Passive Detection of Correlated Subspace Signals in Two MIMO Channels

I. Santamaria, *Senior Member, IEEE,* L. L. Scharf, *Life Fellow, IEEE,* J. Via, *Senior Member, IEEE,* H. Wang, Y. Wang

***Abstract*—In this work, we consider a two-channel multiple-input multiple-output (MIMO) passive detection problem, in which there is a surveillance array and a reference array. The reference array is known to carry a linear combination of broadband noise and a subspace signal of known dimension, but unknown basis. The question is whether the surveillance channel carries a linear combination of broadband noise and a subspace signal of the same dimension, but unknown basis, *which is correlated with the subspace signal in the reference channel*. We consider a second-order detection problem where these subspace signals are structured by an unknown, but common, p-dimensional random vector of symbols transmitted from sources of opportunity, and then received through unknown $M \times p$ matrices at each of the M-element arrays. The noises in each channel have spatial correlation models ranging from arbitrarily correlated to independent with identical variances. We provide a unified framework to derive the generalized likelihood ratio test (GLRT) for these different noise models. In the most general case of arbitrary noise covariance matrices, the test statistic is a monotone function of canonical correlations between the reference and surveillance channels.**

***Index Terms*—Passive detection, MIMO channels, passive radar, generalized likelihood ratio, canonical coordinates, geometric mean of eigenvalues, arithmetic mean of eigenvalues.**

I. Introduction

This paper is motivated by a passive radar [1] application, where the problem is to determine if there are complex demodulations and synchronizations in several surveillance antennas (or antenna arrays) that bring signals in the surveillance antennas into coherence with signals in the reference antennas. This coherence is manifested in the synchronous sharing of transmit symbols from several opportunistic transmitters (e.g. digital television, digital audio broadcast, or mobile communication systems), and as a consequence there is correlation between signals observed at the MIMO surveillance array and the MIMO reference array. So the problem is to detect correlated subspace signals in two MIMO channels. In passive radar the signal paths for the reference and the surveillance channels are typically separated by digital beamforming using directional antennas.

Passive radar systems have been studied for several decades due to their simplicity and low cost of implementation in comparison to systems with dedicated transmitters [2]. The conventional approach for passive detection uses the cross-correlation (CC) between the data received in the reference and surveillance channels as the test statistic [3]. Using also cross-correlations as local test statistics, the authors of [4] consider a decentralized detection approach and propose a linear scheme that fuses local detection statistics to form a global detection statistic with improved performance. The authors in [4] consider a multistatic passive radar system composed of K receivers (each receiver composed in turn of a one-dimensional surveillance channel and a one-dimensional reference channel) paired with K non-cooperative illuminators. However, the noise in the reference signal renders these CC-based detection schemes suboptimal, especially in MIMO scenarios for which the inherent subspace structure of the received signals can be exploited [5].

Passive MIMO target detection with a noisy reference channel has recently been considered in [6], where the transmitted waveform is considered to be deterministic, but unknown. The authors of [6] derive the generalized likelihood ratio test (GLRT) for this deterministic target model under spatially white noise of *known* variance. The work in [7] derives the GLRT in a passive radar problem that models the received signal as a deterministic rank-one waveform scaled by an unknown single-input single-output (SISO) channel. The noise is white of either known or unknown variance. In another line of work, a passive detector that exploits the subspace structure of the received signal has been proposed in [8]. Instead of computing the cross-correlation between the surveillance and reference channel measurements, the ad-hoc detector proposed in [8] cross-correlates the dominant left singular vectors of the matrices containing the observations acquired at both channels.

Detection of a subspace signal of dimension-one with a single array of sensors under white noise of unknown level has been addressed in [9], [10] and extensions to diagonal noise covariance matrices and dimension-p signals can be found in [11], and [12], [13], respectively. Other variants of this problem, motivated by cognitive radio and multi-static radio applications, have been considered in [14]–[20]. References [16], [17] are noteworthy for their use of a noninformative prior, in this case the Haar measure on the space of dimension-p subspaces, followed by integration for a marginal measurement density. Different from these detection problems, which

I. Santamaria and J. Via are with the Department of Communications Engineering, University of Cantabria, Spain (e-mail: {javier.via, i.santamaria}@unican.es). L. L. Scharf is with the Dept. of Mathematics, Colorado State University, Fort Collins, CO, USA 80523. H. Wang is with the Dept. of Statistics, Colorado State University, Fort Collins, CO, USA 80523. Y. Wang is with the Dept. of Mathematics and Statistics, Washington State University, Pullman, WA, USA, 99163.

I. Santamaria and J. Via have received funding from Ministerio de Economía y Competitividad (MINECO) of Spain, and AEI/FEDER funds of the E.U. under projects TEC2013-47141-C4-3-R (RACHEL), TEC2016-75067-C4-4-R (CARMEN) and TEC2016-81900-REDT (KERMES). The research of Haonan Wang was partially supported by NSF grant DMS-1521746.

except for [17] are solved with a single array of sensors at the surveillance channel (for radar applications) or at the secondary user (for cognitive radio applications), the model considered in this paper is solved with the assistance of an additional multi-antenna reference channel which acquires a noisy and distorted version of the transmitted signal.

In this paper, we address the MIMO passive detection problem in a multivariate normal model when the surveillance and reference channels are equipped with M antennas. The received signals are subspace signals of known dimension-p, but unknown basis. The noises at the surveillance and reference channels are uncorrelated between channels, but they may otherwise have spatially-structured covariance models. It turns out that this is a problem in factor analysis [21], where there are constraints on the factor loadings and the factors. The problem may be viewed as a one-channel factor analysis problem with constraints on the factor loadings under the null hypothesis, or as a two-channel factor analysis problem, with constraints on the factor loadings under the null, and with common factors under the alternative.

There are four plausible additive noise models for the problems we study, with spatial correlations ranging from arbitrary correlation to independent and identically distributed (i.i.d.) noises across antennas. All lead to a ratio of determinants of estimated covariance matrices as the generalized likelihood ratio test (GLRT). The covariance matrices are maximum likelihood (ML) estimates of covariance, under the constraints of the measurement model, which is determined by the additive noise model and the dimension of the subspace signal in each array. This result is based on a new result for ML estimation showing that the ML estimate of a covariance matrix, constrained to a cone, forces the trace term in Gaussian likelihood to be a constant equal to $2M$, the measurement dimension. The ML estimates of factor loadings are determined by using a noise-whitening trick, [22], [21] to construct a noise-whitened version of the sample covariance matrix, and then using a result from [21] to optimize over factor loadings. The ML estimate of noise covariance is then found by maximizing the geometric mean of trailing eigenvalues of this covariance matrix, under a constraint that the arithmetic mean of these eigenvalues sums to $2M - p$.

The paper is organized as follows. Section II presents the two-channel passive detection problem. Invariance considerations are advanced in III. A common framework to obtain the ML estimates of the covariance matrices under the four noise models considered in this paper, as well as to derive the corresponding GLRTs, is described in Section IV. Two of the four problems we study have closed-form solutions for the GLRT (which are described in Section V), and two require numerical optimization, for which we use an alternating minimization algorithm which is described in Section VI. Numerical simulations under each of the four additive noise models demonstrate performance of each detector against data that is matched to the detector, and to data that is mismatched to the detector. These results are given in Section VII and reviewed in Section VIII, which concludes the paper.

A. Notation

The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian, respectively. The determinant, trace and Frobenius norm of a matrix $\mathbf{A}$ will be denoted, respectively, as $\det(\mathbf{A})$, $\mathrm{tr}(\mathbf{A})$ and $||\mathbf{A}||_F$. $\mathbf{I}_M$ is the identity matrix of dimensions $M \times M$, and $\mathbf{0}$ denotes either a column vector with M zeros, or the zero matrix of appropriate dimensions (the difference should be clear from the context). We use $\mathbf{A}^{1/2}$ ($\mathbf{A}^{-1/2}$) to denote the square root matrix of the Hermitian matrix $\mathbf{A}$ ($\mathbf{A}^{-1}$); $\mathrm{diag}_M(\mathbf{A})$ is a block-diagonal matrix formed by $M \times M$ blocks on the diagonal of $\mathbf{A}$. The expectation operator will be denoted by $E[\cdot]$, and $\mathbf{x} \sim \mathcal{CN}_M(\mathbf{0}, \mathbf{R})$ indicates that $\mathbf{x}$ is an M-dimensional complex circular Gaussian random vector of zero mean and covariance $\mathbf{R}$.

II. Problem Formulation

A. Signal Model

We consider the problem of target detection in a passive network consisting of a reference channel and a surveillance channel, both equipped with M antennas as shown in Fig. 1. The perceptive reader will note that everywhere we assume M antenna elements at the surveillance array and M antenna elements at the reference array, these may be replaced by M_s in the surveillance array and by M_r in the reference array. Then by replacing M in the resulting detector equations by M_s for the surveillance channel and by M_r for the reference channel, all results remain valid for this more general case. This is clear from the derivations. However, there remains the question of performance. We address this question with simulation results for unequal numbers of sensors in Section VII-E.

The system consists of p non-cooperative illuminators (e.g. digital TV stations), transmitting uncorrelated signals over a common bandwidth. We assume that the target path signals received at the surveillance array (solid lines in Fig. 1) have been synchronized in delay τ and Doppler, ν with respect to the reference signal. Therefore, there is a scanning process in the range-Doppler plane to select the matched (τ, ν) that attempts to bring the reference and surveillance channel into coherence, and hence our test statistics are actually ambiguity scores. This synchronization with respect to delay and Doppler is typically assumed in other recent works on passive detection [4], and maximization can be implemented using the method described in [23].

In this work we also assume that the direct-path signals from the non-cooperative transmitters to the surveillance channel (dotted lines in Fig. 1) have been cancelled by directional antennas or spatial filtering. Admittedly, this is an idealized assumption and, in practice, some direct-path residual due to leakage from beampattern sidelobes or to vibrations of the radar platform may still exist in the surveillance channel [24]. The inclusion of a non-negligible direct signal-path in the surveillance channel has been considered in [4], [23].

With these simplifying assumptions, our two-channel measurement model is

$$\begin{bmatrix} \mathbf{x}_s[n] \\ \mathbf{x}_r[n] \end{bmatrix} = \begin{bmatrix} \theta \mathbf{H}_s \\ \mathbf{H}_r \end{bmatrix} \mathbf{s}[n] + \begin{bmatrix} \mathbf{v}_s[n] \\ \mathbf{v}_r[n] \end{bmatrix}; \quad n = 1, \ldots, N \quad (1)$$

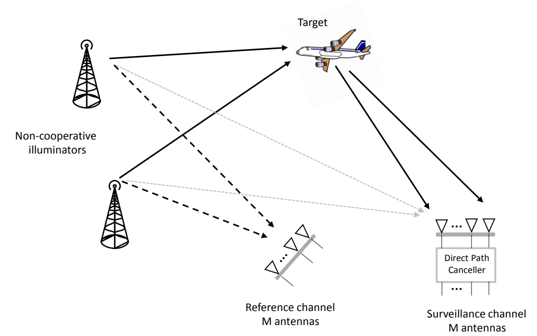

Fig. 1. Passive MIMO radar system: The dashed lines represent the direct link between the non-cooperative sources of opportunity and the reference array, whereas the solid lines represents the illuminator-target-surveillance array path. The dotted lines represent the direct path signals between the non-cooperative illuminators and the surveillance channel, which are assumed to be cancelled.

where $\mathbf{x}_s[n] \in \mathcal{C}^M$ and $\mathbf{x}_r[n] \in \mathcal{C}^M$ are the surveillance and reference measurements; $\mathbf{s}[n] \in \mathcal{C}^p$ contains the signal transmitted by p opportunistic illuminators, $\mathbf{H}_s \in \mathbb{C}^{M\times p}$ and $\mathbf{H}_r \in \mathbb{C}^{M\times p}$ represent the $M \times p$ channels from the transmitter(s) to the surveillance and reference multiantenna receivers, respectively. The parameter $\theta \in \{0, 1\}$ determines whether or not there is a signal $\mathbf{H}_s\mathbf{s}[n]$ in the surveillance channel.

We treat the symbol sequence $\mathbf{s}$ as a sequence of circular, Gaussian random vectors with unknown covariance $E[\mathbf{s}[n]\mathbf{s}^H[m]] = \mathbf{C}\delta[n-m]$. The Gaussian assumption is an accurate approximation of orthogonal frequency-division multiplexing (OFDM) signals with high number of subcarriers [25], as is the case for the European DVB-T (digital video broadcasting-terrestrial) system [26], which has a 2k mode with 1705 subcarriers and a 8k mode with 6817 subcarriers. Although the Gaussian assumption is not realistic when the transmitted sequence belongs to multilevel constellations such as quadrature-amplitude modulation (QAM), it has been shown in [27] that assuming Gaussianity still provides accurate maximum likelihood estimators when the signal-to-noise-ratio is low (which is always the case in passive radar). Simulation results in Section VII show that our detectors, derived for Gaussian symboling, are robust to symboling with OFDM and DVB-T modulations.

As in most works on passive sensing [3], [4], [6], [8], we have assumed in (1) that the channel remains constant over the duration of a sensing period N. This is a reasonable approximation for many signals transmitted by non-cooperative illuminators. Taking again the European DVB-T as an example, the OFDM symbol duration is 256 μsecs and a new full channel estimate is available every 4 OFDM symbols (1.024 msecs), which gives us a rough estimate for the channel coherence time. Typically, the sensing period duration will be less that the channel coherence time even when moving targets are present, and therefore the channel can be safely assumed to remain constant.

The factor loadings $\mathbf{H}_s$ and $\mathbf{H}_r$ are unknown, to be identified in a maximum likelihood procedure. Without loss of generality, the symbol covariance may be absorbed into these factor loadings and thus we assume $\mathbf{C} = \mathbf{I}_p$. The vectors $\mathbf{v}_s[n]$ and $\mathbf{v}_r[n]$ model the additive noise. For notational convenience, the signal, noise, and channel vectors can be stacked as $\mathbf{x}[n] = [\mathbf{x}_s[n]^T, \mathbf{x}_r[n]^T]^T$, $\mathbf{v}[n] = [\mathbf{v}_s[n]^T, \mathbf{v}_r[n]^T]^T$ and $\mathbf{H} = [\mathbf{H}_s^T, \mathbf{H}_r^T]^T$, respectively.

The covariance model for the signal component of equation (1) is

$$E\left[\begin{bmatrix}\theta\mathbf{H}_s\\ \mathbf{H}_r\end{bmatrix}\mathbf{s}[n]\mathbf{s}^H[n]\begin{bmatrix}\theta\mathbf{H}_s^H & \mathbf{H}_r^H\end{bmatrix}\right] = \begin{bmatrix}\theta^2\mathbf{H}_s\mathbf{H}_s^H & \theta\mathbf{H}_s\mathbf{H}_r^H\\ \theta\mathbf{H}_r\mathbf{H}_s^H & \mathbf{H}_r\mathbf{H}_r^H\end{bmatrix}. \tag{2}$$

For the covariance of the noise component, we consider four different models. Under all models, the additive noise is assumed to be temporally white, zero-mean Gaussian distributed, and uncorrelated between the surveillance and reference channels. The noise covariance matrix can then be written as

$$E[\mathbf{v}[n]\mathbf{v}^H[n]] = \mathbf{\Sigma} = \begin{bmatrix}\mathbf{\Sigma}_{ss} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Sigma}_{rr}\end{bmatrix} \in \mathcal{E} \tag{3}$$

$$E[\mathbf{v}[n]\mathbf{v}^H[m]] = \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}, \text{ for } m \neq n, \tag{4}$$

where $\mathcal{E}$ is a set of structured covariances. We study four different structuring sets.

- **Model 1**: Independent and identically distributed (i.i.d.) noises with identical variance at both channels; $\mathbf{\Sigma}_{ss} = \mathbf{\Sigma}_{rr} = \sigma^2\mathbf{I}_M$:
$$\mathcal{E}_1 = \left\{\mathbf{\Sigma} \succ 0 \quad | \quad \mathbf{\Sigma} = \sigma^2\mathbf{I}_{2M}\right\}, \tag{5}$$
where $\sigma^2 > 0$.
- **Model 2**: White noises, but with different variances at the surveillance and reference channels; $\mathbf{\Sigma}_{ss} = \sigma_s^2\mathbf{I}_M, \mathbf{\Sigma}_{rr} = \sigma_r^2\mathbf{I}_M$:
$$\mathcal{E}_2 = \left\{\mathbf{\Sigma} \succ 0 \quad | \quad \mathbf{\Sigma} = \begin{bmatrix}\sigma_s^2\mathbf{I}_M & \mathbf{0}\\ \mathbf{0} & \sigma_r^2\mathbf{I}_M\end{bmatrix}\right\}, \tag{6}$$
with $\sigma_s^2 > 0$, $\sigma_r^2 > 0$.
- **Model 3**: Uncorrelated noises across antennas, thus yielding a diagonal noise covariance matrix with unknown elements along its diagonal; $\mathbf{\Sigma}_{ss}$ and $\mathbf{\Sigma}_{rr}$ are diagonal positive definite (psd) matrices:
$$\mathcal{E}_3 = \{\mathbf{\Sigma} \succ 0 \quad | \quad \mathbf{\Sigma} = \begin{bmatrix}\operatorname{diag}(\sigma_{s,1}^2, \cdots, \sigma_{s,M}^2) & \mathbf{0}\\ \mathbf{0} & \operatorname{diag}(\sigma_{r,1}^2, \cdots, \sigma_{r,M}^2)\end{bmatrix}\right\}, \tag{7}$$
with all $\sigma_{sm}, \sigma_{rm} > 0$.
- **Model 4**: Noises with arbitrary spatial correlation; $\mathbf{\Sigma}_{ss}$ and $\mathbf{\Sigma}_{rr}$ are arbitrary psd matrices:
$$\mathcal{E}_4 = \left\{\mathbf{\Sigma} \succ 0 \quad | \quad \mathbf{\Sigma} = \begin{bmatrix}\mathbf{\Sigma}_{ss} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Sigma}_{rr}\end{bmatrix}\right\}, \tag{8}$$
with $\mathbf{\Sigma}_{ss} \succ 0, \mathbf{\Sigma}_{rr} \succ 0$.

B. Detection problem

The passive detection problem is to test the hypothesis that the surveillance channel contains no signal, versus the alternative that it does:

$$\begin{aligned}\mathcal{H}_0 &: \theta = 0 \\ \mathcal{H}_1 &: \theta = 1\end{aligned} \tag{9}$$

Denote by $\mathcal{R}_{0,j}$ and $\mathcal{R}_{1,j}$ the set of measurement covariance matrices for model j under the null hypothesis and alternative hypothesis, respectively. We have

$$\mathcal{R}_{0,j} = \left\{ \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_r\mathbf{H}_r^H \end{bmatrix} + \mathbf{\Sigma}, \text{ for } \mathbf{\Sigma} \in \mathcal{E}_j \right\} \tag{10}$$

$$\mathcal{R}_{1,j} = \left\{ \begin{bmatrix} \mathbf{H}_s\mathbf{H}_s^H & \mathbf{H}_s\mathbf{H}_r^H \\ \mathbf{H}_r\mathbf{H}_s^H & \mathbf{H}_r\mathbf{H}_r^H \end{bmatrix} + \mathbf{\Sigma}, \text{ for } \mathbf{\Sigma} \in \mathcal{E}_j \right\}. \tag{11}$$

For example, $\mathcal{R}_{1,2}$ is the set of $2M \times 2M$ matrices of structure

$$\mathcal{R}_{1,2} = \left\{ \mathbf{H}\mathbf{H}^H + \begin{bmatrix} \sigma_s^2\mathbf{I}_M & \mathbf{0} \\ \mathbf{0} & \sigma_r^2\mathbf{I}_M \end{bmatrix} \right\}, \tag{12}$$

for some $2M \times p$ matrix $\mathbf{H}$; whereas $\mathcal{R}_{0,4}$ is the set of psd matrices with structure

$$\mathcal{R}_{0,4} = \left\{ \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_r\mathbf{H}_r^H \end{bmatrix} + \begin{bmatrix} \mathbf{\Sigma}_{ss} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{rr} \end{bmatrix} \right\}, \tag{13}$$

with $\mathbf{\Sigma}_{ss}$ and $\mathbf{\Sigma}_{rr}$ arbitrary psd matrices.

This detection problem essentially amounts to testing between two different structures for the composite covariance matrix under the null hypothesis and alternative hypothesis. It can be written as

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{x}[n] \sim \mathcal{CN}_{2M}(\mathbf{0}, \mathbf{R}),\ \mathbf{R} \in \mathcal{R}_{0,j} \\ \mathcal{H}_1 &: \mathbf{x}[n] \sim \mathcal{CN}_{2M}(\mathbf{0}, \mathbf{R}),\ \mathbf{R} \in \mathcal{R}_{1,j}.\end{aligned} \tag{14}$$

There are two possible interpretations of this model: (1) it is a one-channel factor model with special constraints on the loadings under $\mathcal{H}_0$; or (2) it is a two channel factor model with loading constraints under $\mathcal{H}_0$ and common factors in the two channels.

C. The Generalized Likelihood Ratio

Let us now consider N consecutive array snapshots under a model with generic covariance matrix $\mathbf{R}$

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}[1] & \dots & \mathbf{x}[N] \end{bmatrix} \in \mathcal{C}^{2M \times N}, \tag{15}$$

which are assumed to be i.i.d. realizations of $\mathbf{x}[n] \sim \mathcal{CN}_{2M}(\mathbf{0}, \mathbf{R})$. The likelihood may be written as

$$f(\mathbf{X}; \mathbf{R}) = \frac{1}{\pi^{2MN} \det(\mathbf{R})^N} \exp\left\{ -N \operatorname{tr}\left(\mathbf{S}\mathbf{R}^{-1} \right) \right\}, \tag{16}$$

where $\mathbf{S} = \frac{1}{N}\mathbf{X}\mathbf{X}^H$ is the sample covariance matrix, partitioned as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{ss} & \mathbf{S}_{sr} \\ \mathbf{S}_{sr}^H & \mathbf{S}_{rr} \end{bmatrix}. \tag{17}$$

Here $\mathbf{S}_{ss}$ is the sample covariance matrix of the surveillance channel and the other blocks are defined similarly. The likelihood depends on unknown nuisance parameters and consequently standard Neyman-Pearson hypothesis testing does not apply. That is, there is no Neyman-Pearson Lemma for constructing a likelihood ratio. A common approach to derive practical detectors when the distributions under both hypotheses are not completely specified is the generalized likelihood ratio test (GLRT), which replaces the unknowns in the likelihood ratio by their maximum likelihood estimates under each hypothesis [28], [29].

The generalized likelihood ratio (GLR) is

$$\Lambda_j = \frac{f(\mathbf{X}; \hat{\mathbf{R}}_{1,j})}{f(\mathbf{X}; \hat{\mathbf{R}}_{0,j})},$$

where $\hat{\mathbf{R}}_{0,j}$ and $\hat{\mathbf{R}}_{1,j}$ are, respectively, the maximum likelihood (ML) estimates of the covariance matrix for model j under $\mathcal{H}_0$ and $\mathcal{H}_1$. They maximize the log-likelihood function

$$\mathcal{L}(\mathbf{R}) = \operatorname{logdet}(\mathbf{S}\mathbf{R}^{-1}) - \operatorname{tr}\left(\mathbf{S}\mathbf{R}^{-1} \right), \tag{18}$$

The GLRT for noise model j reduces to

$$\log(\Lambda_j) = \log\left(\frac{\det(\hat{\mathbf{R}}_{0,j})}{\det(\hat{\mathbf{R}}_{1,j})} \right) - \operatorname{tr}\left(\mathbf{S}\left(\hat{\mathbf{R}}_{1,j}^{-1} - \hat{\mathbf{R}}_{0,j}^{-1} \right) \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{19}$$

with η a suitable threshold.

III. Invariance Considerations

Consider the random vector $\mathbf{x} \in \mathcal{C}^{2M}$, distributed as $\mathbf{x} \sim \mathcal{CN}_{2M}(\mathbf{0}, \mathbf{R}), \mathbf{R} \in \mathcal{R}$. The data matrix $\mathbf{X} = [\mathbf{x}[1], \dots, \mathbf{x}[N]]$ is a set of independent and identically distributed such vectors. Define the transformation group $\mathcal{G} = \{G \mid G(\mathbf{X}) = \mathbf{T}\mathbf{X}\mathbf{Q}\}$. The group action on the measurement matrix $\mathbf{X}$ is $\mathbf{T}\mathbf{X}\mathbf{Q}$, where $\mathbf{T} \in \mathcal{T}$, the complex linear group of nonsingular $2M \times 2M$ matrices, and $\mathbf{Q} \in \mathcal{Q}$, the unitary group of $N \times N$ unitary matrices. This group action leaves $\mathbf{T}\mathbf{X}\mathbf{Q}$ distributed as iid vectors, each distributed as $\mathcal{CN}_{2M}(\mathbf{0}, \mathbf{T}\mathbf{R}\mathbf{T}^H)$. The distribution of $\mathbf{X}$ is said to be invariant-$\mathcal{G}$, and the transformation group on the parameter space induced by the transformation group $\mathcal{G}$ is $\overline{\mathcal{G}} = \{\overline{G} \mid \overline{G}(\mathbf{R}) = \mathbf{T}\mathbf{R}\mathbf{T}^H\}$.

We are interested in those cases where the group $\overline{G}$ leaves a set $\mathcal{R}$ invariant-$\overline{\mathcal{G}}$, which is to say $\overline{G}(\mathcal{R}) = \mathcal{R}$. We say the hypothesis testing problem for model j is invariant-$\mathcal{G}$ when, for all $\mathbf{R} \in \mathcal{R}_{i,j}, i = 0, 1, \mathbf{T}\mathbf{R}\mathbf{T}^H \in \mathcal{R}_{i,j}$. That is, $\overline{G}(\mathcal{R}_{ij}) = \mathcal{R}_{ij}$. When an hypothesis testing problem is invariant-$\mathcal{G}$, we shall insist that any test of it be invariant-$\mathcal{G}$. It is known that the GLRT will be invariant-$\mathcal{G}$ when the testing problem is. (See, for example, the discussion in [30], based on a standard result like Proposition 7.13 in [31].)

In the itemized paragraphs below, we record the transformation groups that leave each of our hypothesis testing problems invariant-. These results are easy to verify, so we leave it to the reader to do so.

- **Model 1:** The sets $\mathcal{R}_{0,1}$ and $\mathcal{R}_{1,1}$ are invariant-$\overline{\mathcal{G}}$ for group actions

$$G(\mathbf{X}) = \mathbf{T}\mathbf{X}\mathbf{Q},\ \mathbf{T} = \begin{bmatrix} \beta\mathbf{Q}_s & \mathbf{0} \\ \mathbf{0} & \beta\mathbf{Q}_r \end{bmatrix} \tag{20}$$

  where $\beta \neq 0$, and $\mathbf{Q}, \mathbf{Q}_s, \mathbf{Q}_r$ are unitary matrices of respective dimensions $2M \times 2M, M \times M, M \times M$. The corresponding group actions on $\mathbf{R}$ are $\overline{\mathcal{G}}(\mathbf{R}) = \mathbf{T}\mathbf{R}\mathbf{T}^H \in \mathcal{R}_{i,1}, i = 0, 1$ when $\mathbf{R} \in \mathcal{R}_{i,1}, i = 0, 1$.

- **Model 2:** The sets $\mathcal{R}_{0,2}$ and $\mathcal{R}_{1,2}$ are invariant-$\overline{\mathcal{G}}$ for group actions
$$G(\mathbf{X}) = \mathbf{TXQ},\ \mathbf{T} = \begin{bmatrix} \beta_s \mathbf{Q}_s & \mathbf{0} \\ \mathbf{0} & \beta_r \mathbf{Q}_r \end{bmatrix} \tag{21}$$
where $\beta_s, \beta_r \neq 0$, and $\mathbf{Q}, \mathbf{Q}_s, \mathbf{Q}_r$ are unitary matrices of respective dimensions $2M \times 2M, M \times M, M \times M$. The corresponding group actions on $\mathbf{R}$ are $\overline{\mathcal{G}}(\mathbf{R}) = \mathbf{TRT}^H \in \mathcal{R}_{i,2}, i = 0, 1$ when $\mathbf{R} \in \mathcal{R}_{i,2}, i = 0, 1$.
- **Model 3:** The sets $\mathcal{R}_{0,2}$ and $\mathcal{R}_{1,2}$ are invariant-$\overline{\mathcal{G}}$ for group actions
$$G(\mathbf{X}) = \mathbf{TXQ} \tag{22}$$
with
$$\mathbf{T} = \begin{bmatrix} \operatorname{diag}(\beta_{s1}, \ldots, \beta_{sM}) & \mathbf{0} \\ \mathbf{0} & \operatorname{diag}(\beta_{r1}, \ldots, \beta_{rM}) \end{bmatrix}$$
where $\beta_{sm}, \beta_{rm} \neq 0$. The corresponding group actions on $\mathbf{R}$ are $\overline{\mathcal{G}}(\mathbf{R}) = \mathbf{TRT}^H \in \mathcal{R}_{i,3}, i = 0, 1$ when $\mathbf{R} \in \mathcal{R}_{i,3}, i = 0, 1$.
- **Model 4:** The sets $\mathcal{R}_{0,4}$ and $\mathcal{R}_{1,4}$ are invariant-$\overline{\mathcal{G}}$ for group actions
$$G(\mathbf{X}) = \mathbf{TXQ},\ \mathbf{T} = \begin{bmatrix} \mathbf{T}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_r \end{bmatrix} \tag{23}$$
where $\mathbf{T}_s, \mathbf{T}_r$ are nonsingular $M \times M$ matrices. The corresponding group actions on $\mathbf{R}$ are $\overline{\mathcal{G}}(\mathbf{R}) = \mathbf{TRT}^H \in \mathcal{R}_{i,4}, i = 0, 1$ when $\mathbf{R} \in \mathcal{R}_{i,4}, i = 0, 1$.

We shall say a detector Λ_j is invariant-$\mathcal{G}$ if $\Lambda_j(\mathbf{X}) = \Lambda_j(\mathcal{G}(\mathbf{X}))$ for Model j. As a check on our derivations, we shall verify that each of our GLRT detectors is in fact invariant-$\mathcal{G}$.

IV. A Common Optimization Framework for GLR Detection

Recall that $\mathcal{R}_{0,j}$ and $\mathcal{R}_{1,j}$ are sets of structured covariance matrices for the measurements under noise model j, and under the null hypothesis $\mathcal{H}_0$, and alternative hypothesis $\mathcal{H}_1$, respectively.

Proposition 1. *The sets $\mathcal{R}_{i,j}$ are cones.*

Proof. A set $\mathcal{R}$ is a cone [32] if for any $\mathbf{R} \in \mathcal{R}$ and $a \geq 0$, we have
$$a\mathbf{R} \in \mathcal{R}.$$
It is easy to check that this condition is satisfied by all covariance matrices formed by a rank-p signal component plus a noise covariance matrix with the structure specified by any of the models described in Section II; then, all sets $\mathcal{R}_{i,j}$ are cones. □

A. The unified constraint on trace

The following Lemma proves that, for all positive-definite covariance models, the trace term in (19) is zero.

Lemma 1. *Let $\hat{\mathbf{R}}$ be the ML estimate for $\mathbf{R}$ that maximizes the likelihood* (16) *within a cone $\mathcal{R}$, and let $\mathbf{S}$ be the sample covariance matrix. Then,*
$$\operatorname{tr}\left(\mathbf{S}\hat{\mathbf{R}}^{-1}\right) = 2M \text{ and } L(\hat{\mathbf{R}}) = \log\det[\mathbf{S}\hat{\mathbf{R}}^{-1}] - 2M. \tag{24}$$

Proof. Let $\tilde{\mathbf{R}} \in \mathcal{R}$ be an estimate (not necessarily the ML estimate) of a covariance matrix within $\mathcal{R}$. Since the set $\mathcal{R}$ is a cone, we can get a new scaled estimate $a\tilde{\mathbf{R}}$ with $a \geq 0$, which also belongs to the set. The log-likelihood as a function of the scaling factor may be written as
$$\begin{aligned} g(a) &= \operatorname{logdet}(\frac{1}{a}\mathbf{S}\tilde{\mathbf{R}}^{-1}) - \frac{1}{a}\operatorname{tr}\left(\mathbf{S}\tilde{\mathbf{R}}^{-1}\right) \\ &= -2M\log(a) + \operatorname{logdet}(\mathbf{S}\tilde{\mathbf{R}}^{-1}) - \frac{1}{a}\operatorname{tr}\left(\mathbf{S}\tilde{\mathbf{R}}^{-1}\right). \end{aligned} \tag{25}$$
Taking the derivative of (25) with respect to a and equating to zero, we find that the optimal scaling factor that maximizes the likelihood is
$$a^* = \frac{\operatorname{tr}\left(\mathbf{S}\tilde{\mathbf{R}}^{-1}\right)}{2M},$$
and thus $g(a^*) \geq g(a)$ for $a \geq 0$. Let $\mathbf{R}^* = a^*\tilde{\mathbf{R}}$. Plugging this value into the trace term of the likelihood function we have
$$\operatorname{tr}\left(\mathbf{S}\mathbf{R}^{-*}\right) = \frac{1}{a^*}\operatorname{tr}\left(\mathbf{S}\tilde{\mathbf{R}}^{-1}\right) = 2M.$$
Since this result has been obtained *for any* estimate belonging to a cone $\mathcal{R}$, it also holds for the ML estimate, thus proving the lemma. To re-iterate, the $\mathcal{R}_{ij}$ are cones, so the result of this lemma holds for all covariance models considered in this paper. □

The following theorem establishes that the GLRT for a subspace signal of dimension p under all noise models considered in this paper is a ratio of determinants.[1]

Theorem 1. *The GLRT for the detection problem (14) under noise model j is given by*
$$\Lambda_j = \frac{\det(\hat{\mathbf{R}}_{0,j})}{\det(\hat{\mathbf{R}}_{1,j})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{26}$$
where $\hat{\mathbf{R}}_{i,j} = \arg\max_{\mathbf{R} \in \mathcal{R}_{i,j}} \log\det(\mathbf{S}\mathbf{R}^{-1})$ *such that* $\operatorname{tr}(\mathbf{S}\mathbf{R}^{-1}) = 2M$.

Proof. From Lemma 1 we know that the trace term of the likelihood function, when evaluated at the ML estimates, is a constant under both hypotheses. Then, substituting $\operatorname{tr}\left(\mathbf{S}\hat{\mathbf{R}}_{i,j}^{-1}\right) = 2M$ into (18) and taking into account the monotonicity of the log function, (26) follows. □

As we will see in Section V, under noise models 1 and 4 we can obtain closed-form expressions for the ML estimates of the covariance matrix under each hypothesis. However, this is not the case for the alternative hypothesis under noise models 2 and 3, for which we resort to numerical methods (e.g. alternating optimization as described in Section VI). In the subsections to follow we describe two new statements of the ML estimation problem for the structured covariance matrices, which provide interesting insights into the problem.

[1] In fact, this results extends to all models such that the structure of the covariance matrices under both hypotheses is defined by a cone.

B. ML estimation in two-channel passive detection problems

As a by-product of Lemma 1, the ML estimates of covariance may be obtained by solving the following optimization problem:

$$\text{Problem 1:} \quad \underset{\mathbf{R}\in\mathcal{R}_{i,j}}{\text{maximize}} \quad \text{logdet}\left(\mathbf{S}\mathbf{R}^{-1}\right) \quad \text{subject to} \quad \text{tr}\left(\mathbf{S}\mathbf{R}^{-1}\right) = 2M. \tag{27}$$

The following theorem illuminates the problem of determining $\mathbf{R}$ in Problem 1, and leads also to an alternative formulation to be given in Problem 2.

Theorem 2. *For a given block-diagonal noise covariance* $\mathbf{\Sigma}$*, we define the noise-whitened sample covariance matrix*

$$\tilde{\mathbf{S}} = \mathbf{\Sigma}^{-1/2}\mathbf{S}\mathbf{\Sigma}^{-1/2} = \begin{bmatrix} \tilde{\mathbf{S}}_{ss} & \tilde{\mathbf{S}}_{sr} \\ \tilde{\mathbf{S}}_{sr}^H & \tilde{\mathbf{S}}_{rr} \end{bmatrix} \tag{28}$$

with eigenvalue decomposition $\tilde{\mathbf{S}} = \tilde{\mathbf{W}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{W}}^H$*, and* $\tilde{\mathbf{\Lambda}} = \text{diag}\left(\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \ldots \geq \tilde{\lambda}_{2M}\right)$*;* $\tilde{\mathbf{S}}_{rr} = \tilde{\mathbf{W}}_{rr}\tilde{\mathbf{\Lambda}}_{rr}\tilde{\mathbf{W}}_{rr}^H$*, and* $\tilde{\mathbf{\Lambda}}_{rr} = \text{diag}\left(\tilde{\lambda}_{rr,1} \geq \tilde{\lambda}_{rr,2} \geq \ldots \geq \tilde{\lambda}_{rr,M}\right)$*. Then, under the alternative* $\mathcal{H}_1$*, the value of* $\mathbf{H}\mathbf{H}^H$ *that maximizes the likelihood* (18) *is*

$$\mathbf{H}\mathbf{H}^H = \mathbf{\Sigma}^{1/2}\tilde{\mathbf{W}}\tilde{\mathbf{D}}\tilde{\mathbf{W}}^H\mathbf{\Sigma}^{1/2} \tag{29}$$

with $\tilde{\mathbf{D}} = \text{diag}\left(d_1 \geq d_2 \geq d_p \geq 0 \ldots 0\right)$ *and* $d_i = \max(\tilde{\lambda}_i - 1, 0)$.

Under the null $\mathcal{H}_0$ ($\theta \equiv 0$)*, and assuming that the noise covariance matrix in the reference channel,* $\mathbf{\Sigma}_{rr}$*, is given, the value of* $\mathbf{H}_r\mathbf{H}_r^H$ *that maximizes the likelihood* (18) *is*

$$\mathbf{H}_r\mathbf{H}_r^H = \mathbf{\Sigma}_{rr}^{1/2}\tilde{\mathbf{W}}_{rr}\tilde{\mathbf{D}}_{rr}\tilde{\mathbf{W}}_{rr}^H\mathbf{\Sigma}_{rr}^{1/2} \tag{30}$$

with $\tilde{\mathbf{D}}_{rr} = \text{diag}\left(d_{rr,1} \geq d_{rr,2} \geq d_{rr,p} \geq 0 \ldots 0\right)$ *and* $d_{rr,i} = \max(\tilde{\lambda}_{rr,i} - 1, 0)$.

Proof. The proof for $\mathcal{H}_1$ is identical to Theorem 9.4.1 in [21] (cf. pages 264-265). The proof for $\mathcal{H}_0$ is straightforward after we rewrite the log-likelihood function using the blockwise decomposition in (28) and use the fact that the noise covariance $\mathbf{\Sigma}$ is block diagonal. □

Theorem 2 can be used to derive Problem 2 for the ML estimate of covariance, under the alternative $\mathcal{H}_1$. For a given $\mathbf{\Sigma}$, Theorem 2 gives the value of $\mathbf{H}\mathbf{H}^H$ that maximizes the log-likelihood function with respect to $\mathbf{R} = \mathbf{H}\mathbf{H}^H + \mathbf{\Sigma}$. Thus, we have the solution $\mathbf{R} = \mathbf{\Sigma}^{1/2}\tilde{\mathbf{W}}\tilde{\mathbf{D}}\tilde{\mathbf{W}}^H\mathbf{\Sigma}^{1/2} + \mathbf{\Sigma}$. Straightforward calculation shows that $\det(\mathbf{S}\mathbf{R}^{-1}) = \prod_{i=1}^{p}\min(\tilde{\lambda}_i, 1)\prod_{j=p+1}^{2M}\tilde{\lambda}_j$ and $\text{tr}(\mathbf{S}\mathbf{R}^{-1}) = \sum_{i=1}^{p}\min(\tilde{\lambda}_i, 1) + \sum_{j=p+1}^{2M}\tilde{\lambda}_j$. Therefore Problem 1 may be re-written as

$$\text{Prob. 2:} \quad \underset{\mathbf{\Sigma}\in\mathcal{E}_j}{\text{maximize}} \quad \left(\prod_{i=1}^{p}\min(\tilde{\lambda}_i, 1)\prod_{j=p+1}^{2M}\tilde{\lambda}_j\right)^{\frac{1}{2M}} \quad \text{subject to} \quad \frac{1}{2M}\left(\sum_{i=1}^{p}\min(\tilde{\lambda}_i, 1) + \sum_{j=p+1}^{2M}\tilde{\lambda}_j\right) = 1 \tag{31}$$

Recall that $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \ldots \geq \tilde{\lambda}_{2M} \geq 0$ is the set of ordered eigenvalues of the noise-whitened sample covariance matrix. Thus, the trace constraint in (31) directly implies $\tilde{\lambda}_k \geq 1$ for $k = 1, \ldots, p$.[2] In consequence, Problem 2 can be written more compactly as

$$\text{Problem 2:} \quad \underset{\mathbf{\Sigma}\in\mathcal{E}_j}{\text{maximize}} \quad \left(\prod_{i=p+1}^{2M}\tilde{\lambda}_i\right)^{\frac{1}{2M-p}} \quad \text{subject to} \quad \frac{1}{2M-p}\sum_{i=p+1}^{2M}\tilde{\lambda}_i = 1. \tag{32}$$

That is, the ML estimation problem under the alternative hypothesis comes down to finding the noise covariance matrix with the required structure that maximizes the geometric mean of the trailing eigenvalues of the noise-whitened sample covariance matrix, subject to the constraint that the arithmetic mean of these trailing eigenvalues is 1. For some specific structures, Problem 2 may significantly simplify derivation of the ML solution, as shown in the next section.

The general formulation of the ML estimation problem for covariance under the null $\mathcal{H}_0$ is more involved. In particular, a similar derivation will result in

$$\text{Problem 3:} \quad \underset{\mathbf{\Sigma}\in\mathcal{E}_j}{\text{maximize}} \quad \left(\prod_{i=1}^{p}\min(\tilde{\lambda}_{rr,i}, 1)\prod_{j=p+1}^{M}\tilde{\lambda}_{rr,j}\prod_{k=1}^{M}\tilde{\lambda}_{ss,k}\right)^{\frac{1}{2M}} \quad \text{subject to} \quad \frac{1}{2M}\left(\sum_{i=1}^{p}\min(\tilde{\lambda}_{rr,i}, 1) + \sum_{j=p+1}^{M}\tilde{\lambda}_{rr,j} + \sum_{k=1}^{M}\tilde{\lambda}_{ss,k}\right) = 1, \tag{33}$$

where the eigenvalues $\tilde{\lambda}_{ss,k}$ of the surveillance channel are defined, for $k = 1, \ldots, M$, analogously to $\tilde{\lambda}_{rr,k}$. Here, it is important to note that in general we can not get rid of the $\min(\cdot)$ operator, which complicates the analysis and solution of the ML estimates under the null. However, in the case of noise models 2, 3, and 4,[3] the estimation of the unknowns in the surveillance, $\mathbf{\Sigma}_{ss}$, and reference channel, $(\mathbf{\Sigma}_{rr}, \mathbf{H}_r)$, is decoupled. In particular, the ML estimates of $\mathbf{\Sigma}_{ss}$ under the null $\mathcal{H}_0$ for noise models 2, 3, and 4 are given by $\frac{1}{\text{tr}(\mathbf{S}_{ss})}\mathbf{I}_M$, $\text{diag}(\mathbf{S}_{ss})$, and $\mathbf{S}_{ss}$, respectively, whereas the ML estimation of $(\mathbf{\Sigma}_{rr}, \mathbf{H}_r)$ under the null $\mathcal{H}_0$ may be found as the solution to

$$\text{Problem 4:} \quad \underset{\mathbf{\Sigma}_{rr}\in\mathcal{E}_j}{\text{maximize}} \quad \left(\prod_{i=p+1}^{M}\tilde{\lambda}_{rr,i}\right)^{\frac{1}{M-p}} \quad \text{subject to} \quad \frac{1}{M-p}\sum_{i=p}^{M}\tilde{\lambda}_{rr,i} = 1, \tag{34}$$

where we recall that in this case $\tilde{\lambda}_{rr,i}$ are the eigenvalues of the noise-whitened sample covariance matrix in the reference channel $\tilde{\mathbf{S}}_{rr} = \mathbf{\Sigma}_{rr}^{-1/2}\mathbf{S}_{rr}\mathbf{\Sigma}_{rr}^{-1/2}$.

[2]Moreover, a value $\tilde{\lambda}_k = 1$ would mean that the observations could be explained by a covariance model with a lower rank, which under the assumption of data drawn from a continuous distribution and a sample size $N > p$, will have zero probability.

[3]For noise model 1, the unknowns under the null are coupled because $\mathbf{\Sigma}_{rr} = \mathbf{\Sigma}_{ss} = \sigma^2\mathbf{I}_M$. Therefore, we relegate the study of this model to Section V-A.

V. GLRTs for Models 1 and 4

In this section we present closed-form GLRTs for noise models 1 and 4.

A. GLR detector for Model 1

We focus here on the case where $p \leq 2M-1$; otherwise the spatial structure of the target plays no role and the GLRT detector is given by the well-known *sphericity test* [33]. Suppose the sample covariance matrices have these eigen decompositions: $\mathbf{S} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^H$, $\mathbf{S}_{ss} = \mathbf{W}_{ss}\mathbf{\Lambda}_{ss}\mathbf{W}_{ss}^H$ and $\mathbf{S}_{rr} = \mathbf{W}_{rr}\mathbf{\Lambda}_{rr}\mathbf{W}_{rr}^H$. Under noise model 1, Problem 2 in (32) directly gives the ML solution for σ^2 under the alternative hypothesis $\mathcal{H}_1$ by realizing that $(2M-p)^{-1}\sum_{p+1}^{2M}\tilde{\lambda}_i = (2M-p)^{-1}\sum_{p+1}^{2M}\sigma^{-2}\lambda_i$, which returns the ML estimate $\hat{\sigma}_1^2 = (2M-p)^{-1}\sum_{p+1}^{2M}\lambda_i$. Therefore, the ML estimates of the covariance matrix under the alternative $\mathcal{H}_1$ is

$$\hat{\mathbf{R}}_{1,1} = \mathbf{W}\mathbf{D}\mathbf{W}^H + \hat{\sigma}_1^2\mathbf{I}_{2M}, \tag{35}$$

where $\hat{\sigma}_1^2$, the ML estimate of the noise variance under $\mathcal{H}_1$, is

$$\hat{\sigma}_1^2 = \frac{1}{2M-p}\sum_{i=p+1}^{2M}\lambda_i, \tag{36}$$

and $\mathbf{D} = \mathrm{diag}(d_1,\ldots,d_p,0,\ldots,0)$ is an $2M \times 2M$ diagonal matrix with $d_i = \lambda_i - \hat{\sigma}_1^2 \geq 0$, by virtue of the eigenvalue ordering.

Let us now consider the ML estimate of the covariance matrix under the null. For a given $\Sigma_{rr} = \sigma_0^2\mathbf{I}_M$, the result in Theorem 2 gives us the value of $\mathbf{H}_r\mathbf{H}_r^H$ that maximizes the likelihood. Then, we can write $\mathbf{R}_{0,1}$ as a function solely of σ_0^2,

$$\mathbf{R}_{0,1} = \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \hat{\mathbf{H}}_r\hat{\mathbf{H}}_r^H\end{bmatrix} + \sigma_0^2\mathbf{I}_{2M}, \tag{37}$$

where $\hat{\mathbf{H}}_r\hat{\mathbf{H}}_r^H = \mathbf{W}_{rr}\mathbf{D}_0\mathbf{W}_{rr}^H$, $\mathbf{D}_0 = \mathrm{diag}(d_1,\ldots,d_p,0,\ldots,0)$ is an $M\times M$ diagonal matrix, and $d_i = \max(\lambda_{rr,i} - \sigma_0^2, 0)$. Taking the inverse of (37) it is straightforward to show that the trace constraint in (33) is

$$\mathrm{tr}(\mathbf{S}\mathbf{R}_{0,1}^{-1}) = p_r + \frac{1}{\sigma_0^2}\sum_{i=p_r+1}^{M}\lambda_{rr,i} + \frac{1}{\sigma_0^2}\mathrm{tr}(\mathbf{S}_{ss}) = 2M, \tag{38}$$

where $p_r = \min(p, p_0)$, and p_0 is the number of eigenvalues satisfying $\lambda_{rr,i} \geq \sigma_0^2$. Therefore, the ML estimate of the noise variance is readily obtained as

$$\hat{\sigma}_0^2 = \frac{1}{2M-p_r}\left(\sum_{i=1}^{M}\lambda_{ss,i} + \sum_{i=p_r+1}^{M}\lambda_{rr,i}\right), \tag{39}$$

and the covariance matrix under the null is

$$\hat{\mathbf{R}}_{0,1} = \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \hat{\mathbf{H}}_r\hat{\mathbf{H}}_r^H\end{bmatrix} + \hat{\sigma}_0^2\mathbf{I}_{2M}. \tag{40}$$

Plugging (40) - (36) into (26), the GLRT under noise model 1 is given by

$$\Lambda_1 = \frac{\left(\prod_{i=1}^{p_r}\lambda_i(\mathbf{S}_{rr})\right)(\hat{\sigma}_0^2)^{2M-p_r}}{\left(\prod_{i=1}^{p}\lambda_i(\mathbf{S})\right)(\hat{\sigma}_1^2)^{2M-p}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{41}$$

where p_r is the largest value of i between 1 and p such that $\lambda_{rr,i} > \hat{\sigma}_0^2$.

Remark 1. *In practice, the procedure for obtaining the ML estimate of σ_0^2 starts with $p_r = p$ and then checks whether the candidate solution satisfies $\lambda_{rr,p_r} \geq \sigma_0^2$. If the condition is not satisfied, the rank of the signal subspace is decreased to $p_r - 1$, which implies in turn a decrease in the estimate of the noise variance until the condition $\lambda_{rr,p_r} \geq \sigma_0^2$ is satisfied. The intuition behind this behavior, which is caused by the coupling between the estimates in the reference and surveillance channels (see Problem (33)), is clear. If the assumed dimension of the signal subspace is not compatible with the estimated noise variance $\hat{\sigma}_0^2$, that is, if the number of signal mode powers above the estimated noise level $\hat{\sigma}_0^2$ is lower than expected, then the dimension of the signal subspace is reduced and the noise variance is estimated based on a smaller signal subspace, and correspondingly a larger noise subspace. Thus, the potential solutions for the ML estimates under the null range from the case $p_r = p$ (meaning that it is possible to estimate a signal subspace of dimension p in the reference channel), to the case $p_r = 0$ when the sample variance in the surveillance channel is larger than the sample variance in the reference channel, which makes us conclude that all the energy in the reference channel is due to the effect of noise.*

Remark 2. *Scaling of the surveillance and reference channels by a* common *β scales $\mathbf{S}$ accordingly. Consequently all of the eigenvalues in the formula for Λ_1 scale commonly, making Λ_1 invariant-$\mathcal{G}$. The detector is CFAR with respect to common scaling of the surveillance and reference channels. This scaling scales the noise power in the surveillance channel and signal-plus-noise power in the reference channel.*

B. GLR detector for Model 4

Under noise model 4, the ML estimate of the covariance matrix under the null is given by

$$\hat{\mathbf{R}}_{0,4} = \mathrm{diag}_M(\mathbf{S}) = \begin{bmatrix}\mathbf{S}_{ss} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}_{rr}\end{bmatrix}. \tag{42}$$

Under the alternative, the ML estimate has been derived for $p=1$ in [5] and for general p in [34]. To present this result, let $\mathbf{C} = \mathbf{S}_{ss}{}^{-1/2}\mathbf{S}_{sr}\mathbf{S}_{rr}{}^{-H/2}$ be the sample coherence matrix between the surveillance and reference channels, and let $\mathbf{C} = \mathbf{F}\mathbf{K}\mathbf{G}^H$ be its singular value decomposition (SVD), where the matrix $\mathbf{K} = \mathrm{diag}\,(k_1,\cdots,k_M)$ contains the sample canonical correlations $1 \geq k_1 \geq \cdots \geq k_M \geq 0$ along its diagonal. With these preliminaries, the ML estimate of the covariance matrix under $\mathcal{H}_1$ is given by

$$\hat{\mathbf{R}}_{1,4} = \begin{bmatrix}\mathbf{S}_{ss} & \mathbf{S}_{ss}^{1/2}\mathbf{C}_p\mathbf{S}_{rr}^{1/2}\\ \mathbf{S}_{rr}^{1/2}\mathbf{C}_p^H\mathbf{S}_{ss}^{1/2} & \mathbf{S}_{rr}\end{bmatrix} = \tag{43}$$

$$= \begin{bmatrix}\mathbf{S}_{ss}^{1/2} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}_{rr}^{1/2}\end{bmatrix}\begin{bmatrix}\mathbf{F} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}\end{bmatrix}\begin{bmatrix}\mathbf{I}_M & \mathbf{K}_p\\ \mathbf{K}_p & \mathbf{I}_M\end{bmatrix}\begin{bmatrix}\mathbf{F}^H & \mathbf{0}\\ \mathbf{0} & \mathbf{G}^H\end{bmatrix}\begin{bmatrix}\mathbf{S}_{ss}^{1/2} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}_{rr}^{1/2}\end{bmatrix}.$$

where $\mathbf{K}_p = \text{diag}(k_1, \cdots, k_p, 0, \ldots, 0)$ is a rank-p truncation of $\mathbf{K}$. Plugging the ML estimates into (26), it is easy to check that the GLRT under noise model 4 is

$$\Lambda_4 = \det(\mathbf{I} - \mathbf{K}_p^2)^{-1} = \prod_{i=1}^{p} \frac{1}{(1-k_i^2)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (44)$$

where k_i is the i-th sample canonical correlation between the surveillance and reference channels, and η is a suitable threshold. This formula was also derived in [35] for a different problem. Equation (44) has an interesting interpretation: $1-\Lambda_4{}^{-1}$ is the coherence statistic, $0 \leq 1-\prod_{i=1}^{p}(1-k_i^2) \leq 1$, which has the interpretation of a soft OR detector, based on squared canonical correlations.

Remark 3. *Independent transformation of the surveillance channel by a non-singular transformation $\mathcal{T}_s$ and the reference channel by a non-singular transformation $\mathcal{T}_r$, leaves the coherence matrix C invariant. Consequently its singular values k_i are invariant, and as a consequence the detector Λ_4 is invariant-$\mathcal{G}$. As a special case, Λ_4 is CFAR with respect to noise power in the surveillance channel and signal-plus-noise power in the reference channel.*

Remark 4. *From the identified model for $\hat{\mathbf{R}}_{1,4}$ in equation (43) it is a standard result in the theory of MMSE estimation that the estimator of a measurement $\mathbf{x}_s$ in the surveillance channel can be orthogonally decomposed as $\mathbf{x}_s = \hat{\mathbf{x}}_s + \hat{\mathbf{e}}_s$, where $\hat{\mathbf{x}}_s = \mathbf{S}_{ss}^{1/2}\mathbf{F}\mathbf{K}_p\mathbf{G}^H\mathbf{S}_{rr}^{-1/2}\mathbf{x}_r \sim \mathcal{CN}_M(\mathbf{0}, \mathbf{S}_{ss}^{1/2}\mathbf{F}\mathbf{K}_p\mathbf{K}_p^H\mathbf{F}^H\mathbf{S}_{ss}^{H/2})$, and $\hat{\mathbf{e}}_s \sim \mathcal{CN}_M(\mathbf{0}, \mathbf{S}_{ss}^{1/2}[\mathbf{I}_M - \mathbf{F}\mathbf{K}_p\mathbf{K}_p^H\mathbf{F}^H]\mathbf{S}_{ss}^{H/2})$. The matrix $\mathbf{S}_{ss}^{1/2}\mathbf{F}\mathbf{K}_p\mathbf{G}^H\mathbf{S}_{rr}^{-1/2}$ is the MMSE filter in canonical coordinates, and the matrix $\mathbf{S}_{ss}^{1/2}[\mathbf{I}_M - \mathbf{F}\mathbf{K}_p\mathbf{K}_p^H\mathbf{F}^H]\mathbf{S}_{ss}^{H/2}$ is the error covariance matrix in canonical coordinates. The matrix $\mathbf{K}_p$ is the MMSE filter for estimating the canonical coordinates $\mathbf{F}^H\mathbf{S}_{ss}^{-1/2}\mathbf{x}_s$ from the canonical coordinates $\mathbf{G}^H\mathbf{S}_{rr}^{-1/2}\mathbf{x}_r$, and the matrix $[\mathbf{I}_M - \mathbf{F}\mathbf{K}_p\mathbf{K}_p^H\mathbf{F}^H]$ is the error covariance matrix when doing so. As a consequence, we may interpret the coherence, or canonical coordinate detector Λ_4^{-1} as the volume of the error concentration ellipse when predicting the canonical coordinates of the surveillance channel signal from the canonical coordinates of the reference channel signal. When the channels are highly correlated, then this prediction is accurate, the volume of the error concentration ellipse is small, and $1 - \Lambda_4^{-1}$ is near to one, indicating a detection.*

Remark 5. Connection to Generalized Coherence. *If the covariance matrix under $\mathcal{H}_0$ were assumed only block diagonal and under $\mathcal{H}_1$ it were assumed an arbitrary psd matrix, the GLRT statistic would be the following generalized Hadamard ratio:*

$$H = \frac{\det(\mathbf{S})}{\det(\mathbf{S}_{ss})\det(\mathbf{S}_{rr})} = \prod_{i=1}^{M}(1-k_i^2) \quad (45)$$

Notice also that $1 - H$ is the Generalized Coherence (GC) originally defined in [36], and widely applied to multi-channel detection problems, as in [37], [38]. So under noise model 4 the net of prior knowledge of dimension p is to replace M by p in the coherence statistic.

VI. GLRTs for Models 2 and 3

Under noise models 2 and 3 closed-form GLRTs do not exist in general and one needs to resort to numerical methods. In this section, we briefly comment on three alternative approaches to the ML estimation problem. We give more specific details for the simplest and best performing algorithm, which is based on a careful reparameterization of the problem, and the application of the alternating minimization approach. The resulting algorithms can be used to obtain the ML estimates of $\mathbf{R}_{0,j}$ and $\mathbf{R}_{1,j}$ under noise models $j = 2, 3$ [4]. These estimates can then be plugged into the general expression (26) to get the corresponding GLR detector.

Let us start by introducing an important property of the sets of structured covariance matrices considered in this paper, which allows us to obtain relatively simple ML estimation algorithms:

Proposition 2. *The structure of the sets $\mathcal{E}_i$ is preserved under matrix inversion. That is*

$$\mathbf{\Sigma} \in \mathcal{E}_i \quad \Leftrightarrow \quad \mathbf{\Sigma}^{-1} \in \mathcal{E}_i. \quad (46)$$

Proof. The result directly follows from the (block)-diagonal structure of the matrices in the sets $\mathcal{E}_i$. □

Taking into account this property, and focusing on the alternative hypothesis $\mathcal{H}_1$ (the null can be treated similarly), we are ready to introduce two alternative approaches to the ML estimation problem:

1) Taking into account the monotonicity of the objective function in Problem (32), the equality constraint can be replaced by a $\leq$ constraint. If we also write the optimization problem in terms of the matrix $\mathbf{\Sigma}^{-1}$, one can readily see that the problem reduces to maximizing a concave function (the geometric mean of the $2M - p$ smallest eigenvalues) subject to an upper bound constraint on a concave function (the arithmetic mean of those eigenvalues) [32]. Since the problem is not convex due to the constraint, we can resort to a successive convex approximations approach [39] based on a linear approximation of the function in the constraint. This procedure results in an iterative algorithm with guaranteed convergence to a solution that satisfies the KKT conditions [32] of the original problem. However, the convex problems to be solved in each outer iteration of the algorithm do not admit a closed form solution, which results in a relatively slow convergence and high computational complexity.
2) A second approach, which can be seen as a generalization of the alternating optimization algorithm proposed in [12], consists in writing the log-likelihood function in eq. (18) in terms of the matrices $\mathbf{\Sigma}^{-1/2}$ and $\tilde{\mathbf{H}} = \mathbf{\Sigma}^{-1/2}\mathbf{H}$. Thus, for fixed $\mathbf{\Sigma}^{-1/2}$, the matrix $\tilde{\mathbf{H}}$ maximizing the log-likelihood can be easily found, and for fixed $\tilde{\mathbf{H}}$, the log-likelihood function is concave in $\mathbf{\Sigma}^{-1/2}$, whose optimal value can be obtained by means of a Newton method [12]. This approach also guarantees the convergence to

[4] Actually, the ML estimate of $\mathbf{R}_{0,2}$ can be obtained in closed form. For simplicity in the exposition, however, we consider in this section the iterative estimation of the covariance matrices under both hypotheses.

a stationary solution of the original problem. Its main drawback consists in the need for an (inner) algorithm for obtaining $\boldsymbol{\Sigma}^{-1/2}$, as well as a coupling between the two parameters ($\boldsymbol{\Sigma}^{-1/2}$ and $\tilde{\mathbf{H}}$), which also results into a relatively slow convergence.

In order to obtain a simpler, more intuitive, and faster algorithm, we rely on the following property of the sets of inverse covariance matrices associated with $\mathcal{R}_{i,j}$:

Proposition 3. *The sets of inverse covariance matrices* $\mathcal{P}_{i,j} = \left\{\mathbf{R}^{-1}, \textit{for } \mathbf{R} \in \mathcal{R}_{i,j}\right\}$ *can be written as*

$$\mathcal{P}_{i,j} = \left\{\mathbf{D} - \mathbf{G}\mathbf{G}^H, \textit{for } \mathbf{D} \in \mathcal{E}_j \textit{ and } \mathbf{D} \succeq \mathbf{G}\mathbf{G}^H\right\}. \quad (47)$$

In particular, $\mathbf{D} = \boldsymbol{\Sigma}^{-1}$ *and* $\mathbf{G}\mathbf{G}^H = \boldsymbol{\Sigma}^{-1}\mathbf{H}\left(\mathbf{I} + \mathbf{H}^H\boldsymbol{\Sigma}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^H\boldsymbol{\Sigma}^{-1}$, *or equivalently,* $\boldsymbol{\Sigma} = \mathbf{D}^{-1}$ *and* $\mathbf{H}\mathbf{H}^H = \mathbf{D}^{-1/2}\mathbf{U}\left(\mathbf{E}^{-1} - \mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{D}^{-1/2}$, *where* $\mathbf{U}$ *and* $\mathbf{E}$ *are the eigenvector and eigenvalue matrices in the EV decomposition* $\mathbf{D}^{-1/2}\mathbf{G}\mathbf{G}^H\mathbf{D}^{-1/2} = \mathbf{U}\mathbf{E}\mathbf{U}^H$.

Proof. Applying the Matrix Inversion Lemma [40], we can write

$$\begin{aligned}\mathbf{R}^{-1} &= \left(\mathbf{H}\mathbf{H}^H + \boldsymbol{\Sigma}\right)^{-1} \\ &= \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{H}\left(\mathbf{I} + \mathbf{H}^H\boldsymbol{\Sigma}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^H\boldsymbol{\Sigma}^{-1}, \quad (48)\end{aligned}$$

which allows us to easily identify $\mathbf{D} = \boldsymbol{\Sigma}^{-1}$ and $\mathbf{G}\mathbf{G}^H = \boldsymbol{\Sigma}^{-1}\mathbf{H}\left(\mathbf{I} + \mathbf{H}^H\boldsymbol{\Sigma}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^H\boldsymbol{\Sigma}^{-1}$. In order to recover $\mathbf{H}$ from $\mathbf{D}$ and $\mathbf{G}$, let us write $\tilde{\mathbf{H}} = \mathbf{D}^{1/2}\mathbf{H}$, which yields

$$\mathbf{D}^{-1/2}\mathbf{G}\mathbf{G}^H\mathbf{D}^{-1/2} = \mathbf{U}\mathbf{E}\mathbf{U}^H = \tilde{\mathbf{H}}\left(\mathbf{I} + \tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\right)^{-1}\tilde{\mathbf{H}}^H, \quad (49)$$

where in the first equality we have used the EV decomposition. Finally, writing the EV decomposition of $\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H$ as $\mathbf{U}_{\tilde{\mathbf{H}}}\mathbf{E}_{\tilde{\mathbf{H}}}\mathbf{U}_{\tilde{\mathbf{H}}}^H$ allows us to identify

$$\mathbf{U}_{\tilde{\mathbf{H}}} = \mathbf{U}, \qquad \mathbf{E}_{\tilde{\mathbf{H}}} = \left(\mathbf{E}^{-1} - \mathbf{I}\right)^{-1}, \quad (50)$$

which obviously requires $\mathbf{I} \succeq \mathbf{E}$, or equivalently $\mathbf{D} \succeq \mathbf{G}\mathbf{G}^H$. □

Thanks to Proposition 3, the ML estimation problem can be formulated in terms of the matrices $\mathbf{D}$ and $\mathbf{G}$ as

$$\begin{aligned}\underset{\mathbf{D},\mathbf{G}}{\text{maximize}} \quad & \operatorname{logdet}(\mathbf{D} - \mathbf{G}\mathbf{G}^H) - \operatorname{tr}\left[(\mathbf{D} - \mathbf{G}\mathbf{G}^H)\mathbf{S}\right], \\ \text{subject to} \quad & \mathbf{D} - \mathbf{G}\mathbf{G}^H \succeq \mathbf{0}, \\ & \mathbf{D} \in \mathcal{E}_j.\end{aligned} \quad (51)$$

Although this problem is still non-convex, it is formulated in a form suitable for applying the alternating optimization approach. Thus, for a fixed inverse noise covariance matrix $\mathbf{D} = \boldsymbol{\Sigma}^{-1}$, the problem of finding the optimal $\mathbf{G}$ reduces to

$$\underset{\mathbf{G}}{\text{maximize}} \quad \operatorname{logdet}(\mathbf{I} - \mathbf{D}^{-1/2}\mathbf{G}\mathbf{G}^H\mathbf{D}^{-1/2}) + \operatorname{tr}\left[\mathbf{G}\mathbf{G}^H\mathbf{S}\right], \quad (52)$$

or in terms of $\tilde{\mathbf{G}} = \mathbf{D}^{-1/2}\mathbf{G}$ and $\tilde{\mathbf{S}} = \mathbf{D}^{1/2}\mathbf{S}\mathbf{D}^{1/2}$

$$\underset{\tilde{\mathbf{G}}}{\text{maximize}} \quad \operatorname{logdet}(\mathbf{I} - \tilde{\mathbf{G}}\tilde{\mathbf{G}}^H) + \operatorname{tr}\left[\tilde{\mathbf{G}}\tilde{\mathbf{G}}^H\tilde{\mathbf{S}}\right]. \quad (53)$$

The solution of (53) can be found in a straightforward manner, and is given by any $\tilde{\mathbf{G}}$ of the form $\tilde{\mathbf{G}} = \tilde{\mathbf{W}}_p\left[\mathbf{I} - \tilde{\boldsymbol{\Lambda}}_p^{-1}\right]_+^{1/2}\mathbf{Q}$, where $\tilde{\mathbf{W}}_p$ and $\tilde{\boldsymbol{\Lambda}}_p$ are the matrices with the p principal eigenvectors and eigenvalues of the whitened sample covariance matrix $\tilde{\mathbf{S}}$, $\mathbf{Q}$ is an arbitrary unitary matrix, and $[\cdot]_+$ denotes the element-wise operation $\max(\cdot, 0)$. Finally, using Proposition 3, the optimal matrix $\mathbf{H}$ satisfies

$$\hat{\mathbf{H}}\hat{\mathbf{H}}^H = \boldsymbol{\Sigma}^{1/2}\tilde{\mathbf{W}}_p\left[\tilde{\boldsymbol{\Lambda}}_p - \mathbf{I}\right]_+\tilde{\mathbf{W}}_p^H\boldsymbol{\Sigma}^{1/2}. \quad (54)$$

Fixing the matrix $\mathbf{G}$, the optimization problem in (51) reduces to

$$\underset{\mathbf{D}\in\mathcal{E}_j}{\text{maximize}} \quad \operatorname{logdet}(\mathbf{D} - \mathbf{G}\mathbf{G}^H) - \operatorname{tr}\left[\mathbf{D}\mathbf{S}\right], \quad (55)$$

which is a convex optimization problem. Thus, taking the (constrained) gradient with respect to $\mathbf{D}$ yields

$$\nabla_{\mathbf{D}} = \boldsymbol{\Theta}_j\left[(\mathbf{D} - \mathbf{G}\mathbf{G}^H)^{-1} - \mathbf{S}\right], \quad (56)$$

where $\boldsymbol{\Theta}_j[\cdot]$ is an operator imposing the structure in $\mathcal{E}_j$. In particular, for a psd block matrix

$$\mathbf{X} = \begin{bmatrix}\mathbf{X}_{1,1} & \mathbf{X}_{1,2} \\ \mathbf{X}_{2,1} & \mathbf{X}_{2,2}\end{bmatrix} \quad (57)$$

with blocks of size M. The operators $\boldsymbol{\Theta}_j[\cdot]$ are

$$\boldsymbol{\Theta}_1[\mathbf{X}] = \frac{\operatorname{tr}(X)}{2M}\mathbf{I}_{2M}, \quad (58)$$

$$\boldsymbol{\Theta}_2[\mathbf{X}] = \frac{1}{M}\begin{bmatrix}\operatorname{tr}(\mathbf{X}_{1,1})\mathbf{I}_M & \mathbf{0}_M \\ \mathbf{0}_M & \operatorname{tr}(\mathbf{X}_{2,2})\mathbf{I}_M\end{bmatrix} \quad (59)$$

$$\boldsymbol{\Theta}_3[\mathbf{X}] = \begin{bmatrix}\operatorname{diag}(\mathbf{X}_{1,1}) & \mathbf{0}_M \\ \mathbf{0}_M & \operatorname{diag}(\mathbf{X}_{2,2})\end{bmatrix} \quad (60)$$

$$\boldsymbol{\Theta}_4[\mathbf{X}] = \begin{bmatrix}\mathbf{X}_{1,1} & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{X}_{2,2}\end{bmatrix} \quad (61)$$

and $\operatorname{diag}(\mathbf{A})$ denotes the diagonal matrix obtained from the diagonal elements of $\mathbf{A}$.

Going back to eq. (56), and noting that $(\mathbf{D} - \mathbf{G}\mathbf{G}^H)^{-1} = \mathbf{H}\mathbf{H}^H + \boldsymbol{\Sigma}$, we can conclude that the gradient is zero when $\boldsymbol{\Theta}_j\left[\mathbf{H}\mathbf{H}^H + \boldsymbol{\Sigma} - \mathbf{S}\right] = \mathbf{0}_{2M}$, and therefore the optimal $\boldsymbol{\Sigma}$ is given by

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Theta}_j\left[\mathbf{S} - \hat{\mathbf{H}}\hat{\mathbf{H}}^H\right]. \quad (62)$$

Finally, this overall alternating optimization approach for the ML estimation of the matrices $\boldsymbol{\Sigma}$ and $\mathbf{H}$ is summarized in Algorithm 1. The procedure can be initialized at $\hat{\boldsymbol{\Sigma}} = \mathbf{I}$ and typically it converges in a few iterations. Since at each step the value of the objective function can only increase, the method is guaranteed to converge to a (possibly local) maximum. While the alternating minimization approach does not guarantee that the global maximizer of the log-likelihood has been found, in the simulation experiments the resulting detector showed good performance, and we believe it can be safely taken as the GLRT for noise models 2 and 3.

Remark 6. *The GLRT for Model 2 is invariant to independent scaling of the surveillance and reference channels. As a special case of this invariance the detector* Λ_2 *is CFAR with respect to the noise power in the surveillance channel, and to the signal-plus-noise power in the reference chanel. The GLRT for Model 3 is invariant to independent diagonal scaling of the components in the surveillance and reference channels.*

Input: Sample covariance matrix $\mathbf{S}$, noise structure index j, and rank p.
Output: Estimates ($\hat{\mathbf{H}}$, $\hat{\mathbf{\Sigma}}$) of the channel $\mathbf{H}$ and $\mathbf{\Sigma} \in \mathcal{E}_j$.
Initialize: $\mathbf{\Sigma} = \mathbf{I}_{2M}$.
repeat
Fix $\mathbf{\Sigma}$ and update the estimate of the channel $\mathbf{H}$ with (54).
Fix the channel estimate $\hat{\mathbf{H}}$ and update $\hat{\mathbf{\Sigma}} = \mathbf{\Theta}_j \left[\mathbf{S} - \hat{\mathbf{H}}\hat{\mathbf{H}}^H \right]$.
until Convergence.

Algorithm 1: Proposed alternating optimization algorithm.

As a special case of this invariance, the detector Λ_3 is CFAR with respect to arbitrary unequal noise powers at the elements of the surveillance channel, and to unequal signal-plus-noise powers at the elements of the reference channel.

VII. Simulation Results

In this section we evaluate the performance of the GLR detectors for noise models 1-4 by means of Monte Carlo simulations. The input signal-to-noise-ratio (SNR) for both the surveillance and reference channels is defined as

$$\text{SNR}_i = 10 \log_{10} \frac{\text{tr}(\mathbf{H}_i^H \mathbf{H}_i)}{\text{tr}(\mathbf{\Sigma}_{ii})}, \qquad i = \{s, r\}. \qquad (63)$$

The noise at each channel follows a Gaussian distribution with covariance matrices ($\mathbf{\Sigma}_{ss}$,$\mathbf{\Sigma}_{rr}$), whose structure is determined by the spatial correlation model. For given values of SNR_s and SNR_r, the probability of detection, P_d, and probability of false alarm, P_{fa}, are estimated by averaging 10^4 independent simulations. For each choice of the channel matrices $(\mathbf{H}_s, \mathbf{H}_\mathbf{r})$ and noise covariance matrices $(\mathbf{\Sigma}_{ss}, \mathbf{\Sigma}_{rr})$, N values of $\mathbf{x}[n]$ are generated with $\mathbf{s}[n] = 0, n = 1, \ldots, N$, and N realizations of $\mathbf{x}[n]$ are generated with $\mathbf{s}[n], n = 1, \ldots, N$ drawn from unit normals. From these N values under each hypothesis, detection statistics are computed. As thresholds are scanned, a false alarm (or not) and a detection (or not) is recorded. This is repeated 10^4 times, but at each repeat, a different set of the $(\mathbf{H}_s, \mathbf{H}_\mathbf{r})$ and $(\mathbf{\Sigma}_{ss}, \mathbf{\Sigma}_{rr})$ are drawn. The elements of $(\mathbf{H}_s, \mathbf{H}_\mathbf{r})$ are drawn as unit normals and scaled to give the desired SNR's at each channel as in (63), and the $(\mathbf{\Sigma}_{ss}, \mathbf{\Sigma}_{rr})$ are constructed according to one of the assumed noise models. We use the following generation models for the noise covariance matrices:

- Model 1: $\mathbf{\Sigma}_{ss} = \mathbf{\Sigma}_{rr} = \mathbf{I}$.
- Model 2: $\mathbf{\Sigma}_{ss} = \sigma_s^2 \mathbf{I}$, $\mathbf{\Sigma}_{rr} = \sigma_r^2 \mathbf{I}$; with $\sigma_r^2 \sim \mathcal{U}(0, 1)$ and $\sigma_s^2 \sim \mathcal{U}(0, 1)$ independent uniform random variables.
- Model 3: $\mathbf{\Sigma}_{ss} = \text{diag}(\sigma_{s1}^2, \ldots, \sigma_{sM}^2)$, $\mathbf{\Sigma}_{rr} = \text{diag}(\sigma_{r1}^2, \ldots, \sigma_{rM}^2)$; with $\sigma_{ri}^2 \sim \mathcal{U}(0, 1)$ and $\sigma_{si}^2 \sim \mathcal{U}(0, 1)$ for $i = 1, \ldots, M$.
- Model 4: $\mathbf{\Sigma}_{ss} = \mathbf{A}^H \mathbf{A}$ and $\mathbf{\Sigma}_{rr} = \mathbf{B}^H \mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ are random matrices with independent elements drawn from unit complex normals.

A. GLRT performance

We first evaluate the performance of the GLR tests under different noise models. For models 1 and 4, we used the closed-form GLRTs in Sections V-A and V-B, respectively; whereas for models 2 and 3 we used the iterative solution described in Section VI. The results shown in this subsection involve a scenario with $p = 1$ (dimension-one subspace signal), $M = 5$ antennas and $N = 50$ snapshots. Fig. 2 depicts the Receiver Operating Characteristic (ROC) curve when the noise is generated according to model 1 and the channel matrices are scaled to give $\text{SNR}_s = -6$ dB at the surveillance channel and $\text{SNR}_r = 10$ dB at the reference channel. In this situation, the GLRT for model 1 is matched to the measurement, and model 2-4 detectors are mismatched.

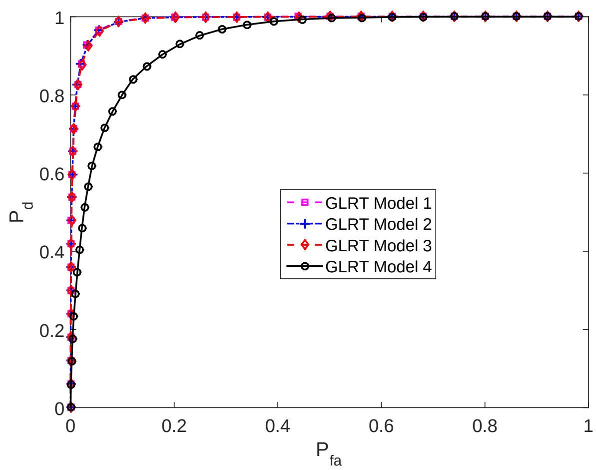

Fig. 2. ROC for a scenario under model 1 with $M = 5$ antennas, $p = 1$, $N = 50$ snapshots, $\text{SNR}_s = -6$ dB and $\text{SNR}_r = 10$ dB.

Fig. 3 depicts the ROC when the noise follows model 2 (spatially white noises but with different variance at each channel), $\text{SNR}_s = -5$ dB and $\text{SNR}_r = 5$ dB; hence, the model 2 detector is matched to the measurement and the others are mismatched. Similarly, Figs. 4 and 5 show the ROC curves for noise models 3 and 4, respectively, at the SNRs indicated in the figures.

As one would expect, the GLRTs matched to the actual noise model that generates the measurements outperform the mismatched detectors. In terms of robustness against mismatched noise models, the GLRT for model 3 seems to be the preferred option.

B. Distribution under the null

An important aspect regarding the applicability of the proposed tests is selecting thresholds that achieve a desired P_{fa}. A rigorous solution to this problem involves deriving the distribution of the proposed test statistics under the null hypothesis. Although one may find in the literature exact solutions for a few specific cases (mainly when the GLRT reduces to a Hadamard ratio as in [36], [38]), deriving the exact null distribution is in general not possible and one has to resort to asymptotic approximations.

A conventional approach is provided by the Wilks theorem [41], which proves that, for nested hypotheses and under some regularity conditions, when $N \to \infty$ the test statistic

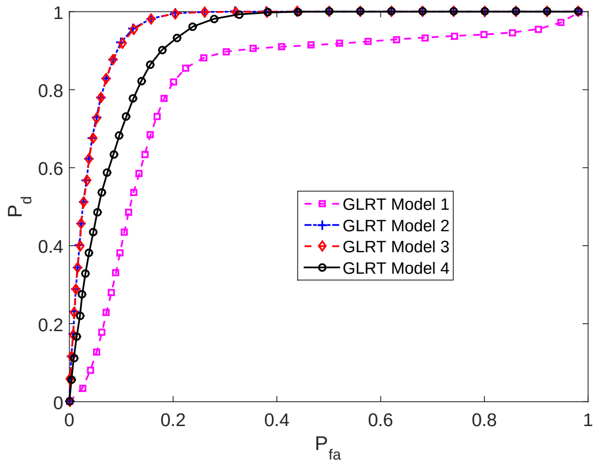

Fig. 3. ROC for a scenario under model 2 with $M = 5$ antennas, $p = 1$, $N = 100$ snapshots, $\text{SNR}_s = -5$ dB and $\text{SNR}_r = 5$ dB.

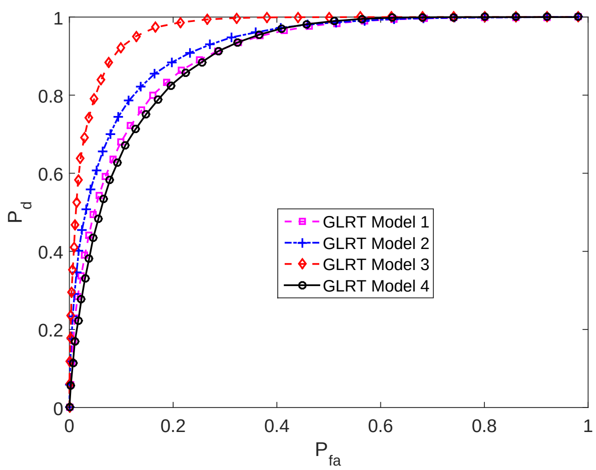

Fig. 4. ROC for a scenario under model 3 with $M = 5$ antennas, $p = 1$, $N = 50$ snapshots, $\text{SNR}_s = -6$ dB and $\text{SNR}_r = 10$ dB.

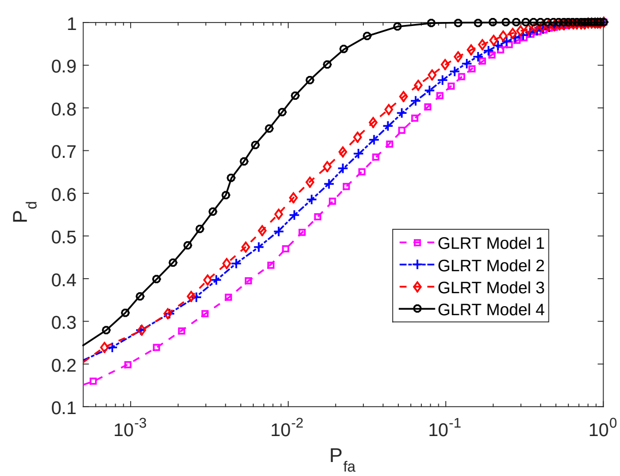

Fig. 5. ROC for a scenario under model 4 with $M = 5$ antennas, $p = 1$, $N = 50$ snapshots, $\text{SNR}_s = -10$ dB and $\text{SNR}_r = 0$ dB.

$2 \log \Lambda$ converges to a chi-squared distribution with degrees of freedom equal to the difference in dimensionality of the parameters in H_1 and H_0. For noise models 1-3, the Wilks approximation reduces to $2 \log \Lambda \sim \chi^2_{2Mp}$, which is only accurate for large values of N. To illustrate this point, Fig. 6 compares the empirical cumulative distribution function (CDF) for the GLRT under model 1 when $\text{SNR}_r = 0$ dB with a chi-square distribution with $2Mp$ degrees of freedom. The rate of convergence to the chi-squared distribution depends on the subspace dimension, as well as on the signal-to-noise-ratio of the reference channel, SNR_r, with faster convergence rates for lower values of p and higher values of SNR_r. A similar result is depicted in Fig. 7 for the test statistic under the noise model 2.

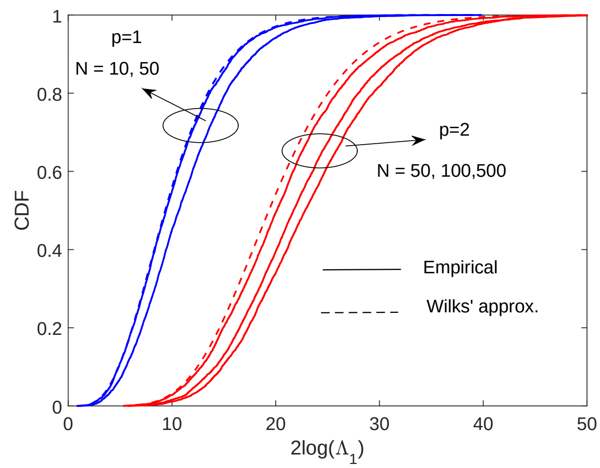

Fig. 6. Empirical CDF for the GLRT under the null for noise model 1 in solid line for different values of p and N with $\text{SNR}_r = 0$ dB. The dashed line is the Wilks approximation given by χ^2_{2Mp}.

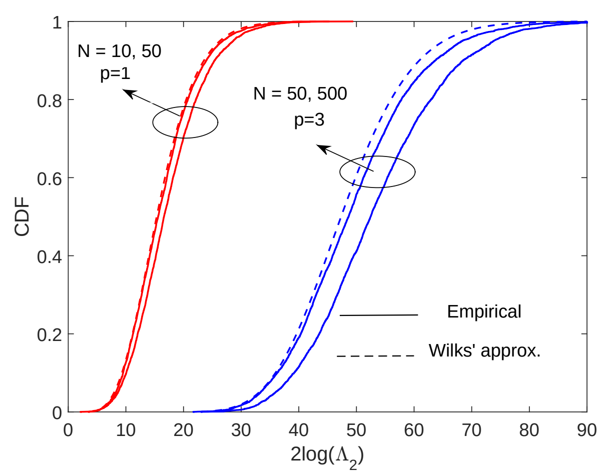

Fig. 7. Empirical CDF for the GLRT under the null for noise model 2 in solid line for different values of p and N with $\text{SNR}_r = 5$ dB. The dashed line is the Wilks approximation given by χ^2_{2Mp}.

For noise model 4, the classical Wilks approximation, including a correction term as proposed by Bartlett in [42], is accurate only for a full-rank signal model ($p = M$). For low-rank signal models, however, alternative approaches to approximate the null distribution are needed [43], [44]. The case of rank-one signals ($p = 1$) has been discussed in [5], where a random matrix result by Johnstone [45] was exploited: after an appropriate transformation, the distribution of the largest canonical correlation under the null converges to a Tracy-Widom law of order 2. In particular, let $l_1 = \log(k_1^2/(1-k_1^2))$, be the logit transform of the largest squared canonical correlation. Then, as $N \to \infty$, $M \to \infty$, $M/N \to \text{constant}$, the limiting distribution is

$$\mathcal{P}\left\{\frac{l_1 - \mu_{N,M}}{\sigma_{N,M}} \leq x\right\} \to F_2(x), \tag{64}$$

where $F_2(x)$ is the distribution function for the Tracy-Widom law of order 2. The centering and scaling constants are given by

$$\mu_{N,M} = \frac{\sigma_1^{-1} u_1 + \sigma_2^{-1} u_2}{\sigma_1^{-1} + \sigma_2^{-1}}, \quad \text{and} \quad \sigma_{N,M}^{-1} = \frac{1}{2}\left(\sigma_1^{-1} + \sigma_2^{-1}\right), \tag{65}$$

where $u_1 = 2\log\tan(2\alpha)$, $u_2 = 2\log\tan(2\beta)$, and

$$\sigma_1^3 = \frac{16}{(N+1)^2 \sin^2(4\alpha)\sin^2(2\alpha)},$$
$$\sigma_2^3 = \frac{16}{(N-1)^2 \sin^2(4\beta)\sin^2(2\beta)},$$

with

$$\sin^2(\alpha) = \frac{M + 1/2}{N+1}, \quad \text{and} \quad \sin^2(\beta) = \frac{M - 1/2}{N-1}.$$

The accuracy of the Tracy-Widom approximation for $p = 1$ is verified in Fig. 8, which shows the CDF of the random variable $\frac{l_1 - \mu_{N,M}}{\sigma_{N,M}}$ under the null, and the unitary Tracy-Widom distribution $F_2(x)$ in dashed line. In this example the number of antennas is $M = 5$ and the number of snapshots is $N = 20$. For values $p > 1$, the GLRT involves a function of the p largest sample canonical correlations and, to the best of our knowledge, there are no accurate approximations of the null distribution in the literature. This is left for future work.

As a final remark, notice that the Wilks approximation for noise models 1-3 and the Tracy-Widom approximation for model 4 depend not on the unknowns (e.g., channels matrices or noise covariance matrices), but only on known parameters such as the number of antennas M, the dimension p of the signal subspace, or the number of samples N (for the Tracy-Widom approximation). These results are consistent with the invariances established for each of the detectors in Section III.

C. Comparison with other detectors

In this subsection we compare the performance of the GLRT for noise model 4 and $p = 1$ with the following suboptimal detectors:

1) Covariance-matching detector: this ad-hoc detector uses structured estimates for $\mathbf{R}_0$ and $\mathbf{R}_1$ that minimize the Frobenius norm between the sample covariance and the estimate: $||\mathbf{S} - \mathbf{R}||_F^2$. For $\mathbf{R}_0$ the covariance-matching estimate coincides with the ML estimate in (42), whereas for $\mathbf{R}_1$ it is given by (for $p = 1$)

$$\hat{\mathbf{R}}_1 = \begin{bmatrix} \mathbf{S}_{ss} & \lambda_1 \mathbf{u}_1 \mathbf{v}_1^H \\ \lambda_1 \mathbf{v}_1 \mathbf{u}_1^H & \mathbf{S}_{rr} \end{bmatrix}, \tag{66}$$

where λ_1 is the maximum singular value of $\mathbf{S}_{sr}$, and $\mathbf{u}_1$ and $\mathbf{v}_1$ are the corresponding left and right singular vectors. After obtaining the covariance-matching estimates for $\mathbf{R}_0$ and $\mathbf{R}_1$, the test statistic is computed as a ratio of determinants.

2) Cross-correlation (Cross-Corr) detector:

$$|\operatorname{tr}(\mathbf{S}_{sr}^H \mathbf{S}_{sr})| \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{67}$$

which is a natural extension to the multiantenna case of the cross-correlation detector typically used in passive radar systems [3].

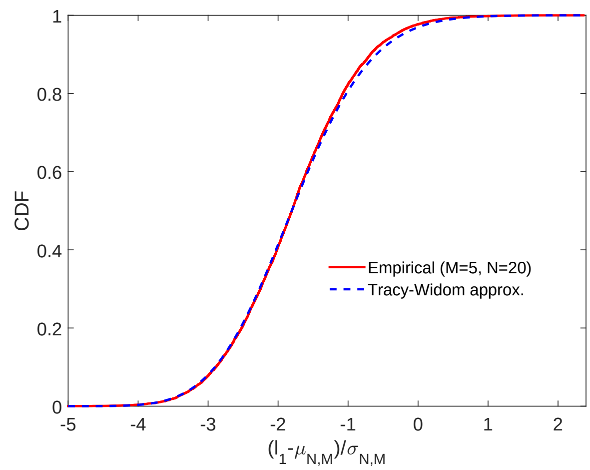

Fig. 8. Empirical CDF for the random variable $\frac{l_1 - \mu_{N,M}}{\sigma_{N,M}}$ under the null for noise model 4; $p = 1$, $M = 5$, $N = 20$ and $\text{SNR}_r = 0$ dB. The dashed line is the approximation given by Tracy-Widom distribution of order 2.

We consider a scenario with $M = 4$ antennas, $p = 1$, and $N = 100$ snapshots. Fig. 9 depicts the probability of detection P_d versus the signal-to-noise-ratio (for simplicity, in this example we assume $\text{SNR}_r = \text{SNR}_s$) for a fixed $P_{fa} = 10^{-3}$. The threshold to achieve the desired P_{fa} for the GLRT has been obtained from the Tracy-Widom approximation presented previously. We observe that the GLRT outperforms the covariance matching and the Cross-Corr detectors.

D. Performance with QPSK and OFDM signals

The GLR tests in this paper have been derived under the assumption that the signals transmitted by the p illuminators of opportunity follow a zero-mean circular complex Gaussian distribution. While this assumption was made for mathematical tractability, it is also an accurate approximation when the non-cooperative illuminators transmit orthogonal frequency-division multiplexing (OFDM) signals as in the European Digital Video Broadcasting-Terrestrial (DVB-T) system. To validate this approximation, in this subsection we compare the

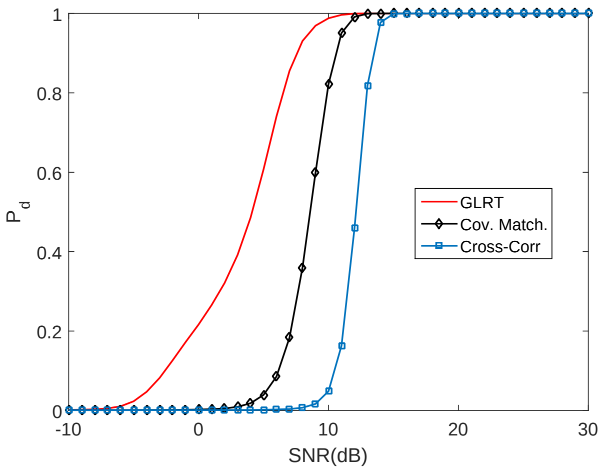

Fig. 9. P_d curves versus SNR for a scenario with $p = 1$, $M = 4$ antennas and $N = 100$ snapshots; $P_{fa} = 10^{-3}$.

performance of the GLRT for noise model 4 when $s[n]$ is: i) a zero-mean complex Gaussian signal, ii) an OFDM-modulated DVB-T signal in 2k mode (1705 subcarriers), and iii) a single-carrier quadrature phase-shift keying (QPSK) signal. The results are shown in Fig. 10 for a scenario with $M = 5$, $p = 1$, and $N = 50$ snapshots. The ROC curves for Gaussian, OFDM, and QPSK symboling are nearly indistinguishable.

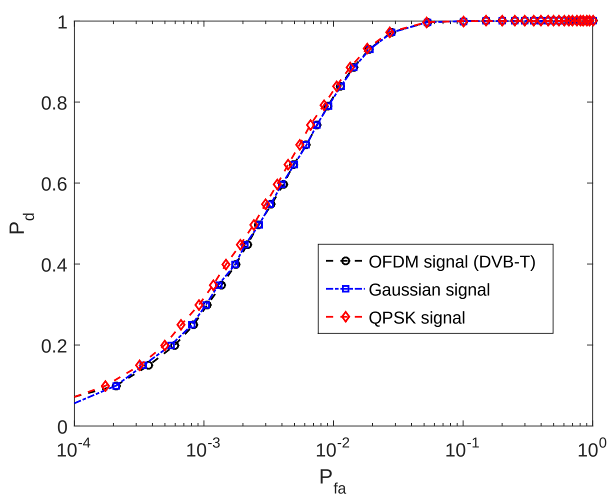

Fig. 10. ROC under model 4 with Gaussian, OFDM (DVB-T), and QPSK signals. The scenario has $M = 5$ antennas, $p = 1$, $N = 50$ snapshots, $\text{SNR}_s = -10$ dB and $\text{SNR}_r = -5$ dB.

E. Performance for unequal number of sensors at the surveillance and reference channels

To ease the exposition and simplify the notation, throughout this paper we also made the assumption that the reference and the surveillance channels have the same number of antennas, M. It is however clear that all results in the paper remain valid when the surveillance and the reference channels have M_s and M_r antennas, respectively. In this subsection, we evaluate the performance of the GLR tests in this situation. We consider a scenario with a fixed number of $M_r = 4$ antennas for the reference channel and a varying number of antennas ranging from $M_s = 3$ to $M_s = 12$ for the surveillance channel. The rank of the transmitted signal is $p = 1$, the noise is generated according to model 4, and the channel matrices are scaled to give $\text{SNR}_s = -10$ dB and $\text{SNR}_r = 0$ dB. The P_{fa} is fixed to $1e-2$ and the GLR tests are computed from $N = 50$ snapshots. Fig. 11 shows the P_d for the 4 GLR tests under these conditions for an increasing number of antennas at the surveillance channel. The GLRT for noise model 4 is matched to the generated measurements and thus is the best performing detector, achieving a probability of detection close to one for $M_s = 6$ antennas.

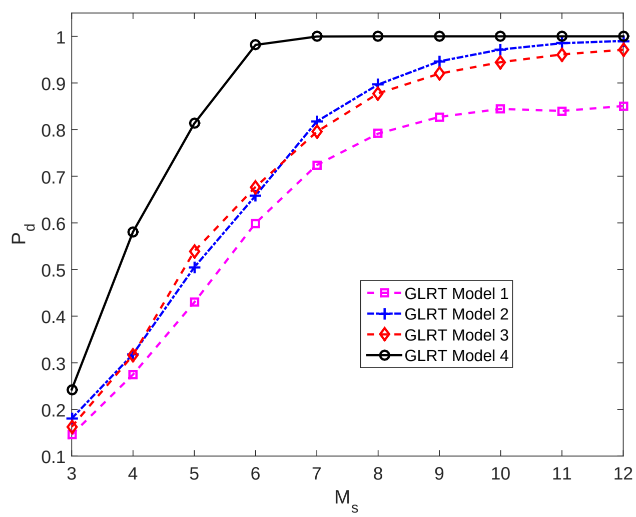

Fig. 11. Probability of detection of a rank-one signal for a fixed $P_{fa} = 1e-2$ under noise model 4 with $M_r = 4$ antennas in the reference channel and a varying number of antennas, M_s, in the surveillance channel.

VIII. Conclusion

In this paper we have addressed a problem motivated by passive radar. The problem is to detect a common subspace signal in two MIMO channels. It turns out that the problem is a problem in factor analysis, where there are constraints on the factor loadings and the factors. The problem may be viewed as a one-channel factor analysis problem with constraints on the factor loadings under the null hypothesis, or as a two-channel factor analysis problem, with constraints on the factor loadings under the null, and with common factors under the alternative.

There are four plausible additive noise models for the hypothesis testing problems we have studied, but each may be formulated in a common framework, using a noise-whitening trick that leads to a common problem of choosing a whitening matrix that minimizes the geometric mean of what might be called constrained canonical coordinates, under a constraint on their arithmetic mean. Two of the four problems have closed-form solutions, and two require numerical optimization, based on alternating minimizations. For each noise model, the invariances of the hypothesis testing problem and its GLRT are established.

A new result has been derived for maximum likelihood (ML) estimation of structured covariance matrices in the multivariate normal model: the ML estimate for a structured covariance matrix $\mathbf{R}$ in a cone class always satisfies the constraint $\text{tr}[\hat{\mathbf{R}}^{-1}\mathbf{S}] = 2M$, leading to the result that the GLRT for covariance testing is a ratio of determinants in estimated covariance matrices.

For the case of unstructured noise covariance, the GLRT compares the product $\prod_1^p(1-k_i^2)$ to a threshold, where the k_i's are squared canonical correlations of the two channel sample covariance matrix. The product may be replaced by $1-\prod_1^p(1-k_i^2)$, which is coherence, so that the detector is a coherence detector. There is a filtering interpretation of this result, showing that $\prod_1^p(1-k_i^2)$ is the volume of the error concentration ellipse when estimating canonical coordinates in the surveillance channel from canonical coordinates in the reference channel. When the channels are highly correlated, this concentration ellipse has small volume and coherence is near to one. Numerical simulations suggest that the detector based on common white noise variances in the surveillance and reference channels is badly mismatched to the other models, and should not be considered. In terms of robustness against mismatched noise models, the GLRT for model 3 should be the preferred detector.

IX. Acknowledgments

The authors wish to acknowledge fruitful discussions of this problem with Douglas Cochran, Stephen Howard, Songsri Sirianunpiboon, Dan Hack, and Lee Patton.

References

[1] H. D. Griffiths and C. J. Baker, "Passive coherent location radar systems. Part 1: Performance prediction," *IEE Proc. Radar Sonar Navig.*, vol. 152, no. 3, pp. 124–132, Jun. 2005.

[2] H. D. Griffiths and N. R. W. Long, "Television based bistatic radar," *IEE Proc. F. Commun. Radar Signal Process.*, vol. 133, pp. 649–657, 1986.

[3] J. Liu, H. Li, and B. Himed, "On the performance of the cross-correlation detector for passive radar applications," *Signal Process.*, vol. 113, pp. 32–37, 2015.

[4] H. Zhao, J. Liu, Z. Zhang, H. Liu, and S. Zhou, "Linear fusion for target detection in passive multistatic radar," *Signal Processing*, vol. 130, pp. 175–182, 2017.

[5] I. Santamaria, L. L. Scharf, D. Cochran, and J. Via, "Passive detection of rank-one signals with a multiantenna reference signal," in *European Signal Processing Conference (EUSIPCO)*, Budapest, Hungary, Sep. 2016.

[6] D. E. Hack, L. K. Patton, B. Himed, and M. A. Saville, "Detection in passive MIMO radar networks," *IEEE Trans. Signal Process.*, vol. 62, no. 11, pp. 2999–3012, Jun. 2014.

[7] G. Cui, J. Liu, H. Li, and B. Himed, "Target detection for passive radar with noisy reference channel," in *Proceedings of 2014 IEEE Radar Conference*, Cincinnati (OH), USA, May 2014.

[8] S. Gogineni, P. Setlur, M. Rangaswamy, and R. R. Nadakuditi, "Random matrix theory inspired passive bistatic radar detection of low-rank signals," in *IEEE Int. Conf. Acoust., Speech and Signal Proc. (ICASSP)*, Brisbane, Australia, Apr. 2015.

[9] O. Besson, S. Kraut, and L. L. Scharf, "Detection of an unknown rank-one component in white noise," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2835–2839, Jul. 2006.

[10] O. Besson and L. L. Scharf, "CFAR matched direction detector," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2840–2844, Jul. 2006.

[11] D. Hack, L. Patton, and B. Himed, "Multichannel detection of an unknown rank-one signal with uncalibrated receivers," in *IEEE Int. Conf. Acoust., Speech and Signal Proc. (ICASSP)*, Florence, Italy, May 2014.

[12] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, and I. Santamaria, "Detection of rank-p signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.

[13] D. E. Hack, C. W. Rossler, and L. K. Patton, "Multichannel detection of an unknown rank-N signal using uncalibrated receivers," *IEEE Signal Process. Lett.*, vol. 21, no. 8, pp. 998–1002, Aug. 2014.

[14] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 814–823, Feb. 2010.

[15] J. Sala-Alvarez, G. Vazquez-Vilar, and R. Lopez-Valcarce, "Multi-antenna GLR detection of rank-one signals with unknown power spectrum in white noise with unknown spatial correlation," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3065–3078, Jun. 2012.

[16] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Multiple-channel detection of signals having known rank," in *IEEE Int. Conf. Acoust., Speech and Signal Proc. (ICASSP)*, Vancouver, Canada, May 2013, pp. 6536–6540.

[17] S. D. Howard, S. Sirianunpiboon, and D. Cochran, "An exact Bayesian detector for multistatic passive radar," in *50th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2016.

[18] J. K. Tugnait, "On multiple antenna spectrum sensing under noise variance uncertainty and flat fading," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1823–1832, Apr. 2012.

[19] A. Leshem and A.-J. Van der Veen, "Multichannel detection of Gaussian signals with uncalibrated receivers," *IEEE Signal Process. Lett.*, vol. 8, no. 4, pp. 120–122, Apr. 2001.

[20] R. Zhang, T. Lim, Y.-C. Liang, and Z. Y., "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. on Comm.*, vol. 58, no. 1, pp. 84–88, Jan. 2010.

[21] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. Academic Press, 1979.

[22] K. G. Joreskog, "Some contributions to maximum likelihood factor analysis," *Psychometrika*, vol. 32, pp. 443–482, 1967.

[23] X. Zhang, H. Li, J. Liu, and B. Himed, "Joint delay and Doppler estimation for passive sensing with direct-path interference," *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 630–640, Feb. 2016.

[24] R. Tao, Z. Wu, and T. Shan, "Direct-path supression by spatial filtering in digital television terrestrial broadcasting-based passive radar," *IET Radar, Sonar, Navig.*, vol. 4, no. 6, pp. 791–805, Dec. 2010.

[25] P. Banelli and S. Cacopardi, "Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels," *IEEE Trans. Comm.*, vol. 48, no. 3, pp. 430–441, Mar. 2000.

[26] *ETSI, Digital Video Broadcasting (DVB); Frame Structure Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2)*, vol. ETSI EN 302 755, 2009.

[27] J. Villares and G. Vazquez, "The Gaussian assumption in second-order estimation problems in digital communications," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4994–5002, Oct. 2007.

[28] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.

[29] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol II: Detection Theory*. Upper Saddle River, NJ: Prentice Hall, 1998.

[30] S. Kay and R. Gabriel, "An invariance property of the generalized likelihood ratio test," *IEEE Signal Processing Letters*, vol. 10, no. 12, pp. 352–355, Dec. 2003.

[31] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach*. Wiley Interscience, 1983.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[33] J. Mauchly, "Significance test for sphericity of a normal n-variate distribution," *Ann. Math. Statist.*, vol. 11, pp. 204–209, 1940.

[34] Y. Wang, I. Santamaria, L. L. Scharf, and H. Wang, "Canonical coordinates for target detection in a passive radar network," in *50th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove (CA), USA, Nov. 2016.

[35] P. Stoica, K. M. Wong, and Q. Wu, "On a nonparametric detection method for array signal processing in correlated noise fields," *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 1030–1032, Apr. 1996.

[36] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple channel signal detection," *IEEE Trans. Signal Process.*, vol. 43, pp. 2049–2057, Sep. 1995.

[37] D. Ramirez, J. Via, I. Santamaria, and L. L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5006–5015, Oct. 2010.

[38] N. Klausner, M. R. Azimi-Sadjadi, and L. L. Scharf, "Detection of spatially correlated time series from a network of sensor arrays," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1396–1407, Mar. 2014.

[39] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.

[40] M. A. Woodbury, "Inverting modified matrices," *Memorandum report*, vol. 42, p. 106, 1950.

[41] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses," *Ann. Math. Statist.*, vol. 9, no. 1, pp. 60–62, 1938.

[42] M. S. Bartlett, "The statistical significance of canonical correlations," *Biometrika*, vol. 32, no. 1, pp. 29–37, Jan. 1941.

[43] X. Mestre, P. Vallet, and W. Hachem, "Correlation test for high dimensional data with application to signal detection in sensor networks," in *Proceedings of the European Signal Processing Conference 2014 (EUSIPCO)*, Lisbon, Portugal, Sep. 2014.

[44] N. Klausner, M. R. Azimi-Sadjadi, and L. L. Scharf, "Saddlepoint approximations for correlation testing among multiple Gaussian random vectors," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 703–707, May 2016.

[45] I. M. Johnstone, "Multivariate analysis and Jacobi ensembles: Largest eigenvalue, Tracy-Widom limits and rates of convergence," *Ann. Statist.*, vol. 36, no. 6, pp. 2638–2716, 2008.