It is shown that the ML estimator for the alternative hypothesis is nonlinear and there exists no closed-form expression. To address this issue, an asymptotic ML (AML) estimator is presented, which yields asymptotically optimum parameter estimates at reduced complexity. The performance of the parametric GLRT is studied by considering challenging cases with *limited or no training signals* for parameter estimation. Such cases (especially when training is unavailable) are of great interest in detecting signals in heterogeneous, fast changing, or dense-target environments, but generally cannot be handled by most existing multichannel detectors which rely more heavily on training at an adequate level. Compared with the recently introduced parametric adaptive matched filter (PAMF) and parametric Rao detectors, the parametric GLRT achieves higher data efficiency, offering improved detection performance in general.**

***Index Terms*—Generalized likelihood ratio test (GLRT), maximum likelihood (ML) parameter estimation, multichannel signal detection, parametric models, space-time adaptive processing (STAP).**

I. Introduction

DETECTING a multichannel signal in disturbance is of great importance in numerous applications including radar [1], [2], wireless communications [3], hyperspectral imaging [4], [5], and others. Multichannel signal detection based on space-time adaptive processing (STAP), which is capable of handling strong spatially and temporally colored disturbance, has received significant attention over the past few decades. A multitude of STAP based detectors have been proposed and used to mitigate clutter and jamming in radar, remote sensing, and communication systems [1]–[5].

STAP detectors can be classified based on whether an estimate of the *space-time covariance matrix*, denoted by $\mathbf{R}$ (see Section II), of the multichannel disturbance signal is needed for interference suppression. This class of STAP detectors are henceforth referred to as *covariance matrix based detectors*, which are among the most extensively studied multichannel detectors, including the Reed, Mallett, and Brennan detector [6], Kelly's generalized likelihood ratio test (GLRT) [7], the adaptive matched filter (AMF) detector [8]–[10], the adaptive coherence estimator (ACE) detector [11], among others. These detectors involve estimating $\mathbf{R}$ from target-free training data and inverting it, which may impose excessive training and computational burdens when the joint space-time dimension is large. At a minimum, we need $K \geq JN$ training signals[1] to ensure a full-rank estimate of the $JN \times JN$ matrix $\mathbf{R}$, where J denotes the number of spatial channels and N the number of temporal observations. Moreover, the Reed-Brennan rule [6] suggests that, in average, $K \geq (2JN-3)$ training signals are needed to obtain performance within 3 dB from the optimum bound. Such conditions may not be satisfied, especially in heterogeneous or dense-target environments that offer limited training, thus rendering covariance matrix based techniques inapplicable.

The above difficulty is primarily because for large JN and in the absence of any specific structure, $\mathbf{R}$ involves an enormous number of unknowns. In many applications, however, the disturbance usually exhibits certain spatial, temporal, and/or spectral structures that can be exploited to reduce the number of unknowns and ease the training/computational burden. Among other alternatives, one general structured approach is to model the disturbance as a multichannel autoregressive (AR) process, which has been found to be very useful in representing the spatial and temporal correlation of radar signals [12]–[15]. Using both simulated and real data, the so-called *parametric adaptive matched filter* (PAMF) [12], [13] has been shown to significantly outperform the aforementioned covariance matrix based detectors for small training size at a reduced complexity. While covariance matrix based detectors perform joint space-time whitening for interference mitigation, the PAMF detector adopts a two-step approach, involving temporal whitening via an inverse moving-average (MA) filter followed by spatial whitening. Recently, the PAMF detector has been shown to be equivalent to a *parametric Rao detector* [16], [17].

In this paper, we develop a *parametric GLRT*. It is natural to extend the results of [16] and [17] and consider the parametric GLRT for several reasons. First, as shown in Section II, the problem of interest is a two-sided parameter testing problem that admits no uniformly most powerful (UMP) solution [18]. A GLRT approach is widely used in such cases due to its good

Manuscript received May 10, 2006; revised February 16, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Marcelo G. S. Bruno. This work was supported by the Air Force Research Laboratory (AFRL) under Contract FA8750-05-2-0001. This paper was presented in part at the 4th IEEE Workshop on Sensor Array and Multi-channel Processing (SAM 2006),Waltham, MA, July 2006.

K. J. Sohn and H. Li are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: ksohn@stevens.edu; hli@stevens.edu).

B. Himed was with the Air Force Research Laboratory/SNRT, Rome, NY 13441 USA. He is now with Signal Labs, Inc., Reston, VA 20191 USA (e-mail: braham.himed@signal-labs.com).

Digital Object Identifier 10.1109/TSP.2007.896068

[1] $K \geq JN - 1$ training signals are needed if both the test and training signals are used to estimate $\mathbf{R}$.

asymptotic properties including asymptotic CFAR and consistency. Second, the parametric GLRT may yield improved performance than the parametric Rao detector, especially when the data is limited. This is because the latter is an asymptotic (large-sample) parametric GLRT ([18, App. 6B]). Third, all Rao tests, including the parametric Rao detector, are obtained based on a further approximation that is valid only for weak signals ([18, p. 238]). As such, the parametric Rao detector is expected to degrade when the weak signal assumption is violated. These observations motivate us to consider the parametric GLRT, in hope of finding a better solution to the problem.

The parametric GLRT to be discussed is different from Kelly's GLRT [7]. The latter does not utilize a parametric model to model the disturbance. For this reason, our solution is referred to as the *parametric* GLRT. Our parametric GLRT is also different from the GLRT developed in [14], where a different detection problem is addressed that involves unknown nonlinear signal parameters associated with the signal to be detected. We follow the direction of [6]–[13] and consider a detection problem whereby the signal to be detected is known up to an unknown amplitude. The data model and assumptions for this problem are further discussed in Section II.

The parametric GLRT relies on maximum likelihood (ML) parameter estimation for both the null and alternative hypotheses. The null hypothesis estimation problem is addressed in [16], where the ML estimator is obtained in closed-form. We show in Section III.A that the ML estimator under the alternative hypothesis is nonlinear and requires searches on a two-dimensional parameter space. To address this issue, we introduce an asymptotic ML (AML) estimator that is considerably simpler, yielding estimates in a noniterative fashion, and asymptotically coincides with the optimum ML estimator. The AML estimator is related to an iterative alternating least-squares (ALS) estimator developed in [19], but with several notable distinctions (see Section III-C). The Cramér-Rao bound (CRB) for the estimation problem is also derived, offering a baseline for comparing various (unbiased) estimators.

To examine the performance of the parametric GLRT, we consider scenarios with *very limited or even no training signals*. The less challenging case with more training is extensively considered in, e.g., [13], [16], and [17], for the PAMF and parametric Rao detectors, which are equivalent to the parametric GLRT with a large amount of training. It should be noted that the parametric GLRT and Rao detectors utilize both test and training signals for parameter estimation; as such, they are functional even without training. *The capability to handle the training-free case is a unique and desirable attribute of the parametric GLRT and Rao detectors*. Although the performance of the parametric GLRT and Rao detectors degrades in the absence of training, such degradation can be remedied by using a larger N, i.e., increasing temporal observations of the test signal. We show that the parametric GLRT outperforms the parametric Rao detector when N is small and, overall, the former yields a better detection performance.

The rest of the paper is organized as follows. Section II contains the problem statement. Parameter estimation is addressed in Section III, including the ML estimators for both hypotheses, the AML estimator for the alternative hypothesis, and the CRB. The test statistic, implementation, and asymptotic analysis for the parametric GLRT are discussed in Section IV. Numerical results are presented in Section V, followed by our conclusions in Section VI.

Notation: Vectors (matrices) are denoted by boldface lower (upper) case letters, all vectors are column vectors, superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and complex conjugate transpose, respectively, $\mathbf{I}_M$ denotes the $M \times M$ identity matrix (with subscript suppressed sometimes), $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ denotes the multivariate complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$, $\|\cdot\|$ takes the Frobenius norm of a matrix/vector, $\Re\{\cdot\}$ takes the real part while $\Im\{\cdot\}$ the imaginary part, and finally, $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse.

II. Data Model and Problem Statement

The problem of interest is to detect a known multichannel signal with unknown amplitude in the presence of spatially and temporally correlated disturbance (e.g., [1])

$$\begin{aligned} H_0&: \mathbf{x}_0(n) = \mathbf{d}_0(n), \quad n = 0, 1, \ldots, N-1 \\ H_1&: \mathbf{x}_0(n) = \alpha \mathbf{s}(n) + \mathbf{d}_0(n), \quad n = 0, 1, \ldots, N-1 \end{aligned} \tag{1}$$

where all vectors are $J \times 1$ vectors, J denotes the number of spatial channels, and N is the number of temporal observations. In the sequel, $\mathbf{x}_0(n)$ is referred to as the *test signal*, $\mathbf{s}(n)$ is the signal to be detected with amplitude α, and $\mathbf{d}_0(n)$ is the *disturbance signal* that may be correlated in space and time. In addition to the test signal $\mathbf{x}_0(n)$, there may be a set of *training* or *secondary* signals $\mathbf{x}_k(n)$, $k = 1, 2, \ldots, K$, to assist in the signal detection:

$$\mathbf{x}_k(n) = \mathbf{d}_k(n), \quad n = 0, 1, \ldots, N-1. \tag{2}$$

In radar systems, training data may be obtained from range cells adjacent to the test cell. However, training data is generally limited or may even be unavailable. In the training-free case, we have $K = 0$.

Define the following $JN \times 1$ space-time vectors:

$$\begin{aligned} \mathbf{s} &= \left[\mathbf{s}^T(0), \quad \mathbf{s}^T(1), \quad \ldots, \quad \mathbf{s}^T(N-1)\right]^T \\ \mathbf{d}_k &= \left[\mathbf{d}_k^T(0), \quad \mathbf{d}_k^T(1), \quad \ldots, \quad \mathbf{d}_k^T(N-1)\right]^T \\ \mathbf{x}_k &= \left[\mathbf{x}_k^T(0), \quad \mathbf{x}_k^T(1), \quad \ldots, \quad \mathbf{x}_k^T(N-1)\right]^T \\ & \qquad\qquad\qquad\qquad k = 0, 1, \ldots, K. \end{aligned} \tag{3}$$

It follows that (1) can be more compactly written as

$$\begin{aligned} H_0&: \mathbf{x}_0 = \mathbf{d}_0 \\ H_1&: \mathbf{x}_0 = \alpha \mathbf{s} + \mathbf{d}_0. \end{aligned} \tag{4}$$

Clearly, the composite hypothesis testing problem (1) or (4) is a *two-sided parameter testing problem* that tests $\alpha = 0$ against $\alpha \neq 0$. The above signal detection problem occurs in an airborne STAP radar system with J array channels and a coherent processing interval (CPI) of N pulse repetition intervals (PRIs). The disturbance $\mathbf{d}_k$ consists of ground clutter, jamming, and thermal noise, while $\mathbf{s}$ is the *target space-time steering vector* [13].

The general assumptions in the literature are [1], [6]–[11]:

AS1: The signal vector $\mathbf{s}$ is deterministic and *known* to the detector;

AS2: The signal amplitude α is complex-valued, deterministic, and *unknown*;

AS3: The disturbance signals $\{\mathbf{d}_k\}_{k=0}^{K}$ are independent and identically distributed (i.i.d.) with distribution $\mathcal{CN}(\mathbf{0},\mathbf{R})$, where $\mathbf{R}$ is the *unknown* space-time covariance matrix.

While AS1 to AS3 are standard [1], [6]–[11], we follow a parametric approach as in [12]–[17]:

AS4: The disturbance signal $\mathbf{d}_k(n)$, $k=0,1,\ldots,K$, can be modeled as a J-channel $\mathrm{AR}(P)$ process with *known*[2] model order P:

$$\mathbf{d}_k(n) = -\sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{d}_k(n-p) + \boldsymbol{\varepsilon}_k(n) \tag{5}$$

where $\{\mathbf{A}^H(p)\}_{p=1}^{P}$ denote the *unknown* $J\times J$ AR coefficient matrices, $\boldsymbol{\varepsilon}_k(n)$ denote the driving J-channel spatial noise vectors that are temporally white but spatially colored Gaussian noise: $\boldsymbol{\varepsilon}_k(n) \sim \mathcal{CN}(\mathbf{0},\mathbf{Q})$, where $\mathbf{Q}$ denotes the *unknown* $J\times J$ spatial covariance matrix.

The *problem of interest* is to develop a GLRT based on Assumptions AS1 to AS4 for the above composite hypothesis testing problem, using the test signal $\mathbf{x}_0$ and training signals $\{\mathbf{x}_k\}_{k=1}^{K}$ if any. The likelihood functions under both hypotheses are parameterized by the signal parameter α as well as nuisance parameters $\mathbf{Q}$ and $\mathbf{A}$, where

$$\mathbf{A}^H = [\mathbf{A}^H(1), \quad \mathbf{A}^H(2), \quad \ldots, \quad \mathbf{A}^H(P)] \in \mathbb{C}^{J\times JP}. \tag{6}$$

For simplicity, we write the likelihood functions as

$$f_i(\alpha,\mathbf{A},\mathbf{Q}), \quad i=0 \quad \text{or} \quad 1 \tag{7}$$

where $\alpha=0$ under H_0 (i.e., $i=0$) and $\alpha\neq 0$ under H_1 ($i=1$), and the dependence on the test/training signals $\{\mathbf{x}_k\}_{k=0}^{K}$ is omitted. While the test statistic of the GLRT is well known, which is given by the generalized likelihood ratio (GLR) [18]

$$\mathrm{GLR} = \frac{\max_{\alpha,\mathbf{A},\mathbf{Q}} f_1(\alpha,\mathbf{A},\mathbf{Q})}{\max_{\mathbf{A},\mathbf{Q}} f_0(0,\mathbf{A},\mathbf{Q})} \tag{8}$$

finding the ML estimates of the unknown parameters is nontrivial. We first address the estimation problem before examining the GLR test statistic in more details.

III. Parameter Estimation

Parameter estimators required by the parametric GLRT as well as the CRB are developed in Appendices I–IV. The main results are summarized here.

A. ML Estimation Under H_1

The ML estimate of α under H_1 is given by (see Appendix I)

$$\hat{\alpha}_{\mathrm{ML}} = \arg\min_{\alpha} \left| \hat{\mathbf{R}}_{xx}(\alpha) - \hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\hat{\mathbf{R}}_{yx}(\alpha) \right| \tag{9}$$

[2]If P is unknown, it can be estimated using a variety of model order selection techniques (e.g., ([20, App. C]).

where the correlation matrices conditioned on α are given by

$$\hat{\mathbf{R}}_{xx}(\alpha) = \sum_{n=P}^{N-1} [\mathbf{x}_0(n) - \alpha\mathbf{s}(n)][\mathbf{x}_0(n) - \alpha\mathbf{s}(n)]^H + \sum_{n=P}^{N-1}\sum_{k=1}^{K} \mathbf{x}_k(n)\mathbf{x}_k^H(n), \tag{10}$$

$$\hat{\mathbf{R}}_{yy}(\alpha) = \sum_{n=P}^{N-1} [\mathbf{y}_0(n) - \alpha\mathbf{t}(n)]\,[\mathbf{y}_0(n) - \alpha\mathbf{t}(n)]^H + \sum_{n=P}^{N-1}\sum_{k=1}^{K} \mathbf{y}_k(n)\mathbf{y}_k^H(n), \tag{11}$$

$$\hat{\mathbf{R}}_{yx}(\alpha) = \sum_{n=P}^{N-1} [\mathbf{y}_0(n) - \alpha\mathbf{t}(n)][\mathbf{x}_0(n) - \alpha\mathbf{s}(n)]^H + \sum_{n=P}^{N-1}\sum_{k=1}^{K} \mathbf{y}_k(n)\mathbf{x}_k^H(n) \tag{12}$$

and where $\mathbf{y}_k(n) = [\mathbf{x}_k^T(n-1),\ldots,\mathbf{x}_k^T(n-P)]^T$ and $\mathbf{t}(n) = [\mathbf{s}^T(n-1),\ldots,\mathbf{s}^T(n-P)]^T$. Once $\hat{\alpha}_{\mathrm{ML}}$ is available, the ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ under H_1 are obtained as

$$\hat{\mathbf{A}}_{\mathrm{ML},1}^H = \hat{\mathbf{A}}^H(\alpha)|_{\alpha=\hat{\alpha}_{\mathrm{ML}}} \tag{13}$$

$$\hat{\mathbf{Q}}_{\mathrm{ML},1} = \hat{\mathbf{Q}}(\alpha)|_{\alpha=\hat{\alpha}_{\mathrm{ML}}} \tag{14}$$

where the conditional estimates are given by

$$\hat{\mathbf{A}}^H(\alpha) = -\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha), \tag{15}$$

$$\hat{\mathbf{Q}}(\alpha) = \frac{1}{L}\left[\hat{\mathbf{R}}_{xx}(\alpha) - \hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\hat{\mathbf{R}}_{yx}(\alpha)\right] \tag{16}$$

with

$$L = (K+1)(N-P). \tag{17}$$

Remark 1: Although statistically optimum, $\hat{\alpha}_{\mathrm{ML}}$ has no closed-form expression. The cost function (9) is highly nonlinear. A brute-force exhaustive search over the two-dimensional parameter space (i.e., the real and imaginary part of α) is generally impractical. Alternatively, we can resort to Newton-like iterative nonlinear searches, provided an initial estimate of α is available. Hence, there is a need for suboptimum estimators with reduced computational complexity. One such suboptimum estimator is discussed in Section III-C.

B. ML Estimation Under H_0

This is a special case of the one addressed in Section III-A (see Appendix I). The ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ under H_0 are given by

$$\hat{\mathbf{A}}_{\mathrm{ML},0}^H = -\hat{\mathbf{R}}_{yx,0}^H\hat{\mathbf{R}}_{yy,0}^{-1} \tag{18}$$

$$\hat{\mathbf{Q}}_{\mathrm{ML},0} = \frac{1}{L}\left(\hat{\mathbf{R}}_{xx,0} - \hat{\mathbf{R}}_{yx,0}^H\hat{\mathbf{R}}_{yy,0}^{-1}\hat{\mathbf{R}}_{yx,0}\right) \tag{19}$$

where the correlation matrices $\hat{\mathbf{R}}_{xx,0}$, $\hat{\mathbf{R}}_{yy,0}$, $\hat{\mathbf{R}}_{yx,0}$ are obtained from (10)–(12), respectively, by setting $\alpha=0$.

C. Asymptotic ML Estimation Under H_1

We now introduce a computationally more efficient estimator that is asymptotically equivalent to the ML estimator. The estimator is, henceforth, referred to as the AML estimator. The

idea is to replace $\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)$ in the cost function of (9) by a statistically consistent estimate $\hat{\mathbf{A}}_{\rm con}$ (how to obtain such a consistent estimate is discussed next). The resulting cost function $C_1(\alpha)$, which can be shown to be asymptotically equivalent to the cost function of (9) (e.g., [21], [22]), can be written as

$$C_1(\alpha)=\Bigg|\sum_{n=P}^{N-1}\Big\{\tilde{\mathbf{x}}_0(n;\hat{\mathbf{A}}_{\rm con})-\alpha\tilde{\mathbf{s}}(n;\hat{\mathbf{A}}_{\rm con})\Big\}\cdot\Big\{\tilde{\mathbf{x}}_0(n;\hat{\mathbf{A}}_{\rm con})-\alpha\tilde{\mathbf{s}}(n;\hat{\mathbf{A}}_{\rm con})\Big\}^H+\sum_{n=P}^{N-1}\sum_{k=1}^{K}\tilde{\mathbf{x}}_k(n;\hat{\mathbf{A}}_{\rm con})\tilde{\mathbf{x}}_k^H(n;\hat{\mathbf{A}}_{\rm con})\Bigg| \quad (20)$$

where $\hat{\tilde{\mathbf{x}}}_k(n)$ and $\hat{\tilde{\mathbf{s}}}(n)$ are the temporally whitened versions of $\mathbf{x}_k(n)$ and $\mathbf{s}(n)$, respectively, using the consistent AR coefficient estimate $\hat{\mathbf{A}}_{\rm con}$

$$\tilde{\mathbf{x}}_k(n;\hat{\mathbf{A}}_{\rm con})=\mathbf{x}_k(n)+\hat{\mathbf{A}}_{\rm con}^H\mathbf{y}_k(n), \quad (21)$$

$$\tilde{\mathbf{s}}(n;\hat{\mathbf{A}}_{\rm con})=\mathbf{s}(n)+\hat{\mathbf{A}}_{\rm con}^H\mathbf{t}(n). \quad (22)$$

Note that the matrix inside the determinant of (20) is a quadratic form of α. An asymptotic solution is obtained in Appendix II, which is given by

$$\hat{\alpha}_{\rm AML}=\frac{\mathrm{tr}(\hat{\tilde{\mathbf{S}}}^H\boldsymbol{\Psi}^{-1}\hat{\tilde{\mathbf{X}}}_0)}{\mathrm{tr}(\hat{\tilde{\mathbf{S}}}^H\boldsymbol{\Psi}^{-1}\hat{\tilde{\mathbf{S}}})} \quad (23)$$

where

$$\hat{\tilde{\mathbf{S}}}=[\tilde{\mathbf{s}}(P;\hat{\mathbf{A}}_{\rm con}),\quad\ldots,\quad\tilde{\mathbf{s}}(N-1;\hat{\mathbf{A}}_{\rm con})] \quad (24)$$

$$\hat{\tilde{\mathbf{X}}}_k=[\tilde{\mathbf{x}}_k(P;\hat{\mathbf{A}}_{\rm con}),\quad\ldots,\quad\tilde{\mathbf{x}}_k(N-1;\hat{\mathbf{A}}_{\rm con}) \quad (25)$$

$$\boldsymbol{\Psi}=\hat{\tilde{\mathbf{X}}}_0\mathbf{P}^{\perp}\hat{\tilde{\mathbf{X}}}_0^H+\sum_{k=1}^{K}\hat{\tilde{\mathbf{X}}}_k\hat{\tilde{\mathbf{X}}}_k^H \quad (26)$$

with $\mathbf{P}^{\perp}$ denoting the projection matrix projecting to the orthogonal complement of the range of $\hat{\tilde{\mathbf{S}}}^H$

$$\mathbf{P}^{\perp}=\mathbf{I}-\mathbf{P}=\mathbf{I}-\hat{\tilde{\mathbf{S}}}^H(\hat{\tilde{\mathbf{S}}}^H)^{\dagger}. \quad (27)$$

The above AML estimator requires a consistent estimate $\hat{\mathbf{A}}_{\rm con}$, which can be obtained by using a consistent estimate of α in (15). One such estimate is obtained by the least-squares (LS) amplitude estimator:

$$\hat{\alpha}_{\rm LS}=\frac{\mathbf{s}^H\mathbf{x}_0}{\mathbf{s}^H\mathbf{s}} \quad (28)$$

which ignores the fact that the disturbance signal is colored. We show in Appendix III that $\hat{\alpha}_{\rm LS}$ is statistically unbiased and consistent.

To summarize, the AML estimator can be implemented as follows.

Step 1: Determine a consistent estimate $\hat{\mathbf{A}}_{\rm con}$. This can be obtained by first computing the LS amplitude estimate $\hat{\alpha}_{\rm LS}$ as in (28), and using $\hat{\alpha}_{\rm LS}$ in (15)

$$\hat{\mathbf{A}}_{\rm con}^H=-\hat{\mathbf{R}}_{yx}^H(\hat{\alpha}_{\rm LS})\hat{\mathbf{R}}_{yy}^{-1}(\hat{\alpha}_{\rm LS}). \quad (29)$$

Step 2: Compute the AML amplitude estimate $\hat{\alpha}_{\rm AML}$ using (23).

Step 3: Find the AML estimates of the AR coefficients and spatial covariance matrix by substituting $\hat{\alpha}_{\rm AML}$ for α in (15) and (16), respectively.

Remark 2: The AML estimator is obtained based on multiple approximations. The first involves approximating the likelihood function by dropping out the initial samples of the AR process, as shown in Appendix I. The second approximates the cost function in (9) by $C_1(\alpha)$ as in (20), which replaces $\mathbf{A}$ with a consistent estimate $\hat{\mathbf{A}}_{\rm con}$. $C_1(\alpha)$ is further shown to be equivalent to $C_2(\alpha)$ in Appendix II. The third approximation is to replace the nonlinear $C_2(\alpha)$ by a quadratic $C_3(\alpha)$ in Appendix II, which admits a closed form solution. All three approximations are valid in the large-sample case.

Remark 3: The above AML estimator is related to an alternating LS (ALS) estimator discussed in [19], but there are several notable differences. First, AML covers both training ($K\neq 0$) and training-free ($K=0$) cases, whereas ALS, which was introduced to solve an explosive detection problem, considers only the case without training. Second, ALS is an iterative approach, whereas iteration is not required by AML. Finally, by using an asymptotic approximation of the ML cost function, AML is directly related to the ML estimator and asymptotically coincides with the latter. Such an asymptotic relation was not established for ALS. Numerical results in Section V indicate that AML and ML yield nearly identical estimation performance.

D. CRB

From Section III-A, an amplitude estimate is obtained first and then used to produce the nuisance parameter estimates. As such, amplitude estimation is the most critical step in the estimation process. Next, we provide the CRB for amplitude estimation. The CRB specifies a lower bound on the variance of any unbiased amplitude estimator, thus offering a baseline for comparison. The CRB for α is derived in Appendix IV, which is given by

$$\mathrm{CRB}(\alpha)=\left[\sum_{n=P}^{N-1}\tilde{\mathbf{s}}^H(n;\mathbf{A})\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n;\mathbf{A})\right]^{-1}. \quad (30)$$

Like $\hat{\tilde{\mathbf{s}}}(n)$ in (22), $\tilde{\mathbf{s}}(n;\mathbf{A})$ is the temporally whitened version of $\mathbf{s}(n)$, but by using the true AR coefficient matrix $\mathbf{A}$ (the dependence on $\mathbf{A}$ is explicitly shown)

$$\tilde{\mathbf{s}}(n;\mathbf{A})=\mathbf{s}(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{s}(n-p). \quad (31)$$

The CRB for the nuisance parameters can be obtained in a similar fashion, but skipped for brevity.

IV. Parametric GLRT

A. Test Statistic

With the ML parameter estimates obtained in Sections III-A and III-B, the GLR reduces to

$$\mathrm{GLR}=\frac{\max_{\alpha,\mathbf{A},\mathbf{Q}}f_1(\alpha,\mathbf{A},\mathbf{Q})}{\max_{\mathbf{A},\mathbf{Q}}f_0(0,\mathbf{A},\mathbf{Q})}=\frac{\left|\hat{\mathbf{Q}}_{\mathrm{ML},0}\right|^L}{\left|\hat{\mathbf{Q}}_{\mathrm{ML},1}\right|^L} \quad (32)$$

where $\hat{\mathbf{Q}}_{\mathrm{ML},0}$ and $\hat{\mathbf{Q}}_{\mathrm{ML},1}$ are given by (19) and (14), respectively. Equivalently, taking a logarithm (with a scaling constant 2) yields[3]

$$T_{\mathrm{GLRT}} = 2L \ln \frac{|\hat{\mathbf{Q}}_{\mathrm{ML},0}|}{|\hat{\mathbf{Q}}_{\mathrm{ML},1}|} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\mathrm{GLRT}} \tag{33}$$

where γ_{GLRT} denotes the test threshold (see Section IV-B for discussion on the setting of γ_{GLRT}).

Remark 4: The final test statistic is a ratio of two matrix determinants. Note that the two covariance matrix estimates $\hat{\mathbf{Q}}_{\mathrm{ML},0}$ and $\hat{\mathbf{Q}}_{\mathrm{ML},1}$ have an identical form given by (15), except that $\alpha = 0$ for the former and $\alpha = \hat{\alpha}_{\mathrm{ML}}$ for the latter. Hence, once $\hat{\alpha}_{\mathrm{ML}}$ is obtained, the remaining steps involved in calculating $\hat{\mathbf{Q}}_{\mathrm{ML},1}$ are very similar to those needed for $\hat{\mathbf{Q}}_{\mathrm{ML},0}$, which can be performed by the same computing algorithm or hardware, thus simplifying implementation.

Remark 5: Instead of using the nonlinear ML amplitude estimate $\hat{\alpha}_{\mathrm{ML}}$, we can employ the computationally more efficient AML amplitude estimate $\hat{\alpha}_{\mathrm{AML}}$ in calculating the GLRT test statistic. As shown in Section V, the two different versions of GLRT offer nearly identical detection performance.

B. Asymptotic Analysis

As shown in Appedix V, the asymptotic distribution of the parametric GLRT statistic in (33) is given by

$$T_{\mathrm{GLRT}} \overset{a}{\sim} \begin{cases} \chi_2^2, & \text{under } H_0 \\ \chi_2^2(\lambda), & \text{under } H_1 \end{cases} \tag{34}$$

where χ_2^2 denotes the central Chi-squared distribution with 2 degrees of freedom (i.e., exponential distribution) and $\chi_2'^2(\lambda)$ the noncentral Chi-squared distribution with 2 degrees of freedom and noncentrality parameter λ

$$\lambda = 2|\alpha|^2 \sum_{n=P}^{N-1} \tilde{\mathbf{s}}^H(n;\mathbf{A})\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n;\mathbf{A}). \tag{35}$$

Note that λ is related to the signal-to-interference-plus-noise ratio (SINR) at the output of the temporal whitening filter. Using the above result, we can write the asymptotic detection and false alarm probabilities as

$$P_{\mathrm{d}} = \int_{\gamma_{\mathrm{GLRT}}}^{\infty} \frac{1}{2} \exp\left[-\frac{1}{2}(x+\lambda)\right] I_0(\sqrt{\lambda x})\, dx \tag{36}$$

$$P_{\mathrm{f}} = \exp\left(-\frac{1}{2}\gamma_{\mathrm{GLRT}}\right) \tag{37}$$

where $I_0(u)$ is the modified Bessel function of the first kind and zeroth order [18].

Remark 6: The asymptotic distribution under H_0 is independent of the unknown parameters. The probability of false alarm in (37) depends only on the test threshold, which is a design parameter. It is evident that the parametric GLRT asymptotically achieves CFAR.

Remark 7: The above analysis holds under Assumptions AS1 to AS4 of Section II with one exception. In particular, since the ML parameter estimates are asymptotically Gaussian irrespective of the distribution of the observed data, the above analysis still holds if the Gaussian assumption in AS3 is dropped.

V. Numerical Results

In this section, we present simulation results to illustrate the performance of the proposed detection and estimation techniques. The disturbance signal is generated as a multichannel AR(2) process with AR coefficients $\mathbf{A}$ and a spatial covariance matrix $\mathbf{Q}$. These parameters are set to ensure that the AR process is stable and $\mathbf{Q}$ is a valid covariance matrix, but otherwise randomly selected. The signal vector $\mathbf{s}$ corresponds to a uniform equispaced linear array with $J = 4$ antenna elements and randomly selected normalized spatial and Doppler frequencies (see [13]). The SINR is defined as

$$\mathrm{SINR} = |\alpha|^2 \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s} \tag{38}$$

where the $JN \times JN$ space-time covariance matrix can be uniquely determined once $\mathbf{A}$ and $\mathbf{Q}$ are selected. Note that the above SINR can be considered as an *overall* SINR that takes into account all spatial and temporal signals observed within one CPI. A different SINR that is also frequently used is defined based on one snapshot of the array output; see, e.g., [19].

A. Estimation

The estimation results are presented for the estimators discussed in Section III, namely, the LS (28), AML (23), and ML (9) estimators. The ML estimator is implemented via local nonlinear iterative searching, initialized by the AML estimate. We first consider the *training-free* case with $K = 0$. This is also the case considered by the ALS estimator [19] and, thus, we include it for comparison. Fig. 1 shows the mean-squared error (MSE) of the amplitude estimate $\hat{\alpha}$ obtained by each estimator, along with the CRB (30), versus the SINR. It is seen that even for a moderate value of $N = 32$, the AML amplitude estimate is nearly identical to the ML estimate, and both are very close to the CRB and considerably better than the simple LS estimate. It is also observed that the ALS estimate is nearly identical to the AML estimate in this case.

Fig. 2 depicts the results for a *limited-training* case with $K = 1$. The LS and ALS estimators are not included since they do not utilize any training signal for estimation. It is seen that both the AML and ML estimates are nearly identical and close to the CRB for all values of SINR. It is observed that use of training data slightly improves the estimation performance.

B. Detection

For the parametric GLRT (33), the test statistic can be computed using either the ML or AML parameter estimates, as indicated in Remark 5 of Section IV-A. The resulting tests, which are denoted as *parametric GLRT/ML* and *parametric GLRT/AML*, respectively, are compared with the parametric

[3]Although the factor of $2L$ and logarithm in the test statistic can be absorbed by the test threshold, it is retained to keep the asymptotic distribution more compact. See Section IV-B.

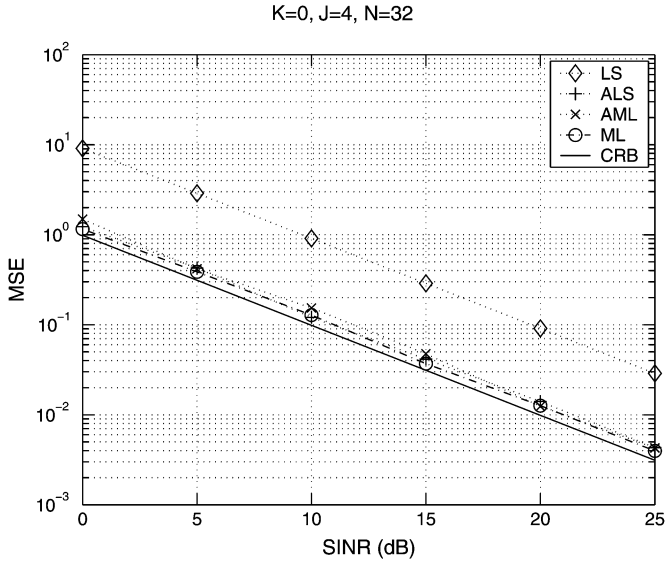


Fig. 1. MSE of amplitude estimate $\hat{\alpha}$ versus SINR when $J = 4$, $N = 32$, and $K = 0$ (no training data).

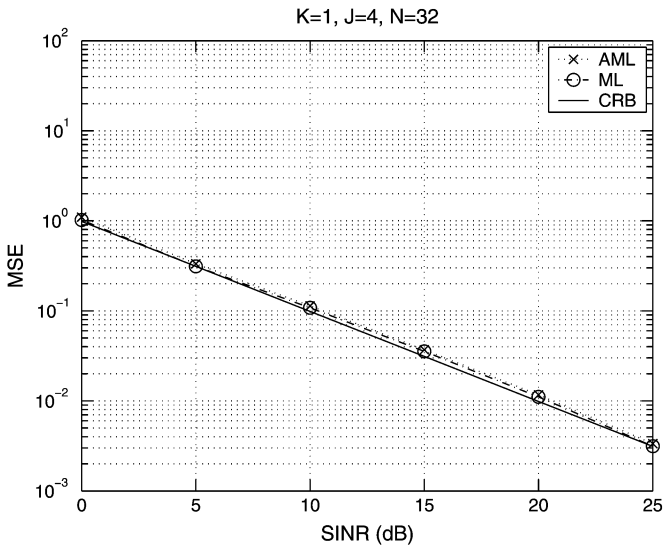


Fig. 2. MSE of amplitude estimate $\hat{\alpha}$ versus SINR when $J = 4$, $N = 32$, and $K = 1$ (limited training data).

Rao detector [16], [17], which is a large-sample approximation of the parametric GLRT. Also included in the comparison are the asymptotic analysis for the parametric GLRT given in Section IV-B, the ideal matched filter (MF) [10] which assumes exact knowledge of $\mathbf{R}$ and, therefore, cannot be used in practice but offers a baseline for comparison, and Kelly's GLRT [7] which is included to show the gain offered by parametric detection.[4] In all examples, we set $J = 4$ and the probability of false alarm $P_{\mathrm{f}} = 0.01$.

The *training-free* $(K = 0)$ case is considered in Figs. 3 to 5, which show the probability of detection versus SINR for various detectors when the number of temporal observations varies from $N = 32, 64$, to 128. Meanwhile, the *limited-training* $(K = 1)$ case is considered in Figs. 6 and 7 for $N = 32$ and 64, respectively. An examination of these figures reveals the following.

- The parametric GLRT/AML yields nearly identical detection performance to that of the parametric GLRT/ML, and may be preferred to the latter due to its reduced complexity.
- For the training-free case, the parametric GLRT is about 3 to 4 dB from the optimum MF bound at $N = 32$; the gap reduces to about 1 dB at $N = 64$ and a fractional dB at $N = 128$. Training, even modest, helps improving the detection, which can be seen by comparing Fig. 3 with Fig. 6, or Fig. 4 with Fig. 7. However, *the degradation incurred by lack of training can be remedied by increasing temporal observations of the test signal*, as seen in Figs. 3 to 5.
- *For small* N (e.g., $N = 32$), *the parametric GLRT outperforms the parametric Rao detector*. At larger values of N, the two detectors exhibit similar performance, especially at the low SINR region.
- The parametric Rao detector is seen to degrade dramatically as the SINR increases. This is not surprising since all Rao tests, including the parametric Rao detector, are based on a *weak* signal approximation of the GLRT ([18, App. 6B]). This has also been observed in ([23, Fig. 3]) for a single-channel detection problem. Such degradation may not be critical in applications where weak signal detection is of primary interest.
- Compared to Kelly's GLRT, both parametric detectors can produce better detection performance with significantly less training or even no training, when N is not too small.

[4]Recall that Kelly's GLRT is a covariance matrix based detector that cannot handle the limited-training or training-free case. In the following examples, we use either $K = 0$ or $K = 1$ for the parametric detectors; but for Kelly's GLRT, K is chosen significantly larger so that $K \geq JN - 1$ to ensure a nonsingular estimate of $\mathbf{R}$ (see discussions in Section I).

VI. Conclusion

We have developed a new parametric GLRT for multichannel adaptive signal detection. The parametric GLRT is obtained by exploiting multichannel AR modeling for the disturbance signal. We have investigated the underlying parameter estimation problem. The ML estimator has been derived, but not in a closed form. An AML estimator has been introduced as an asymptotically optimum but computationally more efficient alternative. We have examined the detection performance of the parametric GLRT as well as a recently proposed parametric Rao detector. We have shown that while both parametric detectors are significantly less dependent on training than conventional covariance matrix based detectors, the parametric GLRT is the better solution of the two, especially when temporal observations of the test signal are limited.

One most interesting feature of the parametric GLRT and Rao detectors is that both use the test and training signals for parameter estimation and can handle the training-free case. We have shown that the performance degradation caused by the lack of training can be remedied by increasing the temporal observations of the test signal. Such a tradeoff may be of interest and exploited in some applications, such as radars, when the environment is highly heterogeneous such that using neighboring range cells for training becomes impossible. In particular, the i.i.d. assumption AS3 concerning the test cell and the neighboring range cells will be seriously violated in that case.

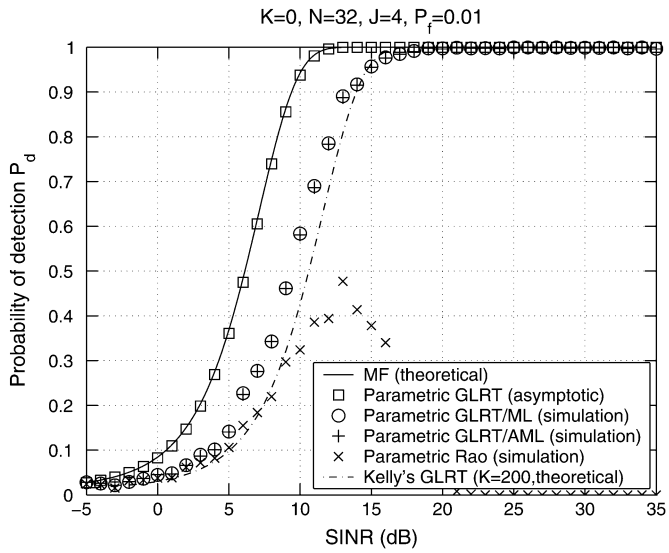


Fig. 3. Probability of detection $P_{\rm d}$ versus SINR when $P_{\rm f} = 0.01, J = 4, N = 32$, and $K = 0$ (no training data).

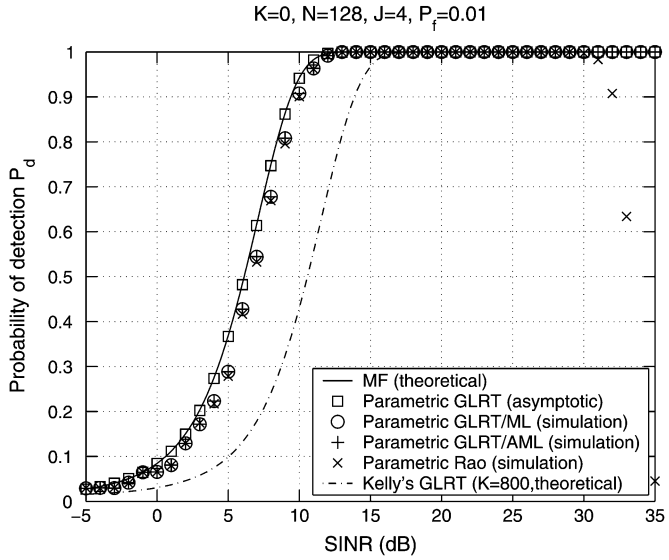


Fig. 5. Probability of detection $P_{\rm d}$ versus SINR when $P_{\rm f} = 0.01, J = 4, N = 128$, and $K = 0$ (no training data).

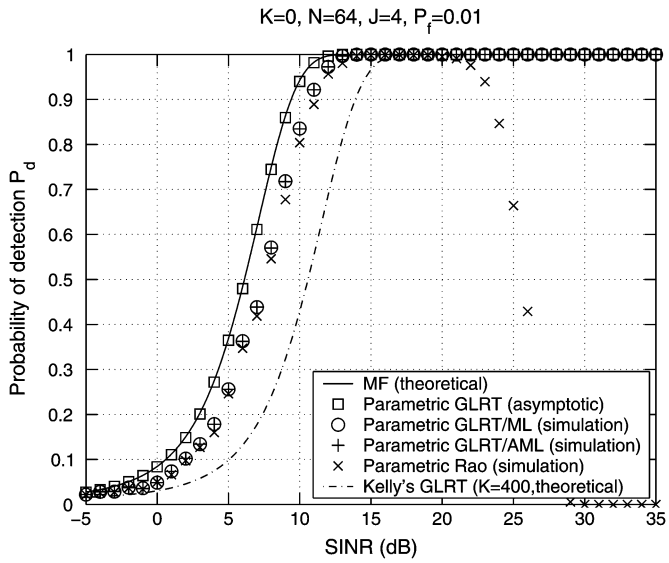


Fig. 4. Probability of detection $P_{\rm d}$ versus SINR when $P_{\rm f} = 0.01, J = 4, N = 64$, and $K = 0$ (no training data).

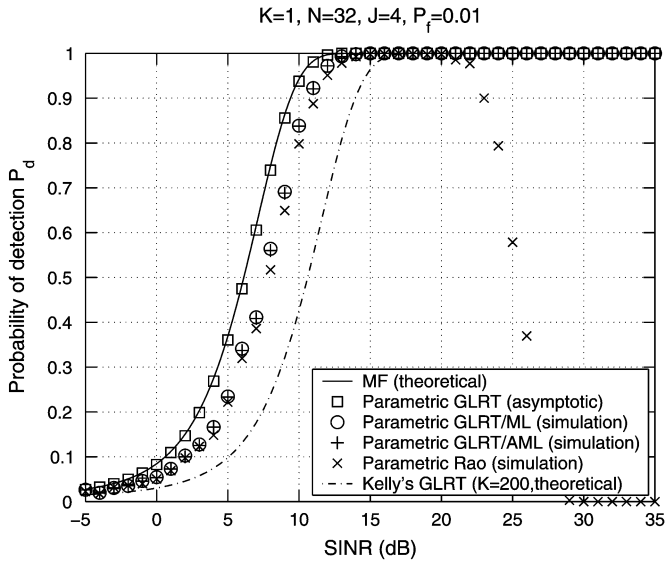


Fig. 6. Probability of detection $P_{\rm d}$ versus SINR when $P_{\rm f} = 0.01, J = 4, N = 32$, and $K = 1$ (limited training data).

APPENDIX I
ML PARAMETER ESTIMATION

In this Appendix, we develop the ML parameter estimators under both hypotheses. Recall that the likelihood functions under both hypotheses differ only in the value of α, that is, $\alpha = 0$ under H_0 and $\alpha \neq 0$ under H_1. We will show that the ML estimates under H_0 can be obtained by setting $\alpha = 0$ in the ML estimates under H_1.

Let $\tilde{\mathbf{x}}_k(n;\mathbf{A})$ denote the temporally whitened version of $\mathbf{x}_{k_{\rm S}}(n)$

$$\tilde{\mathbf{x}}_k(n;\mathbf{A}) = \mathbf{x}_k(n) + \sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{x}_k(n-p). \tag{39}$$

Conditioned on the first P values $\{\mathbf{x}_k(n)\}_{n=0}^{P-1}, k = 0, 1, \ldots, K$, the log-likelihood function is proportional to (within an additive constant) [16]

$$\begin{aligned} &- L\ln|\mathbf{Q}| - \sum_{k=1}^{K}\sum_{n=P}^{N-1} \tilde{\mathbf{x}}_k^H(n;\mathbf{A})\mathbf{Q}^{-1}\tilde{\mathbf{x}}_k(n;\mathbf{A}) \\ &\quad - \sum_{n=P}^{N-1} \{\tilde{\mathbf{x}}_0(n;\mathbf{A}) - \alpha\tilde{\mathbf{s}}(n;\mathbf{A})\}^H\mathbf{Q}^{-1} \\ &\quad \times \{\tilde{\mathbf{x}}_0(n;\mathbf{A}) - \alpha\tilde{\mathbf{s}}(n;\mathbf{A})\}. \end{aligned} \tag{40}$$

It is noted that for the large-sample case, the likelihood function can be well approximated by the above conditional distribution [24]. We, therefore, use (40) for ML estimation. Taking

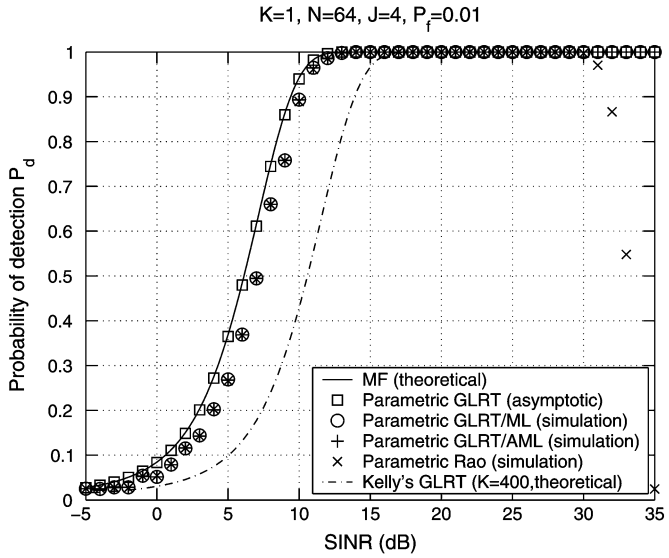


Fig. 7. Probability of detection $P_{\rm d}$ versus SINR when $P_{\rm f} = 0.01, J = 4, N = 64$, and $K = 1$ (limited training data).

the derivative of (40) with respect to $\mathbf{Q}$ and equating it to zero produce the ML estimates of $\mathbf{Q}$ conditioned on α and $\mathbf{A}$:

$$\hat{\mathbf{Q}}(\alpha,\mathbf{A}) = \frac{1}{L}\sum_{n=P}^{N-1}\sum_{k=1}^{K}\big[\tilde{\mathbf{x}}_k(n;\mathbf{A})\tilde{\mathbf{x}}_k^H(n;\mathbf{A}) + \{\tilde{\mathbf{x}}_0(n;\mathbf{A})-\alpha\tilde{\mathbf{s}}(n;\mathbf{A})\} \times \{\tilde{\mathbf{x}}_0(n;\mathbf{A})-\alpha\tilde{\mathbf{s}}(n;\mathbf{A})\}^H\big]. \quad (41)$$

Substituting the above $\hat{\mathbf{Q}}(\alpha,\mathbf{A})$ back in (40), we find that maximizing (40) reduces to minimizing $|\hat{\mathbf{Q}}(\alpha,\mathbf{A})|$. Therefore, the ML estimates of α and $\mathbf{A}$ can be obtained by minimizing $|\hat{\mathbf{Q}}(\alpha,\mathbf{A})|$ with respect to α and $\mathbf{A}$. In turn, we can get the ML estimate of $\mathbf{Q}$ by replacing α and $\mathbf{A}$ with their ML estimates in (41). Next, observe that

$$\begin{aligned} L\hat{\mathbf{Q}}(\alpha,\mathbf{A}) &= \hat{\mathbf{R}}_{xx}(\alpha)+\mathbf{A}^H\hat{\mathbf{R}}_{yx}(\alpha)+\hat{\mathbf{R}}_{yx}^H(\alpha)\mathbf{A}+\mathbf{A}^H\hat{\mathbf{R}}_{yy}(\alpha)\mathbf{A} \\ &= \left(\mathbf{A}^H+\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\right)\hat{\mathbf{R}}_{yy}(\alpha) \\ &\quad\times\left(\mathbf{A}^H+\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\right)^H \\ &\quad+\hat{\mathbf{R}}_{xx}(\alpha)-\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\hat{\mathbf{R}}_{yx}(\alpha) \end{aligned} \quad (42)$$

where the α-dependent correlation matrices are defined in (10)–(12). Since $\hat{\mathbf{R}}_{yy}(\alpha)$ is nonnegative definite and the remaining terms in (42) do not depend on $\mathbf{A}$, it follows that[5]

$$\hat{\mathbf{Q}}(\alpha,\mathbf{A}) \geq \hat{\mathbf{Q}}(\alpha,\mathbf{A})|_{\mathbf{A}=\hat{\mathbf{A}}(\alpha)} \quad (43)$$

where

$$\hat{\mathbf{A}}^H(\alpha) = -\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha). \quad (44)$$

[5]For two nonnegative definite matrices $\mathbf{A}$ and $\mathbf{B}$, we have $\mathbf{A} \geq \mathbf{B}$ if $\mathbf{A}-\mathbf{B}$ is nonnegative definite [22].

When $\hat{\mathbf{Q}}(\alpha,\mathbf{A})$ is minimized, the estimate $\hat{\mathbf{A}}(\alpha)$ of $\mathbf{A}$ will minimize any nondecreasing function including the determinant of $\hat{\mathbf{Q}}(\alpha,\mathbf{A})$ [22]. It should be noted that in finding the estimate of $\mathbf{A}$, we did not impose the constraint that the underlying AR process is stable for the sake of obtaining a simple solution. Hence, the unconstrained ML estimate of $\mathbf{A}$ and $\mathbf{Q}$ conditioned on α are given by (15) and (16), respectively.

Replacing $\mathbf{A}$ in (42) by $\hat{\mathbf{A}}(\alpha)$ followed by minimizing $|\hat{\mathbf{Q}}(\alpha,\hat{\mathbf{A}}(\alpha))|$ yields the ML amplitude estimator of α given by (9). Once the ML estimate $\hat{\alpha}_{\rm ML}$ of the signal amplitude α is obtained, substituting $\hat{\alpha}_{\rm ML}$ in (15) and (16) yields the ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ under H_1, which are given by (13) and (14), respectively.

Since $\alpha = 0$ under H_0, substituting $\alpha = 0$ in (43) and (44) leads to the ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ under H_0, which are given by (18) and (19), respectively.

Appendix II
Derivation of the AML Estimator

Using definitions in (24)–(27), (20) can be written as

$$\begin{aligned} &\left|(\hat{\tilde{\mathbf{X}}}_0-\alpha\hat{\tilde{\mathbf{S}}})(\hat{\tilde{\mathbf{X}}}_0-\alpha\hat{\tilde{\mathbf{S}}})^H+\sum_{k=1}^{K}\hat{\tilde{\mathbf{X}}}_k\hat{\tilde{\mathbf{X}}}_k^H\right| \\ &\quad=\left|(\hat{\tilde{\mathbf{X}}}_0-\alpha\hat{\tilde{\mathbf{S}}})(\mathbf{P}+\mathbf{P}^{\perp})(\hat{\tilde{\mathbf{X}}}_0-\alpha\hat{\tilde{\mathbf{S}}})^H+\sum_{k=1}^{K}\hat{\tilde{\mathbf{X}}}_k\hat{\tilde{\mathbf{X}}}_k^H\right| \\ &\quad=|(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})^H+\mathbf{\Psi}| \\ &\quad=|(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})^H\mathbf{\Psi}^{-1}+\mathbf{I}||\mathbf{\Psi}|. \end{aligned} \quad (45)$$

Next, observe that minimizing

$$C_2(\alpha) = |(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})^H\mathbf{\Psi}^{-1}+\mathbf{I}| \quad (46)$$

is asymptotically equivalent to minimizing [19], [22]:

$$C_3(\alpha) = \mathrm{tr}\{(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})^H\mathbf{\Psi}^{-1}(\hat{\tilde{\mathbf{X}}}_0\mathbf{P}-\alpha\hat{\tilde{\mathbf{S}}})\} \quad (47)$$

which is a quadratic function in α. Minimizing (47) with respect to α leads to the AML estimate $\hat{\alpha}_{\rm AML}$ given by (23) (see also [19]).

Appendix III
Unbiasedness and Consistency of the LS Estimator

First, note that $E[\hat{\alpha}_{\rm LS}] = E[(\mathbf{s}^H\mathbf{x}_0)/(\mathbf{s}^H\mathbf{s})] = E[\alpha + (\mathbf{s}^H\mathbf{d}_0)/(\mathbf{s}^H\mathbf{s})] = \alpha$, which indicates that the LS estimator is unbiased. Moreover, the variance is given by $\mathrm{var}[\hat{\alpha}_{\rm LS}] = E[((\mathbf{s}^H\mathbf{d}_0)/(\mathbf{s}^H\mathbf{s}))((\mathbf{s}^H\mathbf{d}_0)/(\mathbf{s}^H\mathbf{s}))^H] = (\mathbf{s}^H\mathbf{R}\mathbf{s})/(\mathbf{s}^H\mathbf{s})^{\mathbf{2}}$. Next, we show that the variance vanishes as the number of observations N increases. Note that the numerator can be written as

$$\mathbf{s}^H\mathbf{R}\mathbf{s} = \mathbf{s}^H\mathbf{R}^{1/2}\mathbf{R}^{1/2}\mathbf{s} \quad (48)$$

where $\mathbf{R}^{1/2}$ denotes the Hermitian square-root of $\mathbf{R}$. It is known that multiplying by $\mathbf{R}^{1/2}$ is a *coloring* linear transform. Under

Assumption AS4, such a coloring transform using the square-root of the joint space-time covariance matrix $\mathbf{R}$ is asymptotically equivalent to a cascade of an AR filer, which performs temporal coloring, followed by a spatial coloring filter [24]. As such, (48) can be approximated (for large N) as

$$\mathbf{s}^H\mathbf{R}\mathbf{s} \approx \sum_{n=0}^{N-1} \breve{\mathbf{s}}^H(n)\mathbf{Q}\breve{\mathbf{s}}(n) \tag{49}$$

where $\breve{\mathbf{s}}(n) \in \mathbb{C}^{J\times 1}$ denotes the output of the multichannel AR filter as specified in AS4, given the input signal $\{\mathbf{s}(n)\}_{n=0}^{N-1}$.

Let the eigenvalue decomposition of $\mathbf{Q}$ be expressed as: $\mathbf{Q} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where $\mathbf{\Lambda}$ is a diagonal matrix containing all eigenvalues and $\mathbf{U}$ is composed of the corresponding eigenvectors. Let $\lambda_{\max}$ denote the largest eigenvalue of $\mathbf{Q}$. We have

$$\begin{aligned}&\sum_{n=0}^{N-1} \breve{\mathbf{s}}^H(n)\mathbf{Q}\breve{\mathbf{s}}(n)\\ &\quad\le \lambda_{\max}\sum_{n=0}^{N-1} \breve{\mathbf{s}}^H(n)\mathbf{U}\mathbf{U}^H\breve{\mathbf{s}}(n)\\ &= \lambda_{\max}\sum_{n=0}^{N-1} \|\breve{\mathbf{s}}^H(n)\|^2.\end{aligned} \tag{50}$$

It follows that

$$\mathrm{var}[\hat{\alpha}_{\mathrm{LS}}] \le \frac{\lambda_{\max}\sum_{n=0}^{N-1}\|\breve{\mathbf{s}}(n)\|^2}{\left(\sum_{n=0}^{N-1}\|\mathbf{s}(n)\|^2\right)^2} \tag{51}$$

since $(\mathbf{s}^H\mathbf{s})^2 = (\sum_{n=0}^{N-1}\|\mathbf{s}(n)\|^2)^2$. Assuming that the AR filter is stable, we have (e.g., [25])

$$\sum_{n=0}^{N-1}\|\breve{\mathbf{s}}(n)\|^2 \le C\sum_{n=0}^{N-1}\|\mathbf{s}(n)\|^2 \tag{52}$$

for some bounded constant C. Hence, for a given AR filter and spatial covariance $\mathbf{Q}$, the right-hand side (RHS) of (51) vanishes as N goes to infinity. This proves that the LS amplitude estimate is statistically consistent.

Appendix IV
Derivation of CRB

Let $\boldsymbol{\theta} = [\boldsymbol{\theta}_r^T, \boldsymbol{\theta}_s^T]^T$, where $\boldsymbol{\theta}_r = [\Re\{\alpha\}, \Im\{\alpha\}]^T$, and $\boldsymbol{\theta}_s$ contains all nuisance parameters in $\{\mathbf{A}^H(p)\}_{p=1}^P$ and $\mathbf{Q}$. It is shown in [16] that the Fisher information matrix (FIM) for $\boldsymbol{\theta}$ is block diagonal with respect to $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_s$. Therefore, the CRB for the signal amplitude estimate is given by the FIM associated with $\boldsymbol{\theta}_r$, which is given by [16]

$$\mathbf{J}(\boldsymbol{\theta})_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} = \left[2\sum_{n=p}^{N-1}\tilde{\mathbf{s}}^H(n;\mathbf{A})\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n;\mathbf{A})\right]\mathbf{I}_2. \tag{53}$$

By inverting (53) and using $\mathrm{CRB}(\alpha) = \mathrm{CRB}(\Re\{\alpha\}) + \mathrm{CRB}(\Im\{\alpha\})$, we have the CRB given by (30).

Appendix V
Asymptotic Distribution of the Parametric GLRT Statistic

Using the asymptotic results for the GLRT [18], the asymptotic distribution of our parametric GLRT statistic is given by

$$T_{\mathrm{GLRT}} \overset{a}{\sim} \begin{cases}\chi_2^2 & \text{under } H_0,\\ \chi_2^2(\lambda) & \text{under } H_1\end{cases} \tag{54}$$

where the noncentrality parameter λ is given by

$$\lambda = (\boldsymbol{\theta}_{r_1} - \boldsymbol{\theta}_{r_0})^T([\mathbf{J}^{-1}([\boldsymbol{\theta}_{r_0},\boldsymbol{\theta}_s])]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r})^{-1}(\boldsymbol{\theta}_{r_1} - \boldsymbol{\theta}_{r_0}) \tag{55}$$

where $\boldsymbol{\theta}_{r_0}$ and $\boldsymbol{\theta}_{r_1}$ are $\boldsymbol{\theta}_r$ under H_0 and H_1, respectively; $[\mathbf{J}^{-1}([\boldsymbol{\theta}_{r_0},\boldsymbol{\theta}_s])]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}$ is the 2×2 upper-left partition of $\mathbf{J}^{-1}([\boldsymbol{\theta}_{r_0},\boldsymbol{\theta}_s])$. Using the observations $\boldsymbol{\theta}_{r_1} - \boldsymbol{\theta}_{r_0} = [\alpha_R,\ \alpha_I]$ and [cf. (53)]

$$\begin{aligned}&[\mathbf{J}^{-1}([\boldsymbol{\theta}_{r_0},\boldsymbol{\theta}_s])]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}\\ &\qquad= \left[2\sum_{n=P}^{N-1}\tilde{\mathbf{s}}^H(n;\mathbf{A})\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n;\mathbf{A})\right]^{-1}\mathbf{I}_2\end{aligned} \tag{56}$$

we have the asymptotic distribution of the parametric GLRT statistic as shown in (34).

Acknowledgment

The authors would like to express their gratitude to the anonymous reviewers and the associate editor for their constructive comments on the paper.

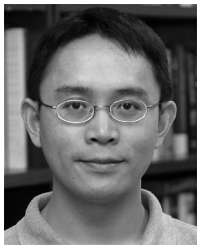
References

[1] J. Ward, "Space-time adaptive processing for airborne radar," Lincoln Lab., MIT, Tech. Rep. 1015, Dec. 1994.
[2] R. Klemm, *Principles of Space-Time Adaptive Processing*. London, U.K.: Inst. Elect. Eng., 2002.
[3] A. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Process. Mag.*, vol. 14, pp. 49–83, Nov. 1997.
[4] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12–16, Jan. 2002.
[5] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.
[6] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 10, no. 6, pp. 853–863, 1974.
[7] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, pp. 115–127, Mar. 1986.
[8] L. Cai and H. Wang, "On adaptive filtering with the CFAR feature and its performance sensitivity to non-Gaussian interference," in *Proc. 24th Ann. Conf. Inf. Sci. Syst.*, Princeton, NJ, Mar. 1990, pp. 558–563.
[9] W. Chen and I. S. Reed, "A new CFAR detection test for radar," *Digit. Signal Process.*, vol. 1, no. 4, pp. 198–214, 1991.
[10] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.
[11] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2538–2541, Sep. 1999.
[12] M. Rangaswamy and J. H. Michels, "A parametric multichannel detection algorithm for correlated non-Gaussian random processes," in *Proc. 1997 IEEE Nat. Radar Conf.*, Syracuse, NY, May 1997, pp. 349–354.
[13] J. R. Román, M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 677–692, Apr. 2000.

[14] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proc. IEEE*, vol. 86, pp. 421–441, Feb. 1998.
[15] J. Li, G. Liu, N. Jiang, and P. Stoica, "Moving target feature extraction for airborne high-range resolution phased-array radar," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 277–289, Feb. 2001.
[16] H. Li, K. J. Sohn, and B. Himed, "The PAMF detector is a parametric Rao test," in *Proc. 39th Asilomar Conf. Signals, Syst., Comp.*, Pacific Grove, CA, Nov. 2005.
[17] K. J. Sohn, H. Li, and B. Himed, "Multichannel parametric Rao detector," presented at the 2006 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Toulouse, France, May 2006.
[18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice-Hall, 1998.
[19] Y. Jiang, P. Stoica, and J. Li, "Array signal processing in the known waveform and steering vector case," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 23–35, Jan. 2004.
[20] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Pearson/Prentice-Hall, 2005.
[21] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall Int., 1989.
[22] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2154–2163, Sep. 1995.
[23] D. Snegupta and S. M. Kay, "Parameter estimation and GLRT detection in colored non-Gaussian autoregressive processes," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1661–1676, Oct. 1990.
[24] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[25] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

Kwang June Sohn (S'07) received the B.E. degree in electronics engineering from Kyungpook National University, Daegu, Korea, in 1995. From 1995 to 1997, he was a Research Assistant with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejon. He received the M.E. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, in 2005.

He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, where he is a Research Assistant. In 1997, he joined the Digital Media Laboratory, LG Electronics, Seoul, Korea, where he was involved in the development of the DVD player and recorder. From 1999 to 2001, he was a Senior Software Engineer with Korea Telecom Freetel (KTF), Seoul, where he was involved in the development of the CDMA handset. From 2001 to 2003, he was a Senior Member of Technical Staff with KTF where he was responsible for the development and assessment of W-CDMA radio access network and handset. His current research interests lie in the area of statistical signal processing with the applications to radars and wireless communications including detection, estimation, multichannel adaptive signal processing, and space-time adaptive processing.



Hongbin Li (M'99) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China, Chengdu, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant with the Department of Electrical and Computer Engineering, University of Florida. He was a Summer Visiting Faculty Member of the Air Force Research Laboratory, Rome, NY, in summers 2003 and 2004. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, where he is an Associate Professor. His current research interests include wireless communications and networking, statistical signal processing, and radars.

Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society. He is an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (1/2003 to 12/2006), the IEEE SIGNAL PROCESSING LETTERS (1/2005 to 12/2006), and the IEEE TRANSACTIONS ON SIGNAL PROCESSING (10/2006 to present), and serves as a Guest Editor for *EURASIP Journal on Applied Signal Processing* Special Issue on Distributed Signal Processing Techniques for Wireless Sensor Networks.

Braham Himed (S'88–M'90–SM'01–F'07) received the B.S. degree in electrical engineering from Ecole Nationale Polytechnique of Algiers, Algeria, in 1984, and the M.S. and Ph.D. degrees from Syracuse University, Syracuse, NY, in 1987 and 1990, respectively, all in electrical engineering.

From 1990 to 1991, he was an Assistant Professor in the Electrical Engineering Department, Syracuse University. In 1991, he joined Adaptive Technology, Inc., Syracuse, where he was responsible for radar systems analyses and the analysis and design of antenna measurement systems. In 1994, he joined Research Associates for Defense Conversion, Marcy, NY, where he was responsible for radar systems analyses and radar signal processing, including the development of techniques for the detection and estimation of targets in severe clutter and jamming environments. From March 1999 to August 2006, he was with the U.S. Air Force Research Laboratory, Sensors Directorate, Radar Signal Processing Branch, Rome, NY, where he was involved with several aspects of airborne and spaceborne phased-array radar systems. Since August 2006, he has been with Signal Labs inc., Reston, VA, where he serves as Chief Research Officer. His research interests include detection, estimation, multichannel adaptive signal processing, time series analyses, array processing, space-time adaptive processing, ground penetrating radar, and medical imaging.

Dr. Himed is the recipient of the 2001 IEEE region I award for his work on bistatic radar systems, algorithm development, and phenomenology. He is a member of the Radar Systems Panel.