Adaptive Detection With Bounded Steering Vectors Mismatch Angle

Olivier Besson, *Senior Member, IEEE*

***Abstract*—We address the problem of detecting a signal of interest (SOI), using multiple observations in the primary data, in a background of noise with unknown covariance matrix. We consider a situation where the signal's signature is not known perfectly, but its angle with a nominal and known signature is bounded. Furthermore, we consider a possible scaling inhomogeneity between the primary and the secondary noise covariance matrix. First, assuming that the noise covariance matrix is known, we derive the generalized-likelihood ratio test (GLRT), which involves solving a semidefinite programming problem. Next, we substitute the unknown noise covariance matrix for its estimate obtained from secondary data, to yield the final detector. The latter is compared with a detector that assumes a known signal's signature.**

***Index Terms*—Array processing, detection, generalized-likelihood ratio test (GLRT), steering vector mismatch.**

I. Introduction and Problem Statement

The problem of detecting the presence of a signal of interest (SOI) against a background of colored noise is fundamental, especially in radar applications. This problem has been studied extensively in the literature (see, e.g., [1] for a list of references). Usually, the presence of a target is sought in a single vector under test, assuming that training samples, which contain noise only, are available (see, e.g., [2] and [3]). Furthermore, the space or space–time signature of the target is assumed to be known. Herein, we consider a slightly different framework, namely we assume that the primary data contains multiple observations and that the SOI signature is not known perfectly. The first assumption arises, for instance, when a high-resolution radar attempts to detect a range-spread target [4]. In such a case, the primary data consists of the array outputs in the range cells in which the target is likely to be present. Uncertainties about the SOI signature can be attributed to many possible causes, including uncalibrated arrays, pointing errors, multipath propagation, and wavefront distortions [5], [6].

Detection with multiple observations in the primary data and partly known signals of interest has been considered, among others in [4], [5], and [7]–[9]. Reference [5] considers detecting uncertain rank-one waveforms when both the space and time signatures of the SOI are assumed to belong to known linear subspaces. Bose and Steinhardt derive the maximal invariant along with the generalized-likelihood ratio test (GLRT) for this general framework. An extension to partially homogeneous environments is considered in [9]. Reference [4] considers the detection of a range-spread target using a high-resolution radar, assuming that the steering vector is known. The authors derive and analyze a two-step GLRT for both the homogeneous and the partially homogeneous case. In [7] the SOI signature is considered as unknown and arbitrary. The theory of invariance is invoked to obtain a most powerful invariant test and a suboptimal constant false alarm rate (CFAR) detector. Reference [8] proposes a CFAR detector based on a two-step GLRT when the signals of interest are Gaussian and belong to a known subspace.

Manuscript received January 11, 2006; revised June 28, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Daniel Fuhrman.

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

Digital Object Identifier 10.1109/TSP.2006.890820

For most of the above-mentioned studies, the fact that the SOI lies in a subspace facilitates the derivation of the detectors, as one usually ends up with closed-form detectors. However, the choice of this subspace is delicate since one must ensure that the SOI really belongs to it. Otherwise, there is some loss of performance. In this paper, we investigate a different approach. We assume that we have knowledge of the nominal value of the SOI signature and that the actual signature is "close" to its nominal value. More precisely, we assume that the angle between these two vectors is bounded. This approach was already advocated by the author in [10]. However, the latter reference only considers the case of a single vector under test, viz. $N_p = 1$: the present paper is an extension of [10] to the case where the primary data contains multiple snapshots. As will be illustrated below, considering $N_p > 1$ involves considerable complications, notably in the derivation of the maximum-likelihood estimate (MLE) of the SOI's signature. When $N_p = 1$, the technique of Lagrange multipliers (with a single Lagrange multiplier) is invoked in [10] to obtain the maximum-likelihood (ML) estimator. It is shown in [10] that the Lagrange multiplier is the unique solution to a secular equation. When $N_p > 1$, finding the MLEs turns out to be more complicated as it involves maximizing a quadratic form with an equality constraint and a nonconvex quadratic inequality constraint. In this case, the problem can be written as a semidefinite program, as will be shown in the next section.

We now define formally our detection problem. It consists of deciding between the following two hypotheses:

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

We denote by $\boldsymbol{X} = [\boldsymbol{x}(1) \ \cdots \ \boldsymbol{x}(N_p)]$ the $m \times N_p$ primary data array and $\boldsymbol{Y} = [\boldsymbol{y}(1) \ \cdots \ \boldsymbol{y}(N_s)]$, the $m \times N_s$ secondary data array. $\boldsymbol{x}(k)$ stands either for a space snapshot—in which case k denotes a time index—or a space–time snapshot, with m being the number of array antennas times the number of pulses. In the latter case, k corresponds to a range index.

In (1), $\boldsymbol{n}_p(k)$ and $\boldsymbol{n}_s(k)$ stand for the noise in the primary and secondary data, respectively. We assume that they are proper, zero-mean, independent, and Gaussian-distributed random vectors with covariance matrices

$$\mathcal{E}\{\boldsymbol{n}_p(k)\boldsymbol{n}_p(k)^H\} = \gamma \boldsymbol{M} \tag{2a}$$

$$\mathcal{E}\{\boldsymbol{n}_s(k)\boldsymbol{n}_s(k)^H\} = \boldsymbol{M}. \tag{2b}$$

Hence, we consider a partially homogeneous environment in which the covariance matrices of the primary and the secondary data have the same structure, but possibly different power. Both γ and $\boldsymbol{M}$ are unknowns.

Under H_1, the SOI is a rank-one matrix $\boldsymbol{s}\boldsymbol{t}^H$, with $\boldsymbol{t} = [t(1) \ \cdots \ t(N_p)]^T$. $\boldsymbol{s}$ is either the spatial or the space-time signature of interest, referred to as the steering vector, while $\boldsymbol{t}$ denotes the amplitude of the SOI in the time or range domain. We assume that $\boldsymbol{t}$ is arbitrary and unknown. Although $\boldsymbol{s}$ is unknown, we assume that it is close to a nominal (and therefore known) steering vector $\bar{\boldsymbol{a}}$. To account for this uncertainty, we assume that $\boldsymbol{s}$ is mostly aligned with $\bar{\boldsymbol{a}}$, and that the fraction of its energy outside $\mathcal{R}(\bar{\boldsymbol{a}})$ is bounded. More precisely, it is assumed that

$$\boldsymbol{s} \in \mathcal{C} = \left\{ \boldsymbol{s}; \frac{|\boldsymbol{s}^H \bar{\boldsymbol{a}}|^2}{(\boldsymbol{s}^H \boldsymbol{s})(\bar{\boldsymbol{a}}^H \bar{\boldsymbol{a}})} \geq \rho \right\} \tag{3}$$

where $0 < \rho < 1$ is a scalar that sets how much of the energy is allowed to be outside $\mathcal{R}(\bar{\boldsymbol{a}})$. The constraint (3) means that the square

of the cosine angle between $\boldsymbol{s}$ and $\bar{\boldsymbol{a}}$ must be above ρ. In other words, $\boldsymbol{s}$ belongs to a cone, whose axis is $\bar{\boldsymbol{a}}$ and whose angle θ_c is such that $\cos^2\theta_c = \rho$. In the limiting case $\rho = 1$, the actual steering vector is aligned with the presumed steering vector. Observe that such modeling of the steering vector uncertainties was also considered in [11], but in a slightly different way, namely a cone is defined for the real and the imaginary parts of the steering vector. Moreover, [11] considers the homogeneous case only, viz. $\gamma = 1$, and a single vector in the primary data, i.e., the case $N_p = 1$.

II. Detection

Our goal is therefore to detect the presence of the rank-one matrix $\boldsymbol{s}\boldsymbol{t}^H$, with $\boldsymbol{t}$ arbitrary and under the constraint defined in (3). Before deriving the detector, a few words are in order regarding the invariances of the detection problem at hand. When the steering vector is known to be $\bar{\boldsymbol{a}}$, the hypothesis testing problem corresponding to (1) is invariant under the group of transformations G defined by [12]

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

where α and β are arbitrary scalars, $\boldsymbol{T}$ is a full-rank matrix such that $\boldsymbol{T}\bar{\boldsymbol{a}} \propto \bar{\boldsymbol{a}}$, and $\boldsymbol{B}$ and $\boldsymbol{D}$ are unitary matrices. The rationale behind (4) can be briefly explained as follows. Only linear transformations are considered so as to retain Gaussianity of the measurements. Postmultiplication of $\boldsymbol{X}$ and $\boldsymbol{Y}$ by unitary matrices ensure that the columns of these matrices remain independent. Arbitrary scaling of $\boldsymbol{X}$ and $\boldsymbol{Y}$ by α and β is due to the unknown scaling factor γ between the noise covariance matrices of the primary and the secondary data. Finally, the transformation matrix $\boldsymbol{T}$ is constrained i) to be full rank in order for the transformed covariance matrix to be full rank and ii) to retain the structure of the SOI data matrix under H_1. More precisely, the steering vector of the transformed data should satisfy the same hypotheses as those of the original steering vector. This is why when $\boldsymbol{s}$ is known to be proportional to $\bar{\boldsymbol{a}}$, the matrix $\boldsymbol{T}$ should be such that $\boldsymbol{T}\bar{\boldsymbol{a}}$ is proportional to $\bar{\boldsymbol{a}}$. Under the framework considered here, the group of transformations is the same as in (4), except that the matrix $\boldsymbol{T}$ should now be chosen in such a way that the cone $\mathcal{C}$ is invariant to $\boldsymbol{T}$. In other words, the set $\{\boldsymbol{T}\boldsymbol{s}; \boldsymbol{s} \in \mathcal{C}\}$ should be $\mathcal{C}$. However, the natural invariances of a cone are scaling, rotation around its axis, and symmetry with respect to the hyperplane orthogonal to $\bar{\boldsymbol{a}}$. Therefore, the transformations $\boldsymbol{T}$ that leave the cone invariant are of the form [13]

$$\boldsymbol{T} = \xi \left[\boldsymbol{P}_{\bar{\boldsymbol{a}}} + \boldsymbol{U}_{\bar{\boldsymbol{a}}}^{\perp}\boldsymbol{Q}\boldsymbol{U}_{\bar{\boldsymbol{a}}}^{\perp\,H}\right] \tag{5}$$

where ξ is an arbitrary scalar, $\boldsymbol{P}_{\bar{\boldsymbol{a}}}$ is the orthogonal projection onto $\bar{\boldsymbol{a}}$, $\boldsymbol{U}_{\bar{\boldsymbol{a}}}^{\perp}$ is a $m \times (m-1)$ matrix whose columns form an orthonormal basis for $\mathcal{R}(\bar{\boldsymbol{a}})^{\perp}$, and $\boldsymbol{Q}$ is an unitary matrix. Indeed, every vector $\boldsymbol{s} \in \mathcal{C}$ can be written as a component aligned with $\bar{\boldsymbol{a}}$ and a component orthogonal to $\bar{\boldsymbol{a}}$, i.e., $\boldsymbol{s} \propto \bar{\boldsymbol{a}}\cos\theta + \bar{\boldsymbol{a}}_{\perp}\sin\theta$, with $\bar{\boldsymbol{a}}_{\perp}$ orthogonal to $\bar{\boldsymbol{a}}$, and $|\theta| \leq \theta_c$. Premultiplying by $\boldsymbol{T}$ (and ignoring ξ), the component along $\bar{\boldsymbol{a}}$ is not modified while the orthogonal component is rotated around $\bar{\boldsymbol{a}}$. It is straightforward to show that the square of the cosine angle between $\boldsymbol{s}$ and $\bar{\boldsymbol{a}}$ is unchanged, and hence $\mathcal{C}$ is invariant to $\boldsymbol{T}$. Accordingly, every vector in the one can be written as the product of $\boldsymbol{T}$ by another vector in the cone. To summarize, the hypothesis testing problem (1) with the constraint that $\boldsymbol{s} \in \mathcal{C}$ is invariant under the group of transformations G defined in (4), with $\boldsymbol{T}$ given by (5). The group of transformations induced on the parameter space is

$$\bar{g} : \begin{pmatrix} \boldsymbol{s} \\ \boldsymbol{t} \\ \gamma \\ \boldsymbol{M} \end{pmatrix} \rightarrow \begin{pmatrix} \alpha_1 \boldsymbol{T}\boldsymbol{s} \\ \alpha_2 \boldsymbol{B}\boldsymbol{t} \\ \gamma|\alpha|^2|\beta|^{-2} \\ |\beta|^2 \boldsymbol{T}\boldsymbol{M}\boldsymbol{T}^H \end{pmatrix} \tag{6}$$

with $\alpha_1\alpha_2^* = \alpha$. Once the invariances of the hypothesis testing problem have been found, it is natural to restrict the attention to invariant tests (we will see that the detector derived below is invariant). All of them will be a function of the maximal invariant statistic [5], [14]. Furthermore, the distribution of the maximal invariant statistic only depends on the so-called induced maximal invariant. In our case, we were not able to identify the maximal invariant statistic. Note that, even in the simple case where $\boldsymbol{s}$ belongs to a known linear subspace, derivation of the maximal invariant statistic is quite involved (see [5] for homogeneous environment, and [9], [15] for partially homogeneous environments). In our case, the problem is still more complicated as the steering vector $\boldsymbol{s}$ is defined through an inequality involving a quadratic form in $\boldsymbol{s}$. Therefore, in the sequel, we will simply prove invariance of the detector.

Let us now turn to the derivation of our detector. Towards this end, we will proceed in two steps. First, we assume that the covariance matrix $\boldsymbol{M}$ is known, and we derive the GLRT using the primary data only. Next, we substitute $\boldsymbol{M}$ for its MLE based on secondary data. Observe that, in principle, the GLRT based on the whole set of measurements $[\boldsymbol{X} \ \boldsymbol{Y}]$ could be used. However, as mentioned in [4], this results in a complicated estimator for the scaling factor γ, as soon as $N_p \geq 3$. Moreover, this one-step GLRT does not result in any significant improvement compared to the two-step GLRT considered herein. Therefore, we only consider the latter in the sequel. Hence, let us first assume that $\boldsymbol{M}$ is known. Under the assumptions made, the probability density function (PDF) of $\boldsymbol{X}$ is given by [13]

$$f(\boldsymbol{X}) = \frac{\mathrm{etr} - \{\gamma^{-1}(\boldsymbol{X} - \mu\boldsymbol{s}\boldsymbol{t}^H)(\boldsymbol{X} - \mu\boldsymbol{s}\boldsymbol{t}^H)^H\boldsymbol{M}^{-1}\}}{\pi^{mN_p}|\gamma\boldsymbol{M}|^{N_p}}. \tag{7}$$

In the previous equations, $\mathrm{etr}\{\cdot\} = \exp\mathrm{Tr}\{\cdot\}$, where $\mathrm{Tr}\{\cdot\}$ stands for the trace, $|\cdot|$ denotes the determinant, and $\mu = 0$ under H_0, $\mu = 1$ under H_1. It is straightforward to show that the maximum of $f(\boldsymbol{X})$ with respect to γ is attained when

$$\gamma = \frac{1}{mN_p}\mathrm{Tr}\{(\boldsymbol{X} - \mu\boldsymbol{s}\boldsymbol{t}^H)(\boldsymbol{X} - \mu\boldsymbol{s}\boldsymbol{t}^H)^H\boldsymbol{M}^{-1}\}. \tag{8}$$

Therefore

$$\begin{aligned}\max_{\gamma} f(\boldsymbol{X}) &= \frac{e^{-mN_p}}{\pi^{mN_p}|\gamma\boldsymbol{M}|^{N_p}} \\ &\propto [\mathrm{Tr}\{(\boldsymbol{X} - \mu\boldsymbol{s}\boldsymbol{t}^H)(\boldsymbol{X} - \mu\boldsymbol{s}\boldsymbol{t}^H)^H\boldsymbol{M}^{-1}\}]^{-mN_p}.\end{aligned} \tag{9}$$

Consequently, the generalized-likelihood ratio (GLR) can be written as

$$\mathrm{GLR} = \left\{\frac{\mathrm{Tr}\{\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{M}^{-1}\}}{\min_{\boldsymbol{s},\boldsymbol{t}}\mathrm{Tr}\{(\boldsymbol{X} - \boldsymbol{s}\boldsymbol{t}^H)(\boldsymbol{X} - \boldsymbol{s}\boldsymbol{t}^H)^H\boldsymbol{M}^{-1}\}}\right\}^{mN_p}. \tag{10}$$

The last step to complete the derivation of the GLRT consists of solving the minimization problem at the denominator of (10), with the constraint that $\boldsymbol{s} \in \mathcal{C}$. First, we show that this minimization problem can be solved for $\boldsymbol{t}$ explicitly, assuming that $\boldsymbol{s}$ is known. Indeed, observe that

$$\begin{aligned} J &= \mathrm{Tr}\{(\boldsymbol{X} - \boldsymbol{s}\boldsymbol{t}^H)(\boldsymbol{X} - \boldsymbol{s}\boldsymbol{t}^H)^H\boldsymbol{M}^{-1}\} \\ &= \mathrm{Tr}\{\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{M}^{-1}\} - \boldsymbol{t}^H\boldsymbol{X}^H\boldsymbol{M}^{-1}\boldsymbol{s} \\ &\quad - \boldsymbol{s}^H\boldsymbol{M}^{-1}\boldsymbol{X}\boldsymbol{t} + (\boldsymbol{t}^H\boldsymbol{t})(\boldsymbol{s}^H\boldsymbol{M}^{-1}\boldsymbol{s}) \\ &= \left[\boldsymbol{t} - \frac{\boldsymbol{X}^H\boldsymbol{M}^{-1}\boldsymbol{s}}{\boldsymbol{s}^H\boldsymbol{M}^{-1}\boldsymbol{s}}\right]^H(\boldsymbol{s}^H\boldsymbol{M}^{-1}\boldsymbol{s})\left[\boldsymbol{t} - \frac{\boldsymbol{X}^H\boldsymbol{M}^{-1}\boldsymbol{s}}{\boldsymbol{s}^H\boldsymbol{M}^{-1}\boldsymbol{s}}\right] \\ &\quad + \mathrm{Tr}\{\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{M}^{-1}\} - \frac{\boldsymbol{s}^H\boldsymbol{M}^{-1}\boldsymbol{X}\boldsymbol{X}^H\boldsymbol{M}^{-1}\boldsymbol{s}}{\boldsymbol{s}^H\boldsymbol{M}^{-1}\boldsymbol{s}}. \end{aligned} \tag{11}$$

Therefore, for any $\boldsymbol{s}$, the vector $\boldsymbol{t}$ that minimizes J is given by

$$\boldsymbol{t} = \frac{\boldsymbol{X}^H \boldsymbol{M}^{-1} \boldsymbol{s}}{\boldsymbol{s}^H \boldsymbol{M}^{-1} \boldsymbol{s}}. \tag{12}$$

Substituting this value in (11), we are left with the following maximization problem

$$\max_{\boldsymbol{s} \in \mathcal{C}} \frac{\boldsymbol{s}^H \boldsymbol{M}^{-1} \boldsymbol{X} \boldsymbol{X}^H \boldsymbol{M}^{-1} \boldsymbol{s}}{\boldsymbol{s}^H \boldsymbol{M}^{-1} \boldsymbol{s}}. \tag{13}$$

Note that both the objective function and the constraint are invariant to arbitrary scaling of $\boldsymbol{s}$. This stems from the fact that there exits an inherent scaling ambiguity between $\boldsymbol{s}$ and $\boldsymbol{t}$ from the formulation of the problem. Hence, we must somehow enforce an additional constraint on the norm of $\boldsymbol{s}$. For the sake of simplicity, we impose that $\boldsymbol{s}^H \boldsymbol{M}^{-1} \boldsymbol{s} = 1$ and the maximization problem can be reformulated as

$$\max_{\boldsymbol{s}} \boldsymbol{s}^H \boldsymbol{M}^{-1} \boldsymbol{X} \boldsymbol{X}^H \boldsymbol{M}^{-1} \boldsymbol{s} \text{ subject to } \begin{cases} \boldsymbol{s}^H \boldsymbol{M}^{-1} \boldsymbol{s} = 1 \\ \rho \boldsymbol{s}^H \boldsymbol{s} - \boldsymbol{s}^H \boldsymbol{P}_{\bar{\boldsymbol{a}}} \boldsymbol{s} \leq 0 \end{cases} \tag{14}$$

where $\boldsymbol{P}_{\bar{\boldsymbol{a}}} = (\bar{\boldsymbol{a}}^H \bar{\boldsymbol{a}})^{-1} \bar{\boldsymbol{a}} \bar{\boldsymbol{a}}^H$ is the orthogonal projection on $\bar{\boldsymbol{a}}$. Let us introduce the whitened vector $\tilde{\boldsymbol{s}} = \boldsymbol{M}^{-1/2} \boldsymbol{s}$ where $\boldsymbol{M}^{-1/2}$ is the inverse of the square-root of $\boldsymbol{M}$. Then, we need to solve

$$\max_{\tilde{\boldsymbol{s}}} \tilde{\boldsymbol{s}}^H \boldsymbol{A} \tilde{\boldsymbol{s}} \text{ subject to } \tilde{\boldsymbol{s}}^H \tilde{\boldsymbol{s}} = 1 \quad \text{and} \quad \tilde{\boldsymbol{s}}^H \boldsymbol{Q} \tilde{\boldsymbol{s}} \leq 0 \tag{15}$$

with

$$\boldsymbol{A} = \boldsymbol{M}^{-1/2} \boldsymbol{X} \boldsymbol{X}^H \boldsymbol{M}^{-1/2} \tag{16a}$$

$$\boldsymbol{Q} = \boldsymbol{M}^{1/2} [\rho \boldsymbol{I} - \boldsymbol{P}_{\bar{\boldsymbol{a}}}] \boldsymbol{M}^{1/2}. \tag{16b}$$

The problem consists of maximizing a quadratic form with the constraints that the solution lies on a sphere and within a (transformed) cone. Equivalently, it is a quadratic problem with a quadratic equality constraint and a quadratic inequality constraint. Note that the objective function is convex, but the inequality constraint is not convex. Indeed, $\boldsymbol{Q}$ and $\rho \boldsymbol{I} - \boldsymbol{P}_{\bar{\boldsymbol{a}}}$ have the same inertia [16]. However, $\rho \boldsymbol{I} - \boldsymbol{P}_{\bar{\boldsymbol{a}}}$ has $m-1$ positive eigenvalues equal to ρ and one negative eigenvalue equal to $\rho - 1$. Therefore, $\boldsymbol{Q}$ is not positive semidefinite.

The usual and widely used method to solve the aforementioned problem is the Lagrange multiplier technique [17]. The Lagrangian associated with the maximization problem in (15) can be written as [17]

$$L(\tilde{\boldsymbol{s}}, \lambda, \mu) = -\tilde{\boldsymbol{s}}^H \boldsymbol{A} \tilde{\boldsymbol{s}} + \lambda (\tilde{\boldsymbol{s}}^H \tilde{\boldsymbol{s}} - 1) + \mu \tilde{\boldsymbol{s}}^H \boldsymbol{Q} \tilde{\boldsymbol{s}} \tag{17}$$

where $\mu \geq 0$ and λ are the Lagrange multipliers, and $-\boldsymbol{A} + \lambda \boldsymbol{I} + \mu \boldsymbol{Q} \geq 0$ so that the Lagrangian can be effectively minimized. Differentiating (17) with respect to $\tilde{\boldsymbol{s}}$ and equating the result to zero yields

$$[-\boldsymbol{A} + \lambda \boldsymbol{I} + \mu \boldsymbol{Q}] \tilde{\boldsymbol{s}} = \boldsymbol{0}. \tag{18}$$

Therefore, the Lagrange multipliers must be found such that the matrix $\boldsymbol{A} - \lambda \boldsymbol{I} - \mu \boldsymbol{Q}$ is negative semidefinite with one eigenvalue equal to zero. However, it is not clear how to obtain simply λ and μ from this equation. Further insights can be gained by considering the dual problem. Prior to that, note that there is a zero duality gap between the primal problem and the dual problem if there exist two vectors $\tilde{\boldsymbol{s}}_1$ and $\tilde{\boldsymbol{s}}_2$ such that $\tilde{\boldsymbol{s}}_1^H \boldsymbol{Q} \tilde{\boldsymbol{s}}_1 < 0$, $\tilde{\boldsymbol{s}}_2^H \boldsymbol{Q} \tilde{\boldsymbol{s}}_2 < 0$ and $\tilde{\boldsymbol{s}}_1^H \tilde{\boldsymbol{s}}_1 > 1$, $\tilde{\boldsymbol{s}}_2^H \tilde{\boldsymbol{s}}_2 < 1$ (see [18]). In our case, observe that $\bar{\boldsymbol{a}}^H \boldsymbol{M}^{-1/2} \boldsymbol{Q} \boldsymbol{M}^{-1/2} \bar{\boldsymbol{a}} = |\bar{\boldsymbol{a}}|^2 [\rho - 1] < 0$. Hence, $\tilde{\boldsymbol{s}}_1 = \gamma_1 \boldsymbol{M}^{-1/2} \bar{\boldsymbol{a}}$ and $\tilde{\boldsymbol{s}}_2 = \gamma_2 \boldsymbol{M}^{-1/2} \bar{\boldsymbol{a}}$ with $\gamma_1 > 1/\|\boldsymbol{M}^{-1/2} \bar{\boldsymbol{a}}\|$ and $\gamma_2 < 1/\|\boldsymbol{M}^{-1/2} \bar{\boldsymbol{a}}\|$ are such vectors. It follows that the optimal value of the primal problem (15) will be identical to the optimal value of the dual problem. Therefore, we turn to the derivation of the latter. The dual function associated with $L(\tilde{\boldsymbol{s}}, \lambda, \mu)$ is [17]

$$\begin{aligned} g(\lambda, \mu) &= \inf_{\tilde{\boldsymbol{s}}} L(\tilde{\boldsymbol{s}}, \lambda, \mu) \\ &= \inf_{\tilde{\boldsymbol{s}}} \tilde{\boldsymbol{s}}^H [-\boldsymbol{A} + \lambda \boldsymbol{I} + \mu \boldsymbol{Q}] \tilde{\boldsymbol{s}} - \lambda \\ &= \begin{cases} -\lambda, & \text{if } -\boldsymbol{A} + \lambda \boldsymbol{I} + \mu \boldsymbol{Q} \geq 0 \\ -\infty, & \text{otherwise} \end{cases}. \end{aligned} \tag{19}$$

The dual problem is thus feasible if $\mu \geq 0$ and $-\boldsymbol{A} + \lambda \boldsymbol{I} + \mu \boldsymbol{Q} \geq 0$, in which case we need to solve

$$\max_{\lambda, \mu} -\lambda \text{ subject to } \begin{cases} \mu \geq 0 \\ -\boldsymbol{A} + \lambda \boldsymbol{I} + \mu \boldsymbol{Q} \geq 0 \end{cases} \tag{20a}$$

$$\Leftrightarrow \min_{\lambda, \mu} \lambda \text{ subject to } \begin{cases} \mu \geq 0 \\ -\boldsymbol{A} + \lambda \boldsymbol{I} + \mu \boldsymbol{Q} \geq 0 \end{cases}. \tag{20b}$$

The problem in (20) is a convex optimization problem, more precisely a semidefinite program [17]. It can be solved efficiently, in polynomial time, using interior-point methods. Moreover, these methods are now available through software packages such as SeDuMi [19].

Let us denote $[\hat{\lambda} \ \hat{\mu}]^T$ as the so-obtained solution; observe that $\hat{\lambda}$ is the largest eigenvalue of $\boldsymbol{A} - \hat{\mu} \boldsymbol{Q}$. Let $\hat{\tilde{\boldsymbol{s}}}$ denote the corresponding eigenvector, i.e., $(\boldsymbol{A} - \hat{\mu} \boldsymbol{Q} - \hat{\lambda} \boldsymbol{I}) \hat{\tilde{\boldsymbol{s}}} = \boldsymbol{0}$. Since there is a zero duality gap between the problems (17) and (20), $\hat{\lambda}$ is the maximum value of $\tilde{\boldsymbol{s}}^H \boldsymbol{A} \tilde{\boldsymbol{s}}$ under the constraints $\tilde{\boldsymbol{s}}^H \tilde{\boldsymbol{s}} = 1$ and $\tilde{\boldsymbol{s}}^H \boldsymbol{Q} \tilde{\boldsymbol{s}} \leq 0$. Furthermore, the maximum value is attained when $\tilde{\boldsymbol{s}} = \hat{\tilde{\boldsymbol{s}}}$. This implies, in particular, that $\hat{\tilde{\boldsymbol{s}}}^H \boldsymbol{Q} \hat{\tilde{\boldsymbol{s}}} = 0$, and hence the solution lies on the boundary of the cone.

The GLRT for known $\boldsymbol{M}$ can thus be written as

$$\text{GLRT} = \frac{\hat{\tilde{\boldsymbol{s}}}^H \boldsymbol{A} \hat{\tilde{\boldsymbol{s}}}}{\text{Tr}\{\boldsymbol{A}\}} = \frac{\hat{\boldsymbol{s}}^H \boldsymbol{M}^{-1} \boldsymbol{X} \boldsymbol{X}^H \boldsymbol{M}^{-1} \hat{\boldsymbol{s}}}{\text{Tr}\{\boldsymbol{X} \boldsymbol{X}^H \boldsymbol{M}^{-1}\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{21}$$

where $\hat{\boldsymbol{s}} = \boldsymbol{M}^{1/2} \hat{\tilde{\boldsymbol{s}}}$. Note the numerator in (21) is simply $\hat{\lambda}$. When $\boldsymbol{M}$ is unknown, it is replaced by its MLE based on the secondary data, namely

$$\boldsymbol{S} = N_s^{-1} \boldsymbol{Y} \boldsymbol{Y}^H. \tag{22}$$

Doing so, the two-step GLRT is given by

$$\frac{\hat{\boldsymbol{s}}^H \boldsymbol{S}^{-1} \boldsymbol{X} \boldsymbol{X}^H \boldsymbol{S}^{-1} \hat{\boldsymbol{s}}}{\text{Tr}\{\boldsymbol{X} \boldsymbol{X}^H \boldsymbol{S}^{-1}\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{23}$$

It should be understood that $\hat{\boldsymbol{s}}$ in (23) differs from $\hat{\boldsymbol{s}}$ in (21) as the former is the solution of (14) with $\boldsymbol{S}$ replacing $\boldsymbol{M}$. However, for the sake of simplicity, we keep the same notation. We point out that the detector above is invariant to the group of transformations defined in (4). In fact, the numerator of (23), which coincides with the maximal value of the criterion in (14), will be multiplied by a factor $|\alpha|^2 |\beta|^{-2}$ (and the solution of (14) becomes $\boldsymbol{T} \hat{\boldsymbol{s}}$). This scaling factor is canceled in the GLR as the denominator of (23) will also be scaled by the same quantity. Hence, our two-step GLRT enjoys the natural invariances of the problem. However, we cannot assert that the present detector possesses the CFAR property with respect to the noise covariance matrix. In fact, a detector may be invariant but not CFAR (see [14]), although the most encountered situation is that invariance implies CFARness. Indeed, to prove CFARness, one should prove that the induced maximal invariant is a function of parameters that are relevant under H_1 only. In other words, one should prove that the distribution of the maximal invariant statistic is parameter-free under H_0 (see, e.g., [7] and [14] for a comprehensive discussion). In our case, since the induced maximal invariant is unknown, we cannot assess theoretically the performance

of the two-step GLRT. Another route would be to derive directly the distribution of the test statistic in (23). However, this seems quite complicated as we do not have any closed-form expression for $\hat{\boldsymbol{s}}$. Therefore, in the next section, the performance of the detector will be evaluated by numerical simulations.

Before closing this section, we make the following observation. When $\boldsymbol{s}$ is known to be aligned with $\bar{\boldsymbol{a}}$, derivation of the GLRT ends with (11), and the two-step GLRT becomes

$$\frac{\bar{\boldsymbol{a}}^H \boldsymbol{S}^{-1} \boldsymbol{X}\boldsymbol{X}^H \boldsymbol{S}^{-1} \bar{\boldsymbol{a}}}{(\bar{\boldsymbol{a}}^H \boldsymbol{S}^{-1} \bar{\boldsymbol{a}}) \mathrm{Tr}\{\boldsymbol{X}\boldsymbol{X}^H \boldsymbol{S}^{-1}\}} \underset{H_0}{\overset{H_1}{\gtrless}} \zeta. \tag{24}$$

The detector in (24) coincides with the two-step GLRT derived in [4, eq. (26)], where it is referred to as the generalized adaptive subspace detector (GASD). Note that, when $\boldsymbol{s}$ is known to be aligned with $\bar{\boldsymbol{a}}$, the GASD is CFAR with respect to the noise covariance matrix, i.e., for a given probability of false alarm, the threshold ζ can be set independently of $\boldsymbol{M}$. In the present case where $\boldsymbol{s}$ belongs to a cone, the CFAR property can no longer be claimed for the GASD. Finally, observe that, as θ_c goes to 0, the cone reduces to its axis $\bar{\boldsymbol{a}}$, and the two-step GLRT converges to the GASD.

III. Numerical Examples

In this section, we assess the performance of the detector in (23) and compare it with that of the GASD (24). We consider different scenarios depending on wether $\boldsymbol{s}$ is aligned with $\bar{\boldsymbol{a}}$ or not. We also investigate the influence of the choice of the cone angle θ_c. More precisely, we let θ denote the actual angle between $\boldsymbol{s}$ and $\bar{\boldsymbol{a}}$, and we consider cases where $\theta \leq \theta_c$ and cases where $\theta \geq \theta_c$. In other words, the cone angle can overestimate or underestimate the true angle between the steering vector and its nominal value. This enables us to test for the robustness of our detector and to assess the sensitivity of the detection performance towards the user-defined parameter θ_c.

In all simulations, we consider an array with $m = 8$ elements and an exponentially shaped noise covariance matrix. The (k, ℓ)th element of $\boldsymbol{M}$ is

$$\boldsymbol{M}_{k,\ell} = \alpha^{|k-\ell|} \tag{25}$$

with $\alpha = 0.9$ in the simulations below. In order to set the thresholds η and ζ for a given probability of false alarm P_{fa}, Monte Carlo counting techniques are used. 10^6 simulations of the data under the null hypothesis were run, and the test statistics in (23) and (24) were computed and sorted. The thresholds were set from the $1 - P_{fa}$ quantile. In the simulations shown below, $P_{fa} = 10^{-3}$. The probability of detection P_d is obtained from 10^5 independent trials. P_d is plotted as a function of the signal-to-noise ratio (SNR), which is defined as $\mathrm{SNR} = (\boldsymbol{t}^H \boldsymbol{t})[\boldsymbol{s}^H (\gamma \boldsymbol{M})^{-1} \boldsymbol{s}]$. The number of snapshots in the primary and secondary data is set to $N_p = 4$ and $N_s = 20$, respectively. In all simulations, SeDuMi was used to obtain the ML estimates.

First, we consider a scenario for which $\theta = 0$, i.e., the actual steering vector is aligned with the presumed steering vector $\bar{\boldsymbol{a}}$. In this case, GASD is expected to perform better than the two-step GLRT, which does not know $\boldsymbol{s}$. In Fig. 1, we display the probability of detection of both detectors, with different values of the cone angle, namely $\cos\theta_c = 0.97, \cos\theta_c = 0.95$, and $\cos\theta_c = 0.9$. Observe that, as $\cos\theta_c$ decreases, the cone angle increases, and, hence, one may expect that the performance of the two-step GLRT degrades. From inspection of these figures, it turns out that the GASD performs slightly better than the two-step GLRT: the difference is about 1 dB at $P_d = 0.8$ for $\cos\theta_c = 0.97$ and $\cos\theta_c = 0.95$. Obviously, with $\cos\theta = 1$ and $\cos\theta_c = 0.9$, the cone is chosen too large and a detection loss is incurred by the two-step GLRT.

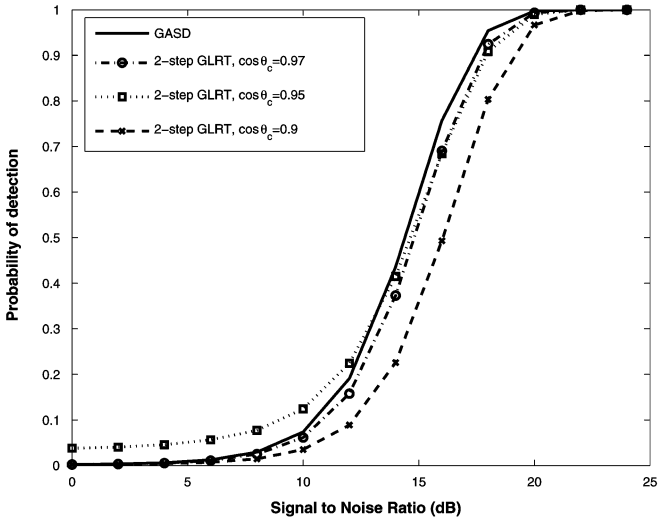

Fig. 1. P_d of the two-step GLRT and GASD versus SNR. $\boldsymbol{s}$ is aligned with $\bar{\boldsymbol{a}}$, $(\cos\theta = 1)$, varying cone angle θ_c.

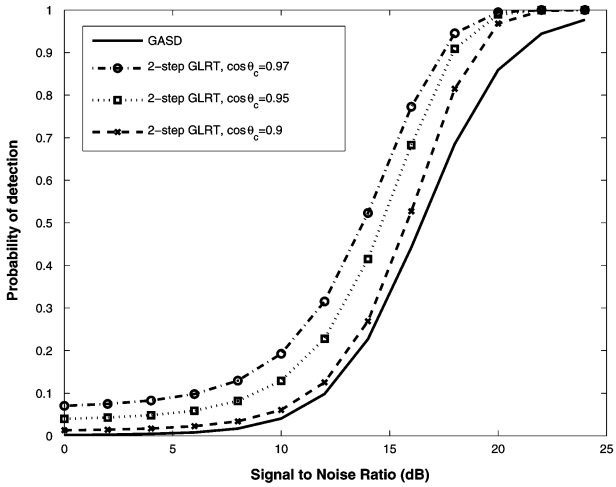

Fig. 2. P_d of the two-step GLRT and GASD versus SNR. $\boldsymbol{s}$ is not aligned with $\bar{\boldsymbol{a}}$, $(\cos\theta = 0.95)$, varying cone angle θ_c.

Next, we consider a scenario where $\cos\theta = 0.95$, i.e., the actual steering vector is not aligned with $\bar{\boldsymbol{a}}$. The results for this case are shown in Fig. 2. It can be observed that the two-step GLRT now performs better than the GASD; the difference is about 3, 2, and 1 dB at $P_d = 0.8$ for $\cos\theta_c = 0.97, \cos\theta_c = 0.95$ and $\cos\theta_c = 0.9$, respectively. However, if the cone angle was chosen too large, the GASD would finally perform better than the two-step GLRT. Interestingly enough, the improvement is more pronounced when $\cos\theta_c = 0.97$, indicating that it is not necessary to have a cone perfectly suited to the actual mismatch between the true and the presumed steering vector. In fact, it seems preferable to slightly underestimate the cone angle than to overestimate it. In any case, the two-step GLRT provides a significant robustness improvement compared to the GASD. This improvement increases when θ increases.

IV. Conclusion

We considered the problem of detecting a signal whose unknown signature lies in a cone, using multiple observations in the primary data, and under a possible scaling inhomogeneity between the samples under test and the training samples. A two-step GLRT was derived, which

involves solving a semidefinite programming problem. The new detector offers additional robustness compared to a detector that assumes a known signature, and is a relevant alternative whenever there exists a possible mismatch between the actual signature and the presumed one.

REFERENCES

[1] F. Gini, A. Farina, and M. Greco, "Selected list of references on radar signal processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 1, pp. 329–359, Jan. 2001.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] O. Besson, L. L. Scharf, and S. Kraut, "Adaptive detection of a signal known only to lie on a line in a known subspace, when primary and secondary data are partially homogeneous," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4698–4705, Dec. 2006.

[10] O. Besson, "Detection of a signal in linear subspace with bounded mismatch," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 3, pp. 1131–1139, Jul. 2006.

[11] A. De Maio, "Robust adaptive radar detection in the presence of steering vector mismatches," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1322–1337, Oct. 2005.

[12] 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.

[13] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

[14] 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.

[15] E. Conte and A. De Maio, "An invariant framework for adaptive detection in partially homogeneous environment," *WSEAS Trans. Circuits Syst.*, vol. 2, no. 1, pp. 282–287, Jan. 2003.

[16] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[18] A. Beck and Y. C. Eldar, "Doubly constrained robust Capon beamformer with ellipsoidal uncertainty sets," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 753–758, Feb. 2007.

[19] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimiz. Methods Softw.*, vol. 11–12, pp. 625–653, Aug. 1999.

ISI-Free Block Transceivers for Unknown Frequency Selective Channels

Chih-Hao Liu, See-May Phoong, and Yuan-Pei Lin

***Abstract*—The orthogonal frequency-division multiplexing (OFDM) transceiver has enjoyed great success in many wideband communication systems. It has low complexity and robustness against channel-induced intersymbol interference (ISI). When the channel order does not exceed the length of cyclic prefix, any frequency-selective channel is converted to a set of frequency-nonselective subchannels. This channel-independent ISI-free property is useful for many applications. In this correspondence, we study general block transceiver with such a property. We will show that the solutions of channel-independent ISI-free block transceivers are given in a closed form. It is found that except for some special cases, the solutions are identical to the Lagrange–Vandermonde and Vandermonde–Lagrange transceivers.**

***Index Terms*—Filter bank, multicarrier, multitone, orthogonal frequency-division multiplexing (OFDM), transceiver, transmultiplexer.**

I. INTRODUCTION

In recent years, the orthogonal-frequency-division-multiplexing (OFDM) system has been widely adopted for wideband communications [1]. One of the advantages of OFDM systems is their ability to combat channel-induced intersymbol interference (ISI). In an OFDM system, the transmitter and receiver perform respectively M point inverse discrete Fourier transform (IDFT) and discrete Fourier transform (DFT) operations. By adding a cyclic prefix of length L, any frequency-selective channel of order L is converted to a set of M parallel frequency-nonselective subchannels. Symbol recovery can be obtained by using simple one-tap equalizers at the receiver. Such a channel-independent ISI-free property is useful for many applications.

Recently, there has been some interest in finding other transceivers with channel-independent ISI-free property [2]–[8]. The first non-DFT-based transceiver with such a property was proposed [2]. By judiciously selecting the zeros of the transmit filters, the authors showed that when the number of trailing zeros is larger than or equal to the channel order, ISI can be eliminated completely by using a channel-independent receiver. The transmit filters are M Lagrange interpolation polynomials, whereas the receive filters are M Vandermonde filters, and therefore such a transceiver is called a Lagrange–Vandermonde (LV) transceiver. A dual system called a Vandermonde–Lagrange (VL) transceiver, where the transmit filters are Vandermonde filters and the receive filters are Lagrange filters, was derived in [3]. The LV and VL systems were generalized to the so-called mutually orthogonal user code receiver (AMOUR) system [4]. In [5] and [6], these LV and VL systems were studied using a different framework. Using the multirate technique, it was demonstrated that given an exponential vector input, the output of the Toeplitz channel matrix is also an exponential vector. Exploiting this

Manuscript received November 14, 2005; revised June 27, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zoran Cvetkovic. This work was supported by the National Science Council, Taiwan, R.O.C., under Grant NSC94-2752-E-002-006-PAE, NSC94-2213-E-002-075, and NSC94-2213-E-009-038.

C.-H. Liu and S.-M. Phoong are with the Department of Electrical Engineering and the Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan 106, R.O.C. (e-mail: smp@cc.ee.ntu.edu.tw).

Y.-P. Lin is with the Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan 300, R.O.C.

Digital Object Identifier 10.1109/TSP.2006.890826