Adaptive Detection and Steering Estimation in the Presence of Useful Signal and Interference Mismatches

Antonio De Maio, Silvio De Nicola, Yongwei Huang, Shuzhong Zhang, Alfonso Farina

Abstract

This paper considers adaptive detection and steering estimation in the presence of useful signal and interference mismatches. We assume a homogeneous environment where the random disturbance components from the primary and secondary data share the same covariance matrix. Moreover the data under test contains a deterministic interference vector in addition to the possible useful signal. We focus on the situation where an energy fraction of both the useful signal and the deterministic interference may lie outside the range span of their nominal directions. Under these conditions, we devise a procedure for the computation of the joint Maximum Likelihood (ML) estimators of the useful signal and interference vectors, resorting to a suitable rank-one decomposition of a Semidefinite Program (SDP) problem optimal solution. Hence we use the aforementioned estimators for the synthesis of adaptive receivers based on different Generalized Likelihood Ratio Test (GLRT) criteria. At the analysis stage we assess the performance of the new detectors also in comparison with two decision rules, available in open literature, which assume the useful signal and interference components perfectly aligned with their nominal directions.

Keywords. Adaptive Detection, Non-Convex Quadratic Optimization, Semidefinite Program Relaxation.

A. De Maio and S. De Nicola are with Università degli Studi di Napoli “Federico II”, Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni, Via Claudio 21, I-80125 Napoli, Italy. E-mail: ademaio@unina.it, silviodenicola@gmail.com

Y. Huang and S. Zhang are with Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. E-mail: ywhuang@se.cuhk.edu.hk, zhang@se.cuhk.edu.hk

A. Farina is with Selex - Sistemi Integrati Via Tiburtina Km. 12.4, I-00131, Roma, Italy. E-mail: afarina@selex-si.com

This work has been performed during the visit of A. De Maio and S. De Nicola at the Department of Systems Engineering and Engineering Management of the Chinese University of Hong Kong.

I. INTRODUCTION

Adaptive detection of a signal vector in the presence of homogeneous Gaussian environment (namely disturbance components from the test vector and the training data with the same spectral properties) with unknown covariance matrix has been extensively addressed in open literature under different models for the useful signal. If the quoted vector is known up to a scaling factor, an elegant framework based on the theory of invariance, is developed in [1]. Therein the maximal invariant statistic is derived and it is shown that no Uniformly Most Powerful Invariant (UMPI) test exists. Due to the lack of an optimum receiver, sub-optimum decision rules, which exhibit the invariance property, have been proposed. In [2] the one-step GLRT criterion is exploited, in [3] the Adaptive Matched Filter (AMF) is synthesized resorting to the two-step GLRT technique[1], in [6] the Adaptive Beamforming Orthogonal Rejection Test (ABORT) is devised, and in [7] the Rao test receiver is introduced and analyzed. Two-stage decision rules (equipped with two detection thresholds) have also been studied. In [8] the Adaptive Sidelobe Blanker (ASB) is proposed, in [9] the AMF and Kelly's receiver [2] are connected through an *and logic*, and in [7] the Rao test is used as second stage of the aforementioned connection in place of Kelly's GLRT. Another detector which still possesses invariant properties is the Adaptive Normalized Matched Filter (ANMF) [10], also known as Adaptive Coherence Estimator (ACE) [12]. Actually it also exhibits (in addition to the invariance under the group specified in [1]) a further invariance with respect to a common scaling of the secondary data relative to the primary one.

The first generalization of the aforementioned model for the useful signal (where only a scale factor is unknown) relies on considering a subspace model. This is tantamount to assuming that the useful component of the received signal belongs to a known subspace. In this case the invariance group and the maximal invariant statistic is derived in [13]. Again there is not a UMPI test but the one-step and two-step GLRT are invariant decision rules [13]. Other subspace detectors can be found in [11] with reference to the case of known disturbance covariance matrix, in [14] where an extensive analysis on real Over The Horizon (OTH) radar data is also conducted, in [15] where a scale-mismatch between the primary and secondary data is accounted for, in [16] and [17] for radar Space Time Adaptive Processing (STAP) applications.

[1] This receiver is also known as the Modified Sample Matrix Inversion (MSMI) receiver [4] and coincides with the Wald test [5].

If the useful signal component is modeled as a completely unknown vector, the maximal invariant statistic is one-dimensional and a UMPI test, coinciding with the GLRT, exists [18].

Non-linear signal models for the useful vector have been introduced in [19], in [20], and in [21]. They permit to handle situations where the useful signal is not actually contained a nominal subspace. However, exept for the detector in [21], the decision statistics have no longer a closed form expression like under the linear subspace modeling for the useful signal.

In this paper we consider the detection of a useful signal component within the primary data vector in the presence of noise, modeled as complex normal with unknown covariance matrix, plus deterministic interference. Moreover we suppose that a set of noise-only additional data (the secondary data) is available and the environment is homogeneous. We explicitly highlight that the idea of introducing a deterministic interference component in the primary data has already been proposed in [22], with reference to the detection of range-spread targets (multiple primary data), and in [23] for the design of matched direction detectors. The main novelty of this paper is the assumption that an energy fraction of both the useful signal component and the deterministic interference vector may lie outside the range span of their nominal directions. Actually the resulting detection problem doesn't admit a Uniformly Most Powerful (UMP) test because the optimum receiver, according to the Neyman-Pearson criterion, requires the knowledge of the useful and interfering signals as well as of the disturbance covariance matrix. Hence, to cope with this uncertainty, we synthesize GLRT-based receivers for the aforementioned hypothesis test. Toward this goal, it is first necessary to evaluate the Maximum Likelihood (ML) estimates of the unknown parameters under the hypothesis of useful signal presence (H_1). In this context we prove that, under H_1, joint ML estimation of the useful and interfering signatures is tantamount to solving a non-convex Quadratic optimization Program (QP) problem which, interestingly, possesses a so called *hidden convexity* property. Otherwise stated, solving the QP problem is equivalent to solving an SDP problem, which is convex, and its optimal value can be found in polynomial time. Moreover, after the optimal value of the SDP problem is proven to be attainable, we design a procedure for the construction of an optimal solution of the QP problem starting from an optimal solution of SDP one. The proposed technique resorts to the rank-one decomposition theorem of [24] and still exhibits a polynomial computational complexity. Exploiting this finding we design several receivers based on different GLRT criteria, one-step GLRT, two-step GLRT, modified one-step GLRT, and modified two-step GLRT. Moreover we evaluate the computational complexity for the implementation of the decision statistics.

At the analysis stage we assess the performance of the devised receivers also in comparison with the one-step GLRT and two-step GLRT of [22] which assume the coincidence of the nominal and the actual steering vector (both with reference to the useful signal and the interference). The results show that, when a mismatch is present, the new receivers outperforms their counterparts and the performance gain depends on the entity of the mismatch.

The paper is organized as follows. In Section II we formulate the problem and specify the mismatch sets for both the useful signal and the interference. In Section III, the one-step GLRT is considered and the procedure for the derivation of the ML estimators of the unknown signatures (under H_1) is designed. In Section IV the two-step GLRT criterion is exploited, whereas the use of the modified one-step and two-step GLRT strategies is proposed in Section V. The performance analysis of the new decision rules is handled in Section VI. Finally concluding remarks and hints for possible future research tracks are given in Section VII.

II. Problem Formulation

We assume that data are collected from N sensors and denote by $\boldsymbol{z}$ the complex vector of the samples where the presence of the useful signal is sought (primary data). As in [2], we also suppose that a secondary data set $\boldsymbol{z}_t$, $t = 1, \dots, K$, is available ($K \geq N$), that each of such snapshots does not contain any useful target echo, and exhibits the same covariance matrix as the primary data.

The detection problem to be solved can be formulated in terms of the following binary hypotheses test:

$$\begin{cases} H_0 : \begin{cases} \boldsymbol{z} = \boldsymbol{q} + \boldsymbol{w} \\ \boldsymbol{z}_t = \boldsymbol{w}_t \qquad t = 1, \dots, K \end{cases} \\ \\ H_1 : \begin{cases} \boldsymbol{z} = \boldsymbol{p} + \boldsymbol{q} + \boldsymbol{w} \\ \boldsymbol{z}_t = \boldsymbol{w}_t \qquad t = 1, \dots, K \end{cases} \end{cases} \tag{1}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ denote respectively the actual steering vector and the interference signature, which due to several effects such as source waveform distortion, pointing errors, and imperfect array calibration might not be aligned with the nominal directions $\boldsymbol{u}_p$ and $\boldsymbol{u}_q$ assumed, without loss of generality, with unitary norm and such that $\boldsymbol{u}_p \neq \boldsymbol{u}_q$. As to the noise components, we suppose that $\boldsymbol{w}$ and $\boldsymbol{w}_t$'s, $t = 1, \ldots, K$, are independent, zero-mean complex circular Gaussian vectors with positive definite covariance matrix given by

$$E[\boldsymbol{w}\boldsymbol{w}^\dagger] = E[\boldsymbol{w}_t \boldsymbol{w}_t^\dagger] = \boldsymbol{M}, \ \ t = 1, \ldots, K \, , \tag{2}$$

where $E[\cdot]$ denotes statistical expectation and $(\cdot)^\dagger$ conjugate transpose.

In the perfect matching case it is assumed that $\boldsymbol{p} = \alpha \boldsymbol{u}_p$ and $\boldsymbol{q} = \beta \boldsymbol{u}_q$ with α and β unknown complex parameters accounting for the channel propagation effects, the target reflectivity (α), and the complex amplitude of the directional interference (β). In practical situations mismatches may occur which cause deviations of $\boldsymbol{p}$ and $\boldsymbol{q}$ from the nominal directions $\boldsymbol{u}_p$ and $\boldsymbol{u}_q$. A possible way to cope with these scenarios is to assume more uncertainty about $\boldsymbol{p}$ and $\boldsymbol{q}$ than the case of perfect matching where only the proportionality factors α and β are considered unknown. According to this guideline direction mismatches can be accounted for, at the design stage, assuming that the actual steering vector and the interference signature lie respectively within the conic regions C_p and C_q defined as

$$C_p = \left\{ \boldsymbol{p} \in \mathcal{C}^N \ \ : \ \ \|(\boldsymbol{I} - \boldsymbol{u}_p \boldsymbol{u}_p^\dagger)\boldsymbol{p}\| \leq \gamma_p |\boldsymbol{p}^\dagger \boldsymbol{u}_p| \ \right\} \, , \tag{3}$$

and

$$C_q = \left\{ \boldsymbol{q} \in \mathcal{C}^N \ \ : \ \ \|(\boldsymbol{I} - \boldsymbol{u}_q \boldsymbol{u}_q^\dagger)\boldsymbol{q}\| \leq \gamma_q |\boldsymbol{q}^\dagger \boldsymbol{u}_q| \ \right\} \, , \tag{4}$$

where $\gamma_p > 0$ and $\gamma_q > 0$ are design parameters which rule the size of the two uncertainty regions, $\|\cdot\|$ denotes the Euclidean norm of a complex vector, and $|\cdot|$ the modulus of a complex number. This is tantamount to assuming that the fraction of energy of $\boldsymbol{p}$ ($\boldsymbol{q}$) outside the range span of $\boldsymbol{u}_p$ ($\boldsymbol{u}_q$) is bounded, and the parameter γ_p (γ_q) rules how much of the total energy is allowed to be outside the subspace spanned by $\boldsymbol{u}_p$ ($\boldsymbol{u}_q$). The sets C_p and C_q can also be written more compactly as

$$C_p = \left\{ \boldsymbol{p} \in \mathcal{C}^N \ \ : \ \ \boldsymbol{p}^\dagger \boldsymbol{R}_p \boldsymbol{p} \leq 0, \quad \boldsymbol{R}_p = \boldsymbol{I} - (1+\gamma_p^2)\boldsymbol{u}_p \boldsymbol{u}_p^\dagger \ \right\} \, , \tag{5}$$

and

$$C_q = \left\{ \boldsymbol{q} \in \mathcal{C}^N \ \ : \ \ \boldsymbol{q}^\dagger \boldsymbol{R}_q \boldsymbol{q} \leq 0, \quad \boldsymbol{R}_q = \boldsymbol{I} - (1+\gamma_q^2)\boldsymbol{u}_q \boldsymbol{u}_q^\dagger \ \right\} \, . \tag{6}$$

Moreover they are supposed to comply with the condition

$$\nexists \ \boldsymbol{c} \in \mathcal{C}^N \ \ : \ \ \boldsymbol{c} \neq 0, \ \ \boldsymbol{c} \in C_p \ \text{ and } \ \boldsymbol{c} \in C_q \, . \tag{7}$$

According to the Neyman-Pearson criterion, the optimum solution to the hypotheses testing problem (1), is the Likelihood Ratio Test (LRT); but, for the case at hand, it cannot be implemented since it requires the knowledge of the parameters $\boldsymbol{p}$, $\boldsymbol{q}$, and $\boldsymbol{M}$ which in practical situations are unknown. A possible way to circumvent this drawback is to resort to GLRT techniques which are equivalent to replacing the unknown parameters with their maximum likelihood estimates. Following this guideline in the sequel we design the Robust one-step GLRT (R-1S-GLRT) the Robust two-step GLRT (R-2S-GLRT), the Modified R-1S-GLRT (MR-1S-GLRT), and the Modified R-2S-GLRT (MR-1S-GLRT) for the problem at hand.

III. Robust One-Step GLRT (R-1S-GLRT) and ML Steering Estimators

This criterion is equivalent to replacing into the LRT the unknown parameters with their maximum likelihood estimates under each hypothesis, based upon the entirety of data [25]. For the present problem the one-step GLRT can be written as

$$\frac{\max_{(\boldsymbol{p},\boldsymbol{q})\in C_p\times C_q}\max_{\boldsymbol{M}} f(\boldsymbol{z},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K|\boldsymbol{p},\boldsymbol{q},\boldsymbol{M},H_1)}{\max_{\boldsymbol{q}\in C_q}\max_{\boldsymbol{M}} f(\boldsymbol{z},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K|\boldsymbol{q},\boldsymbol{M},H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} G \tag{8}$$

where $f(\boldsymbol{z},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K|\boldsymbol{q},\boldsymbol{M},H_0)$ and $f(\boldsymbol{z},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K|\boldsymbol{p},\boldsymbol{q},\boldsymbol{M},H_1)$ denote respectively the complex multivariate probability density functions (pdf's) of the vectors $\boldsymbol{z},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K$ under the H_0 and the H_1 hypotheses. Further developments require specifying the quoted pdf's. Previous assumptions imply that

$$f(\boldsymbol{z},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K|\boldsymbol{q},\boldsymbol{M},H_0) = \left\{\frac{1}{\pi^N \det(\boldsymbol{M})} e^{-\mathrm{tr}\left(\boldsymbol{M}^{-1}\boldsymbol{T}_0\right)}\right\}^{K+1} \tag{9}$$

and

$$f(\boldsymbol{z},\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K|\boldsymbol{p},\boldsymbol{q},\boldsymbol{M},H_1) = \left\{\frac{1}{\pi^N \det(\boldsymbol{M})} e^{-\mathrm{tr}\left(\boldsymbol{M}^{-1}\boldsymbol{T}_1\right)}\right\}^{K+1} \tag{10}$$

where $\det(\cdot)$ and $\mathrm{tr}(\cdot)$ denote respectively the determinant and the trace of a square matrix while $\boldsymbol{T}_0$ and $\boldsymbol{T}_1$ are defined as

$$\begin{cases} \boldsymbol{T}_0 = \dfrac{1}{K+1}\left[(\boldsymbol{z}-\boldsymbol{q})(\boldsymbol{z}-\boldsymbol{q})^\dagger + \displaystyle\sum_{t=1}^{K}\boldsymbol{z}_t\boldsymbol{z}_t^\dagger\right] \\ \boldsymbol{T}_1 = \dfrac{1}{K+1}\left[(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})^\dagger + \displaystyle\sum_{t=1}^{K}\boldsymbol{z}_t\boldsymbol{z}_t^\dagger\right] \end{cases}$$

Substituting (9) and (10) in (8) and maximizing the numerator and the denominator over $\boldsymbol{M}$, after some algebraic manipulations, we can recast the GLRT as

$$\frac{\min_{\boldsymbol{q}\in C_q} \det\left[(\boldsymbol{z}-\boldsymbol{q})(\boldsymbol{z}-\boldsymbol{q})^\dagger + \sum_{t=1}^K \boldsymbol{z}_t\boldsymbol{z}_t^\dagger\right]}{\min_{(\boldsymbol{p},\boldsymbol{q})\in C_p\times C_q} \det\left[(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})^\dagger + \sum_{t=1}^K \boldsymbol{z}_t\boldsymbol{z}_t^\dagger\right]} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \tag{11}$$

where G denotes a suitable modification of the original threshold in (8).

Exploiting the equality [26, p. 594 Corollary A.3.1] the above expression can be also written as

$$\frac{1+\min\limits_{\boldsymbol{q}\in C_q}(\boldsymbol{z}-\boldsymbol{q})^\dagger\boldsymbol{S}^{-1}(\boldsymbol{z}-\boldsymbol{q})}{1+\min\limits_{(\boldsymbol{p},\boldsymbol{q})\in C_p\times C_q}(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})^\dagger\boldsymbol{S}^{-1}(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \tag{12}$$

where $\boldsymbol{S}=\sum_{t=1}^K \boldsymbol{z}_t\boldsymbol{z}_t^\dagger$ is K times the secondary data sample covariance matrix.

The minimization at the numerator can be performed exploiting the results in [20]. Precisely the optimal value of the numerator can be evaluated as follows

- Denote by $\boldsymbol{U}\boldsymbol{\Xi}\boldsymbol{U}^\dagger$ the eigenvalue decomposition of $\boldsymbol{S}^{\frac{1}{2}}\frac{\boldsymbol{R}_q}{1+\gamma_q^2}\boldsymbol{S}^{\frac{1}{2}}$, where the eigenvalues are arranged in decreasing order, i.e. $\xi_1>\xi_2>\ldots\xi_{N-1}>0>\xi_N$, and let $\boldsymbol{x}=\boldsymbol{U}^\dagger\boldsymbol{S}^{-\frac{1}{2}}\boldsymbol{z}$.
- Evaluate the unique solution $\hat{\lambda}$, in the interval $[0,-1/\xi_N]$, of the equation

$$\sum_{k=1}^N \frac{\xi_k|x_k|^2}{(1+\lambda\xi_k)^2}=0\,.$$

- The optimal value of the numerator is given by

$$1+\boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\boldsymbol{z}-\boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\left(\boldsymbol{S}^{-1}+\hat{\lambda}\frac{\boldsymbol{R}_q}{1+\gamma_q^2}\right)^{-1}\boldsymbol{S}^{-1}\boldsymbol{z}\,. \tag{13}$$

The procedure required to minimize the denominator is shown in the next subsection.

A. Denominator Minimization and ML estimators of $\boldsymbol{p}$ *and* $\boldsymbol{q}$

We are faced with the following non-convex quadratic optimization problem

$$\begin{cases} \min_{(\boldsymbol{p},\boldsymbol{q})} & (\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})^\dagger\boldsymbol{S}^{-1}(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q}) \\ \text{s.t.} & \boldsymbol{p}^\dagger\boldsymbol{R}_p\boldsymbol{p}\le 0 \\ & \boldsymbol{q}^\dagger\boldsymbol{R}_q\boldsymbol{q}\le 0. \end{cases} \tag{14}$$

In this subsection we prove that solving the above non-convex problem is tantamount to solving a convex SDP problem. Otherwise stated problem (14) possesses a so-called hidden convexity property. To this end, it is not hard to rewrite (14) into the equivalent matrix form

$$(\mathrm{QP})\begin{cases} \min_{(\boldsymbol{p},\boldsymbol{q})} & [1,\boldsymbol{p}^\dagger,\boldsymbol{q}^\dagger]\begin{bmatrix} \boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\boldsymbol{z} & -\boldsymbol{z}^\dagger\boldsymbol{S}^{-1} & -\boldsymbol{z}^\dagger\boldsymbol{S}^{-1} \\ -\boldsymbol{S}^{-1}\boldsymbol{z} & \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1} \\ -\boldsymbol{S}^{-1}\boldsymbol{z} & \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1}\end{bmatrix}\begin{bmatrix} 1 \\ \boldsymbol{p} \\ \boldsymbol{q}\end{bmatrix} \\ \text{s.t.} & [1,\boldsymbol{p}^\dagger,\boldsymbol{q}^\dagger]\begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix} 1 \\ \boldsymbol{p} \\ \boldsymbol{q}\end{bmatrix} \leq 0 \\ & [1,\boldsymbol{p}^\dagger,\boldsymbol{q}^\dagger]\begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{R}_q\end{bmatrix}\begin{bmatrix} 1 \\ \boldsymbol{p} \\ \boldsymbol{q}\end{bmatrix} \leq 0 \end{cases} \tag{15}$$

where $\mathbf{0}$ denotes a matrix of suitable size with all zero entries.

Moreover the optimal value of problem (15) equals that of its homogenized version

$$(\mathrm{HQP})\begin{cases} \min_{(r,\boldsymbol{p},\boldsymbol{q})} & [r^\dagger,\boldsymbol{p}^\dagger,\boldsymbol{q}^\dagger]\begin{bmatrix} \boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\boldsymbol{z} & -\boldsymbol{z}^\dagger\boldsymbol{S}^{-1} & -\boldsymbol{z}^\dagger\boldsymbol{S}^{-1} \\ -\boldsymbol{S}^{-1}\boldsymbol{z} & \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1} \\ -\boldsymbol{S}^{-1}\boldsymbol{z} & \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1}\end{bmatrix}\begin{bmatrix} r \\ \boldsymbol{p} \\ \boldsymbol{q}\end{bmatrix} \\ \text{s.t.} & [r^\dagger,\boldsymbol{p}^\dagger,\boldsymbol{q}^\dagger]\begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix} r \\ \boldsymbol{p} \\ \boldsymbol{q}\end{bmatrix} \leq 0 \\ & [r^\dagger,\boldsymbol{p}^\dagger,\boldsymbol{q}^\dagger]\begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{R}_q\end{bmatrix}\begin{bmatrix} r \\ \boldsymbol{p} \\ \boldsymbol{q}\end{bmatrix} \leq 0 \\ & [r^\dagger,\boldsymbol{p}^\dagger,\boldsymbol{q}^\dagger]\begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix} r \\ \boldsymbol{p} \\ \boldsymbol{q}\end{bmatrix} = 1 \end{cases} \tag{16}$$

In fact, assuming that $[r,\boldsymbol{p}^T,\boldsymbol{q}^T]^T$ (the operator $(\cdot)^T$ denotes transpose) is an optimizer of (HQP), then $[1,r^\dagger\boldsymbol{p}^T,r^\dagger\boldsymbol{q}^T]^T$ is a feasible solution of (QP). Since the optimal value of (QP) is greater than or equal to that of (HQP), and the objective function of (HQP), evaluated at $[r,\boldsymbol{p}^T,\boldsymbol{q}^T]^T$, is equal to the objective function of (QP) at $[1, r^\dagger \boldsymbol{p}^T, r^\dagger \boldsymbol{q}^T]^T$, it follows that $[1, r^\dagger \boldsymbol{p}^T, r^\dagger \boldsymbol{q}^T]^T$ is an optimal solution for (QP).

Denoting by,

$$\boldsymbol{x} = \begin{bmatrix} r \\ \boldsymbol{p} \\ \boldsymbol{q} \end{bmatrix}, \boldsymbol{Q} = \begin{bmatrix} \boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} & -\boldsymbol{z}^\dagger \boldsymbol{S}^{-1} & -\boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \\ -\boldsymbol{S}^{-1}\boldsymbol{z} & \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1} \\ -\boldsymbol{S}^{-1}\boldsymbol{z} & \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1} \end{bmatrix},$$

and

$$\boldsymbol{Q}_0 = \begin{bmatrix} 0 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_p & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \boldsymbol{Q}_1 = \begin{bmatrix} 0 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_q \end{bmatrix}, \boldsymbol{Q}_2 = \begin{bmatrix} 1 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix},$$

the problem (HQP) can be recast as

$$(\text{HQP}) \begin{cases} \min_{\boldsymbol{x}} & \boldsymbol{x}^\dagger \boldsymbol{Q} \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{x}^\dagger \boldsymbol{Q}_0 \boldsymbol{x} \leq 0 \\ & \boldsymbol{x}^\dagger \boldsymbol{Q}_1 \boldsymbol{x} \leq 0 \\ & \boldsymbol{x}^\dagger \boldsymbol{Q}_2 \boldsymbol{x} = 1. \end{cases} \tag{17}$$

Now we claim that the SDP relaxation of (17) is tight, namely, the optimal value of (HQP) coincides with the optimal value of its SDP relaxation. In other words, this last problem has a rank-one optimal solution. To prove the statement we consider the SDP relaxation of (HQP) and its dual problem, respectively given by

$$(\text{SDP}) \begin{cases} \min_{\boldsymbol{X}} & \text{tr}(\boldsymbol{Q}\boldsymbol{X}) \\ \text{s.t.} & \text{tr}(\boldsymbol{Q}_0\boldsymbol{X}) \leq 0 \\ & \text{tr}(\boldsymbol{Q}_1\boldsymbol{X}) \leq 0 \\ & \text{tr}(\boldsymbol{Q}_2\boldsymbol{X}) = 1, \\ & \boldsymbol{X} \succeq 0, \end{cases} \tag{18}$$

and

$$(\text{DSDP}) \begin{cases} \max_{y_0, y_1, y_2} & y_2 \\ & y_0 \boldsymbol{Q}_0 + y_1 \boldsymbol{Q}_1 + y_2 \boldsymbol{Q}_2 - \boldsymbol{Q} \preceq 0, \\ & y_0 \leq 0, y_1 \leq 0. \end{cases} \tag{19}$$

Notice that (SDP) problem has an interior point in its feasible region because $[1, \boldsymbol{u}_p^T, \boldsymbol{u}_q^T]^T$ is a strictly feasible solution of (QP) (see Appendix A). Therefore the optimal values of both (SDP) and (DSDP) are equal to each others and the optimal value of (DSDP) is attainable [27, p. 423]. However, it is not clear that there is an interior point in the dual feasible region, thus the attainability of the optimal value of problem (SDP) cannot be concluded by dual theory of SDP. Nonetheless, we can prove the following lemma which belongs to a Frank-Wolfe type lemma for SDP (see [28] and references therein).

Lemma I. The optimal value of (SDP) is attainable.

The proof is given in Appendix B. However, before proceeding, it is helpful to recall the rank-one decomposition theorem of [24], which has been extensively used for the proof of Lemma I, and which will be also exploited for constructing a rank-one optimal solution of (SDP).

Rank-one Decomposition Theorem. Suppose that $\boldsymbol{X}$ is a $N \times N$ complex Hermitian positive semidefinite matrix of rank R, and $\boldsymbol{A}, \boldsymbol{B}$ are two $N \times N$ given Hermitian matrices. There is a rank-one decomposition of $\boldsymbol{X}$ (synthetically denoted as $\mathcal{D}(\boldsymbol{X}, \boldsymbol{A}, \boldsymbol{B})$), i.e. $\boldsymbol{X} = \sum_{s=1}^{R} \boldsymbol{x}_s \boldsymbol{x}_s^\dagger$, such that

$$\boldsymbol{x}_s^\dagger \boldsymbol{A} \boldsymbol{x}_s = \frac{\text{tr}(\boldsymbol{X}\boldsymbol{A})}{R} \text{ and } \boldsymbol{x}_s^\dagger \boldsymbol{B} \boldsymbol{x}_s = \frac{\text{tr}(\boldsymbol{X}\boldsymbol{B})}{R}, \quad \forall s = 1, \ldots, R. \tag{20}$$

In the last part of this sub-section we derive the ML estimators fo the useful and interfering signature. To this end, denote by $\bar{\boldsymbol{X}}, \bar{y_0}, \bar{y_1}, \bar{y_2}$ optimal solutions for (SDP) and (DSDP), and let $\bar{\boldsymbol{Z}} = \bar{y_0}\boldsymbol{Q}_0 + \bar{y_1}\boldsymbol{Q}_1 + \bar{y_2}\boldsymbol{Q}_2 - \boldsymbol{Q}$. Then, by complementarity conditions, it follows that

$$\text{tr}\left[\left(\bar{y_0}\boldsymbol{Q}_0 + \bar{y_1}\boldsymbol{Q}_1 + \bar{y_2}\boldsymbol{Q}_2 - \boldsymbol{Q}\right)\bar{\boldsymbol{X}}\right] = 0 \tag{21}$$

$$\left[\text{tr}(\boldsymbol{Q}_0\bar{\boldsymbol{X}})\right]\bar{y_0} = 0 \tag{22}$$

$$\left[\text{tr}(\boldsymbol{Q}_1\bar{\boldsymbol{X}})\right]\bar{y_1} = 0 \tag{23}$$

Now we show how to construct a rank-one optimal solution of problem (SDP) from $\bar{\boldsymbol{X}}$. We have to consider the four possible cases

1) $\bar{y_0} = 0$ and $\bar{y_1} = 0$
2) $\bar{y_0} < 0$ and $\bar{y_1} < 0$
3) $\bar{y_0} < 0$ and $\bar{y_1} = 0$
4) $\bar{y_0} = 0$ and $\bar{y_1} < 0$

Case 1: According to the aforementioned theorem there exists a rank-one decomposition of

$\bar{\boldsymbol{X}}$, i.e. $\bar{\boldsymbol{X}} = \sum_{s=1}^{R} \boldsymbol{x}_s \boldsymbol{x}_s^{\dagger}$, such that

$$\boldsymbol{x}_s^{\dagger}(\boldsymbol{Q}_0 - \boldsymbol{Q}_2)\boldsymbol{x}_s = \frac{\text{tr}[(\boldsymbol{Q}_0 - \boldsymbol{Q}_2)\bar{\boldsymbol{X}}]}{R} \leq -\frac{1}{R} \tag{24}$$

and

$$\boldsymbol{x}_s^{\dagger}(\boldsymbol{Q}_1 - \boldsymbol{Q}_2)\boldsymbol{x}_s = \frac{\text{tr}[(\boldsymbol{Q}_1 - \boldsymbol{Q}_2)\bar{\boldsymbol{X}}]}{R} \leq -\frac{1}{R}, \tag{25}$$

for all $s = 1, \ldots, R$, where R is the rank of $\bar{\boldsymbol{X}}$. Moreover, since $\text{tr}(\boldsymbol{Q}_2 \bar{\boldsymbol{X}}) = \sum_{s=1}^{R} \boldsymbol{x}_s^{\dagger} \boldsymbol{Q}_2 \boldsymbol{x}_s = 1$ and $\boldsymbol{Q}_2 \succeq 0$, there is an $s_0 \in \{1, \ldots, R\}$, without loss of generality suppose $s_0 = 1$, such that

$$0 < \boldsymbol{x}_1^{\dagger} \boldsymbol{Q}_2 \boldsymbol{x}_1 \leq \frac{1}{R}. \tag{26}$$

Let

$$\bar{\boldsymbol{x}} = \frac{\boldsymbol{x}_1}{\sqrt{\boldsymbol{x}_1^{\dagger} \boldsymbol{Q}_2 \boldsymbol{x}_1}}\,,$$

and observe that $\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^{\dagger}$ is a feasible solution of (SDP). Now let us check that $\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^{\dagger}$ complies with the complementarity conditions (21) - (23). Clearly, (22) and (23) hold due to $\bar{y}_0 = \bar{y}_1 = 0$. As to (21), since

$$0 \geq \boldsymbol{x}_1^{\dagger} \bar{\boldsymbol{Z}} \boldsymbol{x}_1 \geq \boldsymbol{x}_1^{\dagger} \bar{\boldsymbol{Z}} \boldsymbol{x}_1 + \sum_{s=2}^{R} \boldsymbol{x}_s^{\dagger} \bar{\boldsymbol{Z}} \boldsymbol{x}_s = \text{tr}(\bar{\boldsymbol{Z}} \bar{\boldsymbol{X}}) = 0\,,$$

then $\text{tr}(\bar{\boldsymbol{Z}} \boldsymbol{x}_1 \boldsymbol{x}_1^{\dagger}) = 0$ and, as a consequence, $\text{tr}(\bar{\boldsymbol{Z}} \bar{\boldsymbol{x}} \bar{\boldsymbol{x}}^{\dagger}) = 0$. Therefore $\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^{\dagger}$ is a feasible solution of (SDP) and satisfies the complementary conditions (21) - (23). Thus it is an optimal solution of (SDP).

Case 2: In this case, $\text{tr}(\boldsymbol{Q}_0 \bar{\boldsymbol{X}}) = \text{tr}(\boldsymbol{Q}_1 \bar{\boldsymbol{X}}) = 0$. Applying the rank-one decomposition theorem we get a decomposition of $\bar{\boldsymbol{X}}$, namely $\bar{\boldsymbol{X}} = \sum_{s=1}^{R} \boldsymbol{x}_s \boldsymbol{x}_s^{\dagger}$, such that

$$\boldsymbol{x}_s^{\dagger} \boldsymbol{Q}_0 \boldsymbol{x}_s = 0, \text{ and } \boldsymbol{x}_s^{\dagger} \boldsymbol{Q}_1 \boldsymbol{x}_s = 0\,, \;\; \forall s = 1, \ldots, R.$$

Suppose, without loss of generality, that $0 < \boldsymbol{x}_1^{\dagger} \boldsymbol{Q}_2 \boldsymbol{x}_1 \leq \frac{1}{R}$, and set $\bar{\boldsymbol{x}} = \frac{\boldsymbol{x}_1}{\sqrt{\boldsymbol{x}_1^{\dagger} \boldsymbol{Q}_2 \boldsymbol{x}_1}}$. It is easily verified that $\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^{\dagger}$ fulfills both feasibility and complementarity, and thus it is an optimal solution of (SDP).

Case 3: In this case, $\text{tr}(\boldsymbol{Q}_0 \bar{\boldsymbol{X}}) = 0$ and $\text{tr}(\boldsymbol{Q}_1 \bar{\boldsymbol{X}}) \leq 0$. By the rank-one decomposition theorem, it follows that there is a rank-one decomposition of $\bar{\boldsymbol{X}}$, i.e. $\bar{\boldsymbol{X}} = \sum_{s=1}^{R} \boldsymbol{x}_s \boldsymbol{x}_s^{\dagger}$, such that

$$\boldsymbol{x}_s^{\dagger} \boldsymbol{Q}_0 \boldsymbol{x}_s = 0, \text{ and } \boldsymbol{x}_s^{\dagger}(\boldsymbol{Q}_1 - \boldsymbol{Q}_2)\boldsymbol{x}_s = \frac{\text{tr}[(\boldsymbol{Q}_1 - \boldsymbol{Q}_2)\bar{\boldsymbol{X}}]}{R} \leq -\frac{1}{R}\,, \;\; \forall s = 1, \ldots, R.$$

Again, suppose that $0 < \boldsymbol{x}_1^\dagger \boldsymbol{Q}_2 \boldsymbol{x}_1 \leq \frac{1}{R}$, and set $\bar{\boldsymbol{x}} = \frac{\boldsymbol{x}_1}{\sqrt{\boldsymbol{x}_1^\dagger \boldsymbol{Q}_2 \boldsymbol{x}_1}}$. Then $\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\dagger$ complies with $\text{tr}(\boldsymbol{Q}_0 \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\dagger) = 0$, $\text{tr}(\boldsymbol{Q}_1 \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\dagger) \leq 0$, and $\text{tr}(\boldsymbol{Q}_2 \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\dagger) = 1$. Moreover it also satisfies the complementarity conditions (21) - (23). Hence $\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\dagger$ is an optimal solution of (SDP).

Case 4: In this case, $\text{tr}(\boldsymbol{Q}_0 \bar{\boldsymbol{X}}) \leq 0$ and $\text{tr}(\boldsymbol{Q}_1 \bar{\boldsymbol{X}}) = 0$. We apply the rank-one decomposition theorem and get a decomposition $\bar{\boldsymbol{X}} = \sum_{s=1}^{R} \boldsymbol{x}_s \boldsymbol{x}_s^\dagger$ such that

$$\boldsymbol{x}_s^\dagger (\boldsymbol{Q}_0 - \boldsymbol{Q}_2) \boldsymbol{x}_s = \frac{\text{tr}[(\boldsymbol{Q}_0 - \boldsymbol{Q}_2)\boldsymbol{X}]}{R} \leq -\frac{1}{R} \text{ and } \boldsymbol{x}_s^\dagger \boldsymbol{Q}_1 \boldsymbol{x}_s = 0\,, \;\; \forall s = 1, \ldots, R.$$

Now, suppose that $0 < \boldsymbol{x}_1^\dagger \boldsymbol{Q}_2 \boldsymbol{x}_1 \leq \frac{1}{R}$, and set $\bar{\boldsymbol{x}} = \frac{\boldsymbol{x}_1}{\sqrt{\boldsymbol{x}_1^\dagger \boldsymbol{Q}_2 \boldsymbol{x}_1}}$. Working as in Case 3, it can be shown that $\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^\dagger$ is an optimal solution of (SDP).

Summarizing, in all the cases, we have proved that problem (SDP) has a rank-one optimal solution. It follows that the SDP relaxation is tight, namely, if $v(\cdot)$ is the optimal value of the problem ($\cdot$), then $v(\text{QP}) = v(\text{HQP}) = v(\text{SDP})$. Moreover an optimal solution of (14), i.e. the ML estimators of $\boldsymbol{p}$ and $\boldsymbol{q}$ (denoted by $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$ respectively) can be obtained as follows

Computation of the ML Steering Estimators

- Denote by $\bar{\boldsymbol{X}}$ (with rank R), $\bar{y}_0, \bar{y}_1, \bar{y}_2$ optimal solutions for (SDP) and (DSDP).
- if $\bar{y}_0 = 0$ and $\bar{y}_1 = 0$ then evaluate $\mathcal{D}(\bar{\boldsymbol{X}}, \boldsymbol{Q}_0 - \boldsymbol{Q}_2, \boldsymbol{Q}_1 - \boldsymbol{Q}_2)$;
- if $\bar{y}_0 < 0$ and $\bar{y}_1 < 0$ then evaluate $\mathcal{D}(\bar{\boldsymbol{X}}, \boldsymbol{Q}_0, \boldsymbol{Q}_1)$;
- if $\bar{y}_0 < 0$ and $\bar{y}_1 = 0$ then evaluate $\mathcal{D}(\bar{\boldsymbol{X}}, \boldsymbol{Q}_0, \boldsymbol{Q}_1 - \boldsymbol{Q}_2)$;
- if $\bar{y}_0 = 0$ and $\bar{y}_1 < 0$ then evaluate $\mathcal{D}(\bar{\boldsymbol{X}}, \boldsymbol{Q}_0 - \boldsymbol{Q}_2, \boldsymbol{Q}_1)$;
- Determine a vector of the decomposition $\boldsymbol{x}_s$, $s \in \{1, \ldots, R\}$, such that $0 < \boldsymbol{x}_s^\dagger \boldsymbol{Q}_2 \boldsymbol{x}_s \leq 1/R$. The following equalities provide the vectors $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$

$$\begin{aligned} [\tilde{r}, \tilde{\boldsymbol{p}}^T, \tilde{\boldsymbol{q}}^T]^T &= \frac{\boldsymbol{x}_s}{\sqrt{\boldsymbol{x}_s^\dagger \boldsymbol{Q}_2 \boldsymbol{x}_s}}\,, \\ \hat{\boldsymbol{p}} &= \tilde{r}^\dagger \tilde{\boldsymbol{p}}\,, \\ \hat{\boldsymbol{q}} &= \tilde{r}^\dagger \tilde{\boldsymbol{q}}\,. \end{aligned}$$

We explicitly point out that the equivalence between (14) and an SDP problem can be also shown exploiting Theorem 2.2 in [29]. However our approach, in addition to the equivalence, also shows the achievability of the optimal value of (SDP) and provides a procedure for evaluating the ML estimators of the steering vectors $\boldsymbol{p}$ and $\boldsymbol{q}$, namely an optimal solution of (14). These optimal solutions will be explicitly used in the design of the modified GLRT detectors, which we will discuss in Section V.

Exploiting the aforementioned findings we can express the R-1S-GLRT as

$$\frac{1}{1+v(\mathrm{SDP})}\left[1+\boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\boldsymbol{z}-\boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\left(\boldsymbol{S}^{-1}+\hat{\lambda}\frac{\boldsymbol{R}_q}{1+\gamma_q^2}\right)^{-1}\boldsymbol{S}^{-1}\boldsymbol{z}\right]\begin{array}{c}H_1\\ >\\ <\\ H_0\end{array} G\,. \tag{27}$$

The computational complexity connected with the implementation of the decision statistic is $O(N^{3.5}\log(1/\delta)+KN^2)$. In fact, the amount of operations involved in solving (SDP) is $O(N^{3.5}\log(1/\delta))$ [27, p. 250][2], while the construction of the matrix $\boldsymbol{S}$ involves $O(KN^2)$. We finally notice that the computation of the decision statistic does not explicitly require the Maximum Likelihood (ML) estimates of $\boldsymbol{p}$ and $\boldsymbol{q}$ but only the optimal value of (14).

IV. Robust Two-Step GLRT (R-2S-GLRT)

This design technique [3] first assumes that the covariance matrix $\boldsymbol{M}$ is known and derives the GLRT with respect to $\boldsymbol{p}$ and $\boldsymbol{q}$. Then, after the GLRT is derived, the sample covariance matrix, based upon secondary data, is substituted, in place of the true covariance matrix, into the test.

For the present problem the primary data GLRT can be written as

$$\frac{\max_{(\boldsymbol{p},\boldsymbol{q})\in C_p\times C_q} f(\boldsymbol{z}|\boldsymbol{p},\boldsymbol{q},\boldsymbol{M},H_1)}{\max_{\boldsymbol{q}\in C_q} f(\boldsymbol{z}|\boldsymbol{q},\boldsymbol{M},H_0)}\begin{array}{c}H_1\\ >\\ <\\ H_0\end{array} G\,, \tag{28}$$

where $f(\boldsymbol{z}|\boldsymbol{q},\boldsymbol{M},H_0)$ and $f(\boldsymbol{z}|\boldsymbol{p},\boldsymbol{q},\boldsymbol{M},H_1)$ denote the primary data complex multivariate pdf's under the H_0 and the H_1 hypotheses, i.e.

$$f(\boldsymbol{z}|\boldsymbol{q},\boldsymbol{M},H_0)=\frac{1}{\pi^N\det(\boldsymbol{M})}e^{-\mathrm{tr}\left[\boldsymbol{M}^{-1}(\boldsymbol{z}-\boldsymbol{q})(\boldsymbol{z}-\boldsymbol{q})^\dagger\right]}\,, \tag{29}$$

and

$$f(\boldsymbol{z}|\boldsymbol{p},\boldsymbol{q},\boldsymbol{M},H_1)=\frac{1}{\pi^N\det(\boldsymbol{M})}e^{-\mathrm{tr}\left[\boldsymbol{M}^{-1}(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})^\dagger\right]}\,. \tag{30}$$

[2]An SDP problem can be efficiently solved in polynomial time through *interior point methods*, namely iterative algorithms which terminate once a pre-specified accuracy δ is reached [27].

Substituting (29) and (30) in (28), the R-2S-GLRT becomes

$$\min_{\boldsymbol{q}\in C_q} (\boldsymbol{z}-\boldsymbol{q})^\dagger \boldsymbol{M}^{-1}(\boldsymbol{z}-\boldsymbol{q}) - \min_{(\boldsymbol{p},\boldsymbol{q})\in C_p\times C_q} (\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})^\dagger \boldsymbol{M}^{-1}(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q}) \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \tag{31}$$

where G is a suitable modification of the threshold in (28). Performing the minimizations in (31), and plugging $\boldsymbol{S}/K$ in place of $\boldsymbol{M}$, the R-2S-GLRT can be finally written as

$$\boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\left(\boldsymbol{S}^{-1} + \hat{\lambda}\frac{\boldsymbol{R}_q}{1+\gamma_q^2}\right)^{-1}\boldsymbol{S}^{-1}\boldsymbol{z} - v(\text{SDP}) \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \tag{32}$$

where the same symbol has been used to denote the modified threshold. As in the previous case, the computational complexity connected with the implementation of the decision statistic is $O(N^{3.5}\log(1/\delta) + KN^2)$.

V. Modified R-1S-GLRT (MR-1S-GLRT) and Modified R-2S-GLRT (MR-2S-GLRT)

These receivers are designed according to the following criteria. First the one-step GLRT and the two-step GLRT are derived assuming that the directions of the useful signal and of the interference are known. Then after the GLRT's have been derived, the ML estimates of the aforementioned directions, derived under the H_1 hypothesis, are substituted in place of their exact values.

Assuming known signal and interference directions the one-step GLRT and the two-step GLRT are respectively given by [22]

$$\frac{1+\boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{q}(\boldsymbol{q}^\dagger \boldsymbol{S}^{-1}\boldsymbol{q})^{-1}\boldsymbol{q}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}}{1+\boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{W}(\boldsymbol{W}^\dagger \boldsymbol{S}^{-1}\boldsymbol{W})^{-1}\boldsymbol{W}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \tag{33}$$

and

$$\boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{W}(\boldsymbol{W}^\dagger \boldsymbol{S}^{-1}\boldsymbol{W})^{-1}\boldsymbol{W}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{q}(\boldsymbol{q}^\dagger \boldsymbol{S}^{-1}\boldsymbol{q})^{-1}\boldsymbol{q}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \tag{34}$$

where $\boldsymbol{W} = [\boldsymbol{p}, \boldsymbol{q}]$. Then substituting an optimal solution of (14), denoted by $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$, in place of $\boldsymbol{p}$ and $\boldsymbol{q}$ in (33) and (34), we get the MR-1S-GLRT and the MR-2S-GLRT, i.e.

$$\frac{1+\boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\hat{\boldsymbol{q}}(\hat{\boldsymbol{q}}^\dagger \boldsymbol{S}^{-1}\hat{\boldsymbol{q}})^{-1}\hat{\boldsymbol{q}}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}}{1+\boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1}\hat{\boldsymbol{W}}(\hat{\boldsymbol{W}}^\dagger \boldsymbol{S}^{-1}\hat{\boldsymbol{W}})^{-1}\hat{\boldsymbol{W}}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \qquad \text{MR-1S-GLRT} \tag{35}$$

and

$$\boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \hat{\boldsymbol{W}} (\hat{\boldsymbol{W}}^\dagger \boldsymbol{S}^{-1} \hat{\boldsymbol{W}})^{-1} \hat{\boldsymbol{W}}^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \hat{\boldsymbol{q}} (\hat{\boldsymbol{q}}^\dagger \boldsymbol{S}^{-1} \hat{\boldsymbol{q}})^{-1} \hat{\boldsymbol{q}}^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \qquad \text{MR-2S-GLRT} \tag{36}$$

where $\hat{\boldsymbol{W}} = [\hat{\boldsymbol{p}}, \hat{\boldsymbol{q}}]$.

Again the computational complexity connected with the implementation of the decision statistics is $O(N^{3.5}\log(1/\delta) + KN^2)$. However in this last case the ML estimate is explicitly required and thus it is necessary to implement the rank-one decomposition of [24] which exhibits a computational complexity $O(N^3)$.

VI. Performance Analysis

This section is devoted to the performance analysis of the proposed receivers also in comparison with the GLRT-based detectors, derived in [22], which assume both the target and the interference perfectly aligned with their nominal steering directions (modeled as linearly independent complex vectors), i.e.

$$\frac{1 + \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{u}_q (\boldsymbol{u}_q^\dagger \boldsymbol{S}^{-1} \boldsymbol{u}_q)^{-1} \boldsymbol{u}_q^\dagger \boldsymbol{S}^{-1} \boldsymbol{z}}{1 + \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{H} (\boldsymbol{H}^\dagger \boldsymbol{S}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^\dagger \boldsymbol{S}^{-1} \boldsymbol{z}} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \qquad \text{1S-GLRT}\,, \tag{37}$$

and

$$\boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{H} (\boldsymbol{H}^\dagger \boldsymbol{S}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{u}_q (\boldsymbol{u}_q^\dagger \boldsymbol{S}^{-1} \boldsymbol{u}_q)^{-1} \boldsymbol{u}_q^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \qquad \text{2S-GLRT}\,, \tag{38}$$

where $H = [\boldsymbol{u}_p, \boldsymbol{u}_q]$. Moreover we also compare the new detectors with the Unknown $\boldsymbol{p}$-GLRT (U-GLRT) receiver (derived in Appendix C), which models $\boldsymbol{p}$ as a completely unknown N-dimensional complex vector and assumes $\boldsymbol{q}$ perfectly aligned with its nominal direction, i.e.

$$l_{U-GLRT} = \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} - \boldsymbol{z}^\dagger \boldsymbol{S}^{-1} \boldsymbol{u}_q (\boldsymbol{u}_q^\dagger \boldsymbol{S}^{-1} \boldsymbol{u}_q)^{-1} \boldsymbol{u}_q^\dagger \boldsymbol{S}^{-1} \boldsymbol{z} \underset{H_0}{\overset{H_1}{\gtrless}} G\,, \qquad \text{U-GLRT}. \tag{39}$$

Notice that none of the devised and analyzed receivers ensure the CFAR property with respect to the disturbance covariance matrix when mismatches, both in $\boldsymbol{p}$ and $\boldsymbol{q}$, are present. Actually, to our knowledge, there is not a receiver, in open literature, exhibiting a CFAR behavior in the considered mismatched environment.

The procedure for the construction of the two conic regions, complying with (7), around the nominal useful and interfering directions is now given (see also Figure 1). Denoting by

$$\theta = \arccos(|\boldsymbol{u}_p^\dagger \boldsymbol{u}_q|)\,,$$

condition (7) holds if $\gamma_p = \tan(\theta_p)$ and $\gamma_q = \tan(\theta_q)$, where the angles θ_p and θ_q comply with

$$\theta_p = \frac{\theta}{h_p} \qquad h_p > 2$$

$$\theta_q = \frac{\theta}{h_q} \qquad h_q > 2$$

(a sufficient condition could also be $1/h_p + 1/h_q < 1$).

Since closed form expressions for the false alarm Probability (P_{fa}) and the detection Probability (P_d) are not available, we extensively resort to Monte Carlo techniques and, in order to limit the simulation time, we assume $P_{fa} = 10^{-2}$. We also set $N = 8$, cosiders Gaussian disturbance whose covariance matrix is exponentially shaped with one-lag correlation coefficient $\rho = 0.9$, i.e.

$$\boldsymbol{M}_{l,t} = \rho^{|l-t|}\,, \qquad (l,t) \in \{1,\ldots,N\}^2\,,$$

and exploit the convex optimization MATLAB toolbox SeDuMi [30] for solving the SDP problems. Finally

$$\boldsymbol{u}_p = \frac{1}{N}[1,1,\ldots,]^T\,,$$

$$\boldsymbol{u}_q = \frac{1}{N}[1,\exp(j\pi/3),\ldots,\exp(j(N-1)\pi/3)]^T\,,$$

j is the imaginary unit, the angle between $\boldsymbol{p}$ and $\boldsymbol{u}_p$ is denoted by $\theta_T = \arccos(|\boldsymbol{u}_p^\dagger \boldsymbol{p}|)$, whereas $\theta_I = \arccos(|\boldsymbol{u}_q^\dagger \boldsymbol{q}|)$ represents the angle between $\boldsymbol{q}$ and $\boldsymbol{u}_q$.

In Figure 2a, the detection performance of the considered receivers is plotted versus the Signal to Noise Ratio (SNR), i.e.

$$\text{SNR} = \boldsymbol{p}^\dagger \boldsymbol{M}^{-1} \boldsymbol{p}\,,$$

for $K = 16$, $h_p = h_q = 10$, $\theta_T = \theta_p$, θ_I uniformly distributed in $[0, \theta_q]$, and Interference to Noise Ratio (INR), defined as INR $= \boldsymbol{q}^\dagger \boldsymbol{M}^{-1} \boldsymbol{q}$, equal to 20 dB. The curves show that the new robust receivers ensure almost the same performance and outperform their counterparts. Moreover the higher P_d the higher the performance gain. As to the U-GLRT it exhibits, for

$P_d = 0.9$ a performance loss, kept within 3.5 dB, with respect to the robust receivers. However it achieves a better performance level than the GLRT's (37) and (38) for medium-high values of P_d.

A similar performance behavior is shown by the plots in Figure 2b, which assumes the same simulation parameters of Figure 2a, but for the value of K which is set to 32. It can be observed that increasing K there is a reduction in the performance gain of the robust recevers over the 1S-GLRT and 2S-GLRT, even if the aforementioned gain is still noticeable.

The case of perfect matched useful signal and interference is studied in Figures 3a and 3b. They refer to $\theta_T = \theta_I = 0$, $h_p = h_q = 10$, and INR= 20 dB. Moreover Figure 3a assumes $K = 16$ while Figure 3b considers $K = 32$. From Figure 3a, it is interesting to observe that the robust receivers still outperform detectors (37) and (38), even if there is a perfect matching condition. This is actually not a great surprise because a similar performance behavior was also noticed in [20] with reference to the case of useful signal mismatch only. Nevertheless, Figure 3b shows that, when the number of training data increases from 16 to 32, the aforementioned performance gain vanishes and the performance of the robust receivers end up coincident with that of (37) and (38). The U-GLRT performs worse than the other receivers; this was expected since it models the useful signal as a completely unknown vector.

The effect of the INR on the performance of the receivers is analyzed in Figures 4a and 4b which refer to $h_p = h_q = 10$, $K = 32$, $\theta_T = \theta_p$, and θ_I uniformly distributed in $[0, \theta_q]$. The former assumes INR=10 dB whereas the latter INR=30 dB. The figures highlight that, when a mismatch is present, increasing INR leads to higher and higher performance losses of the receivers (37), (38), and (39) with respect to the robust GLRT's. This performance trend has also an intuitive justification since higher values of INR lead to an increase of the interference energy outside the nominal direction.

Figure 5 studies the impact of the parameter h_p on the performance of the robust GLRT's. It considers the following values for the parameters INR= 20 dB, $N = 8$, $h_q = 10$, $K = 32$, $\theta_T = 0$, $\theta_I = 0$, and $h_p \in \{5, 7.5, 10\}$. This behavior can be justified observing that decreasing h_p, the size of the uncertainty region increases and, as a consequence, the performance degrades.

All the previous figures assume a constant θ_T. It is thus on interest to analyze what happens when this parameter varies randomly. This is done in Figure 6 which assumes INR= 20 dB, $N = 8$, $h_p = h_q = 10$, $K = 32$, θ_T uniformly distributed in $[0, \theta_q]$, and θ_I uniformly distributed in $[0, \theta_q]$. The curves highlight that the robust GLRT's still outperform their counterparts even if the performance gain over (37) and (38) is smaller than that in Figure 2b, which refers to $\theta_T = \theta_p$.

VII. Conclusions

In this paper we have considered the problem of detecting a signal vector in the presence of complex circular Gaussian disturbance with unknown covariance matrix plus deterministic interference. We have focused on mismatched scenarios where an energy fraction of both the useful signal and the deterministic interference may lie outside the range span of their nominal directions. In this context we have derived the joint ML estimators of the useful signal and the interference vector. The proposed procedure first relaxes the original non-convex problem into a convex SDP. Then, exploiting the rank-one decomposition theorem of [24], it constructs the ML estimators from an optimal solution of the relaxed problem. The quoted estimates are used to devise robust detectors based on different GLRT criteria (one-step, two-step, modified one-step, and modified two-step). All of them exhibit a polynomial computational complexity.

At the analysis stage we have evaluated the performance of the robust GLRT's, also comparing them with two decision rules already available in open literature which assume the signal and interference signatures perfectly aligned with their nominal directions. Extensive numerical results have shown that the new robust receivers may outperform their counterparts.

Possible future research tracks might concern the extension of the framework to the case of multiple primary observations, such as the detection of range spread targets. Moreover it might be of interest the analysis of the proposed receivers in the presence of real radar OTH data where the presented signal environment might be successfully applied.

VIII. Appendix

A. *Strict Feasibility of (SDP)*

Let $\boldsymbol{u}_1 = [1, \boldsymbol{u}_p^T, \boldsymbol{u}_q^T]^T$. It can be easily verified that $\boldsymbol{u}_1$ is a strictly feasible solution of (HQP) (or (QP)). Now we claim that for some sufficiently small $\epsilon > 0$, $\boldsymbol{X}(\epsilon) = \epsilon \boldsymbol{I} + (1-\epsilon)\boldsymbol{u}_1\boldsymbol{u}_1^\dagger$ is an interior point of (SDP). In fact, $\boldsymbol{X}(\epsilon) \succ 0$ for any $\epsilon \in (0, 1]$, tr$(Q_2\boldsymbol{X}(\epsilon)) = 1$ for any ϵ, and for some sufficiently small positive ϵ,

$$\begin{aligned} \text{tr}(Q_0\boldsymbol{X}(\epsilon)) &= \epsilon\ \text{tr}\boldsymbol{R}_p + (1-\epsilon)\boldsymbol{u}_p^\dagger \boldsymbol{R}_p \boldsymbol{u}_p < 0, \\ \text{tr}(Q_1\boldsymbol{X}(\epsilon)) &= \epsilon\ \text{tr}\boldsymbol{R}_q + (1-\epsilon)\boldsymbol{u}_q^\dagger \boldsymbol{R}_q \boldsymbol{u}_q < 0, \end{aligned}$$

due to $\boldsymbol{u}_p^\dagger \boldsymbol{R}_p \boldsymbol{u}_p < 0$ and $\boldsymbol{u}_q^\dagger \boldsymbol{R}_q \boldsymbol{u}_q < 0$.

B. Proof of Lemma I

Let us denote by $v^* = v(\text{SDP})$, $\mathcal{F} = \{\boldsymbol{X} : \text{tr}(\boldsymbol{Q}_0\boldsymbol{X}) \leq 0, \text{tr}(\boldsymbol{Q}_1\boldsymbol{X}) \leq 0, \text{tr}(\boldsymbol{Q}_2\boldsymbol{X}) = 1, \boldsymbol{X} \succeq 0\}$ the feasible set of (SDP). It is known that if a feasible SDP has a strictly feasible dual problem, then its optimal solution exists (see, e.g., page 72 of [27]). Therefore, $v^* = \min_{\boldsymbol{X}\in\mathcal{F}} \text{tr}(\boldsymbol{Q}\boldsymbol{X})$ due to the feasibility of (DSDP) and the strict feasibility of (SDP)[3]. Notice also that the feasible set $\mathcal{F}$ is closed. If $\mathcal{F}$ is bounded, then the optimal value of (SDP) is attainable because $\mathcal{F}$ is a compact set. If $\mathcal{F}$ is unbounded, then we claim that it is still attainable at some point $\boldsymbol{X} \in \mathcal{F}$. In fact, suppose that the optimal value is not attainable, then there exists a sequence $\{\boldsymbol{X}^k\}_{k=1}^\infty$ within $\mathcal{F}$ such that

$$\text{tr}(\boldsymbol{Q}\boldsymbol{X}^k) \longrightarrow v^*, \text{ and } ||\boldsymbol{X}^k||_F \longrightarrow \infty,$$

when k tends to infinity ($||\cdot||_F$ denotes the Frobenius norm of a matrix). This implies that the sequence $\frac{\boldsymbol{X}^k}{||\boldsymbol{X}^k||_F} \in \mathcal{G}$, where $\mathcal{G} = \{\boldsymbol{X} : ||\boldsymbol{X}||_F = 1\}$, complies with[4]

$$\frac{\boldsymbol{X}^k}{||\boldsymbol{X}^k||_F} \longrightarrow \boldsymbol{X}, \text{ and } \frac{\text{tr}(\boldsymbol{Q}\boldsymbol{X}^k)}{||\boldsymbol{X}^k||_F} = \text{tr}\left(\boldsymbol{Q}\frac{\boldsymbol{X}^k}{||\boldsymbol{X}^k||_F}\right) \longrightarrow 0 = \text{tr}(\boldsymbol{Q}\boldsymbol{X}) \text{ as } k \longrightarrow \infty,$$

and $\boldsymbol{X} \succeq 0$. Observe also that $(\boldsymbol{X}^k)_{11} = 1$ for all k, and, as a consequence, $\boldsymbol{X}_{11} = 0$ and $\boldsymbol{X}_{1j} = \boldsymbol{X}_{j1} = 0$ for all $j = 1, \ldots, N$, i.e.,

$$X = \begin{bmatrix} 0 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}_{22} & \boldsymbol{X}_{23} \\ \boldsymbol{0} & \boldsymbol{X}_{23}^\dagger & \boldsymbol{X}_{33} \end{bmatrix}.$$

Let

$$\hat{\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{X}_{22} & \boldsymbol{X}_{23} \\ \boldsymbol{X}_{23}^\dagger & \boldsymbol{X}_{33} \end{bmatrix}, \hat{\boldsymbol{Q}} = \begin{bmatrix} \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1} \\ \boldsymbol{S}^{-1} & \boldsymbol{S}^{-1} \end{bmatrix}, \hat{\boldsymbol{Q}}_0 = \begin{bmatrix} \boldsymbol{R}_p & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \hat{\boldsymbol{Q}}_1 = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_q \end{bmatrix}.$$

[3] We explicitly remark that if v^* is not attainable, then $v^* = \inf_{\boldsymbol{X}\in\mathcal{F}} \text{tr}(\boldsymbol{Q}\boldsymbol{X})$ (i.e., “min” becomes “inf”).

[4] The convergence of $\frac{\boldsymbol{X}^k}{||\boldsymbol{X}^k||_F}$ is ensured since $\mathcal{G}$ is a compact set.

Since tr$(\boldsymbol{X}^k\boldsymbol{Q}_0) \leq 0$ and tr$(\boldsymbol{X}^k\boldsymbol{Q}_1) \leq 0$, then tr$(\boldsymbol{X}\boldsymbol{Q}_0) \leq 0$ and tr$(\boldsymbol{X}\boldsymbol{Q}_1) \leq 0$, which amounts to tr$(\hat{\boldsymbol{X}}\hat{\boldsymbol{Q}}_0) \leq 0$ and tr$(\hat{\boldsymbol{X}}\hat{\boldsymbol{Q}}_1) \leq 0$. Moreover, since $\hat{\boldsymbol{X}} \succeq 0$, the rank-one decomposition theorem yields

$$\hat{\boldsymbol{X}} = \sum_{s=1}^{R} \begin{bmatrix} \boldsymbol{x}_s \\ \boldsymbol{y}_s \end{bmatrix} [\boldsymbol{x}_s^\dagger, \boldsymbol{y}_s^\dagger],$$

with

$$[\boldsymbol{x}_s^\dagger, \boldsymbol{y}_s^\dagger]\hat{\boldsymbol{Q}}_0 \begin{bmatrix} \boldsymbol{x}_s \\ \boldsymbol{y}_s \end{bmatrix} = \boldsymbol{x}_s^\dagger \boldsymbol{R}_p \boldsymbol{x}_s \leq 0,\ [\boldsymbol{x}_s^\dagger, \boldsymbol{y}_s^\dagger]\hat{\boldsymbol{Q}}_1 \begin{bmatrix} \boldsymbol{x}_s \\ \boldsymbol{y}_s \end{bmatrix} = \boldsymbol{y}_s^\dagger \boldsymbol{R}_q \boldsymbol{y}_s \leq 0, \forall s = 1, \ldots, R, \quad (40)$$

where R is the rank of $\hat{\boldsymbol{X}}$. Since tr$(\boldsymbol{Q}\boldsymbol{X}) =$ tr$(\hat{\boldsymbol{Q}}\hat{\boldsymbol{X}}) = 0$, $\hat{\boldsymbol{Q}} \succeq 0$ and $\hat{\boldsymbol{X}} \succeq 0$, then

$$[\boldsymbol{x}_s^\dagger, \boldsymbol{y}_s^\dagger]\hat{\boldsymbol{Q}} \begin{bmatrix} \boldsymbol{x}_s \\ \boldsymbol{y}_s \end{bmatrix} = 0, \qquad \forall s = 1, \ldots, R.$$

This further implies that

$$\hat{\boldsymbol{Q}} \begin{bmatrix} \boldsymbol{x}_s \\ \boldsymbol{y}_s \end{bmatrix} = \boldsymbol{0}, \qquad \forall s = 1, \ldots, R,$$

or equivalently,

$$\boldsymbol{S}^{-1}(\boldsymbol{x}_s + \boldsymbol{y}_s) = \boldsymbol{0}, \qquad \forall s = 1, \ldots, R.$$

Due to $\boldsymbol{S}^{-1} \succ 0$, it follows that $\boldsymbol{x}_s + \boldsymbol{y}_s = \boldsymbol{0}$, for all $s = 1, \ldots, R$. Then we have $\boldsymbol{x}_s = -\boldsymbol{y}_s$ and $\boldsymbol{x}_s \neq \boldsymbol{0}$ (otherwise $[\boldsymbol{x}_s^T, \boldsymbol{y}_s^T]^T = \boldsymbol{0}$), for all $s = 1, \ldots, R$. Combination with (40) leads to

$$\boldsymbol{x}_s^\dagger \boldsymbol{R}_p \boldsymbol{x}_s \leq 0, \qquad \boldsymbol{x}_s^\dagger \boldsymbol{R}_q \boldsymbol{x}_s \leq 0, \qquad \boldsymbol{x}_s \neq \boldsymbol{0}, \qquad \forall s = 1, \ldots, R.$$

However, this is in contradiction with the assumption (7). Therefore we can conclude that the optimal value of (SDP) is still attainable even when the feasible set is unbounded.

C. U-GLRT Derivation

Let us start with the one-step U-GLRT which is given by

$$\frac{\max_{(\boldsymbol{p}\in\mathcal{C}^N, \beta\in\mathcal{C})} \max_{\boldsymbol{M}} f(\boldsymbol{z}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_K | \boldsymbol{p}, \boldsymbol{q} = \beta \boldsymbol{u}_q, \boldsymbol{M}, H_1)}{\max_{\beta\in\mathcal{C}} \max_{\boldsymbol{M}} f(\boldsymbol{z}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_K | \boldsymbol{q} = \beta \boldsymbol{u}_q, \boldsymbol{M}, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} G \quad (41)$$

Substituting the pdf's, performing the maximization over $\boldsymbol{M}$, applying [26, p. 594 Corollary A.3.1], and after some algebra, we get

$$\frac{1+\min_{\beta\in\mathcal{C}}(\boldsymbol{z}-\beta\boldsymbol{u}_q)^\dagger\boldsymbol{S}^{-1}(\boldsymbol{z}-\beta\boldsymbol{u}_q)}{1+\min_{(\boldsymbol{p}\in\mathcal{C}^N,\beta\in\mathcal{C})}(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})^\dagger\boldsymbol{S}^{-1}(\boldsymbol{z}-\boldsymbol{p}-\boldsymbol{q})} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} G\,. \tag{42}$$

where G denotes a suitable modification of the original threshold in (41). Performing the remaining optimizations, we come up with the test

$$\boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\boldsymbol{z}-\boldsymbol{z}^\dagger\boldsymbol{S}^{-1}\boldsymbol{u}_q(\boldsymbol{u}_q^\dagger\boldsymbol{S}^{-1}\boldsymbol{u}_q)^{-1}\boldsymbol{u}_q^\dagger\boldsymbol{S}^{-1}\boldsymbol{z} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} G\,,$$

where the same symbol is used to denote the modified threshold.

Let us now consider the two-step U-GLRT, i.e.

$$\frac{\max_{(\boldsymbol{p}\in\mathcal{C}^N,\beta\in\mathcal{C})} f(\boldsymbol{z}|\boldsymbol{p},\boldsymbol{q}=\beta\boldsymbol{u}_q,\boldsymbol{M},H_1)}{\max_{\beta\in\mathcal{C}} f(\boldsymbol{z}|\boldsymbol{q}=\beta\boldsymbol{u}_q,\boldsymbol{M},H_0)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} G\,. \tag{43}$$

Substituting the pdf's, maximizing over $\boldsymbol{p}$, and elaborating on, we get

$$\min_{\beta\in\mathcal{C}}(\boldsymbol{z}-\beta\boldsymbol{u}_q)^\dagger\boldsymbol{M}^{-1}(\boldsymbol{z}-\beta\boldsymbol{u}_q) \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} G\,.$$

Performing the minimization and substituting $\boldsymbol{S}/K$ in place of $\boldsymbol{M}$, we get again the test (39). In other words, assuming $\boldsymbol{p}$ unknown and $\boldsymbol{q}$ perfectly aligned with the nominal direction, the one-step GLRT and the two-step GLRT end up coincident.

References

[1] S. Bose and A. O. Steinhardt, "A Maximal Invariant Framework for Adaptive Detection with Structured and Unstructured Covariance Matrices," *IEEE Trans. on Signal Processing*, Vol. 43, No. 3. pp. 2164-2175, September 1995.

[2] E. J. Kelly, "An Adaptive Detection Algorithm," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 22, No. 1, pp. 115-127, March 1986.

[3] F. C. Robey, D. R. Fuhrmann, R. Nitzberg, and E. J. Kelly, "A CFAR Adaptive Matched Filter Detector," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 28, No. 1, pp. 208-216, January 1992.

[4] L. Cai and H. Wang, "On the Adaptive Filtering with the CFAR Feauture and its Performance Sensitivity to Non-Gaussian Interference," *Proc. of the 24th Annual Conference on Information Sciences and Systems*, Princeton, NJ, March 1990, pp. 558-563.

[5] A. De Maio, "A New Derivation of the Adaptive Matched Filter," *IEEE Signal Processing Letters*, Vol. 11, No. 10, pp. 792-793, Ottobre 2004.

[6] N. B. Pulsone and C. M. Rader, "Adaptive Beamformer Orthogonal Rejection Test," *IEEE Trans. on Signal Processing*, Vol. 49, No. 3, pp. 521-529, March 2001.

[7] A. De Maio, "Rao Test for Adaptive Detection in Gaussian Interference with Unknown Covariance Matrix," *IEEE Trans. on Signal Processing*, Vol. 55, No. 7, pp. 3577-3584, July 2007.

[8] C. D. Richmond, "Performance of the Adaptive Sidelobe Blanker Detection Algorithm in Homogeneous Environments," *IEEE Trans. on Signal Processing*, Vol. 48, No. 5, pp. 1235-1247, May 2000.

[9] N. B. Pulsone and M. Zatman, "A Computationally Efficient Two-Step Implementation of the GLRT," *IEEE Trans. on Signal Processing*, Vol. 48, No. 3, pp. 609-616, March 2000.

[10] E. Conte, M. Lops, and G. Ricci, "Asymptotically Optimum Radar Detection in Compound-Gaussian Clutter," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 31, No. 2, pp. 617-625, April 1995.

[11] L. L. Scharf and B. Friedlander, "Matched Subspace Detectors," *IEEE Trans. on Signal Processing*, Vol. 42, No. 8, pp. 2146-2157, August 1994.

[12] S. Kraut and L. L. Scharf: "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. on Signal Processing*, Vol. 47, No. 11, pp. 2538-2541, September 1999.

[13] R. S. Raghavan, N. Pulsone, and D. J. McLaughlin, "Performance of the GLRT for Adaptive Vector Subspace Detection," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 32, No. 4, pp. 1473-1487, October 1996.

[14] G. A. Fabrizio, A. Farina, and M. D. Turley, "Spatial Adaptive Subspace Detection in OTH Radar," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 39, No. 4, pp. 1407-1428, October 2003.

[15] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive Subspace Detectors," *IEEE Trans. on Signal Processing*, Vol. 49, No. 1, pp. 1-16, January 2001.

[16] F. Gini and A. Farina, "Vector Subspace Detection in Compound- Gaussian Clutter. Part I: Survey and New Results," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 4, pp. 1295-1311, October 2002.

[17] F. Gini, A. Farina, and D. Montanari, " "Vector Subspace Detection in Compound-Gaussian Clutter. Part II: Performance Analysis," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 38, No. 4, pp. 1312-1323, October 2002.

[18] R. S. Raghavan, E. Qiu, and D. J McLaughlin, "CFAR Detection in Clutter With Unknown Correlation Properties," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 31, No. 2, pp. 647-657, April 1995.

[19] A. De Maio, "Robust Adaptive Radar Detection in the Presence of Steering Vector Mismatches," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 41, No. 4, pp. 1322-1337, October 2005.

[20] O. Besson, “Detection of a Signal in Linear Subspace with Bounded Mismatch,” *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 42, No. 3, pp. 1131 - 1139, July 2006.

[21] F. Bandiera, A. De Maio, and G. Ricci, “Adaptive CFAR Radar Detection with Conic Rejection”, *IEEE Trans. on Signal Processing*, Vol. 55, No. 6, pp. 2533 - 2541, June 2007.

[22] F. Bandiera, A. De Maio, A.S. Greco, and G. Ricci, “Adaptive Radar Detection of Distributed Targets in Homogeneous and Partially Homogeneous Noise Plus Subspace Interference,” *IEEE Trans. on Signal Processing*, Vol. 55, No. 4, pp. 1223-1237, April 2007.

[23] F. Bandiera, O. Besson, D. Orlando, G. Ricci, and L.L. Scharf, “GLRT-Based Direction Detectors in Homogeneous Noise and Subspace Interference,” *IEEE Trans. on Signal Processing*, Vol. 55, No. 6, pp. 2386-2394, June 2007.

[24] Y. Huang and S. Zhang, “Complex Matrix Decomposition and Quadratic Programming”, *Mathematics of Operations Research*, Vol. 32, No. 3, pp. 758-768, August 2007.

[25] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I, John Wiley & Sons, 1968.

[26] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Second Edition, John Wiley & Sons, 1984.

[27] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Series on Optimization, 2001.

[28] Z. -Q. Luo and S. Zhang, “On Extension of the Frank-Wolfe Theorems”, *Computational Optimization Applications*, Vol. 13, No. 1, pp. 87 - 110, April 1999.

[29] A. Beck and Y. Eldar, “Strong Duality in Nonconvex Optimization with Two Quadratic Constraints,” *SIAM Journal on Optimization*, Vol. 17, No. 3, pp. 844-860, October 2006.

[30] J. F. Sturm, “ Using SeDuMi 1.02, a MATLAB Toolbox for Optimization over Symmetric Cones,” *Optim. Meth. Software*, Vol. 11-12, pp. 625-653, August 1999.

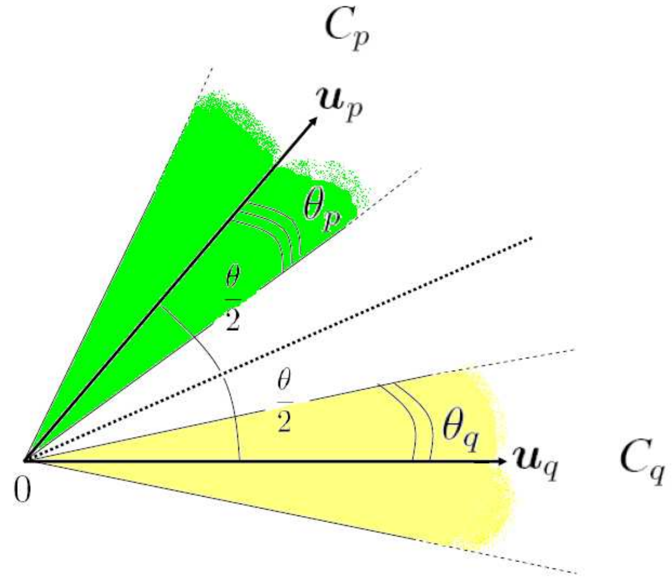

Figure 1: Parts of the Conic Regions C_p (darker cone) and C_q (lighter cone). The dotted curved is the bisector of the angle θ, where $\theta = \arccos|\boldsymbol{u}_p^\dagger \boldsymbol{u}_q|$. The angles θ_q and θ_p are defined by $\theta_i = \theta/h_i$, with $i = q, p$ and $h_i \in \mathbb{R}, h_i > 2$.

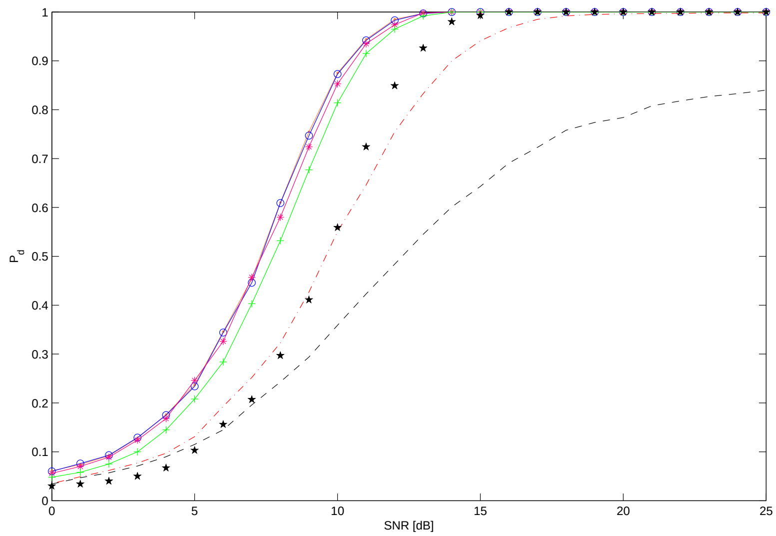

Figure 2a: P_d versus SNR for $P_{fa} = 10^{-2}$, INR= 20 dB, $N = 8$, $h_p = h_q = 10$, $K = 16$, $\theta_T = \theta_p$, θ_I uniformly distributed in $[0, \theta_q]$. R-1S-GLRT (dot-marked curve), R-2S-GLRT (o-marked curve), MR-1S-GLRT (asterisk-marked curve), MR-2S-GLRT (plus-marked curve), 1S-GLRT (dashed curve), 2S-GLRT (dash-dotted curve), U-GLRT (star-marked points).

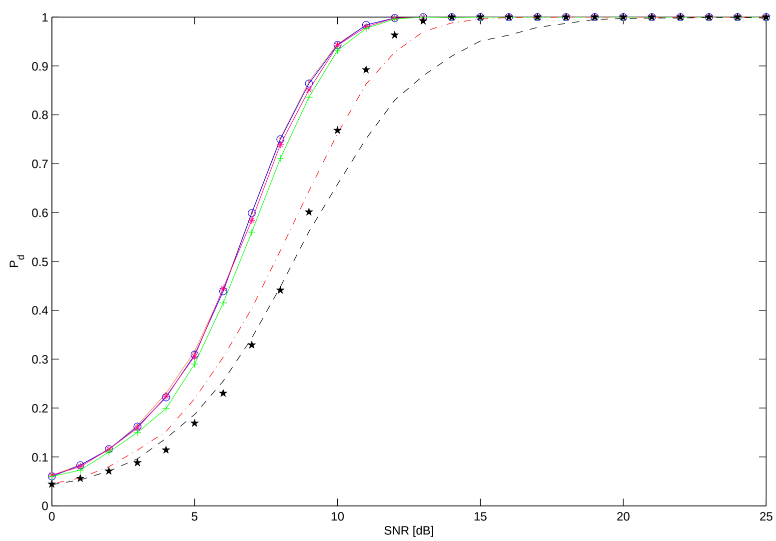

Figure 2b: P_d versus SNR for $P_{fa} = 10^{-2}$, INR= 20 dB, $N = 8$, $h_p = h_q = 10$, $K = 32$, $\theta_T = \theta_p$, θ_I uniformly distributed in $[0, \theta_q]$. R-1S-GLRT (dot-marked curve), R-2S-GLRT (o-marked curve), MR-1S-GLRT (asterisk-marked curve), MR-2S-GLRT (plus-marked curve), 1S-GLRT (dashed curve), 2S-GLRT (dash-dotted curve), U-GLRT (star-marked points).

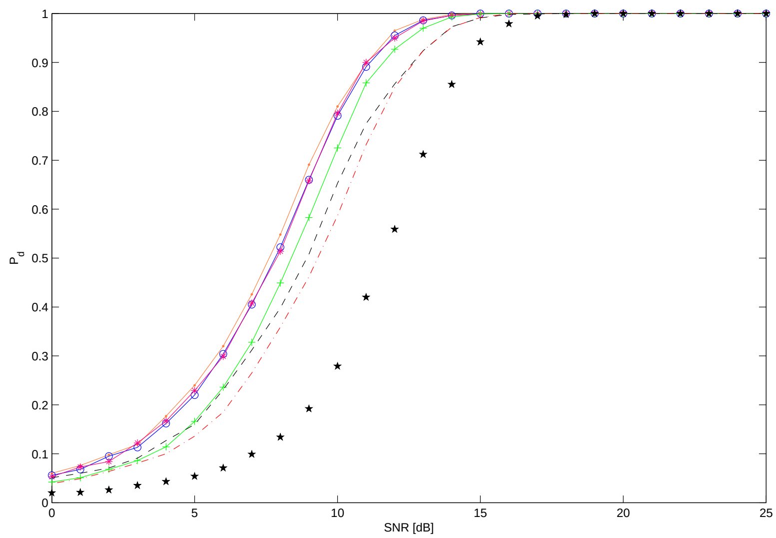

Figure 3a: P_d versus SNR for $P_{fa} = 10^{-2}$, INR= 20 dB, $N = 8$, $h_p = h_q = 10$, $K = 16$, $\theta_T = 0$, $\theta_I = 0$. R-1S-GLRT (dot-marked curve), R-2S-GLRT (o-marked curve), MR-1S-GLRT (asterisk-marked curve), MR-2S-GLRT (plus-marked curve), 1S-GLRT (dashed curve), 2S-GLRT (dash-dotted curve), U-GLRT (star-marked points).

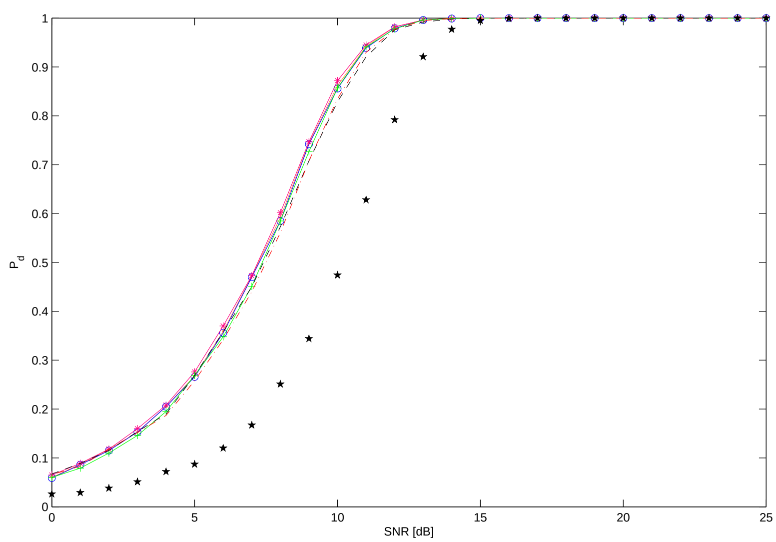

Figure 3b: P_d versus SNR for $P_{fa} = 10^{-2}$, INR= 20 dB, $N = 8$, $h_p = h_q = 10$, $K = 32$, $\theta_T = 0$, $\theta_I = 0$. R-1S-GLRT (dot-marked curve), R-2S-GLRT (o-marked curve), MR-1S-GLRT (asterisk-marked curve), MR-2S-GLRT (plus-marked curve), 1S-GLRT (dashed curve), 2S-GLRT (dash-dotted curve), U-GLRT (star-marked points).

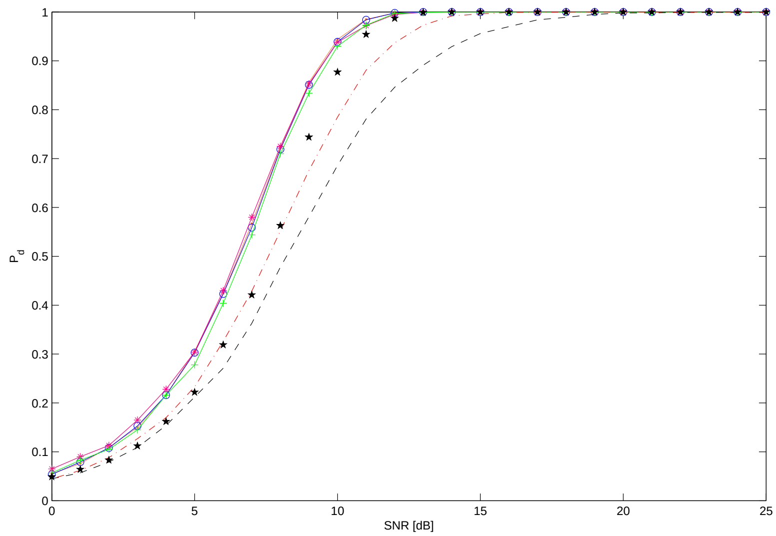

Figure 4a: P_d versus SNR for $P_{fa} = 10^{-2}$, INR= 10 dB, $N = 8$, $h_p = h_q = 10$, $K = 32$, $\theta_T = \theta_p$, θ_I uniformly distributed in $[0, \theta_q]$. R-1S-GLRT (dot-marked curve), R-2S-GLRT (o-marked curve), MR-1S-GLRT (asterisk-marked curve), MR-2S-GLRT (plus-marked curve), 1S-GLRT (dashed curve), 2S-GLRT (dash-dotted curve), U-GLRT (star-marked points).

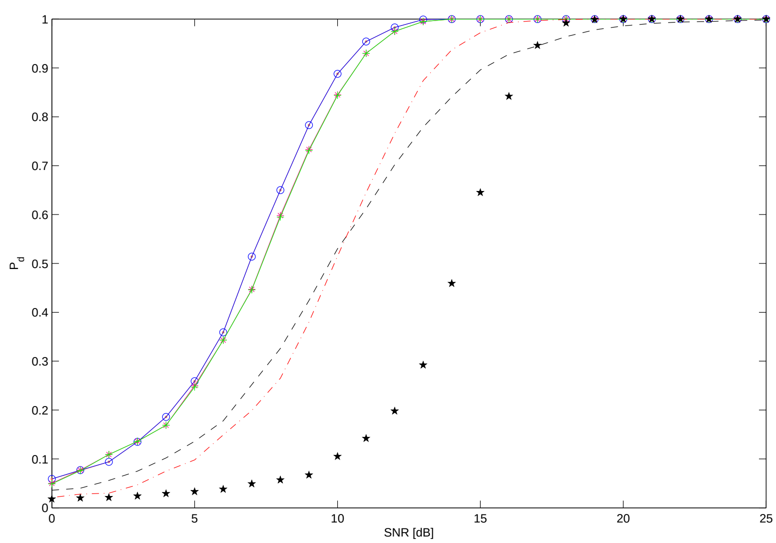

Figure 4b: P_d versus SNR for $P_{fa} = 10^{-2}$, INR= 30 dB, $N = 8$, $h_p = h_q = 10$, $K = 32$, $\theta_T = \theta_p$, θ_I uniformly distributed in $[0, \theta_q]$. R-1S-GLRT (dot-marked curve), R-2S-GLRT (o-marked curve), MR-1S-GLRT (asterisk-marked curve), MR-2S-GLRT (plus-marked curve), 1S-GLRT (dashed curve), 2S-GLRT (dash-dotted curve), U-GLRT (star-marked points).

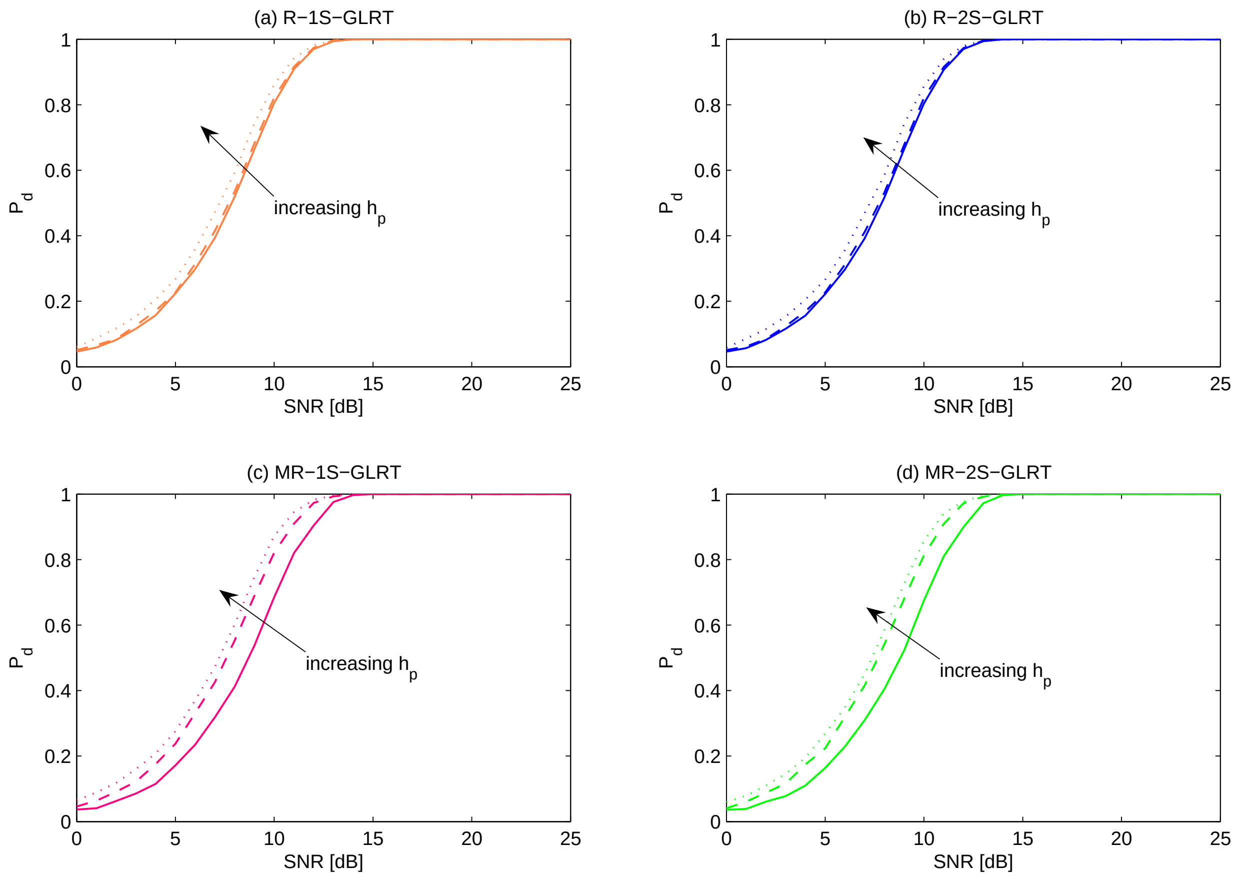

Figure 5: P_d versus SNR for $P_{fa} = 10^{-2}$, INR= 20 dB, $N = 8$, $h_q = 10$, $K = 32$, $\theta_T = 0$, $\theta_I = 0$, and $h_p \in \{5, 7.5, 10\}$, $h_p = 5$ solid curves, $h_p = 7.5$ dashed curves, $h_p = 10$ dotted curves: (a) R-1S-GLRT (b) R-2S-GLRT (c) MR-1S-GLRT (d) MR-2S-GLRT.

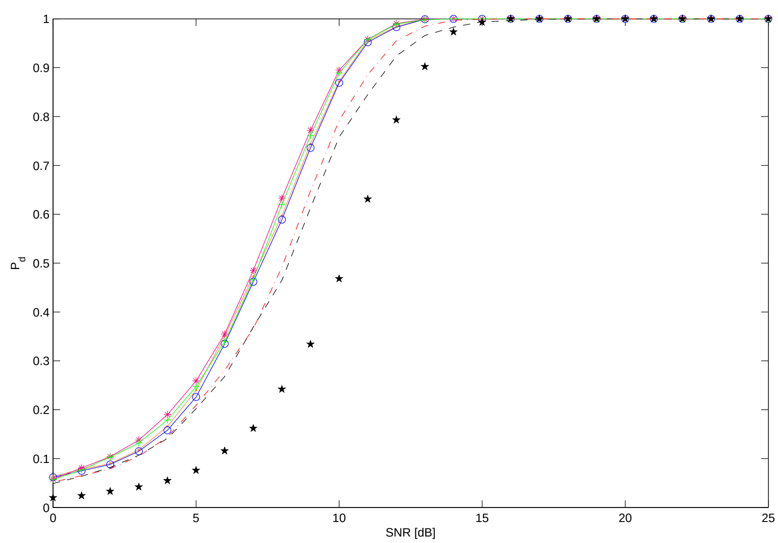

Figure 6: P_d versus SNR for $P_{fa} = 10^{-2}$, INR$=$ 20 dB, $N = 8$, $h_p = h_q = 10$, $K = 32$, θ_T uniformly distributed in $[0, \theta_p]$, θ_I uniformly distributed in $[0, \theta_q]$. R-1S-GLRT (dot-marked curve), R-2S-GLRT (o-marked curve), MR-1S-GLRT (asterisk-marked curve), MR-2S-GLRT (plus-marked curve), 1S-GLRT (dashed curve), 2S-GLRT (dash-dotted curve), U-GLRT (star-marked points).