Adaptive Detection of a Signal Known Only to Lie on a Line in a Known Subspace, When Primary and Secondary Data are Partially Homogeneous

Olivier Besson, *Senior Member, IEEE*, Louis L. Scharf, *Fellow, IEEE*, and Shawn Kraut, *Member, IEEE*

***Abstract*—This paper deals with the problem of detecting a signal, known only to lie on a line in a subspace, in the presence of unknown noise, using multiple snapshots in the primary data. To account for uncertainties about a signal's signature, we assume that the steering vector belongs to a known linear subspace. Furthermore, we consider the partially homogeneous case, for which the covariance matrix of the primary and the secondary data have the same structure but possibly different levels. This provides an extension to the framework considered by Bose and Steinhardt. The natural invariances of the detection problem are studied, which leads to the derivation of the maximal invariant. Then, a detector is proposed that proceeds in two steps. First, assuming that the noise covariance matrix is known, the generalized-likelihood ratio test (GLRT) is formulated. Then, the noise covariance matrix is replaced by its sample estimate based on the secondary data to yield the final detector. The latter is compared with a similar detector that assumes the steering vector to be known.**

***Index Terms*—Array processing, detection, maximal invariant statistic, steering vector uncertainties.**

I. Problem Statement

We consider the problem of detecting a partly unknown rank-one signal using multiple observations from an array of sensors, in the presence of correlated noise with unknown level and covariance matrix. More precisely, the detection problem consists of deciding between the two hypotheses

$$H_0: \begin{cases} \boldsymbol{x}(t) = \boldsymbol{n}_p(t); & t = 1, \ldots, N_p \\ \boldsymbol{y}(t) = \boldsymbol{n}_s(t); & t = 1, \ldots, N_s \end{cases}$$
$$H_1: \begin{cases} \boldsymbol{x}(t) = \boldsymbol{a}s^*(t) + \boldsymbol{n}_p(t); & t = 1, \ldots, N_p \\ \boldsymbol{y}(t) = \boldsymbol{n}_s(t); & t = 1, \ldots, N_s \end{cases} \quad (1)$$

where we have the following.

Manuscript received June 21, 2005; revised December 19, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yuri I. Abramovich. The work of L. L. Scharf was supported by ONR under contract N00014-04-1-0084 and by DARPA under contract FA9550-04-1-0371.

O. Besson is with the Department of Avionics and Systems, ENSICA, 31056 Toulouse, France (e-mail: besson@ensica.fr).

L. L. Scharf is with the Departments of Electrical and Computer Engineering and Statistics, Colorado State University, Fort Collins, CO 80523-1373 USA (e-mail: scharf@engr.colostate.edu).

S. Kraut was with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada. He is now with the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02420 USA (e-mail: kraut@ll.mit.edu).

Digital Object Identifier 10.1109/TSP.2006.881262

- $\boldsymbol{X} = [\boldsymbol{x}(1) \ \cdots \ \boldsymbol{x}(N_p)]$ is the $m \times N_p$ primary data array, and $\boldsymbol{Y} = [\boldsymbol{y}(1) \ \cdots \ \boldsymbol{y}(N_s)]$ is the $m \times N_s$ secondary data array. The data array can be the output of an array of sensors (space-only framework); in such a case, t denotes the time index, and N_p and N_s stand for the number of snapshots collected. The problem formulated in (1) is also relevant in space–time applications, in which case $\boldsymbol{x}(t)$ is a space–time snapshot (whose length is the number of sensors times the number of pulses) and t is the range cell index. In such a situation, the problem amounts to detecting a target which potentially spreads over N_p range cells.
- $\boldsymbol{a}$ is either the spatial or the space–time signature of interest, referred to as the *steering vector*. We consider the case where there exist uncertainties about $\boldsymbol{a}$. This can occur, for instance, if the detection of the target is carried out on a grid of spatial (and possibly Doppler) frequencies while the actual frequencies lie in between the grid. The uncertainties about the actual signature can also be due to an uncalibrated array. Finally, in the case of a Ricean channel, the steering vector is the sum of a (possibly known) line-of-sight component and a random component due to the scattering environment. In order to account for these uncertainties, we assume that $\boldsymbol{a}$ belongs to a *known* p-dimensional subspace spanned by the columns of the $m \times p$ matrix $\boldsymbol{H}$. We refer the reader to [1], [2], and references therein for the rationale of such a model. It should be pointed out that the detection problem considered herein with $\boldsymbol{a} = \boldsymbol{H}\boldsymbol{\theta}$ is rather general and encompasses two special cases of interest, namely the case where $\boldsymbol{a}$ is perfectly known, and the case where $\boldsymbol{a}$ is unknown and arbitrary. The former case corresponds to $p = 1$, $\theta = 1$, and $\boldsymbol{H} = \boldsymbol{a}$ and was considered in [3]. An arbitrary steering vector corresponds to $p = m$, $\boldsymbol{H} = \boldsymbol{I}$ and $\boldsymbol{\theta} = \boldsymbol{a}$.
- $s(t)$ corresponds to the unknown emitted signal waveform—or the amplitude of a target in different range cells in space–time applications—and is considered an unknown deterministic (nonrandom) sequence in this paper.
- $\boldsymbol{n}_p(t)$ and $\boldsymbol{n}_s(t)$ stand for the noise in the primary and secondary data, respectively. They are proper zero-mean independent and Gaussian distributed with $\mathcal{E}\left\{\boldsymbol{n}_s(t)\boldsymbol{n}_s^H(t)\right\} = \boldsymbol{C}$ and $\mathcal{E}\left\{\boldsymbol{n}_p(t)\boldsymbol{n}_p^H(t)\right\} = \alpha\boldsymbol{C}$. The scaling factor α accounts for noise power mismatch between the primary and the secondary data. In other words, the primary and secondary data are only partially homogeneous. In the sequel, we use the terminology "partially homogeneous" to refer

to the case in which the scaling factor α is something other than one.

For the homogeneous case, detecting a signal of interest in a background of noise (with unknown covariance matrix) has been studied extensively in the literature. When $\boldsymbol{a}$ is known, the generalized-likelihood ratio test (GLRT) was derived by Kelly [4], when $N_p = 1$. Extension to a vector signal (i.e., the signal of interest belongs to a known linear subspace) can be found in [5] and [6]. The adaptive matched filter (AMF) [7] proceeds in two steps. First, the GLRT for known noise covariance matrix is derived, and next, the noise covariance matrix is replaced by its estimate obtained from secondary data. Kelly's GLRT and the AMF were shown to be the maximal invariants for this detection problem in [8] and [9].

In the partially homogeneous case, the adaptive coherence estimator (ACE) was derived in [10]–[12] and shown to be a scale-invariant GLRT and the uniformly most powerful invariant test [13]. Extensions of the ACE to a subspace signal were also considered. We refer to [12] and references therein for an overview. Most of the studies cited so far deal with the $N_p = 1$ case, and a known steering vector.

Detection from multiple observations and with partly known signals of interest is rather scarce and was considered, e.g., in [1], [3], [14], and [15]. However, the problem addressed in this paper differs from those in these references, as explained below. In the recent paper [15], a constant false alarm rate (CFAR) detector based on a two-step GLRT is derived under the assumption that the signals of interest belong to a known subspace. However, in [15], the signals of interest evolve in a p-dimensional subspace while here they are constrained to lie in an unknown one-dimensional subspace of a known subspace $\langle \boldsymbol{H} \rangle$. In other words, the signal of interest is $\boldsymbol{H}\boldsymbol{s}(t)$ in [15], with $\boldsymbol{s}(t) \in \mathbb{C}^p$, whereas herein the signal of interest is $\boldsymbol{H}\boldsymbol{\theta}s(t)$. Reference [3] considers the detection of a range-spread target using an high resolution radar. The detection problem is formally equivalent to that in (1), except that the steering vector is assumed to be known in [3]. The authors derive and analyze a two-step GLRT in the homogeneous case (the noise covariance matrix is the same in the primary and secondary data) and in the partially homogeneous case (the noise covariance matrix has the same structure in the primary and secondary data but may differ by a scaling factor, as in the present paper). In contrast to [3] where the steering vector is known, the statistical average of the primary data matrix $\mathcal{E}\{\boldsymbol{X}|H_1\}$ is considered as unknown and arbitrary in [14]. In other words, the signal of interest matrix does not have any structure. The theory of invariance is invoked to obtain a most powerful invariant test and a suboptimal CFAR detector. The detection problem described in (1) is similar to that addressed in [1], where uncertain rank-one waveforms are to be detected. More precisely, in [1], both the space $\boldsymbol{a} = \boldsymbol{A}\boldsymbol{\alpha}$ and time $\boldsymbol{s} = \boldsymbol{B}\boldsymbol{\beta}$ signatures of the signal of interest are assumed to belong to the c-dimensional and r-dimensional linear subspaces spanned by the columns of known matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. Therefore, it is similar to the problem treated here, with $\boldsymbol{A} = \boldsymbol{H}$, $c = p$, $\boldsymbol{B} = \boldsymbol{I}$, and $r = N_p$. Bose and Steinhardt provide insightful derivations and interpretations of the maximal invariant, and derive the GLRT for this very general framework. However, [1] considers the homogeneous case only, meaning that $\alpha = 1$. Hence, the present paper provides an extension of [1] to the partially homogeneous case.

The paper is organized as follows. In Section II, we study the natural invariances of the detection problem along with the maximal invariant statistic. Our detector is derived in Section III. The performance of the detector is evaluated in Section IV and compared with a that of a detector that assumes a perfectly known steering vector. Our conclusions are drawn in Section V. They suggest that the generalized adaptive direction detector (GADD) of this paper is more robust to model mismatch than the generalized adaptive subspace detector (GASD) of [3], and only slightly less powerful when there is no model mismatch.

II. Invariances

Let us consider the natural invariances of the detection problem in (1). Since we wish to preserve zero mean and Gaussianity of the measurements under H_0, we restrict our attention to linear transformations of the data. In order for the transformed noise to be temporally independent between snapshots, these transformations are of the form

$$[\boldsymbol{X} \quad \boldsymbol{Y}] \rightarrow [\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}^H \quad \boldsymbol{A}'\boldsymbol{Y}\boldsymbol{D}^H] \tag{2}$$

with $\boldsymbol{B}$ and $\boldsymbol{D}$ unitary matrices. Furthermore, since the noise covariance matrix must be the same, up to a scaling factor, under each hypothesis, it follows that $\boldsymbol{A}'$ and $\boldsymbol{A}$ must be proportional to each other, and full rank. Finally, since the steering vector must remain in $\langle \boldsymbol{H} \rangle$, $\boldsymbol{A}$ must satisfy

$$\langle \boldsymbol{A}\boldsymbol{H} \rangle = \langle \boldsymbol{H} \rangle . \tag{3}$$

Therefore, the hypothesis testing problem (1) is invariant under the group of transformations G defined by

$$G = \{g : [\boldsymbol{X} \quad \boldsymbol{Y}] \rightarrow [\beta\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}^H \quad \gamma\boldsymbol{A}\boldsymbol{Y}\boldsymbol{D}^H]\} \tag{4}$$

where β and γ are arbitrary scalars, $\boldsymbol{A}$ is a full-rank matrix such that $\boldsymbol{A}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{T}$ with $\boldsymbol{T}$ a $p \times p$ invertible matrix, and $\boldsymbol{B}$ and $\boldsymbol{D}$ are unitary matrices. The group of transformations induced on the parameter space is

$$\bar{g} : \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{s} \\ \alpha \\ \boldsymbol{C} \end{pmatrix} \rightarrow \begin{pmatrix} \beta_1\boldsymbol{T}\boldsymbol{\theta} \\ \beta_2\boldsymbol{B}\boldsymbol{s} \\ \alpha\,|\beta|^2\,|\gamma|^{-2} \\ |\gamma|^2\,\boldsymbol{A}\boldsymbol{C}\boldsymbol{A}^H \end{pmatrix} \tag{5}$$

with $\beta_1\beta_2^* = \beta$. A few remarks are in order. When the steering vector is known, then $\boldsymbol{A}$ is a full-rank matrix such that $\boldsymbol{A}\boldsymbol{a} \propto \boldsymbol{a}$. On the contrary, when $\boldsymbol{a}$ is arbitrary and unknown, $\boldsymbol{A}$ can be any nonsingular matrix.

Now that the group of transformations under which the detection problem is invariant has been defined, we look for test statistics that are invariant to G. All of them will be a function of the *maximal invariant* (see [16] for a theoretical presentation and [8] and [9] for comprehensive overviews and array processing applications of invariance and maximal invariants), which is derived next. In the following, we let $\boldsymbol{H}_\perp$ denote an $m \times m - p$ matrix whose columns form a basis for $\langle \boldsymbol{H} \rangle^\perp$. We also define

$$\boldsymbol{S} = \boldsymbol{Y}\boldsymbol{Y}^H \tag{6}$$

as the sample covariance matrix of the secondary data. Finally, we will denote by $\boldsymbol{\lambda}\{.\}$ the vector whose elements are the eigenvalues of the matrix between parentheses.

Proposition 1: Let

$$\boldsymbol{M} = \boldsymbol{X}^H \boldsymbol{H}_\perp \left(\boldsymbol{H}_\perp^H \boldsymbol{S} \boldsymbol{H}_\perp\right)^{-1} \boldsymbol{H}_\perp^H \boldsymbol{X} \tag{7a}$$

$$= \boldsymbol{U}_M \boldsymbol{\Lambda}_M \boldsymbol{U}_M^H \tag{7b}$$

$$\boldsymbol{E} = \boldsymbol{X}^H \boldsymbol{S}^{-1} \boldsymbol{H} \left(\boldsymbol{H}^H \boldsymbol{S}^{-1} \boldsymbol{H}\right)^{-1} \boldsymbol{H}^H \boldsymbol{S}^{-1} \boldsymbol{X}. \tag{7c}$$

The maximal invariant statistic with respect to the group of transformations (4) is given by the eigenvectors of

$$\boldsymbol{M}_1 = \boldsymbol{U}_M^H \boldsymbol{E} \boldsymbol{U}_M \tag{8}$$

and the scaled eigenvalues

$$\frac{\boldsymbol{\lambda}\{\boldsymbol{M}\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}};\ \frac{\boldsymbol{\lambda}\{\boldsymbol{M}_1\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}}. \tag{9}$$

Proof: The proof is given in Appendix I and builds upon the results of [1] where the maximal invariant for the homogeneous case was derived.

The maximal invariant is a vector-valued function of the data. It consists of the eigenvectors of $\boldsymbol{M}_1$ and the normalized eigenvalues of $\boldsymbol{M}$ and $\boldsymbol{M}_1$. This normalization is mandatory to ensure a constant false alarm rate in the partially homogeneous case. The following observations, concerning the case of a known $\boldsymbol{a}$ and the case of an unknown $\boldsymbol{a}$, can be made.

Remark 1: When $\boldsymbol{a}$ is known, then $p = 1$, $\theta = 1$ and $\boldsymbol{H} = \boldsymbol{a}$ is now the steering vector. In this case, $\boldsymbol{M}_1$, or equivalently $\boldsymbol{E}$, has only one single eigenvalue, and $\boldsymbol{\lambda}\{\boldsymbol{M}_1\}/\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}$ reduces to

$$\frac{\boldsymbol{a}^H \boldsymbol{S}^{-1} \boldsymbol{X}\boldsymbol{X}^H \boldsymbol{S}^{-1} \boldsymbol{a}}{\left(\boldsymbol{a}^H \boldsymbol{S}^{-1} \boldsymbol{a}\right) \mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}}.$$

The maximal invariant then consists of the ratio above, along with the corresponding eigenvector of $\boldsymbol{M}_1$, as well as the N_p eigenvalues of $\boldsymbol{M}$. In the *single snapshot* case, $N_p = 1$, $\boldsymbol{X} = \boldsymbol{x}$ is a vector and $\boldsymbol{M} = m$, $\boldsymbol{M}_1 = m_1$ are scalars. Observing that $\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\} = \mathrm{Tr}\{\boldsymbol{M}\} + \mathrm{Tr}\{\boldsymbol{M}_1\}$, it follows that the maximal invariant reduces to

$$\frac{m_1}{\boldsymbol{x}^H \boldsymbol{S}^{-1} \boldsymbol{x}} = \frac{\left|\boldsymbol{a}^H \boldsymbol{S}^{-1} \boldsymbol{x}\right|^2}{\left(\boldsymbol{a}^H \boldsymbol{S}^{-1} \boldsymbol{a}\right)\left(\boldsymbol{x}^H \boldsymbol{S}^{-1} \boldsymbol{x}\right)}$$

which is exactly the ACE statistic of [10], [11], and [13]. Hence, we recover the fact that the ACE test statistic is a maximal invariant in the partially homogeneous case.

Remark 2: When $\boldsymbol{a}$ is unknown and arbitrary, $p = m$, $\boldsymbol{H} = \boldsymbol{I}$ and $\boldsymbol{\theta} = \boldsymbol{a}$. The maximal invariant then consists of

$$\frac{\boldsymbol{\lambda}\{\boldsymbol{M}_1\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} = \frac{\boldsymbol{\lambda}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}}. \tag{10}$$

III. Detection

In this section, we derive a CFAR detector using a GLRT approach. First note that, in principle, a one-step GLRT can be derived, using the whole array data, including the primary and secondary data. This approach was taken, e.g., by Kelly in [4]. It was also used in [11] to prove that the ACE is a scale invariant GLRT. Unfortunately, in our situation where $N_p > 1$, it is shown in Appendix II (see also [3, Sect. II.A]) that the maximum-likelihood estimates (MLEs) of the unknown parameters cannot be obtained in a simple and closed-form expression, which results in a computationally complex detector. Therefore, this approach is abandoned and we turn to an approach similar to those advocated, e.g., in [3], [14], and [15]. First, we assume that the noise covariance matrix $\boldsymbol{C}$ is known, and we derive the GLRT using the primary data only. Next, we substitute for the unknown covariance matrix $\boldsymbol{C}$ its estimate based on secondary data.

Let us consider first that $\boldsymbol{C}$ is known. Then, the probability density function (PDF) of $\boldsymbol{X}$ is given by

$$f(\boldsymbol{X}|H_k) = \frac{e^{-\mathrm{Tr}\left\{\alpha^{-1}\left(\boldsymbol{X}-\mu_k \boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)^H \boldsymbol{C}^{-1}\left(\boldsymbol{X}-\mu_k \boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)\right\}}}{\pi^{mN_p}\alpha^{mN_p}\left|\boldsymbol{C}\right|^{N_p}} \tag{11}$$

with $\mu_k = 0$ under H_0, $\mu_k = 1$ under H_1. The log-likelihood function is thus given by

$$\Lambda(\boldsymbol{X}|H_k) = \text{const.} - mN_p \ln\alpha - \alpha^{-1}\mathrm{Tr}\left\{\boldsymbol{C}^{-1}\left(\boldsymbol{X}-\mu_k \boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)\left(\boldsymbol{X}-\mu_k \boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)^H\right\}. \tag{12}$$

Differentiating the previous equation with respect to the scaling parameter α and setting the result to zero yields the following estimate of α:

$$\widehat{\alpha}_k = \frac{1}{mN_p}\mathrm{Tr}\left\{\boldsymbol{C}^{-1}\left(\boldsymbol{X}-\mu_k \boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)\left(\boldsymbol{X}-\mu_k \boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)^H\right\}. \tag{13}$$

Under H_0, all unknown parameters are estimated and the estimate of α under H_0 is simply

$$\widehat{\alpha}_0 = \frac{1}{mN_p}\mathrm{Tr}\left\{\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}. \tag{14}$$

Under H_1, we still need to minimize

$$J = \mathrm{Tr}\left\{\boldsymbol{C}^{-1}\left(\boldsymbol{X}-\boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)\left(\boldsymbol{X}-\boldsymbol{H}\boldsymbol{\theta}\boldsymbol{s}^H\right)^H\right\}$$

with respect to $\boldsymbol{\theta}$ and $\boldsymbol{s}$. However

$$\begin{aligned} J &= \mathrm{Tr}\left\{\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\} - \boldsymbol{s}^H\boldsymbol{X}^H\boldsymbol{C}^{-1}\boldsymbol{H}\boldsymbol{\theta} \\ &\quad - \boldsymbol{\theta}^H\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{s} + \left(\boldsymbol{s}^H\boldsymbol{s}\right)\left(\boldsymbol{\theta}^H\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\boldsymbol{\theta}\right) \\ &= \left(\boldsymbol{\theta}^H\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\boldsymbol{\theta}\right)\left\|\boldsymbol{s} - \frac{\boldsymbol{X}^H\boldsymbol{C}^{-1}\boldsymbol{H}\boldsymbol{\theta}}{\boldsymbol{\theta}^H\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\boldsymbol{\theta}}\right\|^2 \\ &\quad + \mathrm{Tr}\left\{\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\} - \frac{\boldsymbol{\theta}^H\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{C}^{-1}\boldsymbol{H}\boldsymbol{\theta}}{\boldsymbol{\theta}^H\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\boldsymbol{\theta}}. \end{aligned} \tag{15}$$

Hence, the MLE of $\boldsymbol{s}$ is given as

$$\widehat{\boldsymbol{s}} = \left(\widehat{\boldsymbol{\theta}}^H\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\widehat{\boldsymbol{\theta}}\right)^{-1}\boldsymbol{X}^H\boldsymbol{C}^{-1}\boldsymbol{H}\widehat{\boldsymbol{\theta}} \tag{16}$$

where $\widehat{\boldsymbol{\theta}}$, the MLE of $\boldsymbol{\theta}$, is given, up to a scaling factor, by

$$\widehat{\boldsymbol{\theta}} = \mathcal{P}\left\{\left(\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{C}^{-1}\boldsymbol{H}\right\} \tag{17}$$

where $\mathcal{P}\{.\}$ stands for the principal eigenvector. Substituting these values in J yields the following estimate of α under H_1:

$$\widehat{\alpha}_1 = \frac{1}{mN_p}\left[\mathrm{Tr}\left\{\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\} - \lambda_{\max}\left\{\left(\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{C}^{-1}\boldsymbol{H}\right\}\right] \quad (18)$$

where $\lambda_{\max}\{.\}$ stands for the principal eigenvalue of the matrix between braces. Therefore, the generalized-likelihood ratio (GLR) for known $\boldsymbol{C}$ is given by

$$\mathrm{GLR} \equiv \frac{\widehat{\alpha}_0}{\widehat{\alpha}_1} \equiv \frac{\lambda_{\max}\left\{\left(\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{C}^{-1}\boldsymbol{H}\right\}}{\mathrm{Tr}\left\{\boldsymbol{C}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \quad (19)$$

where the symbol $\equiv$ means "equivalent to." Our detector consists of replacing $\boldsymbol{C}$ in (19) by its MLE, namely $N_s^{-1}\boldsymbol{S}$, which leads to

$$\frac{\lambda_{\max}\left\{\left(\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{S}^{-1}\boldsymbol{H}\right\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta \quad (20)$$

or, equivalently

$$\frac{\lambda_{\max}\{\boldsymbol{E}\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta. \quad (21)$$

Note that this detector depends on the maximal invariant and is thus CFAR with respect to α and $\boldsymbol{C}$ [16], as its distribution only depends on the signal to noise ratio (see Appendix I). In fact, the detector in (20) only uses the maximal eigenvalue of $\boldsymbol{E}$, which is equal to the square of the largest singular value of $\boldsymbol{P}_{\boldsymbol{S}^{-1/2}\boldsymbol{H}}\boldsymbol{S}^{-1/2}\boldsymbol{X}$. Therefore, it first performs whitening of the data $\boldsymbol{X}$, then looks for a *preferred direction* in the subspace $\left\langle \boldsymbol{S}^{-1/2}\boldsymbol{H}\right\rangle$, namely the direction of maximum energy. This seems logical since, if the signal of interest is present, it should lie along a line in this subspace. The detector in (20) can thus be viewed as an adaptive version of the matched direction detector derived in [2].

Remark 3: In the homogeneous case, Bose and Steinhardt [1] derived analytical expressions for the PDF of the GLRT (see also [17] for a similar analysis). The partially homogeneous case induces a serious complication, namely the normalizing factor at the denominator of (20). This renders the theoretical analysis very difficult, not to say intractable, and we were not able to obtain theoretical expressions for the PDF of (20). Therefore, in the next section, the performance of the detector will be assessed by numerical simulations.

Remark 4: As mentioned above, the detector in (20) is invariant to arbitrary scaling of the measurements $\boldsymbol{X}$ and $\boldsymbol{Y}$. This property is important when the noise is no longer a Gaussian process but a spherically invariant random process. In such a case, under the null hypothesis, the primary and secondary data matrices can be written as [18] and [19] $\boldsymbol{X} = \sqrt{\tau_x}\boldsymbol{X}_G$ and $\boldsymbol{Y} = \sqrt{\tau_y}\boldsymbol{Y}_G$, where $\boldsymbol{X}_G$ and $\boldsymbol{Y}_G$ are Gaussian, and τ_x and τ_y are the so-called textures which might be different between the primary and the secondary data, but constant over the N_p and N_s snapshots or range cells. Usually, the textures are considered as random variables with a given PDF. For fixed τ_x and τ_y, the left-hand side of (20)—let us call it g for the sake of convenience—does not depend on τ_x and τ_y. Therefore, under H_0, its PDF given τ_x and τ_y, $f(g|\tau_x,\tau_y)$, does not depend on τ_x and τ_y. Consequently, under H_0, $f(g)$ does not depend on the PDFs of the textures, which is an interesting feature of the detector.

Remark 5: When $\boldsymbol{a}$ is known, the detector in (20) reduces to

$$\frac{\boldsymbol{a}^H\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{S}^{-1}\boldsymbol{a}}{\left(\boldsymbol{a}^H\boldsymbol{S}^{-1}\boldsymbol{a}\right)\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (22)$$

which corresponds to the detector derived in [3, eq. (26)]. Furthermore,when there is only one snapshot in the primary data, $\boldsymbol{X} = \boldsymbol{x}$ and (20) above reduces to

$$\frac{\left|\boldsymbol{a}^H\boldsymbol{S}^{-1}\boldsymbol{x}\right|^2}{\left(\boldsymbol{a}^H\boldsymbol{S}^{-1}\boldsymbol{a}\right)\left(\boldsymbol{x}^H\boldsymbol{S}^{-1}\boldsymbol{x}\right)} \quad (23)$$

which is exactly the ACE.

Remark 6: When $\boldsymbol{a}$ is unknown and arbitrary, $\boldsymbol{H}$ is the identity matrix, $\boldsymbol{E} = \boldsymbol{X}^H\boldsymbol{S}^{-1}\boldsymbol{X}$ and our detector becomes

$$\frac{\lambda_{\max}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (24)$$

Hence, when an arbitrary rank-one component is to be detected, the detector amounts to comparing the largest eigenvalue of $\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H$ to its trace. The detector first performs a whitening operation of the data, and then tests if the largest eigenvalue is above the others (in which case it decides a rank-one component is present) or if all eigenvalues are of the same order (in which case the data is considered as white noise). This is an interesting variation on the statistic $\prod\lambda_k\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}/\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}^m$, which is the GLRT for sphericity using primary data only, when the measurement model is $\boldsymbol{X} : \mathcal{CN}\left(\boldsymbol{0},\sigma^2\boldsymbol{S}\right)$ under H_0 and $\boldsymbol{X} : \mathcal{CN}(\boldsymbol{0},\boldsymbol{R})$, $\boldsymbol{R} > 0$ arbitrary, under H_1 [20].

IV. Numerical Examples

In this section, we illustrate the performance and the robustness of the detector in (20), and compare it with the detector that assumes that the steering vector is known. For the sake of clarity, the two detectors are repeated below:

$$\frac{\overline{\boldsymbol{a}}^H\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{S}^{-1}\overline{\boldsymbol{a}}}{\left(\overline{\boldsymbol{a}}^H\boldsymbol{S}^{-1}\overline{\boldsymbol{a}}\right)\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (25)$$

$$\frac{\lambda_{\max}\left\{\left(\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{S}^{-1}\boldsymbol{H}\right\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta. \quad (26)$$

In (25), $\overline{\boldsymbol{a}}$ stands for the *presumed* steering vector, which may or may not differ from the *actual* steering vector $\boldsymbol{a}$. In other words, we consider the possibility of a mismatch between $\overline{\boldsymbol{a}}$ and $\boldsymbol{a}$, as could happen for instance with an uncalibrated array or with pointing errors. Following the terminology used in [3], we will refer to (25) as the GASD. The detector in (26) will be referred to as the GADD in the figures. Three different scenarios will be investigated, with various degrees of mismatch.

1) In the first scenario, the presumed and the actual steering vector coincide, i.e., $\overline{\boldsymbol{a}} = \boldsymbol{a}$ and, furthermore, $\boldsymbol{a} \in \langle\boldsymbol{H}\rangle$.

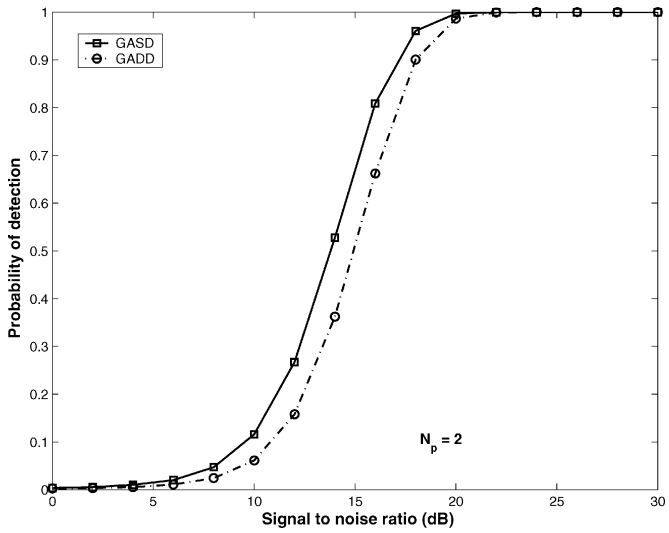


Fig. 1. P_d of GASD and GADD versus SNR. $m = 10, p = 2, N_s = 20$, $N_p = 2$, and $P_{\rm fa} = 10^{-3}$. $\boldsymbol{a} = \overline{\boldsymbol{a}} \in \langle H \rangle$.

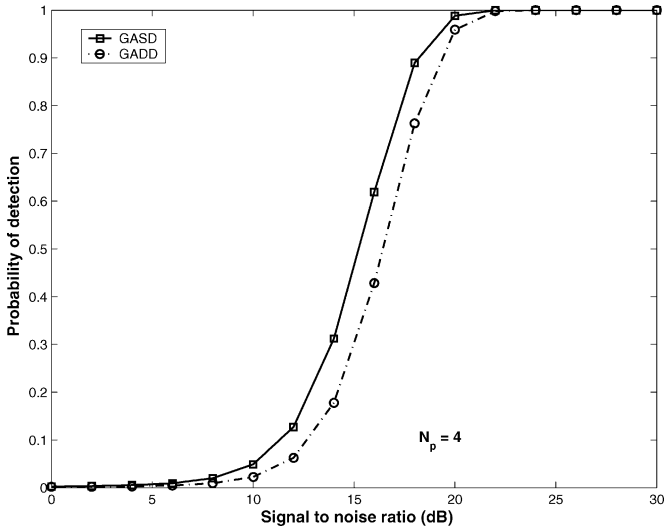


Fig. 2. P_d of GASD and GADD versus SNR. $m = 10, p = 2, N_s = 20$, $N_p = 4$, and $P_{\rm fa} = 10^{-3}$. $\boldsymbol{a} = \overline{\boldsymbol{a}} \in \langle H \rangle$.

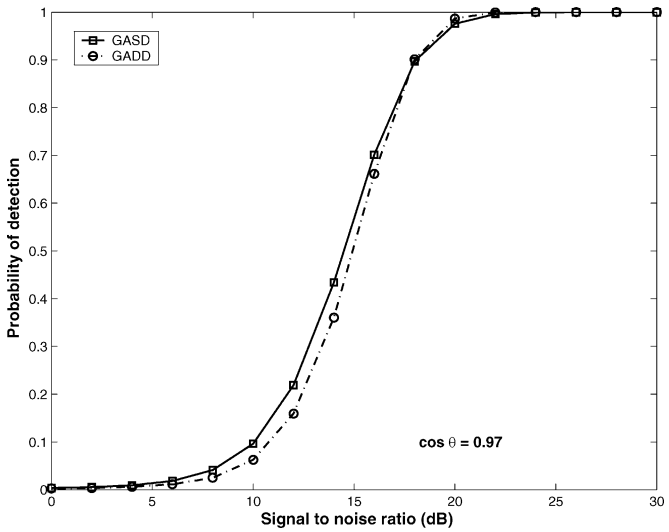


Fig. 3. P_d of GASD and GADD versus SNR. $m = 10, p = 2, N_s = 20$, $N_p = 2$, and $P_{\rm fa} = 10^{-3}$. $\boldsymbol{a} \neq \overline{\boldsymbol{a}}, \boldsymbol{a} \in \langle H \rangle$. $\cos\theta = 0.97$.

This corresponds to the case of a perfectly known steering vector and a matrix $\boldsymbol{H}$ that includes the actual steering vector. Under such conditions, we may expect the detector (25) to perform better than the detector (26).

2) In the second scenario, we consider a mismatch between $\overline{\boldsymbol{a}}$ and $\boldsymbol{a}$ while $\boldsymbol{a}$ still belongs to $\langle \boldsymbol{H} \rangle$. To parameterize the mismatch, we will use the angle between $\boldsymbol{C}^{-1/2}\overline{\boldsymbol{a}}$ and $\boldsymbol{C}^{-1/2}\boldsymbol{a}$, which will be referred to as θ. In this scenario, the detector (26) should not undergo any performance loss while the detector in (25) should. We will investigate the relative performances of the two detectors versus θ.
3) In the third scenario, we still consider a mismatch between $\overline{\boldsymbol{a}}$ and $\boldsymbol{a}$ and, additionally, $\boldsymbol{a}$ no longer belongs to $\langle \boldsymbol{H} \rangle$. This scenario tests the robustness of both detectors. θ will still denote the angle between $\boldsymbol{C}^{-1/2}\overline{\boldsymbol{a}}$ and $\boldsymbol{C}^{-1/2}\boldsymbol{a}$, while γ will stand for the angle between $\boldsymbol{C}^{-1/2}\boldsymbol{a}$ and $\boldsymbol{C}^{-1/2}\boldsymbol{H}$.

In all simulations below, we consider an array with $m = 10$ elements, the dimension of $\langle \boldsymbol{H} \rangle$ is $p = 2$ and the number of snapshots in the secondary data is $N_s = 20$. In order to set the thresholds η and ζ for a given probability of false alarm $P_{\rm fa}$, we resorted to Monte Carlo count techniques. More precisely, $N_{\rm mc} = 10^6$ simulations of the data under the null hypothesis were run, and the two test statistics in (25)–(26) were computed, then sorted in ascending order. The thresholds were computed from the $N_{\rm mc} \times P_{\rm fa}$ most significant values of the test statistics. In the simulations shown below, $P_{\rm fa} = 10^{-3}$. To obtain the probability of detection, 10^5 independent trials were run, and the test statistics are compared to the thresholds in order to obtain P_d. The probability of detection is plotted as a function of the signal-to-noise ratio (SNR), which is defined as

$$\text{SNR} = \left[\boldsymbol{s}^H \boldsymbol{s}\right] \left[\boldsymbol{a}^H \left(\alpha \boldsymbol{C}\right)^{-1} \boldsymbol{a}\right]. \tag{27}$$

Note that the SNR is the induced maximal invariant for the problem at hand (see Appendix I), which means that the distribution of the maximal invariant only depends on SNR [16].

Figs. 1–5 display the results for the three different scenarios mentioned above. From inspection of these figures, the following observations can be made.

- In the first scenario, the detector that assumes $\boldsymbol{a}$ to be known performs better than the detector that assumes that $\boldsymbol{a}$ belongs to the range space of $\boldsymbol{H}$ (see Figs. 1 and 2). This is logical as the actual steering vector perfectly matches the presumed one. Observe that the difference is about 1 dB at $P_d = 0.8$, and hence the performance loss incurred by the GADD is not very important.
- The robustness of the GADD can clearly be seen in Figs. 3 and 4, where P_d is plotted for two different values of θ. As $\cos\theta$ decreases, the performance of the GASD degrades, while that of the GADD remains constant. When $\cos\theta = 0.97$, the two detectors have similar performance, while the GADD performs about 3 dB better at $P_d = 0.8$ than the GASD, for $\cos\theta = 0.9$. Therefore, in a situation where

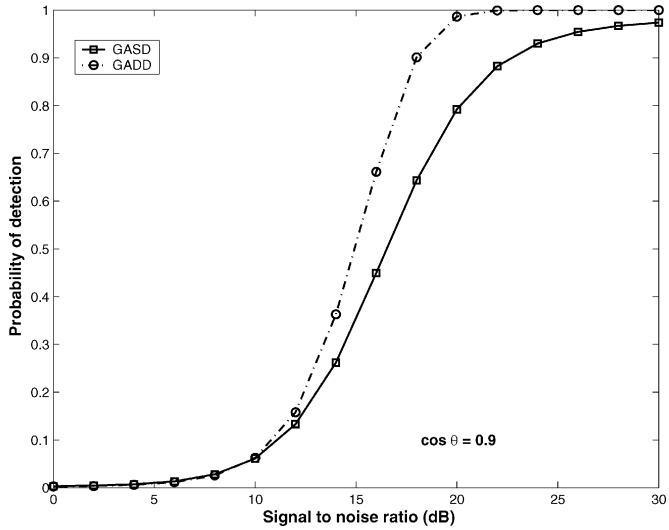


Fig. 4. P_d of GASD and GADD versus SNR. $m = 10, p = 2, N_s = 20$, $N_p = 2$, and $P_{\rm fa} = 10^{-3}$. $\boldsymbol{a} \neq \overline{\boldsymbol{a}}, \boldsymbol{a} \in \langle H \rangle$. $\cos\theta = 0.90$.

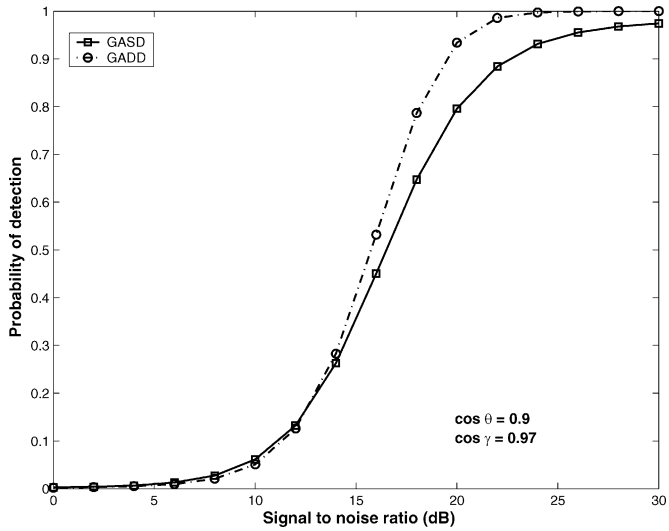


Fig. 5. P_d of GASD and GADD versus SNR. $m = 10, p = 2, N_s = 20$, $N_p = 2$, and $P_{\rm fa} = 10^{-3}$. $\boldsymbol{a} \neq \overline{\boldsymbol{a}}, \boldsymbol{a} \notin \langle H \rangle$. $\cos\theta = 0.9$ and $\cos\gamma = 0.97$.

there exists a mismatch between the presumed and the actual steering vector, the GADD can offer improved robustness.

- Fig. 5 considers a case where the two detectors operate in a mismatched situation. It can be seen that, for low SNR, the two detectors perform the same while, for high SNR, the detector (26) performs better. Again, the GADD offers improved robustness.

V. Conclusion

In this paper, we considered the problem of detecting a partly unknown signal which lies in a known subspace, in the presence of noise which might not be completely homogeneous between the primary and the secondary data. The natural invariances of the problem were studied and a new detector was derived. The detector amounts to searching the direction of maximum energy in a subspace, after whitening of the data. We showed that this detector performs only slightly worse than a detector which knows the steering vector of interest, but offers improved robustness when there is a mismatch between the actual steering vector and the presumed steering vector. Therefore, the choice between the two detectors is dictated by the confidence we have in the presumed steering vector; should there exist uncertainties about the latter, the GADD might be preferred. On the other hand, if a precise knowledge of $\boldsymbol{a}$ is at hand, the GASD remains the best solution.

Appendix I
Proof of Proposition 1

In order to prove Proposition 1, we first need to show that the maximal invariant is indeed invariant to the group of transformations defined in (4). Observing that $\boldsymbol{S}$ is transformed as $|\gamma|^2 \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H$, $\boldsymbol{A}$ is invertible, and $\boldsymbol{A}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{T}$, it follows that

$$\begin{aligned} \boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{H} &\rightarrow |\gamma|^{-2}\boldsymbol{H}^H\boldsymbol{A}^{-H}\boldsymbol{S}^{-1}\boldsymbol{A}^{-1}\boldsymbol{H} \\ &= |\gamma|^{-2}\boldsymbol{T}^{-H}\left[\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{H}\right]\boldsymbol{T}^{-1} \qquad (28) \\ \boldsymbol{X}^H\boldsymbol{S}^{-1}\boldsymbol{H} &\rightarrow \beta^* |\gamma|^{-2}\boldsymbol{B}\boldsymbol{X}^H\boldsymbol{A}^H\boldsymbol{A}^{-H}\boldsymbol{S}^{-1}\boldsymbol{A}^{-1}\boldsymbol{H} \\ &= \beta^* |\gamma|^{-2}\boldsymbol{B}\boldsymbol{X}^H\boldsymbol{S}^{-1}\boldsymbol{H}\boldsymbol{T}^{-1}. \qquad (29) \end{aligned}$$

Therefore, $\boldsymbol{E}$ is transformed as

$$\boldsymbol{E} \rightarrow c\boldsymbol{B}\boldsymbol{X}^H\boldsymbol{S}^{-1}\boldsymbol{H}\left(\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{B}^H = c\,\boldsymbol{B}\boldsymbol{E}\boldsymbol{B}^H \qquad (30)$$

with $c = |\beta|^2 |\gamma|^{-2}$. Accordingly, since $\langle \boldsymbol{H} \rangle$ is invariant for $\boldsymbol{A}$, it follows that $\langle \boldsymbol{H} \rangle^{\perp}$ is invariant for $\boldsymbol{A}^H$, i.e., $\langle \boldsymbol{A}^H\boldsymbol{H}_{\perp} \rangle = \langle \boldsymbol{H}_{\perp} \rangle$, and hence there exists $\boldsymbol{T}'$ such that $\boldsymbol{A}^H\boldsymbol{H}_{\perp} = \boldsymbol{H}_{\perp}\boldsymbol{T}'$. Using this property along with the technique employed above, it can readily be shown that

$$\boldsymbol{M} \rightarrow c\boldsymbol{B}\boldsymbol{M}\boldsymbol{B}^H \qquad (31)$$

from which we have that $\boldsymbol{U}_M \rightarrow \boldsymbol{B}\boldsymbol{U}_M$. Therefore

$$\boldsymbol{M}_1 \rightarrow c\boldsymbol{U}_M^H\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{M}\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{U}_M = c\boldsymbol{M}_1. \qquad (32)$$

The eigenvectors of $\boldsymbol{M}_1$ are left unchanged while its eigenvalues are multiplied by c, as those of $\boldsymbol{M}$. At the same time, $\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}$ is also multiplied by c, which shows that the maximal invariant of Proposition 1 is indeed invariant to G.

Next, we need to show that, if two sets of data $[\boldsymbol{X} \quad \boldsymbol{Y}]$ and $[\boldsymbol{X}' \quad \boldsymbol{Y}']$ have the same maximal invariant, then they are related to one another by a transformation of the type defined in (4). Towards this end, we will build upon the results of [1] which provides the maximal invariant in the homogeneous case. Briefly stated, we will use the same kind of technique as in [13], where the ACE was proven to be a maximal invariant in the partially homogeneous case based on the fact that the AMF and Kelly's GLRT were a maximal invariant in the homogeneous case. For the sake of clarity, we will use the notation $\boldsymbol{E}(\boldsymbol{X},\boldsymbol{Y})$ to emphasize that $\boldsymbol{E}$ depends on $\boldsymbol{X}$ and $\boldsymbol{Y}$. We will also use the same notations for $\boldsymbol{M}$ and $\boldsymbol{M}_1$. Let us assume that $\boldsymbol{M}_1(\boldsymbol{X},\boldsymbol{Y})$ and $\boldsymbol{M}_1(\boldsymbol{X}',\boldsymbol{Y}')$ have the same eigenvectors and that

$$\frac{\boldsymbol{\lambda}\{\boldsymbol{M}(\boldsymbol{X}',\boldsymbol{Y}')\}}{\mathrm{Tr}\left\{\boldsymbol{S}'^{-1}\boldsymbol{X}'\boldsymbol{X}'^H\right\}} = \frac{\boldsymbol{\lambda}\{\boldsymbol{M}(\boldsymbol{X},\boldsymbol{Y})\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^H\right\}} \qquad (33a)$$

$$\frac{\boldsymbol{\lambda}\left\{\boldsymbol{M}_1\left(\boldsymbol{X}',\boldsymbol{Y}'\right)\right\}}{\mathrm{Tr}\left\{\boldsymbol{S}'^{-1}\boldsymbol{X}'\boldsymbol{X}'^{H}\right\}}=\frac{\boldsymbol{\lambda}\left\{\boldsymbol{M}_1\left(\boldsymbol{X},\boldsymbol{Y}\right)\right\}}{\mathrm{Tr}\left\{\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{X}^{H}\right\}}. \tag{33b}$$

Under this hypothesis, one has

$$\boldsymbol{\lambda}\left\{\boldsymbol{M}\left(\boldsymbol{X}',\boldsymbol{Y}'\right)\right\}=c\boldsymbol{\lambda}\left\{\boldsymbol{M}\left(\boldsymbol{X},\boldsymbol{Y}\right)\right\}=\boldsymbol{\lambda}\left\{\boldsymbol{M}\left(\beta\boldsymbol{X},\gamma\boldsymbol{Y}\right)\right\} \tag{34}$$

where β and γ are arbitrary scalars such that $|\beta|^2\,|\gamma|^{-2}=c$. Similarly

$$\boldsymbol{\lambda}\left\{\boldsymbol{M}_1\left(\boldsymbol{X}',\boldsymbol{Y}'\right)\right\}=\boldsymbol{\lambda}\left\{\boldsymbol{M}_1\left(\beta\boldsymbol{X},\gamma\boldsymbol{Y}\right)\right\}. \tag{35}$$

Furthermore, the eigenvectors of $\boldsymbol{M}_1\left(\boldsymbol{X}',\boldsymbol{Y}'\right)$ and $\boldsymbol{M}_1\left(\beta\boldsymbol{X},\gamma\boldsymbol{Y}\right)$ are identical. However, as shown by Bose and Steinhardt, the eigenvectors of $\boldsymbol{M}_1$, as well as the eigenvalues of $\boldsymbol{M}$ and $\boldsymbol{M}_1$ are the maximal invariant in the homogeneous case. Therefore, there must exist a full-rank matrix $\boldsymbol{A}$ such that $\langle\boldsymbol{AH}\rangle=\langle\boldsymbol{H}\rangle$, and unitary matrices $\boldsymbol{B}$ and $\boldsymbol{D}$ such that

$$\boldsymbol{X}'=\beta\boldsymbol{AXB}^H;\ \boldsymbol{Y}'=\gamma\boldsymbol{AYD}^H \tag{36}$$

which concludes the proof. Finally, we note that the *induced maximal invariant* in the homogeneous case is the output array signal to interference and noise ratio, viz. $\left[\boldsymbol{s}^H\boldsymbol{s}\right]\left[\boldsymbol{a}^H\boldsymbol{C}^{-1}\boldsymbol{a}\right]$ (see [1]). Using the same arguments as those used to derive the maximal invariant, it can be shown that the induced maximal invariant in the partially homogeneous case is $\left[\boldsymbol{s}^H\boldsymbol{s}\right]\left[\boldsymbol{a}^H\left(\alpha\boldsymbol{C}\right)^{-1}\boldsymbol{a}\right]$.

Appendix II
A Note on the One-Step GLRT

In this Appendix, we briefly explain the reason why a two-step GLRT-based approach should be preferred to the one-step GLRT. A first reason stems from the fact that the latter does not lead to closed-form and simple expressions of the maximum likelihood estimates of the unknown parameters, as briefly explained below. Under hypothesis H_k the joint distribution of the primary and secondary data is given by

$$f(\boldsymbol{X},\boldsymbol{Y}|H_k)=\left\{\frac{\mathrm{etr}\left\{-\boldsymbol{C}^{-1}\boldsymbol{T}_k\right\}}{\pi^m\alpha^{mN_p/N_t}\,|\boldsymbol{C}|}\right\}^{N_t} \tag{37}$$

with $\mathrm{etr}\{.\}=\exp\mathrm{Tr}\{.\}$, $N_t=N_p+N_s$, and

$$\boldsymbol{T}_k=\frac{1}{N_t}\left\{\alpha^{-1}\left(\boldsymbol{X}-\mu_k\boldsymbol{as}^H\right)\left(\boldsymbol{X}-\mu_k\boldsymbol{as}^H\right)^H+\boldsymbol{YY}^H\right\} \tag{38}$$

with $\mu_k=0$ under H_0, $\mu_k=1$ under H_1. For the sake of convenience, let us first consider the null hypothesis. Then the MLE of α and $\boldsymbol{C}$ are obtained as the maximizers of (37). It is well known that $\boldsymbol{T}_0$ is the MLE of $\boldsymbol{C}$ [4]. Therefore, the MLE of α is obtained by minimizing $\alpha^{mN_p/N_t}\,|\boldsymbol{T}_0|$, or equivalently its logarithm. Let

$$\left(\boldsymbol{YY}^H\right)^{-1}\boldsymbol{XX}^H=\sum_{k=1}^{r}\lambda_k\boldsymbol{u}_k\boldsymbol{u}_k^H \tag{39}$$

be the eigenvalue decomposition of $\left(\boldsymbol{YY}^H\right)^{-1}\boldsymbol{XX}^H$ with $r=\min(m,N_p,N_s)$. Observing that

$$\boldsymbol{T}_0=\frac{1}{N_t}\left(\boldsymbol{YY}^H\right)\left[\boldsymbol{I}+\alpha^{-1}\left(\boldsymbol{YY}^H\right)^{-1}\boldsymbol{XX}^H\right] \tag{40}$$

it follows that we need to minimize

$$J(\alpha)=\frac{mN_p}{N_t}\ln\alpha+\sum_{k=1}^{r}\ln\left(1+\frac{\lambda_k}{\alpha}\right). \tag{41}$$

Differentiating the previous equation with respect to α and equating the result to zero yields

$$\frac{mN_p}{N_t}-\sum_{k=1}^{r}\frac{\lambda_k}{\lambda_k+\alpha}=0. \tag{42}$$

In the case $N_p=1=r$, a closed-form solution for α can be found (see [11]). For small values of r, say $r\leq 3$, an analytic solution is still available [3], but for large r, only a numerical solution can be found. This results in a computationally intensive detector. Furthermore, as shown in [3], this additional complexity does not result in any significant improvement compared to a two-step GLRT. Therefore, we only consider the latter approach in this paper.

References

[1] S. Bose and A. O. Steinhardt, "Adaptive array detection of uncertain rank one waveforms," *IEEE Trans. Signal Process.*, vol. 44, no. 11, pp. 2801–2809, Nov. 1996.

[2] O. Besson, L. L. Scharf, and F. Vincent, "Matched direction detectors and estimators for array processing with subspace steering vector uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4453–4463, Dec. 2005.

[3] E. Conte, A. De Maio, and G. Ricci, "GLRT-based adaptive detection algorithm for range-spread targets," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1336–1348, Jul. 2001.

[4] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 115–127, Mar. 1986.

[5] ——, "Adaptive detection in non-stationary interference, Part III," Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, MA, Tech. Rep. 761, 1987.

[6] R. S. Raghavan, N. Pulsone, and D. J. McLaughlin, "Performance of the GLRT for adaptive vector subspace detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 4, pp. 1473–1487, Oct. 1996.

[7] 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.

[8] S. Bose and A. O. Steinhardt, "A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2164–2175, Sep. 1995.

[9] ——, "Optimum array detector for a weak signal in unknown noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 3, pp. 911–922, Jul. 1996.

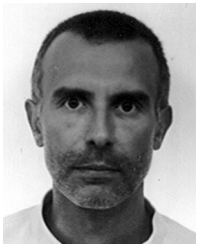
[10] L. L. Scharf and T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proc. 30th Asilomar Conf. Signals Systems Computers*, Pacific Grove, CA, Nov. 3–6, 1996, pp. 1114–1117.

[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] S. Kraut, L. L. Scharf, and T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.

[13] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most powerful invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, Feb. 2005.

[14] E. Conte, A. De Maio, and C. Galdi, "CFAR detection of multidimensional signals: An invariant approach," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 142–151, Jan. 2003.
[15] Y. Jin and B. Friedlander, "A CFAR adaptive subspace detector for second-order Gaussian signals," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 871–884, Mar. 2005.
[16] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York: Springer Verlag, 1986.
[17] C. G. Khatri and C. R. Rao, "Test for a specified signal when the noise covariance matrix is unknown," *J. Multivariate Anal.*, vol. 22, no. 2, pp. 177–188, Aug. 1987.
[18] K. Yao, "A representation theorem and its application to spherically invariant processes," *IEEE Trans. Inf. Theory*, vol. 19, no. 5, pp. 600–608, Sep. 1973.
[19] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.
[20] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Mutivariate Analysis*. New York: Academic, 1979.



Olivier Besson (SM'04) received the Ph.D. degree in signal processing and the "Habilitation à Diriger des Recherches" degree from INP, Toulouse, France, in 1992 and 1998, respectively.

Currently, he is an Associate Professor with the Department of Avionics and Systems of ENSICA, Toulouse, France. His research interests are in the general area of statistical signal and array processing with applications in radar and communications.

Dr. Besson is a member of the IEEE SAM Technical Committee and served as the Co-Technical Chairman of the IEEE SAM 2004 workshop. He was formerly an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and currently serves as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS.

Louis L. Scharf (F'86) received the Ph.D. degree from the University of Washington, Seattle.

From 1971 to 1982, he served as Professor of electrical engineering and statistics at Colorado State University (CSU), Fort Collins. From 1982 to 1985, he was Professor and Chairman of electrical and computer engineering at the University of Rhode Island, Kingston. From 1985 to 2000, he was Professor of electrical and computer engineering at the University of Colorado, Boulder. In January 2001, he rejoined CSU as Professor of electrical and computer engineering, and statistics. Since August 2004 he has served as Chief Scientist for TensorComm. He has held several visiting positions in the United States and abroad: Ecole Supérieure d'Electricité, Gif-sur-Yvette, France; ENST, Paris; EURECOM, Nice, France; the University of La Plata, La Plata, Argentina; Duke University, Durham, NC; the University of Wisconsin, Madison; and the University of Tromso, Tromso, Norway. His interests are in statistical signal processing as it applies to adaptive radar, sonar, and wireless communication. His most important contributions to date are invariance theories for detection and estimation; matched and adaptive subspace detectors and estimators for radar, sonar, and data communication; and canonical decompositions for reduced dimensional filtering and quantizing. His current interests are in rapidly adaptive receiver design for space-time and frequency-time signal processing in the wireless communication channel.

Prof. Scharf was Technical Program Chair for 1980 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Denver, CO; Tutorials Chair for ICASSP 2001, Salt Lake City, UT; and Technical Program Chair for Asilomar 2002, Pacific Grove, CA. He is Past Chair of the Fellow Committee for the IEEE Signal Processing Society and previously served on its Technical Committees for Theory and Methods and for Sensor Arrays and Multichannel Signal Processing. He has received numerous awards for his research contributions to statistical signal processing, including an IEEE Distinguished Lectureship, IEEE Third Millennium Medal, the Technical Achievement Award from the IEEE Signal Processing Society, and its Society Award.

Shawn Kraut (M'00) received the B.S. degree in engineering physics from the University of Arizona, Tucson, in 1993 and the Ph.D. degree in physics from the University of Colorado, Boulder, in 1999. The topic of his Ph.D. dissertation was adaptive subspace detectors used in radar array processing and related areas. In particular, he derived the generalized-likelihood ratio test (GLRT) and related properties of the adaptive coherence estimator (ACE) detection algorithm, which is robust with respect to scaling uncertainty in the noise statistics.

From 1999 to 2002, he was a Postdoctoral Research Associate at Duke University, Raleigh-Durham, NC, where he developed physics-based signal processing algorithms for applications in radar and sonar. From 2002 to 2005, he was an Assistant Professor at Queen's University, Ontario, Canada, where he showed that ACE is a uniformly mostly powerful (UMP)-invariant detection test and developed sequential-Bayesian algorithms for passive-sonar detection of moving sources in shallow water. Currently, he is a member of the technical staff at the Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Lexington, MA.