Parametric Rao Tests for Multichannel Adaptive Detection in Partially Homogeneous Environment

PU WANG, Student Member, IEEE
HONGBIN LI, Senior Member, IEEE
Stevens Institute of Technology

BRAHAM HIMED, Fellow, IEEE
Air Force Research Laboratory

This paper considers the problem of detecting a multichannel signal in partially homogeneous environments, where the disturbances in both test signal and training signals share the same covariance matrix up to an unknown power scaling factor. Two different parametric Rao tests, referred to as the normalized parametric Rao (NPRao) test and the scale-invariant parametric Rao (SI-PRao) test, respectively, are developed by modeling the disturbance as a multichannel autoregressive (AR) process. The NPRao and SI-PRao tests entail reduced training requirements and computational efficiency, compared with conventional fully adaptive, covariance matrix based solutions. The SI-PRao test attains asymptotically a constant false alarm rate (CFAR) that is independent of the covariance matrix and power scaling factor of the disturbance. Comparisons with the covariance matrix based, scale-invariant generalized likelihood ratio test (GLRT), also known as the adaptive coherence estimator (ACE), are included. Numerical results show that the parametric Rao detectors, in particular the SI-PRao test, attain considerably better detection performance and use significantly less training than the ACE detector.

Manuscript received September 24, 2009; revised March 15, 2010; released for publication July, 2010.

IEEE Log No. T-AES/47/3/941767.

Refereeing of this contribution was handled by F. Gini.

This work was supported in part by a subcontract with Dynetics, Inc. for research sponsored by the Air Force Research Laboratory (AFRL) under Contract FA8650-08-D-1303.

Authors' addresses: P. Wang and H. Li, Dept. of Electrical and Computer Engineering, Stevens Institute of Technology, Castel Point on Hudson, Hoboken, NJ 07030, E-mail: (hli@stevens.edu); B. Himed, AFRL/RYMD, 2241 Avionics Circle, Bldg. 620, Dayton, OH 45433.

I. INTRODUCTION

Multichannel adaptive detection is encountered in numerous applications including radar [1, 2], wireless communications [3], hyperspectral imaging [4, 5], and others. Space-time adaptive processing (STAP)-based multichannel adaptive detection has been successfully utilized to mitigate the effect of clutter and/or interference in radar, remote sensing, and communication systems [1–5].

Traditional STAP detectors are developed usually for homogeneous environments, where the disturbances in both the test and training signals are assumed to be independent and identically distributed (IID). Examples include the Reed, Mallet, and Brennan detector [6], Kelly's generalized likelihood ratio test (GLRT) [7], the adaptive matched filter (AMF) detector [8, 9], and Rao test [10], among others. For the above STAP detectors, estimation of the space-time covariance matrix from training signals requires a large number of training signals and excessive computation power, especially when the joint space-time dimension is large. To alleviate these problems, parametric STAP detectors have been developed by modeling the disturbance as a multichannel autoregressive (AR) process, such as the parametric AMF (PAMF) [11], parametric Rao test [12], and parametric GLRT [13].

In this paper, we consider multichannel signal detection in partially homogeneous environments, where the test signal shares the same covariance matrix with the training signals up to an unknown power scaling factor under the null hypothesis [14–17]. This scenario is motivated by the following observation: A number of guard cells used in the STAP implementation to separate the test signal and training signals may lead to a power difference between the test and training signals. Specifically, the training signals over range cells are assumed to be IID $\mathcal{CN}(0,\mathbf{R})$ which denotes a complex Gaussian distribution with zero mean and covariance matrix $\mathbf{R}$, while the disturbance in the test signal is independent with the training signals with distribution $\mathcal{CN}(0,\lambda\mathbf{R})$, where λ denotes the unknown power scaling factor. Depending on a priori knowledge about the power scaling factor λ, or the covariance matrix $\mathbf{R}$, or both, nonadaptive and adaptive STAP detectors have been developed in [15]–[18]. The scale-invariant GLRT, which is also known as the adaptive coherence estimator (ACE), is first introduced in [18] for the compound Gaussian environment and also developed in [16] for the partially homogeneous environment. It is shown that the ACE coincides with the Rao and Wald tests in the partially homogeneous environment [17], which is also a constant false alarm rate (CFAR) detector. As a covariance matrix based STAP detector, the ACE needs to estimate the covariance matrix from the target-free training signals which entails large

training requirement, and to invert the covariance matrix for implementation also leads to a high computational complexity. At a minimum, $K \geq JN$ IID training signals are needed to ensure a full-rank estimate of the $JN \times JN$ covariance matrix $\mathbf{R}$, where J denotes the number of spatial channels and N the number of temporal observations. For example, the KASSPER dataset consists of $J = 11$ spatial channels and $N = 32$ coherent pulses for a total of $JN = 352$ spatial-temporal dimension. Such a demanding training requirement usually cannot be met in practice. It also makes the homogeneous (i.e., IID) assumption across such a broad range of training cells impractical.

To address this issue, we take an approach where the disturbance in the partially homogeneous environment is modeled by a multichannel AR process. This modification results in a parametric STAP detector which requires a relaxed local homogeneous assumption for the training signals: that is, the training signals are assumed to be IID over a small duration, as opposed to the excessive homogeneous training signals required by the ACE. The local homogeneous assumption is reasonable because the effect of clutter variation across a small area is generally negligible. In the compound Gaussian model [18–23], it is often assumed that the texture component is slow varying, which is similar to the local homogeneous assumption.

In this paper, parametric detection by using the Rao principle is considered for the partially homogeneous environment. The first parametric Rao test, which is referred to the normalized parametric Rao (NPRao) test, is developed in a heuristic way. Specifically, it first assumes the knowledge of the power scaling factor, applies the Rao principle and obtains a Rao test that depends on the power scaling factor, and finally replaces the unknown power scaling factor by some estimate. The second Rao test, referred to the scale-invariant parametric Rao (SI-PRao) test, is developed by applying the Rao principle with the joint maximum likelihood (ML) estimates of all unknown parameters in the null hypothesis. The asymptotical distribution of the SI-PRao test is derived in closed form. It is shown that the SI-PRao test is asymptotically independent of the unknown parameters in the null hypothesis, which results in the property of CFAR. Comparisons between the ACE detector show that our parametric Rao tests in the partially homogeneous environment exhibit improved detection performance when the training signals are limited, i.e., when K is too small to yield a reliable estimate of the space-time covariance matrix $\mathbf{R}$ that is required by the ACE detector.

The rest of this paper is organized as follows. Section II contains the data model and problem statement. Prior solutions to the problem of interest are briefly reviewed in Section III. The NPRao and SI-PRao tests are developed and summarized in Section IV. Asymptotical performance of the SI-PRao detector is also included in this section. Numerical results are presented in Section V. Finally, Section VI contains the concluding remarks.

II. DATA MODEL AND PROBLEM STATEMENT

A. Data Model

Consider the problem of detecting a known multichannel signal with unknown amplitude in the presence of spatially and temporally correlated disturbance: (e.g., [1]):

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

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. Besides the test signal $\mathbf{x}_0(n)$, there may be a set of training signals $\mathbf{x}_k(n)$, $k = 1,2,\ldots,K$, to assist in the signal detection:

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

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

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

$$\begin{aligned} \mathbf{s} &= [\mathbf{s}^T(0), \mathbf{s}^T(1), \ldots, \mathbf{s}^T(N-1)]^T \\ \mathbf{d}_k &= [\mathbf{d}_k^T(0), \mathbf{d}_k^T(1), \ldots, \mathbf{d}_k^T(N-1)]^T \\ \mathbf{x}_k &= [\mathbf{x}_k^T(0), \mathbf{x}_k^T(1), \ldots, \mathbf{x}_k^T(N-1)]^T \end{aligned}$$

where $k = 0,1,\ldots,K$. It follows that (1) can be rewritten in a compact form

$$\begin{aligned} H_0 &: \quad \mathbf{x}_0 = \mathbf{d}_0 \\ H_1 &: \quad \mathbf{x}_0 = \alpha \mathbf{s} + \mathbf{d}_0. \end{aligned} \tag{3}$$

The binary composite hypothesis testing problem is to select between $H_0 : \alpha = 0$ and $H_1 : \alpha \neq 0$.

The general assumptions for the STAP in partially homogeneous environment are [15]–[17], [24], [25]:

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 $\mathbf{d}_0$ and $\{\mathbf{d}_k\}_{k=1}^K$ are mutually independent with distribution $\mathcal{CN}(\mathbf{0}, \lambda\mathbf{R})$ and $\mathcal{CN}(\mathbf{0}, \mathbf{R})$, respectively, where $\lambda > 0$ is an unknown power scaling factor.

If $\lambda = 1$, the partially homogeneous environment reduces to the homogeneous case. In this paper, a

multichannel AR process is employed to model the disturbance and the AS3 is modified as follows:

AS3′: The disturbance signal $\mathbf{d}_k(n)$, $k = 0,\ldots,K$, can be modeled as a J-channel AR(P) process with model order P:

$$\mathbf{d}_k(n) = -\sum_{i=1}^{P} \mathbf{A}^H(i)\mathbf{d}_k(n-i) + \varepsilon_k(n) \tag{4}$$

where $\{\mathbf{A}^H(i)\}_{i=1}^{P}$ denotes the unknown $J \times J$ AR coefficient matrices, $\varepsilon_k(n)$ denotes the $J \times 1$ spatial noise vectors that are temporally white but spatially colored: $\varepsilon_0(n) \sim \mathcal{CN}(\mathbf{0}, \lambda\mathbf{Q})$ and $\{\varepsilon_k(n)\}_{k=1}^{K} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$, respectively, where λ is the unknown power scaling factor, and $\mathbf{Q}$ denotes the unknown $J \times J$ spatial covariance matrix.

Note that the power scaling factor λ on the spatial covariance matrix ensures the same power scaling factor on the spatial-temporal covariance matrix. The problem of interest here is to develop a decision rule for the above composite hypothesis testing problem based on assumptions AS1, AS2, and AS3′.

III. PRIOR SOLUTIONS

Depending on the amount of a prior knowledge about the unknown parameters, a number of solutions for the STAP detection in partially homogeneous environments have been proposed. If the space-time covariance matrix $\mathbf{R}$ and the power scaling factor λ are both known exactly, the optimal detector is the phase-invariant matched filter (PIMF) [20, 25]

$$T_{\text{PIMF}} = \frac{|\mathbf{s}^H\mathbf{R}^{-1}\mathbf{x}_0|^2}{\lambda\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\text{PIMF}} \tag{5}$$

where γ_{PIMF} denotes the PIMF threshold subject to a selected probability of false alarm. In the case of unknown λ, the normalized match filter (NMF) replaces λ with an estimate, which is given by

$$T_{\text{NMF}} = \frac{|\mathbf{s}^H\mathbf{R}^{-1}\mathbf{x}_0|^2}{(\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s})(\mathbf{x}_0^H\mathbf{R}^{-1}\mathbf{x}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\text{NMF}} \tag{6}$$

where γ_{NMF} denotes the NMF threshold. It should be noted that both the nonadaptive PIMF and NMF detectors cannot be implemented in practice since $\mathbf{R}$ is unknown, but they can be used as a baseline for performance comparison.

Adaptive detectors are formulated by estimating the covariance matrix $\mathbf{R}$ from the target-free training signals

$$\hat{\mathbf{R}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{x}_k\mathbf{x}_k^H. \tag{7}$$

The normalized adaptive matched filter (NAMF) [18], also known as the ACE [15], is

$$T_{\text{ACE}} = \frac{|\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{x}_0|^2}{(\mathbf{s}^H\hat{\mathbf{R}}^{-1}\mathbf{s})(\mathbf{x}_0^H\hat{\mathbf{R}}^{-1}\mathbf{x}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\text{ACE}} \tag{8}$$

where γ_{ACE} denotes the ACE threshold. It is shown that the ACE is equivalent to the scale-invariant GLRT in the partially homogeneous environment [15].

By utilizing a multichannel AR model, the normalized parametric adaptive matched filter (NPAMF) is reported [20, 21, 26]:

$$T_{\text{NPAMF}} = \frac{\left|\sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}_{\text{P}}^H(n)\hat{\mathbf{Q}}_{\text{P}}^{-1}\hat{\tilde{\mathbf{x}}}_{0,\text{P}}(n)\right|^2}{\left(\sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}_{\text{P}}^H(n)\hat{\mathbf{Q}}_{\text{P}}^{-1}\hat{\tilde{\mathbf{s}}}_{\text{P}}(n)\right)\left(\sum_{n=P}^{N-1} \hat{\tilde{\mathbf{x}}}_{0,\text{P}}^{\text{H}}(n)\hat{\mathbf{Q}}_{\text{P}}^{-1}\hat{\tilde{\mathbf{x}}}_{0,\text{P}}(n)\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\text{NPAMF}} \tag{9}$$

where γ_{NPAMF} denotes the NPAMF threshold, $\hat{\mathbf{Q}}_{\text{P}}$ denotes an estimate of the spatial covariance matrix $\mathbf{Q}$, and $\hat{\tilde{\mathbf{x}}}_{0,\text{P}}(n)$ and $\hat{\tilde{\mathbf{s}}}_{\text{P}}(n)$ are the temporally whitened test and steering vector, respectively. The NPAMF was originally developed as a solution for detection in the compound-Gaussian environment [20, 21]; it can also be used in the partially homogeneous environment.

IV. PARAMETRIC RAO TESTS IN PARTIALLY HOMOGENEOUS ENVIRONMENT

In this section two parametric Rao detectors based on assumptions AS1, AS2, and AS3′ are developed. The NPRao test is developed in a manner similar to the NPAMF, while the SI-PRao test is derived by finding the ML estimates of all nuisance parameters under the null hypothesis.

A. Normalized Parametric Rao Test

The NPRao test is a heuristic detector obtained via a two-step approach. First, by assuming that the power scaling parameter λ is known, a Rao test can be obtained by finding the ML estimates of the AR coefficient matrices $\{\mathbf{A}^H(p)\}$ and spatial covariance matrix $\mathbf{Q}$. The derivation is similar to the Rao test in [12] for a homogeneous environment. Hence, we skip the details and just present the final result. The new detector, conditioned on a known λ, is given by

$$T_{\text{PRao}} = \frac{2\left|\sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}^H(n)\hat{\mathbf{Q}}^{-1}\hat{\tilde{\mathbf{x}}}_0(n)\right|^2}{\lambda\left(\sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}^H(n)\hat{\mathbf{Q}}^{-1}\hat{\tilde{\mathbf{s}}}(n)\right)} \tag{10}$$

where $\hat{\tilde{\mathbf{s}}}$ and $\hat{\tilde{\mathbf{x}}}_0$ are the temporally whitened steering vector and test signal, respectively, by using a set of estimates $\{\hat{\mathbf{A}}^H(p)\}$ of the AR coefficient matrices:

$$\hat{\tilde{\mathbf{s}}}(n) = \mathbf{s}(n) + \sum_{p=1}^{P} \hat{\mathbf{A}}^H(p)\mathbf{s}(n-p) \tag{11}$$

$$\hat{\tilde{\mathbf{x}}}_0(n) = \mathbf{x}_0(n) + \sum_{p=1}^{P} \hat{\mathbf{A}}^H(p)\mathbf{x}_0(n-p) \tag{12}$$

where $\hat{\mathbf{A}}^H = [\hat{\mathbf{A}}^H(1), \hat{\mathbf{A}}^H(2), \ldots, \hat{\mathbf{A}}^H(P)]$ is given by [12]

$$\hat{\mathbf{A}}^H = -\hat{\mathbf{R}}_{yx}^H \hat{\mathbf{R}}_{yy}^{-1}. \tag{13}$$

The $\hat{\mathbf{Q}}$ in (10) is an estimate of the spatial covariance matrix $\mathbf{Q}$ [12]:

$$\hat{\mathbf{Q}} = \frac{1}{L}(\hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{yx}^H \hat{\mathbf{R}}_{yy}^{-1} \hat{\mathbf{R}}_{yx}) \tag{14}$$

where $L = (K+1)(N-P)$ and

$$\hat{\mathbf{R}}_{xx} = \sum_{k=0}^{K} \sum_{n=P}^{N-1} \mathbf{x}_k(n) \mathbf{x}_k^H(n) \tag{15}$$

$$\hat{\mathbf{R}}_{yy} = \sum_{k=0}^{K} \sum_{n=P}^{N-1} \mathbf{y}_k(n) \mathbf{y}_k^H(n) \tag{16}$$

$$\hat{\mathbf{R}}_{yx} = \sum_{k=0}^{K} \sum_{n=P}^{N-1} \mathbf{y}_k(n) \mathbf{x}_k^H(n) \tag{17}$$

and $\mathbf{y}_k(n) = [\mathbf{x}_k^T(n-1), \ldots, \mathbf{x}_k^T(n-P)]^T \in \mathbb{C}^{JP \times 1}$, $k = 0, \ldots, K$.

Second, we replace the power scaling factor λ in (10) by some estimate. One candidate is given by [20], [21] (also see (9))

$$\hat{\lambda} = \sum_{n=P}^{N-1} \hat{\tilde{\mathbf{x}}}_0^H(n) \hat{\mathbf{Q}}^{-1} \hat{\tilde{\mathbf{x}}}_0(n). \tag{18}$$

This leads to our NPRao detector given by

$$T_{\text{NPRao}} = \frac{2 \left| \sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}^H(n) \hat{\mathbf{Q}}^{-1} \hat{\tilde{\mathbf{x}}}_0(n) \right|^2}{\left(\sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}^H(n) \hat{\mathbf{Q}}^{-1} \hat{\tilde{\mathbf{s}}}(n) \right) \left(\sum_{n=P}^{N-1} \hat{\tilde{\mathbf{x}}}_0^H(n) \hat{\mathbf{Q}}^{-1} \hat{\tilde{\mathbf{x}}}_0(n) \right)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\text{NPRao}} \tag{19}$$

where γ_{NPRao} denotes the NPRao test threshold.

It is noted that the NPAMF detector is equivalent to the NPRao test with one exception: while the NPAMF detector uses only training signals for parameter estimation, the NPRao detector uses both training and test signals for estimation. The difference is similar to the one discussed in [12] between the PAMF detector and the parametric Rao detector.

B. Scale-Invariant Parametric Rao Test

The heuristic NPRao test cannot ensure the invariance to the power scaling factor, which is needed for effective clutter mitigation in the partially homogeneous environment [15, 16]. To this end, a parametric Rao test by evaluating the ML estimates of all nuisance parameters including λ, $\mathbf{A}$, and $\mathbf{Q}$ is developed and the resulting detector, the SI-PRao test, is shown to be invariant to the power scaling factor.

Specifically, as shown in the Appendix, the SI-PRao test is given by

$$T_{\text{SI-PRao}} = \frac{2 \left| \sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}^H(n;\hat{\lambda}) \hat{\mathbf{Q}}^{-1}(\hat{\lambda}) \hat{\tilde{\mathbf{x}}}_0(n;\hat{\lambda}) \right|^2}{\hat{\lambda} \sum_{n=P}^{N-1} \hat{\tilde{\mathbf{s}}}^H(n;\hat{\lambda}) \hat{\mathbf{Q}}^{-1}(\lambda) \hat{\tilde{\mathbf{s}}}(n;\hat{\lambda})} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_{\text{SI-PRao}}. \tag{20}$$

The $\hat{\lambda}$ in (20) is the ML estimate of λ, which is obtained by solving the following equation:

$$\frac{J}{K+1} - \sum_{i=1}^{J(P+1)} \frac{1}{1+\lambda\mu_i} + \sum_{i=1}^{JP} \frac{1}{1+\lambda\nu_i} = 0 \tag{21}$$

where $\{\mu_i\}_{i=1}^{J(P+1)}$ and $\{\nu_i\}_{i=1}^{JP}$ are the eigenvalues of the matrices $\hat{\mathbf{R}}_0^{-1/2} \hat{\mathbf{R}}_K \hat{\mathbf{R}}_0^{-1/2}$ and $\hat{\mathbf{R}}_{0,y}^{-1/2} \hat{\mathbf{R}}_{K,y} \hat{\mathbf{R}}_{0,y}^{-1/2}$, respectively, and

$$\hat{\mathbf{R}}_0 = \begin{bmatrix} \sum_{n=P}^{N-1} \mathbf{y}_0(n) \mathbf{y}_0^H(n) & \sum_{n=P}^{N-1} \mathbf{y}_0(n) \mathbf{x}_0^H(n) \\ \sum_{n=P}^{N-1} \mathbf{x}_0(n) \mathbf{y}_0^H(n) & \sum_{n=P}^{N-1} \mathbf{x}_0(n) \mathbf{x}_0^H(n) \end{bmatrix} \tag{22}$$

$$\hat{\mathbf{R}}_K = \begin{bmatrix} \sum_{k=1}^{K} \sum_{n=P}^{N-1} \mathbf{y}_k(n) \mathbf{y}_k^H(n) & \sum_{k=1}^{K} \sum_{n=P}^{N-1} \mathbf{y}_k(n) \mathbf{x}_k^H(n) \\ \sum_{k=1}^{K} \sum_{n=P}^{N-1} \mathbf{x}_k(n) \mathbf{y}_k^H(n) & \sum_{k=1}^{K} \sum_{n=P}^{N-1} \mathbf{x}_k(n) \mathbf{x}_k^H(n) \end{bmatrix} \tag{23}$$

$$\hat{\mathbf{R}}_{0,y} = \sum_{n=P}^{N-1} \mathbf{y}_0(n) \mathbf{y}_0^H(n) \tag{24}$$

$$\hat{\mathbf{R}}_{K,y} = \sum_{k=1}^{K} \sum_{n=P}^{N-1} \mathbf{y}_k(n) \mathbf{y}_k^H(n). \tag{25}$$

REMARK The existence of a nonnegative solution $\hat{\lambda}$ to (21) is guaranteed. To see this, note that $\hat{\mathbf{R}}_0^{-1/2} \hat{\mathbf{R}}_K \hat{\mathbf{R}}_0^{-1/2}$ and $\hat{\mathbf{R}}_{0,y}^{-1/2} \hat{\mathbf{R}}_{K,y} \hat{\mathbf{R}}_{0,y}^{-1/2}$ are positive definite. Therefore, μ_i and ν_i are positive. The function

$$f(\lambda) = \frac{J}{K+1} - \sum_{i=1}^{J(P+1)} \frac{1}{1+\lambda\mu_i} + \sum_{i=1}^{JP} \frac{1}{1+\lambda\nu_i} \tag{26}$$

is a continuous function on $(0, +\infty)$. When $K > 0$, i.e., at least one training signal is available, we have

$$\lim_{\lambda \to 0} f(\lambda) = \frac{J}{K+1} - J < 0 \tag{27}$$

$$\lim_{\lambda \to \infty} f(\lambda) = \frac{J}{K+1} > 0 \tag{28}$$

which implies that there is at least one λ within in the interval $(0, +\infty)$ giving $f(\lambda) = 0$. On the other hand, the uniqueness of the ML estimate of λ is more

difficult to establish due to the nonlinear dependence on the parameters μ_i and ν_i. Numerical examples suggest that the solution is unique. For illustration purposes, Fig. 1 shows the numerical evaluation of $f(\lambda)$ for several sets of system parameters. It is seen that the function $f(\lambda)$ is a monotonically increasing function over $\lambda \in (0, +\infty)$, which, combining with (27) and (28), suggests that the solution is unique. However, a more rigorous proof is not available.

Once the ML estimate of λ, the temporally whitened steering vector $\hat{\tilde{\mathbf{s}}}(n;\hat{\lambda})$ and test signal $\hat{\tilde{\mathbf{x}}}_0(n;\hat{\lambda})$ in (20) are obtained from (11) and (12) with the following ML estimate of $\mathbf{A}$

$$\hat{\mathbf{A}}^H = -\hat{\mathbf{R}}_{yx}^H(\hat{\lambda})\hat{\mathbf{R}}_{yy}^{-1}(\hat{\lambda}). \tag{29}$$

Finally, $\hat{\mathbf{Q}}(\hat{\lambda})$ in (20), the ML estimate of $\mathbf{Q}$, is given by

$$\hat{\mathbf{Q}}(\hat{\lambda}) = \frac{1}{L}(\hat{\mathbf{R}}_{xx}(\hat{\lambda}) - \hat{\mathbf{R}}_{yx}^H(\hat{\lambda})\hat{\mathbf{R}}_{yy}^{-1}(\hat{\lambda})\hat{\mathbf{R}}_{yx}(\hat{\lambda})) \tag{30}$$

where

$$\hat{\mathbf{R}}_{xx}(\hat{\lambda}) = \hat{\lambda}^{-1}\sum_{n=P}^{N-1}\mathbf{x}_0(n)\mathbf{x}_0^H(n) + \sum_{k=1}^{K}\sum_{n=P}^{N-1}\mathbf{x}_k(n)\mathbf{x}_k^H(n) \tag{31}$$

$$\hat{\mathbf{R}}_{yy}(\hat{\lambda}) = \hat{\lambda}^{-1}\sum_{n=P}^{N-1}\mathbf{y}_0(n)\mathbf{y}_0^H(n) + \sum_{k=1}^{K}\sum_{n=P}^{N-1}\mathbf{y}_k(n)\mathbf{y}_k^H(n) \tag{32}$$

$$\hat{\mathbf{R}}_{yx}(\hat{\lambda}) = \hat{\lambda}^{-1}\sum_{n=P}^{N-1}\mathbf{y}_0(n)\mathbf{x}_0^H(n) + \sum_{k=1}^{K}\sum_{n=P}^{N-1}\mathbf{y}_k(n)\mathbf{x}_k^H(n). \tag{33}$$

C. Asymptotic Detection Performance

According to [27], the asymptotic distribution of the Rao test statistics can be obtained as

$$T_{\text{SI-PRao}} \overset{a}{\sim} \begin{cases} \chi_2^2, & \text{under } H_0 \\ \chi_2'^2(\rho), & \text{under } H_1 \end{cases} \tag{34}$$

where χ_2^2 denotes the central chi-squared distribution with 2 degrees of freedom and $\chi_2'^2(\rho)$ the noncentral chi-squared distribution with 2 degrees of freedom and noncentrality parameter ρ

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

where $\tilde{\mathbf{s}}$ is the temporally whitened steering vector given by (11) but with $\hat{\mathbf{A}}$ replaced by the true AR coefficient matrix $\mathbf{A}$. From (35), it is ready to show that, for a given threshold, the probability of false alarm is

$$P_{\text{f}} = \exp\left(-\frac{\gamma_{\text{SI-PRao}}}{2}\right) \tag{36}$$

which shows that the statistic of the SI-PRao test under H_0 is independent of the power scaling factor and the covariance matrix, and further implies the SI-PRao test is a CFAR detector, while the probability of detection is

$$P_{\text{d}} = \int_{\gamma_{\text{SI-PRao}}}^{\infty}\frac{1}{2}\exp\left(-\frac{x+\rho}{2}\right)I_0\left(\sqrt{\rho x}\right)dx \tag{37}$$

where $I_0(x)$ is the modified Bessel function of the first kind and zeroth order.

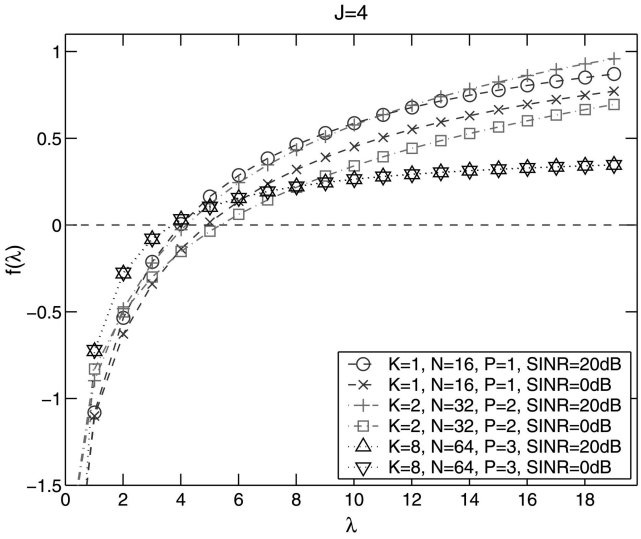


Fig. 1. Numerical evaluation of $f(\lambda)$ for several sets of system parameters when $J = 4$ and $\lambda = 4$.

V. NUMERICAL EXAMPLES

We now report simulation results for the proposed detectors. Throughout this section, the disturbance signal is generated as a multichannel AR(2) process with AR coefficient $\mathbf{A}$ 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 are randomly selected. The signal vector $\mathbf{s}$ corresponds to a uniform equispaced linear array with $J = 4$ antenna elements, N temporal pulses, and randomly selected normalized spatial frequency ω_s and Doppler frequency ω_d. The steering vector is given by

$$\mathbf{s} = \mathbf{s}_t(\omega_d)\otimes\mathbf{s}_s(\omega_s) \tag{38}$$

where $\mathbf{s}_t(\omega_d)$ denotes the $N \times 1$ temporal steering vector

$$\mathbf{s}_t(\omega_d) = \frac{1}{\sqrt{N}}[1, e^{j\omega_d}, \ldots, e^{j(N-1)\omega_d}]^T \tag{39}$$

and $\mathbf{s}_s(\omega_s)$ denotes the $J \times 1$ spatial steering vector

$$\mathbf{s}_s(\omega_s) = \frac{1}{\sqrt{J}}[1, e^{j\omega_s}, \ldots, e^{j(J-1)\omega_s}]^T. \tag{40}$$

The signal-to-interference-plus-noise ratio (SINR) is defined as

$$\text{SINR} = |\alpha|^2\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s} \tag{41}$$

where the $JN \times JN$ covariance matrix $\mathbf{R}$ can be uniquely determined once $\mathbf{A}$ and $\mathbf{Q}$ are selected.

Numerical simulations are provided for various values of K, N, and λ when $J = 4$ and $P = 2$. In particular, we consider two distinct cases: 1) the limited-training case, e.g., $K = 2$ and $K = 4$; and 2) the asymptotic case, e.g., $K = 64$. The simulation results are shown in terms of the probability of detection versus SINR when the probability of false alarm is fixed as $P_f = 0.01$. In each case, the optimum but nonadaptive PIMF detector (5), whose P_d and P_f can be determined analytically [16], is included as a benchmark (note that the PIMF is a clairvoyant detector that has full knowledge of all unknown parameters). The ACE detector, whose performance can also be determined analytically (see [16], [28]), is also included to show how the parametric detectors compare with this popular covariance matrix based detector in partially nonhomogeneous environments. Since the ACE detector needs a full-rank space-time covariance matrix estimate, we assumes it has sufficiently IID training signals with $K = 2JN$. Finally, we include the simulated results for the proposed NPRao and SI-PRao detectors; for the latter, we also include its asymptotic performance as shown in Section IVC.

A. Limited-Training Case

In this case, we first consider two scenarios: 1) $K = 2$ and $N = 32$, where the training signals are limited and the number of temporal samples is moderate; 2) $K = 4$ and $N = 16$, where we have slightly more training signals but the temporal samples are limited. For $K = 2$ and $N = 32$, Figs. 2 and 3 show the probability of detection versus SINR when $\lambda = 4$ and $\lambda = 8$, respectively. It is seen that, for a small value of $\lambda = 4$, the NPRao and SI-PRao tests with $K = 2$ generally outperform the ACE detector with $K = 256$. For a larger value of $\lambda = 8$ as shown in Fig. 3, the SI-PRao test behaves similarly to the PIMF, while the NPRao test degrades considerably, indicating that it is not invariant to the power scaling factor λ.

In the second case with $K = 4$ and $N = 16$, the simulation results are plotted in Figs. 4 and 5, respectively, for $\lambda = 4$ and $\lambda = 8$. It is seen that the SI-PRao test yields the best performance among the adaptive detectors. Both the SI-PRao test and ACE detector are seen to be invariant to the power scaling factor. However, the NPRao test shows much worse detection performance than the SI-PRao test and the ACE detectors for both $\lambda = 4$ and $\lambda = 8$.

We next examine the SINR loss of the proposed Rao detectors with respect to the PIMF in cases of different probabilities of false alarm. The results of the SI-PRao and NPRao detectors are shown in Fig. 6 and Fig. 7, respectively, when several probabilities of false alarm $P_f = \{0.1, 0.01, 0.001\}$ are considered. It is noticed that, for a given probability of detection,

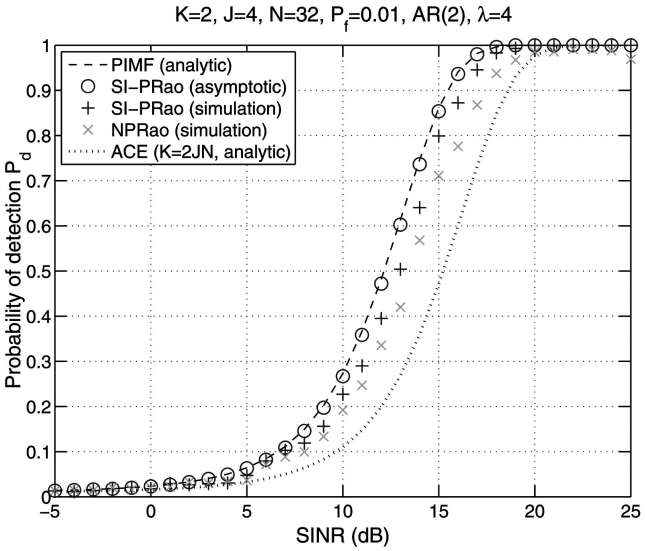


Fig. 2. Probability of detection versus SINR when $P_f = 0.01$, $J = 4$, $N = 32$, $K = 2$, $\lambda = 4$.

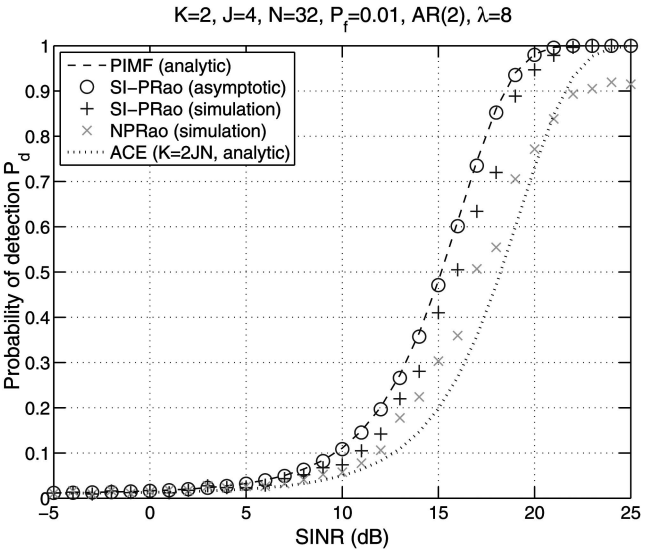


Fig. 3. Probability of detection versus SINR when $P_f = 0.01$, $J = 4$, $N = 32$, $K = 2$, $\lambda = 8$.

both Rao detectors show a larger SINR loss with respect to the PIMF when the probability of false alarm is smaller. However, the performance loss of the SI-PRao test is generally smaller than the NPRao test in all three cases.

B. Asymptotic Case

An asymptotic scenario with large K is simulated to verify the asymptotic performance of the SI-PRao test derived in Section IVC. The simulation parameters are $J = 4$, $N = 16$, $P = 2$, and $K = JN = 64$. The results are shown in Figs. 8 and 9 with $\lambda = 4$ and $\lambda = 8$, respectively. It is seen that the probability of detection obtained by simulation approaches the asymptotic performance of the SI-PRao test in both cases. Also, with plenty of training signals, the NPRao

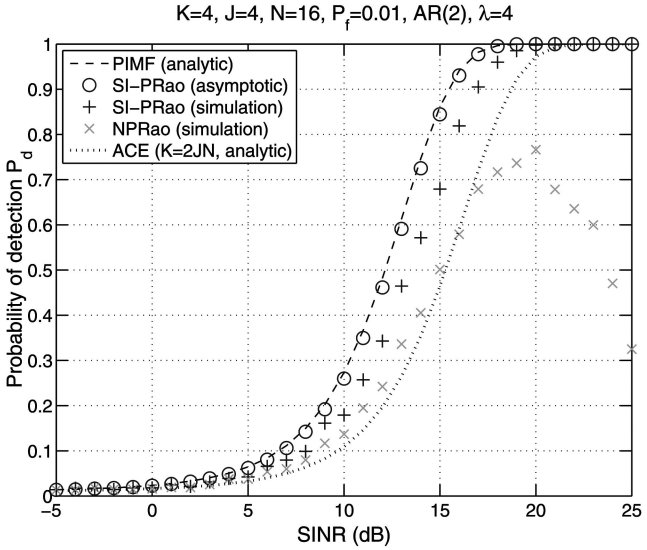


Fig. 4. Probability of detection versus SINR when $P_f = 0.01$, $J = 4$, $N = 16$, $K = 4$, $\lambda = 4$.

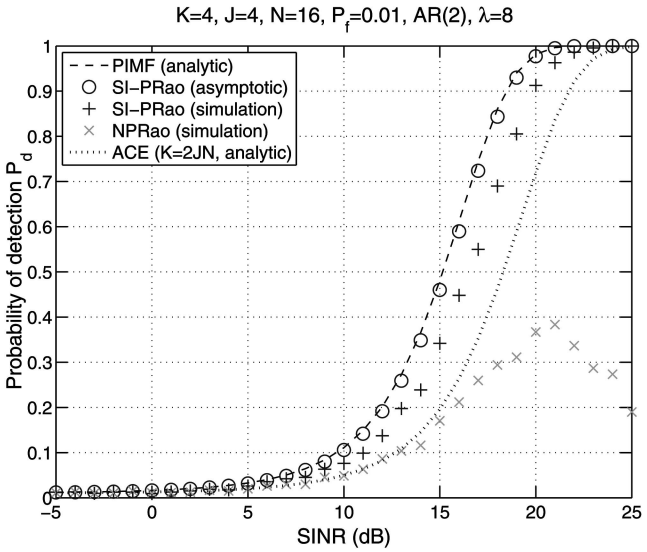


Fig. 5. Probability of detection versus SINR when $P_f = 0.01$, $J = 4$, $N = 16$, $K = 4$, $\lambda = 8$.

test provides detection performance almost identical to that of the SI-PRao test. The performance of the ACE detector with $K = 128$, which is twice that of the Rao tests, is also included for comparison.

It is also interesting to show the impact of the sample size, i.e., the number of pulses N, on the detection performance. The results are shown in Fig. 10 where the detection probability of the SI-PRao detector converges as N increases and approaches the asymptotic performance. The NPRao detector has a similar behavior, which is not shown for space limitation.

C. Effect of Model Mismatch

The above simulation examples are based on two assumptions: 1) the model order is known, and 2) the disturbance is exactly a multichannel AR process. In

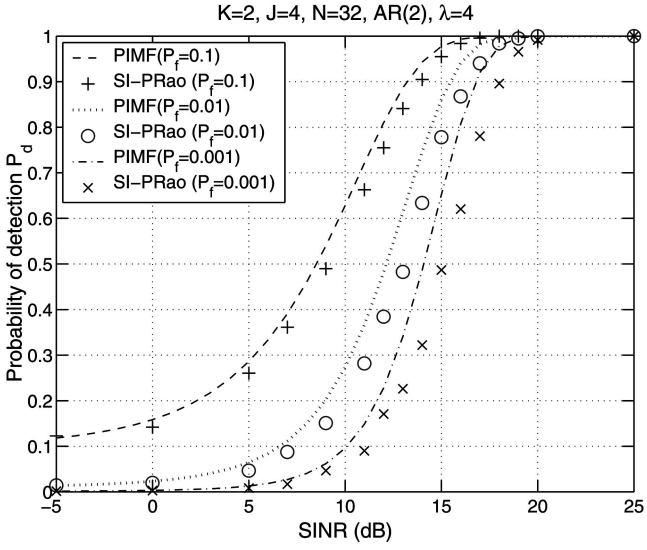


Fig. 6. Comparison of SINR loss of SI-PRao detector with respect to PIMF in cases of $P_f = \{0.1, 0.01, 0.001\}$ when $J = 4$, $N = 32$, $K = 2$, $P = 2$, $\lambda = 4$.

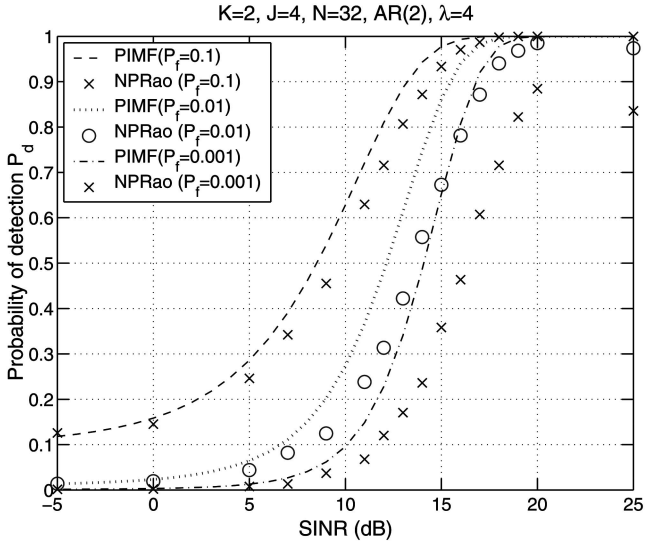


Fig. 7. Comparison of SINR loss of NPRao detector with respect to PIMF in cases of $P_f = \{0.1, 0.01, 0.001\}$when $J = 4$, $N = 32$, $K = 2$, $P = 2$, $\lambda = 4$.

this section, we evaluate the detection performance of the proposed Rao detectors when these assumptions are not met.

We first consider the case when the disturbance is a AR process, but there is an model estimation error. In practice, one needs to estimate the model order P of the multichannel AR before application of the SI-PRao and NPRao detectors. A model order estimation procedure may yield a small estimation error. Fig. 11 depicts the detection performance of the SI-PRao detector when the model order is underestimated and overestimated, respectively. It is seen that the order mismatch causes some performance degradation. However, in both cases, the degradation is not significant. It is also seen that overestimation of the model order has a smaller effect

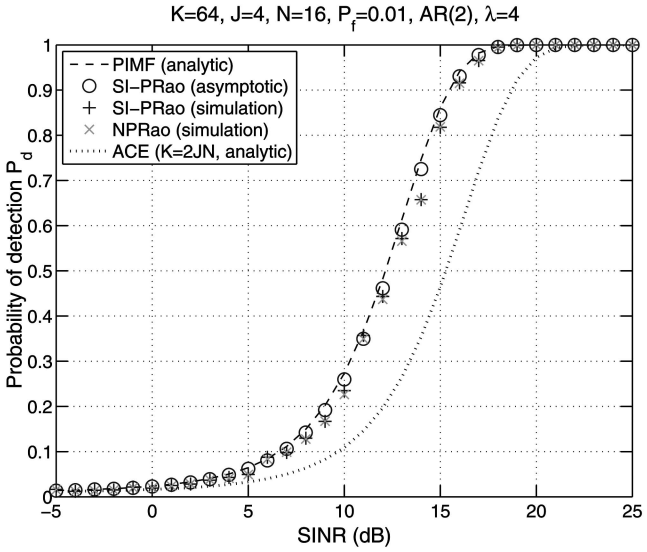


Fig. 8. Probability of detection versus SINR when $P_f = 0.01$, $J = 4$, $N = 16$, $K = 64$, $\lambda = 4$.

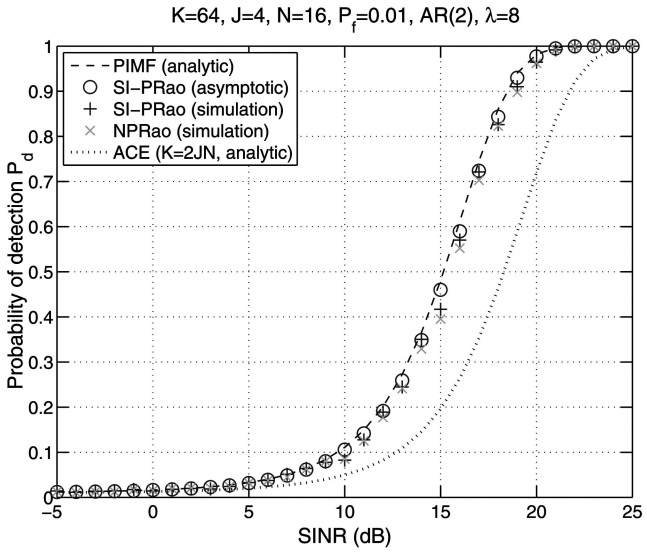


Fig. 9. Probability of detection versus SINR when $P_f = 0.01$, $J = 4$, $N = 16$, $K = 64$, $\lambda = 8$.

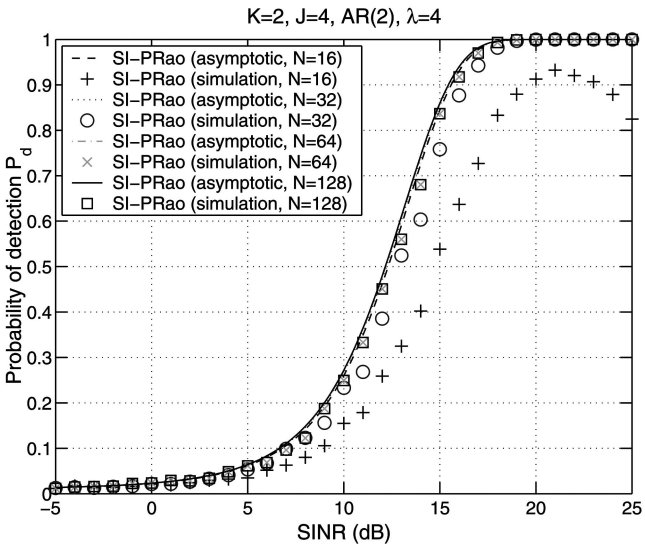


Fig. 10. Effect of number of pulses on detection performance of SI-PRao detector when $P_f = 0.01$, $J = 4$, $K = 2$, $P = 2$, $\lambda = 4$.

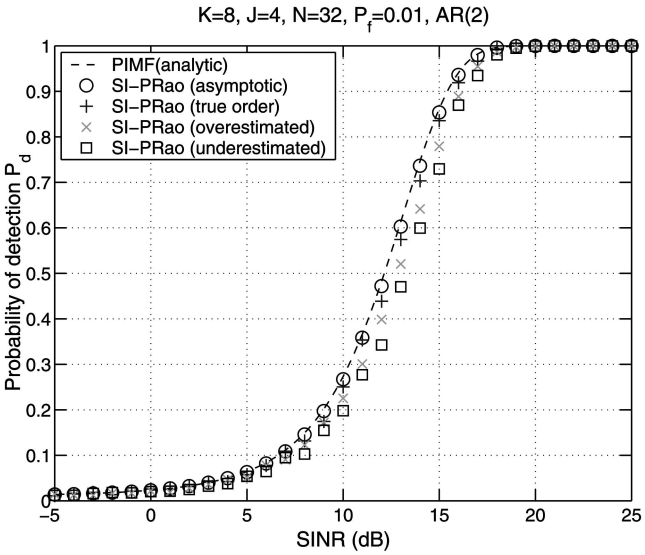


Fig. 11. Probability of detection versus SINR when model order of multichannel AR process used for computing test statistic is true ($P = 2$), overestimated (assuming $P = 3$), and underestimated (assuming $P = 1$), when $P_f = 0.01$, $J = 4$, $N = 32$, $K = 8$, $\lambda = 4$.

on the detection performance than underestimation. This behavior is similar to the case of the standard parametric Rao test in [12].

Next we consider the case when the disturbance is not a multichannel AR process. To show the detection performance in a more realistic environment, we use the KASSPER dataset, which contains many challenging real-world effects, including heterogeneous terrain, array errors, and dense ground targets (see [29] for a detailed description of the KASSPER dataset).

Fig. 12 shows the probability of detection versus SINR for the KASSPER dataset when $P_f = 0.01$, $J = 11$, $N = 32$, $K = 8$, $P = 1$, and $\lambda = 4$. The covariance matrix of the test signal corresponds to range cell $r200$, whereas the covariance matrices of two ($K = 2$) training signals are from range cells $r197$ and $r203$, i.e., with two guard cells between the test cell and training cells. This scenario includes some nonhomogeneous effects since the covariance matrices for the test and training cells are different. As shown in Fig. 12, with two training signals and by modeling the disturbance as an AR(1) process, the SI-PRao and NPRao detectors achieve a close performance that is only 2 to 3 dB away from the PIMF detector.

VI. CONCLUSION

The multichannel adaptive detection in the partially homogeneous environment by modeling the disturbance as a multichannel AR process has

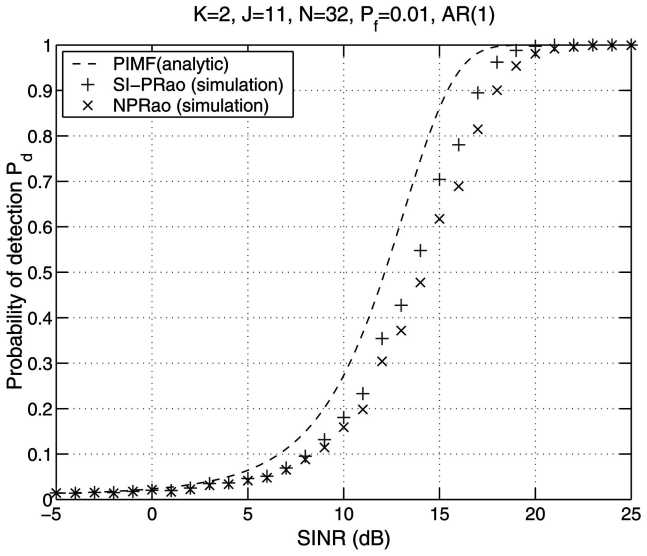


Fig. 12. Probability of detection versus SINR for KASSPER dataset when $P_{\mathrm{f}} = 0.01$, $J = 11$, $N = 32$, $K = 8$, $P = 1$, $\lambda = 4$.

been addressed. Two parametric Rao tests have been developed to relieve the excessive training requirement and reduce the computational complexity of the ACE detector, when the spatial-temporal dimension is large. The test statistic of the SI-PRao test has been asymptotically verified to be independent of the unknown parameters under the null hypothesis, which asymptotically achieves the CFAR. Numerical results verify that the SI-PRao test shows better performance than the NPRao test and the ACE when the training signals are limited.

APPENDIX

In order to facilitate the derivation of the SI-PRao test, we define the following notations.

1) $\boldsymbol{\theta}_r = [\alpha_R, \alpha_I]^T = [\Re\{\alpha\}, \Im\{\alpha\}]^T$ denotes the signal parameter vector, where $\Re$ and $\Im$ denote the real and imaginary parts, respectively. The joint probability density function (pdf) under H_0 and the pdf under H_1 differ only in the value of $\boldsymbol{\theta}_r$, where $\boldsymbol{\theta}_{r_0} = [0,0]^T$ and $\boldsymbol{\theta}_{r_1} = [\alpha_R, \alpha_I]^T$.

2) $\boldsymbol{\theta}_s = [\lambda, \mathbf{q}_R^T, \mathbf{q}_I^T, \mathbf{a}_R^T, \mathbf{a}_I^T]^T$ denotes the nuisance parameter vector with $\mathbf{a}_R^T = \mathrm{vec}(\Re\{\mathbf{A}^H\}), \mathbf{a}_I^T = \mathrm{vec}(\Im\{\mathbf{A}^H\})$, $\mathbf{q}_R^T$ contains the diagonal elements in $\mathbf{Q}$ and the real part of the elements below the diagonal, while $\mathbf{q}_I^T$ contains the imaginary part of the elements below the diagonal.

3) $\boldsymbol{\theta} = [\boldsymbol{\theta}_r^T, \boldsymbol{\theta}_s^T]^T$ contains all unknown parameters.

4) $\tilde{\boldsymbol{\theta}} = [\boldsymbol{\theta}_{r_0}^T, \hat{\boldsymbol{\theta}}_{s_0}^T]^T$ denotes the ML estimate of $\boldsymbol{\theta}$ under H_0.

The Rao test is a general solution to a class of parameter testing problems. It is often simpler than the GLRT, and is also asymptotically equivalent to the latter. A detailed discussion on the attributes of a generic Rao test can be found in [27]. The general Rao test can be expressed as [27]

$$\left.\frac{\partial \ln f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_r}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}^{T} [\mathbf{I}^{-1}(\tilde{\boldsymbol{\theta}})]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} \left.\frac{\partial \ln f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_r}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \tag{42}$$

where

$$[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} = (\mathbf{I}_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}(\boldsymbol{\theta}) - \mathbf{I}_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_s}(\boldsymbol{\theta})\mathbf{I}^{-1}_{\boldsymbol{\theta}_s,\boldsymbol{\theta}_s}(\boldsymbol{\theta})\mathbf{I}_{\boldsymbol{\theta}_s,\boldsymbol{\theta}_r}(\boldsymbol{\theta}))^{-1} \tag{43}$$

which is related to the Fisher information matrix (FIM) [27]

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{I}_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}(\boldsymbol{\theta}) & \mathbf{I}_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_s}(\boldsymbol{\theta}) \\ \mathbf{I}_{\boldsymbol{\theta}_s,\boldsymbol{\theta}_r}(\boldsymbol{\theta}) & \mathbf{I}_{\boldsymbol{\theta}_s,\boldsymbol{\theta}_s}(\boldsymbol{\theta}). \end{bmatrix}. \tag{44}$$

Under both hypothesis, the joint pdf of the test signal and training signal can be written as

$$f(\boldsymbol{\theta}) = \left[\frac{\lambda^{-J/(K+1)}}{\pi^J |\mathbf{Q}|} \exp\{-\mathrm{tr}(\mathbf{Q}^{-1}\mathbf{T}(\lambda, \mathbf{A}))\}\right]^{(K+1)(N-P)} \tag{45}$$

where

$$\mathbf{T}(\lambda, \mathbf{A}) = \frac{\frac{1}{\lambda}\sum_{n=P}^{N-1} \boldsymbol{\varepsilon}_0(n)\boldsymbol{\varepsilon}_0^H(n) + \sum_{k=1}^{K}\sum_{n=P}^{N-1} \boldsymbol{\varepsilon}_k(n)\boldsymbol{\varepsilon}_k^H(n)}{(K+1)(N-P)} \tag{46}$$

$$\boldsymbol{\varepsilon}_0(n) = \tilde{\mathbf{x}}_0(n) - \alpha\tilde{\mathbf{s}}(n) \tag{47}$$

$$\boldsymbol{\varepsilon}_k(n) = \mathbf{x}_k(n) + \sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{x}_k(n-p) \tag{48}$$

with

$$\tilde{\mathbf{x}}_0(n) = \mathbf{x}_0(n) + \sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{x}_0(n-p) \tag{49}$$

$$\tilde{\mathbf{s}}(n) = \mathbf{s}(n) + \sum_{p=1}^{P} \mathbf{A}^H(p)\mathbf{s}(n-p). \tag{50}$$

Equation (45) gives the pdf under H_0 by setting $\alpha = 0$ and the pdf under H_1 for $\alpha \neq 0$.

From (42), the derivation of the SI-PRao test is a two-step process. The first one is to obtain the ML estimates of the nuisance parameters under H_0, and the second one is to evaluate the related terms in (42).

A. ML Estimation Under H_0

From (45), the joint pdf under H_0 is $f(\boldsymbol{\theta})$ with $\alpha = 0$. By taking the derivative of the log likelihood $\ln f(\boldsymbol{\theta})$ with $\alpha = 0$ with respect to (w.r.t.) $\mathbf{Q}$ and equating it to zero results in the ML estimate of $\mathbf{Q}$ as

$$\hat{\mathbf{Q}}_{\mathrm{ML}}(\lambda, \mathbf{A}) = \mathbf{T}(\lambda, \mathbf{A}) \tag{51}$$

where $\mathbf{T}(\lambda,\mathbf{A})$ is given in (46) with $\alpha = 0$. Substituting $\hat{\mathbf{Q}}_{\mathrm{ML}}(\lambda,\mathbf{A})$ into $\ln f(\boldsymbol{\theta})$ under H_0 yields

$$f(\lambda,\mathbf{A},\hat{\mathbf{Q}}_{\mathrm{ML}}) = \left[\frac{(e\pi)^{-J}}{\lambda^{J/(K+1)}|\mathbf{T}(\lambda,\mathbf{A})|}\right]^{(K+1)(N-P)}. \tag{52}$$

The ML estimate of $\mathbf{A}$ is obtained by minimizing $|\mathbf{T}(\lambda,\mathbf{A})|$. Note that $(K+1)(N-P)\mathbf{T}(\lambda,\mathbf{A})$ can be rewritten as

$$\begin{aligned}&(K+1)(N-P)\mathbf{T}(\lambda,\mathbf{A})\\ &\quad = \hat{\mathbf{R}}_{xx}(\lambda) + \hat{\mathbf{R}}_{yx}^{H}(\lambda)\mathbf{A} + \mathbf{A}^{H}\hat{\mathbf{R}}_{yx}(\lambda) + \mathbf{A}^{H}\hat{\mathbf{R}}_{yy}(\lambda)\mathbf{A}.\end{aligned}$$

It can be shown that [12]

$$\mathbf{T}(\lambda,\mathbf{A}) \geq \mathbf{T}(\lambda,\mathbf{A})\big|_{\mathbf{A}=\hat{\mathbf{A}}} \tag{53}$$

where

$$\hat{\mathbf{A}}^{H}(\lambda) = -\hat{\mathbf{R}}_{yx}^{H}(\lambda)\hat{\mathbf{R}}_{yy}^{-1}(\lambda) \tag{54}$$

and $\hat{\mathbf{R}}_{xx}(\lambda)$, $\hat{\mathbf{R}}_{yy}(\lambda)$ and $\hat{\mathbf{R}}_{yx}(\lambda)$ are given by (31), (32), and (33), respectively.

Substituting $\hat{\mathbf{A}}^{H}(\lambda)$ into the log likelihood $\ln f(\boldsymbol{\theta})$ and ignoring the terms independent of λ yields

$$-\ln f(\lambda,\hat{\mathbf{A}},\hat{\mathbf{Q}}_{\mathrm{ML}}) \propto \frac{J}{K+1}\ln\lambda + \ln|\mathbf{T}(\lambda,\hat{\mathbf{A}})| \tag{55}$$

where

$$|\mathbf{T}(\lambda,\hat{\mathbf{A}})| \propto |\hat{\mathbf{R}}_{xx}(\lambda) - \hat{\mathbf{R}}_{yx}^{H}(\lambda)\hat{\mathbf{R}}_{yy}^{-1}(\lambda)\hat{\mathbf{R}}_{yx}(\lambda)| \tag{56}$$

and the symbol $\propto$ means "proportional to."

To derive the ML estimate of λ, we need to evaluate $|\mathbf{T}(\lambda,\hat{\mathbf{A}})|$. Note that the term at the right hand side of (56) is the Schur complement of [30]

$$\hat{\mathbf{R}}(\lambda) = \begin{bmatrix}\hat{\mathbf{R}}_{yy}(\lambda) & \hat{\mathbf{R}}_{yx}(\lambda)\\ \hat{\mathbf{R}}_{yx}^{H}(\lambda) & \hat{\mathbf{R}}_{xx}(\lambda)\end{bmatrix}. \tag{57}$$

Since

$$\frac{|\hat{\mathbf{R}}(\lambda)|}{|\hat{\mathbf{R}}_{yy}(\lambda)|} = |\hat{\mathbf{R}}_{xx}(\lambda) - \hat{\mathbf{R}}_{yx}^{H}(\lambda)\hat{\mathbf{R}}_{yy}^{-1}(\lambda)\hat{\mathbf{R}}_{yx}(\lambda)| \tag{58}$$

we have

$$-\ln f(\lambda,\hat{\mathbf{A}},\hat{\mathbf{Q}}_{\mathrm{ML}}) \propto \frac{J\ln\lambda}{K+1} + \ln|\hat{\mathbf{R}}(\lambda)| - \ln|\hat{\mathbf{R}}_{yy}(\lambda)|.$$

From (57), (31), (32), and (33), it is shown that $\hat{\mathbf{R}}(\lambda)$ can be separated into λ-dependent and λ-independent parts as

$$\hat{\mathbf{R}}(\lambda) = \lambda^{-1}\hat{\mathbf{R}}_0 + \hat{\mathbf{R}}_K \tag{59}$$

where $\hat{\mathbf{R}}_0$ and $\hat{\mathbf{R}}_K$ are given by (22) and (23), respectively. Note that $\hat{\mathbf{R}}_0$ and $\hat{\mathbf{R}}_K$ are positive definite with probability 1 if $N - P \geq J(P+1)$.

To differentiate the determinants of $\hat{\mathbf{R}}(\lambda)$ and $\hat{\mathbf{R}}_{yy}(\lambda)$ w.r.t. λ, we have to resort to the following lemma.

LEMMA 1 *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be* $M \times M$ *positive definite matrices. Then*

$$\frac{\partial}{\partial\lambda}\ln|\lambda^{-1}\mathbf{E} + \mathbf{F}| = \sum_{i=1}^{M}\frac{-1}{\lambda(1+\lambda\xi_i)} \tag{60}$$

where $\{\xi_i\}_{i=1}^{M}$ *are the eigenvalues of* $\mathbf{E}^{-1/2}\mathbf{F}\mathbf{E}^{-1/2}$.

PROOF

$$\begin{aligned}\frac{\partial}{\partial\lambda}\ln|\lambda^{-1}\mathbf{E} + \mathbf{F}| &= \frac{\partial}{\partial\lambda}\ln|\mathbf{E}^{1/2}(\lambda^{-1}\mathbf{I} + \mathbf{E}^{-1/2}\mathbf{F}\mathbf{E}^{-1/2})\mathbf{E}^{1/2}|\\ &= \frac{\partial}{\partial\lambda}\ln|\lambda^{-1}\mathbf{I} + \mathbf{E}^{-1/2}\mathbf{F}\mathbf{E}^{-1/2}|\\ &= \frac{\partial}{\partial\lambda}\ln\left[\prod_{i=1}^{M}(\lambda^{-1}+\xi_i)\right]\\ &= \sum_{i=1}^{M}\frac{\partial}{\partial\lambda}\ln\left(\frac{1+\lambda\xi_i}{\lambda}\right)\\ &= \sum_{i=1}^{M}\frac{-1}{\lambda(1+\lambda\xi_i)}\end{aligned} \tag{61}$$

where we have used the fact that $\{\lambda^{-1}+\xi_i\}_{i=1}^{M}$ are the eigenvectors of the matrix $\lambda^{-1}\mathbf{I} + \mathbf{E}^{-1/2}\mathbf{F}\mathbf{E}^{-1/2}$.

By applying Lemma 1, we get

$$\frac{\partial}{\partial\lambda}\ln|\hat{\mathbf{R}}(\lambda)| = \sum_{i=1}^{J(P+1)}\frac{-1}{\lambda(1+\lambda\mu_i)} \tag{62}$$

$$\frac{\partial}{\partial\lambda}\ln|\hat{\mathbf{R}}_{yy}(\lambda)| = \sum_{i=1}^{JP}\frac{-1}{\lambda(1+\lambda\upsilon_i)} \tag{63}$$

where μ_i and υ_i are eigenvalues of the matrices $\hat{\mathbf{R}}_0^{-1/2}\hat{\mathbf{R}}_K\hat{\mathbf{R}}_0^{-1/2}$ and $\hat{\mathbf{R}}_{0,y}^{-1/2}\hat{\mathbf{R}}_{K,y}\hat{\mathbf{R}}_{0,y}^{-1/2}$, respectively, and $\hat{\mathbf{R}}_{0,y}$ and $\hat{\mathbf{R}}_{K,y}$ are given by (24) and (25). Following that, the ML estimate of λ, denoted as $\hat{\lambda}$, is shown to be the root of (21). As a result, the ML estimates of $\mathbf{Q}$ and $\mathbf{A}$ are given by (51) and (54), which are obtained by replacing λ with $\hat{\lambda}$.

B. Derivation of the SI-PRao Test

From (45), the elements of the first partial derivative of the log likelihood $\ln f$ with respect to $\boldsymbol{\theta}_r$ in (42) are

$$\frac{\partial\ln f(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_r} = \begin{bmatrix}\dfrac{\partial\ln f(\boldsymbol{\theta})}{\partial\alpha_R}\\ \dfrac{\partial\ln f(\boldsymbol{\theta})}{\partial\alpha_I}\end{bmatrix} \tag{64}$$

with

$$\frac{\partial\ln f(\boldsymbol{\theta})}{\partial\alpha_R} = \frac{1}{\lambda}\sum_{n=P}^{N-1}[\tilde{\mathbf{s}}^{H}(n)\mathbf{Q}^{-1}\boldsymbol{\varepsilon}_0(n) + \boldsymbol{\varepsilon}_0^{H}(n)\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n)]$$

$$\frac{\partial\ln f(\boldsymbol{\theta})}{\partial\alpha_I} = \frac{j}{\lambda}\sum_{n=P}^{N-1}[\tilde{\mathbf{s}}^{H}(n)\mathbf{Q}^{-1}\boldsymbol{\varepsilon}_0(n) - \boldsymbol{\varepsilon}_0^{H}(n)\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n)].$$

Since the term $[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{\theta_r,\theta_r}$ is related to the FIM, we first find relevant entries of the FIM:

$$E\left\{\frac{\partial^2 \ln f(\boldsymbol{\theta})}{\partial \alpha_R^2}\right\} = -\frac{2}{\lambda}\sum_{n=P}^{N-1}\tilde{\mathbf{s}}^H(n)\mathbf{Q}^{-1}\tilde{\mathbf{s}}^H(n)$$

$$E\left\{\frac{\partial^2 \ln f(\boldsymbol{\theta})}{\partial \alpha_I^2}\right\} = -\frac{2}{\lambda}\sum_{n=P}^{N-1}\tilde{\mathbf{s}}^H(n)\mathbf{Q}^{-1}\tilde{\mathbf{s}}^H(n)$$

$$E\left\{\frac{\partial^2 \ln f(\boldsymbol{\theta})}{\partial \alpha_R \partial \alpha_I}\right\} = 0$$

$$E\left\{\frac{\partial^2 \ln f(\boldsymbol{\theta})}{\partial \alpha_I \partial \alpha_R}\right\} = 0$$

$$\mathbf{I}_{\theta_r,\theta_s}(\boldsymbol{\theta}) = \mathbf{0}$$

$$\mathbf{I}_{\theta_s,\theta_r}(\boldsymbol{\theta}) = \mathbf{0}.$$

As a result, (43) reduces to

$$[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{\theta_r,\theta_r} = \frac{\lambda}{2\sum_{n=P}^{N-1}\tilde{\mathbf{s}}^H(n)\mathbf{Q}^{-1}\tilde{\mathbf{s}}(n)}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}. \tag{65}$$

Since

$$\left.\frac{\partial \ln f(\boldsymbol{\theta})}{\partial \alpha_R}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} = \frac{1}{\tilde{\lambda}}\sum_{n=P}^{N-1}[\tilde{\mathbf{s}}^H(n)\hat{\mathbf{Q}}^{-1}\tilde{\mathbf{x}}_0(n) + \tilde{\mathbf{x}}_0^H(n)\hat{\mathbf{Q}}^{-1}\tilde{\mathbf{s}}(n)] \tag{66}$$

$$\left.\frac{\partial \ln f(\boldsymbol{\theta})}{\partial \alpha_I}\right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} = \frac{j}{\tilde{\lambda}}\sum_{n=P}^{N-1}[\tilde{\mathbf{s}}^H(n)\hat{\mathbf{Q}}^{-1}\tilde{\mathbf{x}}_0(n) - \tilde{\mathbf{x}}_0^H(n)\hat{\mathbf{Q}}^{-1}\tilde{\mathbf{s}}(n)] \tag{67}$$

the general Rao test in (42) reduces to (20).

REFERENCES

[1] Ward, J.
Space-time adaptive processing for airborne radar.
Lincoln Laboratory, MIT, Technical Report 1015, Dec. 1994.

[2] Klemm, R.
Principles of Space-Time Adaptive Processing.
London: IEE, 2002.

[3] Paulraj, A. and Papadias, C. B.
Space-time processing for wireless communications.
IEEE Signal Processing Magazine, **14** (Nov. 1997), 49–83.

[4] Shaw, G. and Manolakis, D.
Signal processing for hyperspectral image exploitation.
IEEE Signal Processing Magazine, **19**, 1 (Jan. 2002), 12–16.

[5] Manolakis, D. and Shaw, G.
Detection algorithms for hyperspectral imaging applications.
IEEE Signal Processing Magazine, **19**, 1 (Jan. 2002), 29–43.

[6] Reed, I. S., Mallett, J. D., and Brennan, L. E.
Rapid convergence rate in adaptive arrays.
IEEE Transactions on Aerospace and Electronic Systems, **AES-10**, 6 (1974), 853–863.

[7] Kelly, E. J.
An adaptive detection algorithm.
IEEE Transactions on Aerospace and Electronic Systems, **AES-22** (Mar. 1986), 115–127.

[8] Chen, W. and Reed, I. S.
A new CFAR detection test for radar.
Digital Signal Processing, **1**, 4 (1991), 198–214.

[9] Robey, F. C., et al.
A CFAR adaptive matched filter detector.
IEEE Transactions on Aerospace and Electronic Systems, **28**, 1 (Jan. 1992), 208–216.

[10] De Maio, A.
Rao test for adaptive detection in Gaussian interference with unknown covariance matrix.
IEEE Transactions on Signal Processing, **55**, 7 (July 2007), 3577–3584.

[11] Román, J. R., et al.
Parametric adaptive matched filter for airborne radar applications.
IEEE Transactions on Aerospace and Electronic Systems, **36**, 2 (Apr. 2000), 677–692.

[12] Sohn, K. J., Li, H., and Himed, B.
Parametric Rao test for multichannel adaptive signal detection.
IEEE Transactions on Aerospace and Electronic Systems, **43**, 3 (July 2007), 920–933.

[13] Sohn, K. J., Li, H., and Himed, B.
Parametric GLRT for multichannel adaptive signal detection.
IEEE Transactions on Signal Processing, **55**, 11 (Nov. 2007),5351–5360.

[14] Scharf, L. L. and McWhorter, T. L.
Adaptive matched subspace detectors and adaptive coherence estimators.
In *Proceedings of the 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 1997.

[15] Kraut, S. and Scharf, L. L.
The CFAR adaptive subspace detector is a scale-invariant GLRT.
IEEE Transactions on Signal Processing, **47**, 9 (Sept. 1999), 2538–2541.

[16] Kraut, S., Scharf, L. L., and Butler, R. W.
The adaptive coherence estimator:
A uniformly-most-powerful-invariant adaptive detection statistic.
IEEE Transactions on Signal Processing, **53**, 2 (Feb. 2005), 417–438.

[17] De Maio, A. and Iommelli, S.
Coincidence of the Rao test, Wald test, and GLRT in partially homogeneous environment.
IEEE Signal Processing Letters, **15** (2008), 385–388.

[18] Conte, E., Lops, M., and Ricci, G.
Asymptotically optimum radar detection in compound Gaussian clutter.
IEEE Transactions on Aerospace and Electronic Systems, **31**, 2 (Apr. 1995), 617–625.

[19] Rangaswamy, M. and Michels, J. H.
A parametric multichannel detection algorithm for correlated non-Gaussian random processes.
In *Proceedings of the 1997 IEEE National Radar Conference*, Syracuse, NY, May 1997, 349–354.

[20] Michels, J. H., Himed, B., and Rangaswamy, M.
Performance of STAP tests in Gaussian and compound-Gaussian clutter.
Digital Signal Processing, **10**, 4 (Oct. 2000), 309–324.

[21] Michels, J. H., Rangaswamy, M., and Himed, B.
Performance of parametric and covariance based STAP tests in compound-Gaussian clutter.
Digital Signal Processing, **12**, 2, 3 (Apr.–July 2002), 307–328.

[22] Rangaswamy, M.
Statistical analysis of the nonhomogeneity detector for non-Gaussian interference backgrounds.
IEEE Transactions on Signal Processing, **53**, 6 (June 2005), 2101–2111.

[23] Haykin, S.
Adaptive Radar Signal Processing.
New York: Wiley-Interscience, 2005.

[24] McWhorter, L. T. and Scharf, L. L.
Adaptive matched subspace detectors and adaptive coherence estimators.
In *Proceedings of the 30th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1996, 1114–1117.

[25] Kraut, S., McWhorter, T. L., and Scharf, L. L.
A canonical representation for the distributions of adaptive matched subspace detectors.
In *Proceedings of the 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 1997.

[26] Michels, J. H.
Space-time adaptive processing (STAP) in Gaussian and non-Gaussian airborne radar clutter.
In *Proceedings of the Adaptive Radar Clutter Suppression Workshop*, Malvern, UK, July 1999.

[27] Kay, S. M.
Fundamentals of Statistical Signal Processing: Detection Theory.
Upper Saddle River, NJ: Prentice-Hall, 1998.

[28] Richmond, C. D.
Performance of the adaptive sidelobe blanker detection algorithm in homogeneous environments.
IEEE Transactions on Signal Processing, **48**, 5 (May 2000), 1235–1247.

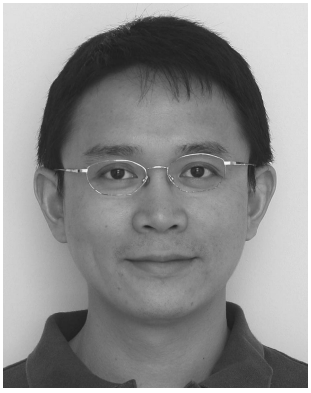
[29] Bergin, J. S. and Techau, P. M.
High-fidelity site-specific radar simulation: KASSPER'02 workshop datacube.
Information Systems Laboratories, Inc., Vienna, VA, Technical Report ISL-SCRD-TR-02-105, May 2002.

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

Pu Wang (S'05) received the B. Eng. and M. Eng. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2003 and 2006, respectively, and the Ph.D. degree in 2011 from the Stevens Institute of Technology, Hoboken, NJ, all in electrical engineering.

He was a research assistant from 2007 to March 2011 and a teaching assistant during 2008 at the Department of Electrical and Computer Engineering of the Stevens Institute of Technology. He was an intern at the Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, in the summer of 2010. Since April 2011, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, where he is a research assistant professor. His current research interests include statistical signal processing for multi-channel/multi-antenna systems, wireless communications, and networks.

Mr. Wang received the Francis T. Boesch Award in 2008 and the Outstanding Research Assistant Award in 2007 both from Stevens Institute of Technology, the Excellent Master Thesis Award of Sichuan Province in 2007 from Sichuan Ministry of Education, China, and the Excellent Master Thesis Award of UESTC in 2006 from UESTC. He is a recipient of the Student Travel Grant twice from the IEEE Radar Conference.



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 in the Department of Electrical and Computer Engineering at the University of Florida. He was a summer visiting faculty member at the Air Force Research Laboratory in the summers of 2003, 2004, and 2009. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, where he is a professor. His current research interests include statistical signal processing, wireless communications, 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 has been an editor or associate editor for the *IEEE Transactions on Wireless Communications*, *IEEE Signal Processing Letters*, and *IEEE Transactions on Signal Processing*, and served 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) was born in Algiers, Algeria. He received his B.S. degree in electrical engineering from Ecole Nationale Polytechnique of Algiers in 1984, and his M.S. and Ph.D. degrees both in electrical engineering, from Syracuse University, Syracuse, NY, in 1987 and 1990, respectively.

He is currently a principal electronics engineer with the Air Force Research Laboratory, Sensors Directorate, Radar Signal Processing Branch, Dayton, OH. His research interests include detection, estimation, multichannel adaptive signal processing, time series analyses, array processing, space-time adaptive processing, hot clutter mitigation, airborne and spaceborne radar, over the horizon radar, and below ground sensing.

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