A Unified Theory of Adaptive Subspace Detection Part I: Detector Designs

Danilo Orlando, *Senior Member, IEEE*, Giuseppe Ricci, *Senior Member, IEEE*, and Louis L. Scharf, *Life Fellow, IEEE*

***Abstract*—This paper addresses the problem of detecting multidimensional subspace signals in noise of unknown covariance. It is assumed that a primary channel of measurements, possibly consisting of signal plus noise, is augmented with a secondary channel of measurements containing only noise. The noises in these two channels share a common covariance matrix, up to a scale, which may be known or unknown. The signal model is a subspace model with variations: the subspace may be known or known only by its dimension; consecutive visits to the subspace may be unconstrained or they may be constrained by a prior distribution. The several original detectors derived in this paper, when organized with previously published detectors, comprise a unified theory of adaptive subspace detection from primary and secondary channels of measurements.**

***Index Terms*—Adaptive detection, subspace model, generalized likelihood ratio test, alternating optimization.**

I. Introduction

IN REAL radar systems equipped with an array of sensors, the array mainbeam is steered by applying specific weights to each tile. However, very often, due to hardware, implementation, and/or architecture issues, setting these weights becomes a difficult task for the presence of unbalanced channels, miscalibration errors, mutual coupling, and so on [1], [2], [3], [4]. As a consequence, an intrinsic uncertainty related to the array pointing direction might exist. The subspace paradigm arises from the need to account for this uncertainty and to control the detection performance degradation due to the presence of mismatched signals [4]. The general problem of matched and adaptive subspace detection of point-like targets in Gaussian and non-Gaussian disturbance has been addressed by many authors, beginning with the seminal work of Kelly and Forsythe [5], [6].

Manuscript received 4 April 2022; revised 14 July 2022; accepted 29 August 2022. Date of publication 12 September 2022; date of current version 25 October 2022. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yao Xie. The work of Louis L. Scharf was supported in part by the US Office of Naval Research under Grant N00014-21-1-2145 and in part by the US Airforce Office of Scientific Research under Grant FA9550-21-1-0169. *(Corresponding author: Giuseppe Ricci.)*

Danilo Orlando is with the Università degli Studi "Niccolò Cusano", 00166 Roma, Italy (e-mail: danilo.orlando@unicusano.it).

Giuseppe Ricci is with the Dipartimento di Ingegneria dell'Innovazione, Università del Salento, 73100 Lecce, Italy (e-mail: giuseppe.ricci@unisalento.it).

Louis L. Scharf is with the Departments of Mathematics and Statistics, Colorado State University, Fort Collins, CO 80521 USA (e-mail: scharf@colostate.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TSP.2022.3205761, provided by the authors.

Digital Object Identifier 10.1109/TSP.2022.3205761

The innovation of [5] was to introduce a *homogeneous* secondary channel of signal-free measurements whose unknown covariance matrix was equal to the unknown covariance matrix of primary (or test) measurements. Likelihood theory was then used to derive what is now called the Kelly detector. In [6], adaptive subspace detection was formulated in terms of the so-called generalized multivariate analysis of variance for complex variables. These papers were followed by the important adaptive detectors of [7], [8]. Then, a scale-invariant *adaptive subspace detector*, now commonly called ACE (adaptive coherence estimator), was introduced. In [9] this detector was derived as an asymptotic approximation to the generalized likelihood ratio (GLR) to detect a coherent signal in compound-Gaussian noise with known spectral properties, and in [10] it was derived as an *estimate-and-plug* version of the scale-invariant *matched subspace detector* [11], [12]. Interestingly, in [13] the authors showed that ACE was a likelihood ratio detector for a *non-homogeneous* secondary channel of measurements whose unknown covariance matrix was a scaled version of the unknown covariance matrix of the primary channel. The scale was unknown. Then, in [14] it was shown that ACE is a uniformly most powerful invariant (UMPI) detector. In subsequent years there has been a flood of important papers. Among published references on adaptive detection we cite here [15], [16], [17], [18], [19], [20], [21], [22], [23], and references therein. All of this work is addressed to adaptive detection in what might be called a first-order (signal) model for measurements. That is, the measurements under test may contain a signal in a known subspace embedded in Gaussian noise of unknown covariance, but no prior distribution is assigned to the location of the signal in the subspace. In particular, in [17], [19] the authors extend adaptive subspace detection to range-spread targets deriving likelihood ratio detectors that were then compared to estimate-and-plug adaptations. The first attempt to replace this model by a second-order (signal) model was made in [24], where the authors used a Gaussian model for the signal. The covariance matrix for the signal was constrained by a known subspace model. The resulting *second-order matched subspace detector* was derived [24], and an estimate-and-plug adaptation from secondary measurements was proposed.

The aim of the current paper is to extend the results of [17], [19], [24], [25], [26] to include all variations on adaptive subspace detection in first- and second-order models for a subspace signal to be detected. These models include signals that lie

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

in a known subspace or in an unknown subspace of known dimension. They will be clarified in due course.

Our results are motivated by the problem of detecting range-spread targets from an active radar system. However, our framework and corresponding results are actually much more generally applicable, as they apply to sonar, communications, hyperspectral imaging, radioastronomy, etc. In all of these applications, measurements in a primary channel may contain signal plus Gaussian noise. Measurements in a secondary channel contain only noise. The noises in the two channels are independent, but they share a common covariance matrix, at least to within an unknown scale. The case of a common covariance matrix in the two channels is typically referred to as a case of *homogeneous environment*, while the more general case of an unknown scale factor is commonly referred to as a case of *partially-homogeneous environment*. As for the signal components, they are determined by a visit to a subspace. According to a *first-order model* for these visits, there is no constraint on their location in the subspace; as a consequence the subspace signal model modulates the mean of a multivariate Gaussian distribution. According to a *second-order model*, the location in the subspace is ruled by a prior distribution, which is taken to be a Gaussian distribution; as a consequence the subspace signal model modulates the covariance matrix of a multivariate Gaussian distribution. For each of these variations on the problem of adaptively detecting a subspace signal, we derive a detector based upon the GLR, or generalized likelihood ratio test (GLRT) (for the definition of GLRT see [27]). Recall that the GLRT compares a GLR statistic to a threshold η, set according to the desired probability of false alarm (P_{fa}), to discriminate between the noise-only hypothesis (H_0) and the signal-plus-noise hypothesis (H_1). Hereafter, η will denote any modification of the original threshold. Taken together, our results comprise a unified theory of adaptive subspace detection.

A. A Preview of the Paper

Before proceeding with the derivations, we summarize below the different variations on a multidimensional subspace signal model addressed in this paper:

- The signal visits a *known subspace, unconstrained by a prior distribution.* We call this a *first-order model*, as the signal appears as a *low-rank component* in the mean of a multivariate Gaussian distribution for the measurements. When there is only one measurement in the primary channel, then the GLRTs are those of [5], [9], [13]. For multiple measurements these results are extended in [6], [17], [19]. The structured interference of [19] is not considered in the present paper. These cases (developed in Section III-A and III-B) are only reviewed, as they form the basis of our extensions to other models.
- The signal visits an *unknown subspace of known dimension, unconstrained by a prior distribution.* Again we call this a *first-order model*. The GLRTs are original to the best of authors' knowledge and are derived in Section III-C and III-D.
- The signal visits a *known subspace, constrained by a Gaussian prior distribution.* We call this a *second-order model*, as the signal model appears as a *constrained, low-rank component* in the covariance matrix of a multivariate Gaussian distribution for the measurements. Adaptive estimate-and-plug GLRTs for this case have been derived in [24], [28]. The approximated GLRTs of the current paper are original and generalize the GLRT designed to detect a rank one signal in homogeneous environment [25].
- The signal visits an *unknown subspace of known dimension, constrained by a Gaussian prior distribution*; this is a *second-order model*. The estimated low-rank covariance matrix for the subspace signal may be called an adaptive factor model. The resulting GLRTs are original and they significantly extend the results of [26] for detecting a dimension-one signal in a homogeneous environment.

B. Notation

In the sequel, vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. Symbols $\det(\cdot)$, $\mathrm{Tr}(\cdot)$, $\mathrm{etr}\{\cdot\}$, $\mathrm{rk}\{\cdot\}$, $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^\dagger$, $(\cdot)^{-1}$, and $(\cdot)^{-\dagger}$ denote the determinant, trace, exponential of the trace, rank, transpose, complex conjugate, conjugate transpose, inverse, and conjugate transpose of the inverse, respectively. As to numerical sets, $\mathbb{C}$ is the set of complex numbers, $\mathbb{C}^{N\times M}$ is the Euclidean space of $(N\times M)$-dimensional complex matrices, and $\mathbb{C}^N$ is the Euclidean space of N-dimensional complex vectors. $\boldsymbol{I}_n$ and $\boldsymbol{0}_{m,n}$ stand for the $n\times n$ identity matrix and the $m\times n$ null matrix. $\langle \boldsymbol{H}\rangle$ denotes the space spanned by the columns of the matrix $\boldsymbol{H}\in\mathbb{C}^{N\times r}$. Given $a_1,\ldots,a_N\in\mathbb{C}$, $\mathrm{diag}\,(a_1,\ldots,a_N)\in\mathbb{C}^{N\times N}$ indicates the diagonal matrix whose ith diagonal element is a_i. We write $\boldsymbol{z}\sim\mathcal{CN}_N(\boldsymbol{x},\boldsymbol{\Sigma})$ to say that the N-dimensional random vector $\boldsymbol{z}$ is a complex normal random vector with mean vector $\boldsymbol{x}$ and covariance matrix $\boldsymbol{\Sigma}$. Moreover, $\boldsymbol{Z}=[\boldsymbol{z}_1\cdots\boldsymbol{z}_K]\sim\mathcal{CN}_{NK}(\boldsymbol{X},\boldsymbol{I}_K\otimes\boldsymbol{\Sigma})$, with $\otimes$ denoting Kronecker product and $\boldsymbol{X}=[\boldsymbol{x}_1\cdots\boldsymbol{x}_K]$, means that $\boldsymbol{z}_k\sim\mathcal{CN}_N(\boldsymbol{x}_k,\boldsymbol{\Sigma})$ and the columns of $\boldsymbol{Z}$ are statistically independent. The acronyms PDF and wp 1 stand for probability density function and with probability 1, respectively. $\widehat{\boldsymbol{R}}_i$ and $\widehat{\gamma}_i$ will denote the (possibly approximated) maximum likelihood (ML) estimates of $\boldsymbol{R}$ and γ, respectively, under the H_i hypothesis, $i=0,1$ (symbols defined in Section II). Finally, $\mathrm{vec}(\cdot)$ is the column vectorizing operator.

II. Four Problems in Adaptive Subspace Detection

For subsequent developments, let us denote by $\boldsymbol{Z}_P=[\boldsymbol{z}_1\cdots\boldsymbol{z}_{K_P}]\in\mathbb{C}^{N\times K_P}$ the matrix of measurements in the primary channel and by $\boldsymbol{Z}_S=[\boldsymbol{z}_{K_P+1}\cdots\boldsymbol{z}_{K_P+K_S}]\in\mathbb{C}^{N\times K_S}$ the matrix of measurements in the secondary channel. In a radar problem the measurements are N-dimensional vectors of space-time samples: the radar system transmits a burst of N_p radio frequency (RF) pulses and the baseband representations of the RF signals collected at the N_a antenna elements are sampled to form range-gate samples for each pulse; it turns out that $N=N_aN_p$. If the signal presence is sought in a subset of K_P range gates, the primary channel consists of NK_P samples. The samples corresponding to any range gate are arranged in a column vector $\boldsymbol{z}_k\in\mathbb{C}^N$. The secondary channel consists of the outputs of K_S properly selected range gates [29]. Finally,

let $\boldsymbol{Z} = [\boldsymbol{Z}_P \ \boldsymbol{Z}_S] \in \mathbb{C}^{N\times K}$ be the overall data matrix with $K = K_P + K_S$.

A. First-Order Models

In a first-order model for measurements, the adaptive detection problem may be formulated as the following test of hypothesis H_0 vs alternative H_1:

$$\begin{aligned} H_0: &\ \begin{cases} \boldsymbol{Z}_P \sim \mathcal{CN}_{NK_P}(\boldsymbol{0}_{N,K_P}, \boldsymbol{I}_{K_P} \otimes \boldsymbol{R}) \\ \boldsymbol{Z}_S \sim \mathcal{CN}_{NK_S}(\boldsymbol{0}_{N,K_S}, \boldsymbol{I}_{K_S} \otimes \gamma\boldsymbol{R}) \end{cases} \\ H_1: &\ \begin{cases} \boldsymbol{Z}_P \sim \mathcal{CN}_{NK_P}(\boldsymbol{H}\boldsymbol{X}, \boldsymbol{I}_{K_P} \otimes \boldsymbol{R}) \\ \boldsymbol{Z}_S \sim \mathcal{CN}_{NK_S}(\boldsymbol{0}_{N,K_S}, \boldsymbol{I}_{K_S} \otimes \gamma\boldsymbol{R}) \end{cases} \end{aligned} \tag{1}$$

where $\boldsymbol{H} \in \mathbb{C}^{N\times r}$ is either a known matrix or an unknown matrix with known rank r, $r \le N$, $\boldsymbol{X} = [\boldsymbol{x}_1 \cdots \boldsymbol{x}_{K_P}] \in \mathbb{C}^{r\times K_P}$ is the matrix of the unknown signal coordinates, $\boldsymbol{R} \in \mathbb{C}^{N\times N}$ is an unknown positive definite covariance matrix, and $\gamma > 0$ is either a known or an unknown parameter. In the following, we suppose that $K_S \ge N$ and, without loss of generality, that $\boldsymbol{H}$ is a slice of a unitary matrix.

B. Second-Order Models

In a second-order model for measurements, the distributions above are treated as conditional distributions, and a prior Gaussian distribution is assumed for the matrix $\boldsymbol{X}$, namely, $\boldsymbol{X} \sim \mathcal{CN}_{rK_P}(\boldsymbol{0}_{r,K_P}, \boldsymbol{I}_{K_P} \otimes \boldsymbol{R}_{xx})$ with $\boldsymbol{R}_{xx} \in \mathbb{C}^{r\times r}$ an unknown positive semidefinite covariance matrix that models sources that may be correlated. The joint distribution of $\boldsymbol{Z}_P$ and $\boldsymbol{X}$ is marginalized for $\boldsymbol{Z}_P$ obtaining that $\boldsymbol{Z}_P \sim \mathcal{CN}_{NK_P}(\boldsymbol{0}_{N,K_P}, \boldsymbol{I}_{K_P} \otimes (\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^\dagger + \boldsymbol{R}))$.

The adaptive detection problem may be formulated as the following test of hypothesis H_0 vs alternative H_1:

$$\begin{aligned} H_0: &\ \begin{cases} \boldsymbol{Z}_P \sim \mathcal{CN}_{NK_P}(\boldsymbol{0}_{N,K_P}, \boldsymbol{I}_{K_P} \otimes \boldsymbol{R}) \\ \boldsymbol{Z}_S \sim \mathcal{CN}_{NK_S}(\boldsymbol{0}_{N,K_S}, \boldsymbol{I}_{K_S} \otimes \gamma\boldsymbol{R}) \end{cases} \\ H_1: &\ \begin{cases} \boldsymbol{Z}_P \sim \mathcal{CN}_{NK_P}(\boldsymbol{0}_{N,K_P}, \boldsymbol{I}_{K_P} \otimes (\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^\dagger + \boldsymbol{R})) \\ \boldsymbol{Z}_S \sim \mathcal{CN}_{NK_S}(\boldsymbol{0}_{N,K_S}, \boldsymbol{I}_{K_S} \otimes \gamma\boldsymbol{R}) \end{cases} \end{aligned} \tag{2}$$

where $\boldsymbol{H} \in \mathbb{C}^{N\times r}$ is either an arbitrary unitary basis for a known subspace $\langle \boldsymbol{H} \rangle$ or an unknown unitary matrix with known rank r, $r \le N$; $\boldsymbol{R} \in \mathbb{C}^{N\times N}$ is an unknown positive definite matrix, and $\gamma > 0$ is either a known or an unknown parameter. Again, we suppose that $K_S \ge N$.

C. Interpretations and Important Statistics

In the derivation of adaptive subspace detectors for first-order models, several data matrices and derived statistics arise. They are summarized and annotated here.

- $\boldsymbol{S}_S = \boldsymbol{Z}_S\boldsymbol{Z}_S^\dagger \in \mathbb{C}^{N\times N}$: K_S times the sample covariance matrix for secondary channel; for $K_S \ge N$, the covariance matrix $\boldsymbol{S}_S$ is positive definite wp 1;
- $\boldsymbol{S}_P = \boldsymbol{Z}_P\boldsymbol{Z}_P^\dagger \in \mathbb{C}^{N\times N}$: K_P times the sample covariance matrix for primary channel; S_P is positive semidefinite with rank $\min(K_P, N)$ wp 1;
- $\boldsymbol{T}_P = \boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\boldsymbol{Z}_P^\dagger\boldsymbol{S}_S^{-1/2} \in \mathbb{C}^{N\times N}$: proportional to the sample covariance matrix for measurements in the primary channel that have been whitened by the square root of the sample covariance matrix computed in the secondary channel; $\boldsymbol{T}_P$ is positive semidefinite of rank $\min(K_p, N)$;
- $\boldsymbol{G} = \boldsymbol{S}_S^{-1/2}\boldsymbol{H} \in \mathbb{C}^{N\times r}$: *whitened* subspace basis; $\boldsymbol{H} \in \mathbb{C}^{N\times r}$ is a unitary basis for the r-dimensional subspace $\langle \boldsymbol{H} \rangle$;
- $\boldsymbol{P}_G^\perp = \boldsymbol{I}_N - \boldsymbol{G}(\boldsymbol{G}^\dagger\boldsymbol{G})^{-1}\boldsymbol{G}^\dagger \in \mathbb{C}^{N\times N}$: projection matrix onto the orthogonal complement of the dimension-r subspace $\langle \boldsymbol{G} \rangle$.

Importantly, the eigenvalues of the statistics $\boldsymbol{T}_P$ and $\boldsymbol{P}_G^\perp\boldsymbol{T}_P\boldsymbol{P}_G^\perp$ are two dramatic compressions of the primary and secondary data that figure prominently in the first-order detectors to be derived in this paper.

III. First-Order Detectors: Derivations

The GLRTs for problem (1) can be obtained by exploiting the results in [19]. Therein, both homogeneous and partially-homogeneous environments are considered, and measurements contain noise plus interference drawn from a subspace that is either known or unknown up to its rank. As a matter of fact, the derivation of the compressed likelihood under the H_0 hypothesis in [19] is the starting point for the derivation of the GLRTs for problem (1).

The joint PDF of primary and secondary data is given by

$$\begin{aligned} f_1(\boldsymbol{Z}; \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{H}, \gamma) = &\frac{\operatorname{etr}\left\{-\frac{1}{\gamma}\boldsymbol{R}^{-1}\boldsymbol{Z}_S\boldsymbol{Z}_S^\dagger\right\}}{\pi^{NK}\gamma^{NK_S}\det^K(\boldsymbol{R})} \\ &\times \operatorname{etr}\left\{-\boldsymbol{R}^{-1}(\boldsymbol{Z}_P - \boldsymbol{H}\boldsymbol{X})(\boldsymbol{Z}_P - \boldsymbol{H}\boldsymbol{X})^\dagger\right\} \end{aligned}$$

under H_1 and under H_0 by

$$f_0(\boldsymbol{Z}; \boldsymbol{R}, \gamma) = \frac{\operatorname{etr}\left\{-\boldsymbol{R}^{-1}\boldsymbol{Z}_P\boldsymbol{Z}_P^\dagger - \frac{1}{\gamma}\boldsymbol{R}^{-1}\boldsymbol{Z}_S\boldsymbol{Z}_S^\dagger\right\}}{\pi^{NK}\gamma^{NK_S}\det^K(\boldsymbol{R})}. \tag{3}$$

A. Known Subspace $\langle \boldsymbol{H} \rangle$, Known γ

Under H_1, the likelihood is maximized through the ML estimates of $\boldsymbol{R}$ and $\boldsymbol{X}$ to produce the partially-compressed likelihood [19]

$$\begin{aligned} &\ell_1(\widehat{\boldsymbol{R}}_1, \widehat{\boldsymbol{X}}, \boldsymbol{H}, \gamma; \boldsymbol{Z}) \\ &= \frac{(K/(e\pi))^{NK}\gamma^{-K_P(K-N)}\det^{-K}(\boldsymbol{S}_S)}{\det^K\left[\frac{1}{\gamma}\boldsymbol{I}_{K_P} + \left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right)^\dagger\boldsymbol{P}_G^\perp\left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right)\right]} \\ &= \frac{(K/(e\pi))^{NK}\gamma^{-K_SN}\det^{-K}(\boldsymbol{S}_S)}{\det^K\left[\frac{1}{\gamma}\boldsymbol{I}_N + \boldsymbol{P}_G^\perp\left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right)\left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right)^\dagger\boldsymbol{P}_G^\perp\right]} \end{aligned} \tag{4}$$

where we have used the identity $\det[\frac{1}{\gamma}\boldsymbol{I}_M + \boldsymbol{A}\boldsymbol{B}] = \gamma^{N-M}\det[\frac{1}{\gamma}\boldsymbol{I}_N + \boldsymbol{B}\boldsymbol{A}]$ with $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ and $\boldsymbol{B} \in \mathbb{C}^{N\times M}$. It is also straightforward to show that compressed likelihood under

H_0 is

$$\ell_0(\widehat{\boldsymbol{R}}_0,\gamma;\boldsymbol{Z}) = \left(\frac{K}{e\pi}\right)^{NK} \frac{\gamma^{-K_P(K-N)}\det^{-K}(\boldsymbol{S}_S)}{\det^{K}\left[\frac{1}{\gamma}\boldsymbol{I}_{K_P}+\boldsymbol{Z}_P^{\dagger}\boldsymbol{S}_S^{-1}\boldsymbol{Z}_P\right]}$$

$$= \left(\frac{K}{e\pi}\right)^{NK} \frac{\gamma^{-K_S N}\det^{-K}(\boldsymbol{S}_S)}{\det^{K}\left[\frac{1}{\gamma}\boldsymbol{I}_{N}+\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\boldsymbol{Z}_P^{\dagger}\boldsymbol{S}_S^{-1/2}\right]}. \quad (5)$$

It follows that the GLRT for homogeneous environment (i.e., $\gamma = 1$) and $r < N$, referred to in the following as first-order known subspace in homogeneous environment (FO-KS-HE) detector, is given by

$$\frac{\det\left[\boldsymbol{I}_{K_P}+\boldsymbol{Z}_P^{\dagger}\boldsymbol{S}_S^{-1}\boldsymbol{Z}_P\right]}{\det\left[\boldsymbol{I}_{K_P}+\left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right)^{\dagger}\boldsymbol{P}_G^{\perp}\left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right)\right]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (6)$$

or, equivalently, as

$$\frac{\det\left[\boldsymbol{I}_N+\boldsymbol{T}_P\right]}{\det\left[\boldsymbol{I}_N+\boldsymbol{P}_G^{\perp}\boldsymbol{T}_P\boldsymbol{P}_G^{\perp}\right]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (7)$$

The expression in (7) illuminates the role of the *secondarily whitened primary data* $\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P$, its corresponding sample covariance $\boldsymbol{T}_P$, and the sample covariance of whitened measurements after their projection onto the subspace $\boldsymbol{P}_G^{\perp}$. The GLRT is a function only of the eigenvalues of $\boldsymbol{T}_P$ and the eigenvalues of $\boldsymbol{P}_G^{\perp}\boldsymbol{T}_P\boldsymbol{P}_G^{\perp}$ (thus implying a massive compression of measurements in primary and secondary channels). For $r = N$ and $\gamma = 1$ the GLRT reduces to

$$\det\left[\boldsymbol{I}_{K_P}+\boldsymbol{Z}_P^{\dagger}\boldsymbol{S}_S^{-1}\boldsymbol{Z}_P\right] = \det\left[\boldsymbol{I}_N+\boldsymbol{T}_P\right] \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (8)$$

These GLRTs are derived for $\gamma = 1$, but generalization to any known value of γ is obviously straightforward. In particular, if γ is known we can normalize the secondary data by the square root of γ, thus obtaining the homogeneous environment. For this reason herafter we will focus on $\gamma = 1$ if γ is known.

B. Known Subspace $\langle \boldsymbol{H} \rangle$, Unknown γ

Determining the GLRT for a partially-homogeneous environment requires one more maximization of the likelihoods with respect to γ, namely the computation of

$$\max_{\gamma>0}\ell_1(\widehat{\boldsymbol{R}}_1,\widehat{\boldsymbol{X}},\boldsymbol{H},\gamma;\boldsymbol{Z}) \quad \text{and} \quad \max_{\gamma>0}\ell_0(\widehat{\boldsymbol{R}}_0,\gamma;\boldsymbol{Z}).$$

For $r = N$ the likelihood under H_1 is unbounded with respect to $\gamma > 0$ and, hence, the GLRT does not exist. Therefore we assume $r < N$. The following result derived in [17], [19] is recalled here for the sake of completeness.

Theorem 1: Let $\boldsymbol{M} \in \mathbb{C}^{K_P\times K_P}$ be a positive semidefinite (Hermitian) matrix of rank t ($1 \le t \le K_P$). Then, the function

$$f(\gamma) = \gamma^{\frac{K_P(K-N)}{K}}\det\left(\frac{1}{\gamma}\boldsymbol{I}_{K_P}+\boldsymbol{M}\right), \quad \gamma > 0, \quad (9)$$

attains its absolute minimum at the unique positive solution of

$$\sum_{k=K_P-t+1}^{K_P}\frac{\lambda_k\gamma}{\lambda_k\gamma+1} = \frac{NK_P}{K} \quad (10)$$

where the λ_ks are the eigenvalues of the matrix $\boldsymbol{M}$ arranged in increasing order ($\lambda_k = 0, k = 1,\ldots,K_P - t$) and provided that $t > \frac{NK_P}{K}$. If $t = \frac{NK_P}{K}$, then $f(\gamma)$ does not possess the absolute minimum over $(0,+\infty)$, but its infimum is positive; finally, if $t < \frac{NK_P}{K}$, the infimum of $f(\gamma)$ over $(0,+\infty)$ is zero.

Proof: See [17], [19]. □

To use this theorem, it is necessary to determine the rank of the matrices $\boldsymbol{M}_0 = \boldsymbol{Z}_P^{\dagger}\boldsymbol{S}_S^{-1}\boldsymbol{Z}_P$ and $\boldsymbol{M}_1 = (\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P)^{\dagger}\boldsymbol{P}_G^{\perp}(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P)$ and whether or not the condition on the rank is satisfied. Preliminarily, we give the following lemma that can be easily proved following the lead of [30, Theorem 3.1.4 pag. 82].

Lemma 1: Let $\boldsymbol{z}_1,\ldots,\boldsymbol{z}_m$ be m independent and complex normal Gaussian vectors with positive definite covariance matrix, i.e., $\boldsymbol{z}_k \sim \mathcal{CN}_N(\boldsymbol{m}_k,\boldsymbol{R}_k)$. The rank of the matrix $[\boldsymbol{z}_1\cdots\boldsymbol{z}_m]$ is equal to the minimum among m and N wp 1.

It is also easy to prove the following theorem concerning the rank of the matrices $\boldsymbol{M}_0$ and $\boldsymbol{M}_1$.

Theorem 2: The rank of

$$\boldsymbol{M}_0 = \boldsymbol{Z}_P^{\dagger}\boldsymbol{S}_S^{-1}\boldsymbol{Z}_P \quad (11)$$

is $m_1 = \min(K_P, N)$ and $m_1 > \frac{NK_P}{K}$ since $K > K_P$ and $K > N$. Similarly, the rank of

$$\boldsymbol{M}_1 = \left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right)^{\dagger}\boldsymbol{P}_G^{\perp}\left(\boldsymbol{S}_S^{-1/2}\boldsymbol{Z}_P\right) \quad (12)$$

is $m_2 = \min(K_P, N-r)$. It follows that $m_2 > NK_P/K$ when $K_P \le N - r$ (since $K > N$); for $N - r < K_P$, the condition is $N - r > NK_P/K$, which requires $r < N(1 - K_P/K)$.

It follows that, under the condition $m_2 > \frac{NK_P}{K}$, the GLRT for partially-homogeneous environment, referred to in the following as first-order known subspace in partially-homogeneous environment (FO-KS-PHE) detector, is given by

$$\frac{\widehat{\gamma}_0^{\frac{K_P(K-N)}{K}}\det\left[\frac{1}{\widehat{\gamma}_0}\boldsymbol{I}_{K_P}+\boldsymbol{M}_0\right]}{\widehat{\gamma}_1^{\frac{K_P(K-N)}{K}}\det\left[\frac{1}{\widehat{\gamma}_1}\boldsymbol{I}_{K_P}+\boldsymbol{M}_1\right]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (13)$$

where $\widehat{\gamma}_i$, $i = 0, 1$, can be computed using Theorem 1 and $\boldsymbol{M}_0$ and $\boldsymbol{M}_1$ are given by (11) and (12), respectively. The equivalent form is more illuminating:

$$\frac{\widehat{\gamma}_0^{N(1-K_P/K)}\det\left[\frac{1}{\widehat{\gamma}_0}\boldsymbol{I}_N+\boldsymbol{T}_P\right]}{\widehat{\gamma}_1^{N(1-K_P/K)}\det\left[\frac{1}{\widehat{\gamma}_1}\boldsymbol{I}_N+\boldsymbol{P}_G^{\perp}\boldsymbol{T}_P\boldsymbol{P}_G^{\perp}\right]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta.$$

Again, the GLRT is a function only of the eigenvalues of $\boldsymbol{T}_P$ and $\boldsymbol{P}_G^{\perp}\boldsymbol{T}_P\boldsymbol{P}_G^{\perp}$. In fact, $\boldsymbol{M}_0$ and $\boldsymbol{T}_P$ share the nonzero eigenvalues. Similarly for $\boldsymbol{M}_1$ and $\boldsymbol{P}_G^{\perp}\boldsymbol{T}_P\boldsymbol{P}_G^{\perp}$.

C. Unknown Subspace $\langle \boldsymbol{H} \rangle$ of Known Dimension, Known γ

The signal subspace $\langle \boldsymbol{H} \rangle$ is unknown, but its rank $r \le N$ is known. To compute the compressed likelihood under H_1, the parameter $\boldsymbol{H}$ is replaced by its ML estimate in (4). The maximization with respect to $\boldsymbol{H}$ can be conducted as shown in [19]. The result is

$$\ell_1(\widehat{\boldsymbol{R}}_1,\widehat{\boldsymbol{X}},\widehat{\boldsymbol{H}},\gamma;\boldsymbol{Z}) = \frac{[K/(e\pi)]^{NK}}{\gamma^{K_P(K-N)}}\frac{1}{\det^{K}(\boldsymbol{S}_S)}\frac{1}{g_1^{K}(\gamma)} \quad (14)$$

where

$$g_1(\gamma) = \begin{cases} \gamma^{N-r-K_P} \prod_{i=1}^{N-r} \left(\frac{1}{\gamma} + \sigma_i^2\right), & m_1 \geq r+1 \\ \left(\frac{1}{\gamma}\right)^{K_P}, & \text{otherwise} \end{cases}.$$

Here $m_1 = \min(N, K_P)$ is the rank of the matrix $\boldsymbol{T}_P = \boldsymbol{S}_S^{-1/2} \boldsymbol{Z}_P \boldsymbol{Z}_P^{\dagger} \boldsymbol{S}_S^{-1/2}$ and σ_i^2, $i = 1, \ldots, N$, are the eigenvalues of $\boldsymbol{T}_P$ arranged in increasing order. Moreover, the compressed likelihood under H_0 can be re-written as

$$\ell_0(\widehat{\boldsymbol{R}}_0, \gamma; \boldsymbol{Z}) = \frac{[K/(e\pi)]^{NK}}{\gamma^{K_P(K-N)}} \frac{1}{\det^K(\boldsymbol{S}_S)} \frac{1}{g_0^K(\gamma)} \tag{15}$$

with $g_0(\gamma) = \gamma^{N-K_P} \prod_{i=1}^{N} (\frac{1}{\gamma} + \sigma_i^2)$.

It follows that, if $\min(N, K_P) \geq r+1$, the GLRT for homogeneous environment, referred to in the following as first-order unknown subspace in homogeneous environment (FO-US-HE) detector, is given by

$$\frac{\prod_{i=1}^{N} (1+\sigma_i^2)}{\prod_{i=1}^{N-r} (1+\sigma_i^2)} = \prod_{i=N-r+1}^{N} (1+\sigma_i^2) \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{16}$$

For $m_1 < r+1$, the GLRT reduces to

$$\prod_{i=1}^{N} (1+\sigma_i^2) = \det(\boldsymbol{I}_N + \boldsymbol{T}_P) \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{17}$$

Notice also that condition $m_1 < r+1$ is equivalent to $N = r$ if $N < K_P$ (recall that $N \geq r$) or to $K_P < r+1$ if $K_P \leq N$.

D. Unknown Subspace $\langle \boldsymbol{H} \rangle$ of Known Dimension, Unknown γ

To obtain the GLRT for partially-homogeneous environment we have to maximize the partially-compressed likelihoods over γ. We focus on $m_1 \geq r+1$; in fact, for $m_1 < r+1$ the likelihood under H_1 is unbounded with respect to γ and, hence, the GLRT does not exist. Equivalently, we have to minimize with respect to γ the following functions

$$f_1(\gamma) = \gamma^{\frac{-K_P N}{K}} \prod_{i=N-m_1+1}^{N-r} (1+\gamma\sigma_i^2)$$

and

$$f_0(\gamma) = \gamma^{\frac{-K_P N}{K}} \prod_{i=N-m_1+1}^{N} (1+\gamma\sigma_i^2).$$

Proceeding as in the proof of Theorem 1, we obtain the following results.

Corollary 1: The function

$$f_1(\gamma) = \gamma^{\frac{-K_P N}{K}} \prod_{i=N-m_1+1}^{N-r} (1+\gamma\sigma_i^2)$$

attains its absolute minimum over $(0, +\infty)$ at the unique positive solution of

$$\sum_{i=N-m_1+1}^{N-r} \frac{\sigma_i^2 \gamma}{\sigma_i^2 \gamma + 1} = \frac{NK_P}{K}, \tag{18}$$

provided that $m_1 - r > \frac{NK_P}{K}$. If $m_1 - r = \frac{NK_P}{K}$, then $f_1(\gamma)$ does not possess the absolute minimum over $(0, +\infty)$, but its infimum is positive; finally, if $m_1 - r < \frac{NK_P}{K}$, the infimum of $f_1(\gamma)$ over $(0, +\infty)$ is zero.

Corollary 2: The function

$$f_0(\gamma) = \gamma^{\frac{-K_P N}{K}} \prod_{i=N-m_1+1}^{N} (1+\gamma\sigma_i^2)$$

attains its absolute minimum over $(0, +\infty)$ at the unique positive solution of

$$\sum_{i=N-m_1+1}^{N} \frac{\sigma_i^2 \gamma}{\sigma_i^2 \gamma + 1} = \frac{NK_P}{K}, \tag{19}$$

provided that $m_1 > \frac{NK_P}{K}$. If $m_1 = \frac{NK_P}{K}$, then $f_0(\gamma)$ does not possess the absolute minimum over $(0, +\infty)$, but its infimum is positive; finally, if $m_1 < \frac{NK_P}{K}$, the infimum of $f_0(\gamma)$ over $(0, +\infty)$ is zero.

It follows that, under the condition $m_1 > \frac{NK_P}{K} + r$, the GLRT, referred to in the following as first-order unknown subspace in partially-homogeneous environment (FO-US-PHE) detector, can be written as

$$\frac{\widehat{\gamma}_0^{N\left(1-\frac{K_P}{K}\right)} \prod_{i=1}^{N} \left(\frac{1}{\widehat{\gamma}_0} + \sigma_i^2\right)}{\widehat{\gamma}_1^{N\left(1-\frac{K_P}{K}\right)-r} \prod_{i=1}^{N-r} \left(\frac{1}{\widehat{\gamma}_1} + \sigma_i^2\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{20}$$

where $\widehat{\gamma}_1$ and $\widehat{\gamma}_0$ can be computed using Corollary 1 and 2, respectively. Notice that the detector is a function of the eigenvalues of the statistic $\boldsymbol{T}_P$ only.

IV. Second-Order Detectors: Derivations

The joint PDF of primary and secondary data is given by

$$f_1(\boldsymbol{Z}; \boldsymbol{R}, \boldsymbol{R}_s, \boldsymbol{H}, \gamma) = \frac{\operatorname{etr}\left\{-\frac{1}{\gamma}\boldsymbol{R}^{-1}\boldsymbol{Z}_S\boldsymbol{Z}_S^{\dagger}\right\}}{\pi^{NK}\gamma^{NK_S}} \times \frac{\operatorname{etr}\left\{-\left(\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^{\dagger} + \boldsymbol{R}\right)^{-1}\boldsymbol{Z}_P\boldsymbol{Z}_P^{\dagger}\right\}}{\det^{K_P}(\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^{\dagger} + \boldsymbol{R}) \det^{K_S}(\boldsymbol{R})} \tag{21}$$

under H_1 and is expressed by (3) under H_0. The compressed likelihood under H_0 has already been computed to implement the GLRTs for first-order models. This result applies also in second-order models.

In this section we reverse the order of derivations by first deriving the GLRTs for an unknown subspace and then deriving the detectors for a known subspace. The justification for this reversal of course is that the optimization results obtained for an unknown subspace of known dimension may then be used for a known subspace.

A. Unknown Subspace $\langle \boldsymbol{H} \rangle$ of Known Dimension, Known γ

In (21) the parameters $\boldsymbol{H}$ and $\boldsymbol{R}_{xx}$ are both unknown, so $\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^{\dagger}$ may be replaced by the unknown covariance matrix

$\tilde{\boldsymbol{R}}_{xx}$. Thus, the log-likelihood under H_1 can be written as

$$L_1(\boldsymbol{R}, \tilde{\boldsymbol{R}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma - K_P \log \det(\tilde{\boldsymbol{R}}_{xx} + \boldsymbol{R}) - \text{Tr}\left[\left(\tilde{\boldsymbol{R}}_{xx} + \boldsymbol{R}\right)^{-1} \boldsymbol{S}_P\right] - K_S \log \det(\boldsymbol{R}) - \text{Tr}\left[\frac{1}{\gamma}\boldsymbol{R}^{-1}\boldsymbol{S}_S\right] \quad (22)$$

where we recall that $\boldsymbol{S}_P = \boldsymbol{Z}_P\boldsymbol{Z}_P^\dagger$ and $\boldsymbol{S}_S = \boldsymbol{Z}_S\boldsymbol{Z}_S^\dagger$ (and the matrix $\boldsymbol{S}_S$ is positive definite since $K_S \geq N$). Notice also that the rank of the matrix $\tilde{\boldsymbol{R}}_{xx}$ is less than or equal to r (in fact, the rank of $\boldsymbol{H}\boldsymbol{R}_{xx}^{1/2}$ is less than or equal to r). The compressed likelihood necessary to obtain the GLRT is given by the following theorem. The focus is on the case $r \leq K_P \leq N$ although extension to $K_P < r$ is straightforward.

Theorem 3: Let $r \leq K_P \leq N$. Denote by $\boldsymbol{\Gamma} = \text{diag}\,(\gamma_1, \ldots, \gamma_N) \in \mathbb{R}^{N\times N}$, $\gamma_1 \geq \ldots \geq \gamma_N \geq 0$, the diagonal matrix containing the eigenvalues of $\boldsymbol{S}_S^{-1/2}\boldsymbol{S}_P\boldsymbol{S}_S^{-1/2}$ and by $\boldsymbol{V} \in \mathbb{C}^{N\times N}$ the unitary matrix of the corresponding eigenvectors. Finally, let $\boldsymbol{K} = \boldsymbol{S}_S^{1/2}\boldsymbol{V} \in \mathbb{C}^{N\times N}$. The maximum of the left-hand side of (22) can be rewritten as

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma - NK - 2\,K \log|\det(\boldsymbol{K})| + \sum_{i=1}^{r-1} K \log \frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)} + \sum_{i=1}^{r-1} K_S \log \widehat{\lambda}_i(\gamma) + \sum_{i=r}^{N} K \log \frac{\gamma K}{\gamma\gamma_i + 1} \quad (23)$$

with $\widehat{\lambda}_i = \max(\frac{K_S\gamma\gamma_i}{K_P}, 1)$, $i = 1, \ldots, r-1$, if $\gamma < \frac{K_P}{K_S}\frac{1}{\gamma_r}$, and as

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma - NK - 2\,K \log|\det(\boldsymbol{K})| + \sum_{i=1}^{r} \left[K \log \frac{K_P}{\gamma_i} + K_S \log \frac{K_S\gamma\gamma_i}{K_P}\right] + \sum_{i=r+1}^{N} K \log \frac{\gamma K}{\gamma\gamma_i + 1} \quad (24)$$

otherwise. More specifically, if $r = 1$ we have to remove the second-to-last and the third-to-last summations from (23), that we recall is valid if $\gamma < \frac{K_P}{K_S}\frac{1}{\gamma_1}$. If instead $r = N$, the compressed likelihood under H_1 is given by (23), for all γ values, after removing the last term and varying the (summation) index i from 1 to N (also in the expression of the $\widehat{\lambda}_i$).

Proof: See Appendix A. □

Notice that $K_P \geq h$ is a "necessary condition" to estimate the eigenvalues of a matrix $\tilde{\boldsymbol{R}}_{xx}$ with rank h.

Finally, the GLRT for homogeneous environment and unknown subspace $\langle \boldsymbol{H} \rangle$, referred to in the following as second-order unknown subspace in homogeneous environment (SO-US-HE) detector, is

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, 1; \boldsymbol{Z}) - L_0(\widehat{\boldsymbol{R}}_0, 1; \boldsymbol{Z}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (25)$$

with $L_0(\widehat{\boldsymbol{R}}_0, 1; \boldsymbol{Z})$ given by the logarithm of (5) (with $\gamma = 1$).

B. Unknown Subspace $\langle \boldsymbol{H} \rangle$ of Known Dimension, Unknown γ

To derive the GLRT for partially-homogeneous environment, we have to maximize the partially-compressed likelihood (under H_1), given by Theorem 3, also with respect to γ. This maximization is summarized by the following theorem.

Theorem 4: Let $r < K_P \leq N$. The maximum with respect to γ of the partially-compressed likelihood, given by Theorem 3, is attained at the unique $\gamma \geq \frac{K_P}{K_S}\frac{1}{\gamma_r}$, say $\widehat{\gamma}$, solving the equation

$$\sum_{i=r+1}^{K_P} \frac{K}{\gamma\gamma_i + 1} = (K_P - r)K_S - (N - K_P)K_P,$$

provided that $(K_P - r)K_S > (N - K_P)K_P$. The compressed likelihood is obtained by plugging $\widehat{\gamma}$ into (24).

Proof: See Appendix B. □

Finally, the GLRT, referred to in the following as second-order unknown subspace in partially-homogeneous environment (SO-US-PHE) detector, is given by

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \widehat{\gamma}_1; \boldsymbol{Z}) - L_0(\widehat{\boldsymbol{R}}_0, \widehat{\gamma}_0; \boldsymbol{Z}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (26)$$

with $L_0(\widehat{\boldsymbol{R}}_0, \widehat{\gamma}_0; \boldsymbol{Z})$ given by the logarithm of the maximum of (5) with respect to γ obtained by using Theorem 1.

C. Known Subspace $\langle \boldsymbol{H} \rangle$, Known γ

As a first step towards the computation of the GLRT, we extend [25] where the case $\boldsymbol{H} \in \mathbb{C}^N$ (rank-one signal) and $\gamma = 1$ (homogeneous environment) is addressed. To this end, we denote by $\boldsymbol{H}_\perp \in \mathbb{C}^{N\times(N-r)}$ a slice of a unitary matrix spanning the orthogonal complement of $\boldsymbol{H} \in \mathbb{C}^{N\times r}$. It follows that the matrix $\boldsymbol{V} = [\boldsymbol{H}\ \boldsymbol{H}_\perp] \in \mathbb{C}^{N\times N}$ is unitary. Then, we rewrite the likelihoods under H_1 and H_0 as

$$\ell_1(\boldsymbol{R}, \boldsymbol{R}_{xx}, \boldsymbol{H}, \gamma; \boldsymbol{Z}) = \frac{\text{etr}\left\{-\left(\tilde{\boldsymbol{R}} + \boldsymbol{E}\boldsymbol{R}_{xx}\boldsymbol{E}^\dagger\right)^{-1}\tilde{\boldsymbol{Z}}_P\tilde{\boldsymbol{Z}}_P^\dagger\right\}}{\pi^{NK}\gamma^{NK_S}\det^{K_P}(\tilde{\boldsymbol{R}} + \boldsymbol{E}\boldsymbol{R}_{xx}\boldsymbol{E}^\dagger)} \times \frac{\text{etr}\left\{-\frac{1}{\gamma}\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{Z}}_S\tilde{\boldsymbol{Z}}_S^\dagger\right\}}{\det^{K_S}(\tilde{\boldsymbol{R}})} \quad (27)$$

and

$$\ell_0(\boldsymbol{R}, \gamma; \boldsymbol{Z}) = \frac{\text{etr}\left\{-\left[\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{Z}}_P\tilde{\boldsymbol{Z}}_P^\dagger + \frac{1}{\gamma}\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{Z}}_S\tilde{\boldsymbol{Z}}_S^\dagger\right]\right\}}{\pi^{NK}\gamma^{NK_S}\det^{K}(\tilde{\boldsymbol{R}})},$$

respectively. Here $\tilde{\boldsymbol{Z}}_P = \boldsymbol{V}^\dagger\boldsymbol{Z}_P = [\tilde{\boldsymbol{Z}}_{P,1}^T \quad \tilde{\boldsymbol{Z}}_{P,2}^T]^T$, with $\tilde{\boldsymbol{Z}}_{P,1} \in \mathbb{C}^{r\times K_P}$ and $\tilde{\boldsymbol{Z}}_{P,2} \in \mathbb{C}^{(N-r)\times K_P}$, $\tilde{\boldsymbol{Z}}_S = \boldsymbol{V}^\dagger\boldsymbol{Z}_S = [\tilde{\boldsymbol{Z}}_{S,1}^T\ \tilde{\boldsymbol{Z}}_{S,2}^T]^T$, with $\tilde{\boldsymbol{Z}}_{S,1} \in \mathbb{C}^{r\times K_S}$ and $\tilde{\boldsymbol{Z}}_{S,2} \in \mathbb{C}^{(N-r)\times K_S}$,

$$\tilde{\boldsymbol{R}} = \boldsymbol{V}^\dagger\boldsymbol{R}\boldsymbol{V} = \begin{bmatrix} \boldsymbol{H}^\dagger\boldsymbol{R}\boldsymbol{H} & \boldsymbol{H}^\dagger\boldsymbol{R}\boldsymbol{H}_\perp \\ \boldsymbol{H}_\perp^\dagger\boldsymbol{R}\boldsymbol{H} & \boldsymbol{H}_\perp^\dagger\boldsymbol{R}\boldsymbol{H}_\perp \end{bmatrix}.$$

Also

$$\boldsymbol{V}^\dagger\left(\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^\dagger + \boldsymbol{R}\right)\boldsymbol{V} = \begin{bmatrix} \boldsymbol{R}_{xx} + \boldsymbol{H}^\dagger\boldsymbol{R}\boldsymbol{H} & \boldsymbol{H}^\dagger\boldsymbol{R}\boldsymbol{H}_\perp \\ \boldsymbol{H}_\perp^\dagger\boldsymbol{R}\boldsymbol{H} & \boldsymbol{H}_\perp^\dagger\boldsymbol{R}\boldsymbol{H}_\perp \end{bmatrix} = \tilde{\boldsymbol{R}} + \boldsymbol{E}\boldsymbol{R}_{xx}\boldsymbol{E}^\dagger$$

with $\boldsymbol{E} = [\boldsymbol{I}_r \ \boldsymbol{0}_{r,N-r}]^T$. We now observe that [31], [32]

$$\tilde{\boldsymbol{R}}^{-1} = \begin{bmatrix} \tilde{\boldsymbol{R}}_{11} & \tilde{\boldsymbol{R}}_{12} \\ \tilde{\boldsymbol{R}}_{21} & \tilde{\boldsymbol{R}}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \tilde{\boldsymbol{R}}_{1.2}^{-1} & -\tilde{\boldsymbol{R}}_{1.2}^{-1}\boldsymbol{\beta}^\dagger \\ -\boldsymbol{\beta}\tilde{\boldsymbol{R}}_{1.2}^{-1} & \tilde{\boldsymbol{R}}_{22}^{-1} + \boldsymbol{\beta}\tilde{\boldsymbol{R}}_{1.2}^{-1}\boldsymbol{\beta}^\dagger \end{bmatrix}$$

$$= \boldsymbol{B}^\dagger \tilde{\boldsymbol{R}}_{1.2}^{-1}\boldsymbol{B} + \begin{bmatrix} \boldsymbol{0}_{r,r} & \boldsymbol{0}_{r,N-r} \\ \boldsymbol{0}_{N-r,r} & \tilde{\boldsymbol{R}}_{22}^{-1} \end{bmatrix}$$

with $\boldsymbol{B} = [\boldsymbol{I}_r \ -\boldsymbol{\beta}^\dagger] \in \mathbb{C}^{r\times N}$, $\tilde{\boldsymbol{R}}_{1.2} = \tilde{\boldsymbol{R}}_{11} - \tilde{\boldsymbol{R}}_{12}\tilde{\boldsymbol{R}}_{22}^{-1}\tilde{\boldsymbol{R}}_{21} \in \mathbb{C}^{r\times r}$, and $\boldsymbol{\beta} = \tilde{\boldsymbol{R}}_{22}^{-1}\tilde{\boldsymbol{R}}_{21} \in \mathbb{C}^{(N-r)\times r}$. Similarly,

$$\left(\tilde{\boldsymbol{R}} + \boldsymbol{E}\boldsymbol{R}_{xx}\boldsymbol{E}^\dagger\right)^{-1} = \begin{bmatrix} \boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{11} & \tilde{\boldsymbol{R}}_{12} \\ \tilde{\boldsymbol{R}}_{21} & \tilde{\boldsymbol{R}}_{22} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1} & -\left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1}\boldsymbol{\beta}^\dagger \\ -\boldsymbol{\beta}\left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1} & \tilde{\boldsymbol{R}}_{22}^{-1} + \boldsymbol{\beta}\left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1}\boldsymbol{\beta}^\dagger \end{bmatrix}$$

$$= \boldsymbol{B}^\dagger \left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1}\boldsymbol{B} + \begin{bmatrix} \boldsymbol{0}_{r,r} & \boldsymbol{0}_{r,N-r} \\ \boldsymbol{0}_{N-r,r} & \tilde{\boldsymbol{R}}_{22}^{-1} \end{bmatrix}.$$

Moreover, we have that [31], [32] $\det\tilde{\boldsymbol{R}} = \det\tilde{\boldsymbol{R}}_{22}\cdot\det\tilde{\boldsymbol{R}}_{1.2}$ and $\det(\tilde{\boldsymbol{R}} + \boldsymbol{E}\boldsymbol{R}_{xx}\boldsymbol{E}^\dagger) = \det\tilde{\boldsymbol{R}}_{22}\cdot\det(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2})$. It follows that

$$\ell_1(\boldsymbol{R}, \boldsymbol{R}_{xx}, \boldsymbol{H}, \gamma; \boldsymbol{Z}) = \frac{1}{\pi^{NK}}\frac{1}{\gamma^{NK_S}}\frac{1}{\det^K(\tilde{\boldsymbol{R}}_{22})}$$
$$\times \frac{\operatorname{etr}\left\{-\left[\tilde{\boldsymbol{R}}_{22}^{-1}\tilde{\boldsymbol{Z}}_{P,2}\tilde{\boldsymbol{Z}}_{P,2}^\dagger + \frac{1}{\gamma}\tilde{\boldsymbol{R}}_{22}^{-1}\tilde{\boldsymbol{Z}}_{S,2}\tilde{\boldsymbol{Z}}_{S,2}^\dagger\right]\right\}}{\det^{K_P}(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2})\det^{K_S}(\tilde{\boldsymbol{R}}_{1.2})}$$
$$\times \operatorname{etr}\left\{-\left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1}\boldsymbol{B}\tilde{\boldsymbol{Z}}_P\tilde{\boldsymbol{Z}}_P^\dagger\boldsymbol{B}^\dagger\right\}$$
$$\times \operatorname{etr}\left\{-\frac{1}{\gamma}\tilde{\boldsymbol{R}}_{1.2}^{-1}\boldsymbol{B}\tilde{\boldsymbol{Z}}_S\tilde{\boldsymbol{Z}}_S^\dagger\boldsymbol{B}^\dagger\right\}.$$

Subsequent developments rely on the fact that we can estimate the parameters $\tilde{\boldsymbol{R}}_{1.2}, \tilde{\boldsymbol{R}}_{22}, \boldsymbol{\beta}, \boldsymbol{R}_{xx}$ in place of $\boldsymbol{R}, \boldsymbol{R}_{xx}$. To this end, first observe that the ML estimate of $\tilde{\boldsymbol{R}}_{22}$, given γ, can be expressed as

$$\hat{\tilde{\boldsymbol{R}}}_{22} = \frac{1}{K}\left(\tilde{\boldsymbol{Z}}_{P,2}\tilde{\boldsymbol{Z}}_{P,2}^\dagger + \frac{1}{\gamma}\tilde{\boldsymbol{Z}}_{S,2}\tilde{\boldsymbol{Z}}_{S,2}^\dagger\right)$$

and the corresponding partially-compressed likelihood becomes

$$\max_{\tilde{\boldsymbol{R}}_{22}} \ell_1(\boldsymbol{R}, \boldsymbol{R}_{xx}, \boldsymbol{H}, \gamma; \boldsymbol{Z})$$
$$= \frac{1}{\pi^{NK}}\frac{1}{\gamma^{NK_S}}\frac{(K/e)^{(N-r)K}}{\det^{K_P}(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2})}$$
$$\times \frac{\operatorname{etr}\left\{-\left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1}\boldsymbol{B}\tilde{\boldsymbol{Z}}_P\tilde{\boldsymbol{Z}}_P^\dagger\boldsymbol{B}^\dagger\right\}}{\det^{K_S}(\tilde{\boldsymbol{R}}_{1.2})\det^K\left(\tilde{\boldsymbol{Z}}_{P,2}\tilde{\boldsymbol{Z}}_{P,2}^\dagger + \frac{1}{\gamma}\tilde{\boldsymbol{Z}}_{S,2}\tilde{\boldsymbol{Z}}_{S,2}^\dagger\right)}$$
$$\times \operatorname{etr}\left\{-\frac{1}{\gamma}\tilde{\boldsymbol{R}}_{1.2}^{-1}\boldsymbol{B}\tilde{\boldsymbol{Z}}_S\tilde{\boldsymbol{Z}}_S^\dagger\boldsymbol{B}^\dagger\right\}. \quad (28)$$

Estimation of the remaining parameters cannot be conducted in closed form to the best of authors' knowledge. For this reason we implement an alternating maximization [33] which estimates a subset of the unknown parameters assuming that the remaining parameters are known and vice versa. In particular, we exploit the following results.

1) Estimate of $\tilde{\boldsymbol{R}}_{1.2}$ and $\boldsymbol{R}_{xx}$, Given $\boldsymbol{\beta}$: We write the logarithm of the partially-compressed likelihood (28) as follows

$$\max_{\tilde{\boldsymbol{R}}_{22}} L_1(\boldsymbol{R}, \boldsymbol{R}_{xx}, \boldsymbol{H}, \gamma; \boldsymbol{Z}) = C - NK_S\log\gamma$$
$$- K_P\log\det(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}) - K_S\log\det(\tilde{\boldsymbol{R}}_{1.2})$$
$$- K\log\det\left(\tilde{\boldsymbol{Z}}_{P,2}\tilde{\boldsymbol{Z}}_{P,2}^\dagger + \frac{1}{\gamma}\tilde{\boldsymbol{Z}}_{S,2}\tilde{\boldsymbol{Z}}_{S,2}^\dagger\right)$$
$$- \operatorname{Tr}\left[\left(\boldsymbol{R}_{xx} + \tilde{\boldsymbol{R}}_{1.2}\right)^{-1}\tilde{\boldsymbol{S}}_P\right] - \frac{1}{\gamma}\operatorname{Tr}\left[\tilde{\boldsymbol{R}}_{1.2}^{-1}\tilde{\boldsymbol{S}}_S\right] \quad (29)$$

where $\tilde{\boldsymbol{S}}_P = \boldsymbol{B}\tilde{\boldsymbol{Z}}_P\tilde{\boldsymbol{Z}}_P^\dagger\boldsymbol{B}^\dagger$ and $\tilde{\boldsymbol{S}}_S = \boldsymbol{B}\tilde{\boldsymbol{Z}}_S\tilde{\boldsymbol{Z}}_S^\dagger\boldsymbol{B}^\dagger$; $C = -NK\log\pi + (N-r)K(\log K - \log e)$ gathers the terms that are irrelevant to the maximization. It can be shown that $K_S \geq N$ makes $\tilde{\boldsymbol{S}}_S$ a non-singular matrix. Exploiting Theorem 6, with $\boldsymbol{R}_{xx}$ and $\tilde{\boldsymbol{R}}_{1.2}$ in place of $\tilde{\boldsymbol{R}}_{xx}$ and $\boldsymbol{R}$, respectively, (notice also that the matrices are $r\times r$ in place of $N\times N$) we obtain that

$$\max_{\tilde{\boldsymbol{R}}_{22}, \tilde{\boldsymbol{R}}_{1.2}, \boldsymbol{R}_{xx}} L_1(\boldsymbol{R}, \boldsymbol{R}_{xx}, \boldsymbol{H}, \gamma; \boldsymbol{Z}) = C$$
$$- K\log\det\left(\tilde{\boldsymbol{Z}}_{P,2}\tilde{\boldsymbol{Z}}_{P,2}^\dagger + \frac{1}{\gamma}\tilde{\boldsymbol{Z}}_{S,2}\tilde{\boldsymbol{Z}}_{S,2}^\dagger\right)$$
$$- NK_S\log\gamma - 2K\log|\det(\boldsymbol{K})| - rK$$
$$+ K\sum_{i=1}^r \log\frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)} + K_S\sum_{i=1}^r\log\widehat{\lambda}_i(\gamma) \quad (30)$$

with $\gamma_1 \geq \ldots \geq \gamma_r \geq 0$ the eigenvalues of $\tilde{\boldsymbol{S}}_S^{-1/2}\tilde{\boldsymbol{S}}_P\tilde{\boldsymbol{S}}_S^{-1/2} \in \mathbb{C}^{r\times r}$, $\boldsymbol{V} \in \mathbb{C}^{r\times r}$ the unitary matrix of the corresponding eigenvectors of $\tilde{\boldsymbol{S}}_S^{-1/2}\tilde{\boldsymbol{S}}_P\tilde{\boldsymbol{S}}_S^{-1/2}$, $\boldsymbol{K} = \tilde{\boldsymbol{S}}_S^{1/2}\boldsymbol{V} \in \mathbb{C}^{r\times r}$, and

$$\widehat{\lambda}_i(\gamma) = \max\left(\frac{K_S}{K_P}\gamma\gamma_i, 1\right), \quad i = 1, \ldots, r.$$

Notice that application of the theorem returns the following estimates of $\tilde{\boldsymbol{R}}_{1.2}$ and $\boldsymbol{R}_{xx}$

$$\hat{\tilde{\boldsymbol{R}}}_{1.2} = \widehat{\boldsymbol{M}}\widehat{\boldsymbol{M}}^\dagger \quad \text{and} \quad \widehat{\boldsymbol{R}}_{xx} = \widehat{\boldsymbol{M}}\widehat{\boldsymbol{\Lambda}}\widehat{\boldsymbol{M}}^\dagger - \hat{\tilde{\boldsymbol{R}}}_{1.2} \quad (31)$$

where $\widehat{\boldsymbol{\Lambda}} = \operatorname{diag}(\widehat{\lambda}_1, \ldots, \widehat{\lambda}_r) \in \mathbb{R}^{r\times r}$; $\widehat{\boldsymbol{M}} = \boldsymbol{K}\widehat{\boldsymbol{D}}^{-1}\widehat{\boldsymbol{\Lambda}}^{-1/2}$ with $\widehat{\boldsymbol{D}}^2 = \operatorname{diag}(\widehat{d}_1^2, \ldots, \widehat{d}_r^2) \in \mathbb{R}^{r\times r}$, and

$$\widehat{d}_i^2(\gamma) = \frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)}, \quad i = 1, \ldots, r.$$

2) Estimate of $\boldsymbol{\beta}$ Given $\tilde{\boldsymbol{R}}_{1.2}$ and $\boldsymbol{R}_{xx}$: First we observe that

$$\boldsymbol{B}\tilde{\boldsymbol{A}}\boldsymbol{B}^\dagger = \tilde{\boldsymbol{A}}_{11} - \boldsymbol{\beta}^\dagger\tilde{\boldsymbol{A}}_{21} - \tilde{\boldsymbol{A}}_{12}\boldsymbol{\beta} + \boldsymbol{\beta}^\dagger\tilde{\boldsymbol{A}}_{22}\boldsymbol{\beta}$$

where

$$\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{Z}}_P\tilde{\boldsymbol{Z}}_P^\dagger = \begin{bmatrix} \tilde{\boldsymbol{A}}_{11} & \tilde{\boldsymbol{A}}_{12} \\ \tilde{\boldsymbol{A}}_{21} & \tilde{\boldsymbol{A}}_{22} \end{bmatrix}$$

with $\tilde{\boldsymbol{A}}_{11} \in \mathbb{C}^{r\times r}$, $\tilde{\boldsymbol{A}}_{12} \in \mathbb{C}^{r\times(N-r)}$, $\tilde{\boldsymbol{A}}_{22} \in \mathbb{C}^{(N-r)\times(N-r)}$, $\tilde{\boldsymbol{A}}_{21} \in \mathbb{C}^{(N-r)\times r}$, and

$$\boldsymbol{B}\tilde{\boldsymbol{B}}\boldsymbol{B}^{\dagger} = \tilde{\boldsymbol{B}}_{11} - \boldsymbol{\beta}^{\dagger}\tilde{\boldsymbol{B}}_{21} - \tilde{\boldsymbol{B}}_{12}\boldsymbol{\beta} + \boldsymbol{\beta}^{\dagger}\tilde{\boldsymbol{B}}_{22}\boldsymbol{\beta}.$$

Here

$$\tilde{\boldsymbol{B}} = \tilde{\boldsymbol{Z}}_S\tilde{\boldsymbol{Z}}_S^{\dagger} = \begin{bmatrix} \tilde{\boldsymbol{B}}_{11} & \tilde{\boldsymbol{B}}_{12} \\ \tilde{\boldsymbol{B}}_{21} & \tilde{\boldsymbol{B}}_{22} \end{bmatrix}$$

with $\tilde{\boldsymbol{B}}_{11} \in \mathbb{C}^{r\times r}$, $\tilde{\boldsymbol{B}}_{12} \in \mathbb{C}^{r\times(N-r)}$, $\tilde{\boldsymbol{B}}_{22} \in \mathbb{C}^{(N-r)\times(N-r)}$, $\tilde{\boldsymbol{B}}_{21} \in \mathbb{C}^{(N-r)\times r}$. Thus, maximization of the right-hand side of (28) with respect to $\boldsymbol{\beta}$ is tantamount to minimizing the following function

$$g(\boldsymbol{\beta}) = \mathrm{Tr}\left[\left(\boldsymbol{R}_{xx}+\tilde{\boldsymbol{R}}_{1.2}\right)^{-1}\left(-\boldsymbol{\beta}^{\dagger}\tilde{\boldsymbol{A}}_{21}-\tilde{\boldsymbol{A}}_{12}\boldsymbol{\beta}+\boldsymbol{\beta}^{\dagger}\tilde{\boldsymbol{A}}_{22}\boldsymbol{\beta}\right)\right.$$
$$\left. + \frac{1}{\gamma}\tilde{\boldsymbol{R}}_{1.2}^{-1}\left(-\boldsymbol{\beta}^{\dagger}\tilde{\boldsymbol{B}}_{21} - \tilde{\boldsymbol{B}}_{12}\boldsymbol{\beta} + \boldsymbol{\beta}^{\dagger}\tilde{\boldsymbol{B}}_{22}\boldsymbol{\beta}\right)\right].$$

Setting to zero the derivative of g with respect to $\boldsymbol{\beta}$, we have that

$$\frac{\partial}{\partial\boldsymbol{\beta}}g(\boldsymbol{\beta}) = \left(\tilde{\boldsymbol{A}}_{22}^{T}\boldsymbol{\beta}^{*} - \tilde{\boldsymbol{A}}_{12}^{T}\right)\left(\boldsymbol{R}_{xx}+\tilde{\boldsymbol{R}}_{1.2}\right)^{-T}$$
$$+ \frac{1}{\gamma}\left(\tilde{\boldsymbol{B}}_{22}^{T}\boldsymbol{\beta}^{*} - \tilde{\boldsymbol{B}}_{12}^{T}\right)\tilde{\boldsymbol{R}}_{1.2}^{-T} = \boldsymbol{0}_{N-r,r}.$$

The above equation can be rewritten as

$$\tilde{\boldsymbol{A}}_{22}^{T}\boldsymbol{\beta}^{*}\left(\tilde{\boldsymbol{R}}_{1.2}+\boldsymbol{R}_{xx}\right)^{-T} + \frac{1}{\gamma}\tilde{\boldsymbol{B}}_{22}^{T}\boldsymbol{\beta}^{*}\tilde{\boldsymbol{R}}_{1.2}^{-T} = \tilde{\boldsymbol{A}}_{12}^{T}$$
$$\times\left(\tilde{\boldsymbol{R}}_{1.2}+\boldsymbol{R}_{xx}\right)^{-T} + \frac{1}{\gamma}\tilde{\boldsymbol{B}}_{12}^{T}\tilde{\boldsymbol{R}}_{1.2}^{-T}.$$

Exploiting the identity 7.2 (7) in [31],

$$\mathrm{vec}\left(\boldsymbol{AXB}\right) = \left(\boldsymbol{B}^{T}\otimes\boldsymbol{A}\right)\mathrm{vec}\boldsymbol{X},$$

this rewriting is

$$\left[\left(\tilde{\boldsymbol{R}}_{1.2}+\boldsymbol{R}_{xx}\right)^{-1}\otimes\tilde{\boldsymbol{A}}_{22}^{T} + \frac{1}{\gamma}\tilde{\boldsymbol{R}}_{1.2}^{-1}\otimes\tilde{\boldsymbol{B}}_{22}^{T}\right]\mathrm{vec}\boldsymbol{\beta}^{*} = \mathrm{vec}\boldsymbol{C}$$

with $\boldsymbol{C} = \tilde{\boldsymbol{A}}_{12}^{T}(\tilde{\boldsymbol{R}}_{1.2}+\boldsymbol{R}_{xx})^{-T} + \frac{1}{\gamma}\tilde{\boldsymbol{B}}_{12}^{T}\tilde{\boldsymbol{R}}_{1.2}^{-T}$. Thus, the solution is given by

$$\mathrm{vec}\boldsymbol{\beta}^{*} = \left[\left(\tilde{\boldsymbol{R}}_{1.2}+\boldsymbol{R}_{xx}\right)^{-1}\otimes\tilde{\boldsymbol{A}}_{22}^{T} + \frac{1}{\gamma}\tilde{\boldsymbol{R}}_{1.2}^{-1}\otimes\tilde{\boldsymbol{B}}_{22}^{T}\right]^{-1}$$
$$\times\,\mathrm{vec}\boldsymbol{C}. \tag{32}$$

Now we observe that the matrix $\gamma\boldsymbol{R}$ can be estimated as the sample covariance matrix of the secondary data only and used to construct an estimate of $\boldsymbol{\beta} = \tilde{\boldsymbol{R}}_{22}^{-1}\tilde{\boldsymbol{R}}_{21}$; denoting by $\widehat{\boldsymbol{\beta}}^{(0)}$ this starting value we can exploit previous results to obtain after n iterations of alternating mazimization $\widehat{\tilde{\boldsymbol{R}}}_{1.2}^{(n)}$, $\widehat{\boldsymbol{R}}_{xx}^{(n)}$, and $\widehat{\boldsymbol{\beta}}^{(n)}$ that together with $\widehat{\tilde{\boldsymbol{R}}}_{22}$ allow us to compute $\widehat{\boldsymbol{R}}_{xx}$ and $\widehat{\boldsymbol{R}}_{1}$.

Finally, the approximated GLRT, referred to in the following as second-order known subspace in homogeneous environment (SO-KS-HE) detector, is given by

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\boldsymbol{R}}_{xx}, \boldsymbol{H}, 1; \boldsymbol{Z}) - L_0(\widehat{\boldsymbol{R}}_0, 1; \boldsymbol{Z}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{33}$$

with $L_0(\widehat{\boldsymbol{R}}_0, 1; \boldsymbol{Z})$ given by the logarithm of (5).

D. Known Subspace $\langle\boldsymbol{H}\rangle$, Unknown γ

Derivation of the approximated GLRT for partially-homogeneous environment is still based on alternating maximization; this time we estimate $\tilde{\boldsymbol{R}}_{1.2}$, $\boldsymbol{R}_{xx}$, and γ, given $\boldsymbol{\beta}$, and again we estimate $\boldsymbol{\beta}$, given $\tilde{\boldsymbol{R}}_{1.2}$, $\boldsymbol{R}_{xx}$, and γ, using (32).

Estimating $\tilde{\boldsymbol{R}}_{1.2}$, $\boldsymbol{R}_{xx}$, and γ, given $\boldsymbol{\beta}$, requires maximizing (30) with respect to γ. To this end, we define $\tilde{\boldsymbol{S}}_{P,2} = \tilde{\boldsymbol{Z}}_{P,2}\tilde{\boldsymbol{Z}}_{P,2}^{\dagger} \in \mathbb{C}^{(N-r)\times(N-r)}$, $\tilde{\boldsymbol{S}}_{S,2} = \tilde{\boldsymbol{Z}}_{S,2}\tilde{\boldsymbol{Z}}_{S,2}^{\dagger} \in \mathbb{C}^{(N-r)\times(N-r)}$, and $\tilde{\boldsymbol{S}}_{S,2}^{-1/2}\tilde{\boldsymbol{S}}_{P,2}\tilde{\boldsymbol{S}}_{S,2}^{-1/2} \in \mathbb{C}^{(N-r)\times(N-r)}$. It follows that

$$\max_{\tilde{\boldsymbol{R}}_{22},\tilde{\boldsymbol{R}}_{1.2},\boldsymbol{R}_{xx}} L_1(\boldsymbol{R}, \boldsymbol{R}_{xx}, \boldsymbol{H}, \gamma; \boldsymbol{Z}) = C - K\log\det\left(\tilde{\boldsymbol{S}}_{S,2}\right)$$
$$- 2\,K\log|\det(\boldsymbol{K})| - rK$$
$$- K\log\det\left(\tilde{\boldsymbol{S}}_{S,2}^{-1/2}\tilde{\boldsymbol{S}}_{P,2}\tilde{\boldsymbol{S}}_{S,2}^{-1/2} + \frac{1}{\gamma}\boldsymbol{I}_{N-r}\right) - NK_S\log\gamma$$
$$+ K\sum_{i=1}^{r}\log\frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)} + K_S\sum_{i=1}^{r}\log\widehat{\lambda}_i(\gamma). \tag{34}$$

Thus, denoting by $\delta_1 \geq \ldots \geq \delta_{N-r} \geq 0$ the eigenvalues of $\tilde{\boldsymbol{S}}_{S,2}^{-1/2}\tilde{\boldsymbol{S}}_{P,2}\tilde{\boldsymbol{S}}_{S,2}^{-1/2}$, we also have that

$$\max_{\tilde{\boldsymbol{R}}_{22},\tilde{\boldsymbol{R}}_{1.2},\boldsymbol{R}_{xx}} L_1(\boldsymbol{R}, \boldsymbol{R}_{xx}, \boldsymbol{H}, \gamma; \boldsymbol{Z}) = C - K\log\det\left(\tilde{\boldsymbol{S}}_{S,2}\right)$$
$$- 2\,K\log|\det(\boldsymbol{K})| - rK$$
$$- K\sum_{i=1}^{N-r}\log\left(\frac{1}{\gamma}+\delta_i\right) - NK_S\log\gamma$$
$$+ K\sum_{i=1}^{r}\log\frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)} + K_S\sum_{i=1}^{r}\log\widehat{\lambda}_i(\gamma). \tag{35}$$

To maximize the partially-compressed likelihood of (35) with respect to γ, we use the following theorem. For simplicity we assume $K_P \geq r$.

Theorem 5: For $K_P \geq r$, the global maximum of the function

$$f(\gamma) = -K\sum_{i=1}^{N-r}\log\left(\frac{1}{\gamma}+\delta_i\right) - NK_S\log\gamma$$
$$+ K\sum_{i=1}^{r}\log\frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)} + K_S\sum_{i=1}^{r}\log\widehat{\lambda}_i(\gamma),$$

$K = K_P + K_S$, is attained at the unique solution $\gamma^{*} \in (0, +\infty)$ of the equation

$$\sum_{i=1}^{\min(K_P, N-r)}\frac{1}{1+\gamma\delta_i} + (N - r - \min(K_P, N-r))$$
$$- \frac{(N-r)K_S}{K} = 0,$$

if $\gamma^{*} \geq \frac{K_P}{K_S}\frac{1}{\gamma_r}$. Otherwise, the stationary points of f and, hence, its global maximum belong to the interval $(\gamma^{*}, \frac{K_P}{K_S}\frac{1}{\gamma_r})$.

Proof: See Appendix C. □

Once we have determined $\widehat{\gamma}$, we can compute the approximated GLRT, referred to in the following as second-order known subspace in partially-homogeneous environment (SO-KS-PHE) detector, written as

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\boldsymbol{R}}_{xx}, \boldsymbol{H}, \widehat{\gamma}_1; \boldsymbol{Z}) - L_0(\widehat{\boldsymbol{R}}_0, \widehat{\gamma}_0; \boldsymbol{Z}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \qquad (36)$$

with $L_0(\widehat{\boldsymbol{R}}_0, \widehat{\gamma}_0; \boldsymbol{Z})$ given by the logarithm of the maximum of (5) with respect to γ obtained by using Theorem 1.

V. Conclusion

The original adaptive detectors of [5], [9], [10], as refined in [13], [14], [34] and extended in [17], [19], have been generalized by considering three new subspace signal models. In the first, the subspace visited by a sequence of symbol transmissions or reflections is assumed to be known only by its dimension; in the previous work of [17], [19] the subspace was known. In the second extension, a known subspace is visited by a sequence of symbol transmissions which are constrained by a Gaussian prior distribution; the result is a second-order adaptive subspace detector. In the third extension, the subspace is known only by its dimension; this extension requires a two-channel extension of standard factor analysis. These extensions, coupled with the results of [17], [19], comprise a unified theory of adaptive subspace detection.

In a companion paper, the performance of the detectors derived in this paper are assessed and compared with ad hoc detectors obtained by plugging the estimate of the noise covariance based on secondary data into detectors derived assuming the covariance matrix is known.

Appendix A
Proof of Theorem 3

First, consider the eigendecomposition of $\boldsymbol{R}^{-1/2}(\tilde{\boldsymbol{R}}_{xx} + \boldsymbol{R})\boldsymbol{R}^{-1/2} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\dagger}$, where $\boldsymbol{\Lambda} = \text{diag}\,(\lambda_1, \ldots, \lambda_N) \in \mathbb{R}^{N\times N}$, $\lambda_1 \geq \ldots \geq \lambda_N \geq 1$, is a diagonal matrix containing the eigenvalues[1] of $\boldsymbol{R}^{-1/2}(\tilde{\boldsymbol{R}}_{xx} + \boldsymbol{R})\boldsymbol{R}^{-1/2}$ and $\boldsymbol{U} \in \mathbb{C}^{N\times N}$ is the unitary matrix of the corresponding eigenvectors. Define $\boldsymbol{M} = \boldsymbol{R}^{1/2}\boldsymbol{U} \in \mathbb{C}^{N\times N}$. Notice that the matrix $\boldsymbol{M}$ can be any $N \times N$ non-singular matrix; in fact, from the eigendecomposition of the non-singular matrix $\boldsymbol{R}^{1/2}$, namely $\boldsymbol{R}^{1/2} = \boldsymbol{W}_1\boldsymbol{\Sigma}\boldsymbol{W}_1^{\dagger}$, it follows that $\boldsymbol{M} = \boldsymbol{W}_1\boldsymbol{\Sigma}\boldsymbol{W}_1^{\dagger}\boldsymbol{U} = \boldsymbol{W}_1\boldsymbol{\Sigma}\boldsymbol{W}_2^{\dagger}$. In addition, we obtain that

$$\boldsymbol{R} = \boldsymbol{M}\boldsymbol{M}^{\dagger} \quad \text{and} \quad \boldsymbol{R} + \tilde{\boldsymbol{R}}_{xx} = \boldsymbol{M}\boldsymbol{\Lambda}\,\boldsymbol{M}^{\dagger} \qquad (37)$$

and (22) can be recast as

$$\begin{aligned} L_1(\boldsymbol{R}, \tilde{\boldsymbol{R}}_{xx}, \gamma; \boldsymbol{Z}) &= -NK\log\pi - NK_S\log\gamma \\ &- 2K_P\log|\det\boldsymbol{M}| - K_P\log\det\boldsymbol{\Lambda} - 2K_S\log|\det\boldsymbol{M}| \\ &- \text{Tr}\left[\boldsymbol{\Lambda}^{-1}\boldsymbol{M}^{-1}\boldsymbol{S}_P\boldsymbol{M}^{-\dagger}\right] - \frac{1}{\gamma}\text{Tr}\left[\boldsymbol{M}^{-1}\boldsymbol{S}_S\boldsymbol{M}^{-\dagger}\right] \end{aligned} \qquad (38)$$

[1]The λ_is are greater than or equal to one since $\boldsymbol{R}^{-1/2}(\tilde{\boldsymbol{R}}_{xx} + \boldsymbol{R})\boldsymbol{R}^{-1/2} = \boldsymbol{I}_N + \boldsymbol{R}^{-1/2}\tilde{\boldsymbol{R}}_{xx}\boldsymbol{R}^{-1/2}$.

where we have used the facts that $\det\boldsymbol{M}^{\dagger} = (\det(\boldsymbol{M}))^*$ and $\text{Tr}(\boldsymbol{A}\boldsymbol{B}) = \text{Tr}(\boldsymbol{B}\boldsymbol{A})$, see for instance [32]. We recall that $\boldsymbol{\Gamma} = \text{diag}\,(\gamma_1, \ldots, \gamma_N) \in \mathbb{R}^{N\times N}$, $\gamma_1 \geq \ldots \geq \gamma_N \geq 0$, is the diagonal matrix containing the eigenvalues of $\boldsymbol{S}_S^{-1/2}\boldsymbol{S}_P\boldsymbol{S}_S^{-1/2}$ and $\boldsymbol{V} \in \mathbb{C}^{N\times N}$ is the unitary matrix of the corresponding eigenvectors; moreover, $\boldsymbol{K} = \boldsymbol{S}_S^{1/2}\boldsymbol{V} \in \mathbb{C}^{N\times N}$. It turns out that $\boldsymbol{S}_S = \boldsymbol{K}\boldsymbol{K}^{\dagger}$ and $\boldsymbol{S}_P = \boldsymbol{K}\boldsymbol{\Gamma}\boldsymbol{K}^{\dagger}$. Thus, we can rewrite (38) as

$$\begin{aligned} L_1(\boldsymbol{R}, \tilde{\boldsymbol{R}}_{xx}, \gamma; \boldsymbol{Z}) &= -NK\log\pi - NK_S\log\gamma \\ &- 2K_P\log|\det\boldsymbol{M}| - K_P\log\det\boldsymbol{\Lambda} - 2K_S\log|\det\boldsymbol{M}| \\ &- \text{Tr}\left[\boldsymbol{\Lambda}^{-1}\boldsymbol{M}^{-1}\boldsymbol{K}\boldsymbol{\Gamma}\boldsymbol{K}^{\dagger}\boldsymbol{M}^{-\dagger}\right] - \frac{1}{\gamma}\text{Tr}\left[\boldsymbol{M}^{-1}\boldsymbol{K}\boldsymbol{K}^{\dagger}\boldsymbol{M}^{-\dagger}\right] \\ &= -NK\log\pi - NK_S\log\gamma - 2K\log|\det\boldsymbol{M}| \\ &- K_P\log\det\boldsymbol{\Lambda} - \text{Tr}\left[\boldsymbol{X}\boldsymbol{\Gamma}\boldsymbol{X}^{\dagger}\right] - \frac{1}{\gamma}\text{Tr}\left[\boldsymbol{\Lambda}^{1/2}\boldsymbol{X}\boldsymbol{X}^{\dagger}\boldsymbol{\Lambda}^{1/2}\right] \end{aligned} \qquad (39)$$

where $\boldsymbol{X} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{M}^{-1}\boldsymbol{K} \in \mathbb{C}^{N\times N}$ and we recall that $\boldsymbol{K}$ and $\boldsymbol{\Gamma}$ are known. Before going further, we also observe that we can maximize over $\boldsymbol{X}$ for any given $\boldsymbol{\Lambda}$ since $\boldsymbol{X}$, given $\boldsymbol{\Lambda}$, is completely specified by $\boldsymbol{M}$. Let us proceed by replacing $\boldsymbol{X}$ in (39) with its singular value decomposition given by $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{D}\boldsymbol{Q}$, where $\boldsymbol{T}, \boldsymbol{Q} \in \mathbb{C}^{N\times N}$ are unitary matrices and[2] $\boldsymbol{D} = \text{diag}\,(d_1, \ldots, d_N) \in \mathbb{R}^{N\times N}$, $0 < d_1 \leq \ldots \leq d_N$. Since

$$\begin{aligned} \log|\det(\boldsymbol{M})| &= \log|\det(\boldsymbol{K})| - \log|\det(\boldsymbol{X})| - \frac{1}{2}\log\det(\boldsymbol{\Lambda}) \\ &= \log|\det(\boldsymbol{K})| - \log\det(\boldsymbol{D}) - \frac{1}{2}\log\det(\boldsymbol{\Lambda}), \end{aligned}$$

we obtain

$$\begin{aligned} L_1(\boldsymbol{R}, \tilde{\boldsymbol{R}}_{xx}, \gamma; \boldsymbol{Z}) &= -NK\log\pi - NK_S\log\gamma \\ &- 2\,K\log|\det(\boldsymbol{K})| + 2K\log\det(\boldsymbol{D}) + K_S\log\det(\boldsymbol{\Lambda}) \\ &- \text{Tr}\left[\boldsymbol{\Gamma}\boldsymbol{Q}^{\dagger}\boldsymbol{D}^2\boldsymbol{Q}\right] - \frac{1}{\gamma}\text{Tr}\left[\boldsymbol{\Lambda}\boldsymbol{T}\boldsymbol{D}^2\boldsymbol{T}^{\dagger}\right]. \end{aligned} \qquad (40)$$

Maximization with respect to $\boldsymbol{R}$ and $\tilde{\boldsymbol{R}}_{xx}$ is tantamount to maximizing with respect to $\boldsymbol{\Lambda}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{Q}$. Therefore, we have that

$$\begin{aligned} L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) &= -NK\log\pi - NK_S\log\gamma \\ &- 2K\log|\det(\boldsymbol{K})| + \max_{\boldsymbol{\Lambda}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{Q}} \Big\{2K\log\det(\boldsymbol{D}) \\ &+ K_S\log\det(\boldsymbol{\Lambda}) - \text{Tr}\left[\boldsymbol{\Gamma}\boldsymbol{Q}^{\dagger}\boldsymbol{D}^2\boldsymbol{Q}\right] - \frac{1}{\gamma}\text{Tr}\left[\boldsymbol{\Lambda}\boldsymbol{T}\boldsymbol{D}^2\boldsymbol{T}^{\dagger}\right]\Big\}. \end{aligned} \qquad (41)$$

Exploiting *Theorem 1* of [35], it follows that

$$\begin{aligned} \min_{\boldsymbol{Q}} \text{Tr}\left[\boldsymbol{\Gamma}\boldsymbol{Q}^{\dagger}\boldsymbol{D}^2\boldsymbol{Q}\right] + \frac{1}{\gamma}\min_{\boldsymbol{T}} \text{Tr}\left[\boldsymbol{\Lambda}\boldsymbol{T}\boldsymbol{D}^2\boldsymbol{T}^{\dagger}\right] \\ = \text{Tr}\left[\left(\boldsymbol{\Gamma} + \frac{1}{\gamma}\boldsymbol{\Lambda}\right)\boldsymbol{D}^2\right] \end{aligned}$$

[2]Notice that the singular values of $\boldsymbol{X}$ have been arranged in ascending order and not, as customary, in descending order [32].

where the maximizers are $\widehat{\boldsymbol{T}} = \text{diag}\,(e^{j\theta_{1T}}, \ldots, e^{j\theta_{NT}})$ and $\widehat{\boldsymbol{Q}} = \text{diag}\,(e^{j\theta_{1Q}}, \ldots, e^{j\theta_{NQ}})$, $\forall \theta_{iT}, \theta_{iQ} \in [0, 2\pi]$, $i = 1, \ldots, N$. Hence[3]

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma$$
$$- 2\,K \log |\det(\boldsymbol{K})| + \max_{\boldsymbol{\Lambda}, \boldsymbol{D}} \left\{ 2K \log\det(\boldsymbol{D}) + K_S \log\det(\boldsymbol{\Lambda}) \right.$$
$$\left. - \text{Tr}\left[\left(\boldsymbol{\Gamma} + \frac{1}{\gamma}\boldsymbol{\Lambda} \right) \boldsymbol{D}^2 \right] \right\} = -NK \log \pi - NK_S \log \gamma$$
$$- 2\,K \log |\det(\boldsymbol{K})| + \max_{\boldsymbol{\Lambda}, \boldsymbol{D}} \left[K \sum_{i=1}^{N} \log d_i^2 + K_S \sum_{i=1}^{N} \log \Lambda_i \right.$$
$$\left. - \sum_{i=1}^{N} d_i^2 \left(\gamma_i + \frac{1}{\gamma}\lambda_i \right) \right]. \tag{42}$$

Now we compute the maximum with respect to d_i^2, given λ_i, of the function

$$g_i(d_i^2) = K \log d_i^2 + K_S \log \lambda_i - d_i^2 \left(\gamma_i + \frac{1}{\gamma}\lambda_i \right).$$

First notice that g_i tends to $-\infty$ as d_i^2 tends to zero or to $+\infty$; moreover, its derivative with respect to d_i^2 is positive iff

$$\frac{K}{d_i^2} - \left(\gamma_i + \frac{1}{\gamma}\lambda_i \right) > 0,$$

implying that the maximizer (given λ_i) is $d_i^2 = \frac{\gamma K}{\gamma\gamma_i + \lambda_i}$. It follows that (42) yields

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma - NK$$
$$- 2\,K \log |\det(\boldsymbol{K})|$$
$$+ \max_{\boldsymbol{\Lambda}} \left[K \sum_{i=1}^{N} \log \frac{\gamma K}{\gamma\gamma_i + \lambda_i} + K_S \sum_{i=1}^{N} \log \lambda_i \right]. \tag{43}$$

Now observe that even though the rank of $\boldsymbol{H}$ is known (and equal to r) that of $\tilde{\boldsymbol{R}}_{xx}$ is unknown (less than or equal to r). Thus, maximization over $\boldsymbol{\Lambda}$ of (43) is limited to $N \times N$ diagonal matrices with (at most) r eigenvalues greater than one and the remaining equal to one.

In order to address the last maximization step we first observe that the function

$$h_i(\lambda_i) = K \log \frac{\gamma K}{\gamma\gamma_i + \lambda_i} + K_S \log \lambda_i$$

tends to $-\infty$ as λ_i tends to $+\infty$; its derivative is positive iff

$$-\frac{K}{\gamma\gamma_i + \lambda_i} + \frac{K_S}{\lambda_i} > 0$$

or equivalently iff $K_P \lambda_i < K_S \gamma\gamma_i$. Thus, it follows that h_i has global maximum over $[1, +\infty)$ at

$$\tilde{\lambda}_i = \max\left(\frac{K_S \gamma\gamma_i}{K_P}, 1 \right).$$

[3]In the following without loss of generality we will assume that $\widehat{\boldsymbol{T}} = \boldsymbol{I}_N$ and $\widehat{\boldsymbol{Q}} = \boldsymbol{I}_N$.

In particular, $\gamma_i = 0$ implies $\tilde{\lambda}_i = 1$. Then

- if $\gamma < \frac{K_P}{K_S}\frac{1}{\gamma_1}$, all maximizers are equal to one and

$$\max_{\lambda_i} h_i(\lambda_i) = K \log \frac{\gamma K}{\gamma\gamma_i + 1}, \quad i = 1, \ldots, N.$$

Thus, the compressed likelihood under H_1 becomes

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma$$
$$- 2\,K \log |\det(\boldsymbol{K})| + K \sum_{i=1}^{N} \log \frac{\gamma K}{\gamma\gamma_i + 1} - NK. \tag{44}$$

- if $\frac{K_P}{K_S}\frac{1}{\gamma_{i-1}} \le \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_i}$, $i = 2, \ldots, r$, the maximizers of h_j are (this case refers to $r > 1$)

$$\tilde{\lambda}_j(\gamma) = \frac{K_S \gamma\gamma_j}{K_P} > 1, \quad j = 1, \ldots, i-1,$$

and the remaining maximizers are $\tilde{\lambda}_j(\gamma) = 1$, $j = i, \ldots, N$. Thus, the compressed likelihood under H_1 becomes

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma$$
$$-2\,K \log |\det(\boldsymbol{K})| + \sum_{j=1}^{i-1} \left[K \log \frac{K_P}{\gamma_j} + K_S \log \frac{K_S \gamma\gamma_j}{K_P} \right]$$
$$+ \sum_{j=i}^{N} K \log \frac{\gamma K}{\gamma\gamma_j + 1} - NK. \tag{45}$$

- if $K_P = r$ and $\gamma \ge \frac{K_P}{K_S}\frac{1}{\gamma_r}$ the maximizers of h_i are

$$\tilde{\lambda}_i(\gamma) = \frac{K_S \gamma\gamma_i}{K_P} > 1, \quad i = 1, \ldots, r,$$

and the remaining maximizers are $\tilde{\lambda}_i(\gamma) = 1$, $i = r+1, \ldots, N$ (as a matter of fact, $\gamma_{K_P+1} = \cdots = \gamma_N = 0$). Thus, the compressed likelihood under H_1 becomes

$$L_1(\widehat{\boldsymbol{R}}_1, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma; \boldsymbol{Z}) = -NK \log \pi - NK_S \log \gamma$$
$$-2\,K \log |\det(\boldsymbol{K})| + \sum_{i=1}^{r} \left[K \log \frac{K_P}{\gamma_i} + K_S \log \frac{K_S \gamma\gamma_i}{K_P} \right]$$
$$+ \sum_{i=r+1}^{N} K \log \frac{\gamma K}{\gamma\gamma_i + 1} - NK. \tag{46}$$

If instead $K_P = r + m$, $m \ge 1$, we have to distinguish the following cases

– if $\frac{K_P}{K_S}\frac{1}{\gamma_r} \le \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_{r+1}}$ the maximizers of h_i are

$$\tilde{\lambda}_i(\gamma) = \frac{K_S \gamma\gamma_i}{K_P} > 1, \quad i = 1, \ldots, r,$$

and the remaining maximizers are $\tilde{\lambda}_i(\gamma) = 1$, $i = r+1, \ldots, N$. The compressed likelihood under H_1 is still given by (46).

– if $\frac{K_P}{K_S}\frac{1}{\gamma_{r+h}} \leq \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_{r+h+1}}$, $h = 1, \ldots, m-1$, the maximizers of h_i are (this case refers to $m > 1$)

$$\tilde{\lambda}_i(\gamma) = \frac{K_S \gamma \gamma_i}{K_P} > 1, \quad i = 1, \ldots, r+h,$$

and the remaining maximizers are $\tilde{\lambda}_i(\gamma) = 1$, $i = r+h+1, \ldots, N$. However, the fact that the maximum number of $\lambda_i > 1$ has to be r (at most) together with the descending order of the λ_is implies that the compressed likelihood under H_1 is still given by (46).

– Finally, if $\gamma \geq \frac{K_P}{K_S}\frac{1}{\gamma_{r+m}}$ the maximizers of h_i are

$$\tilde{\lambda}_i(\gamma) = \frac{K_S \gamma \gamma_i}{K_P} > 1, \quad i = 1, \ldots, r+m,$$

and the remaining maximizers (if any) are $\tilde{\lambda}_i(\gamma) = 1$, $i = r+m+1, \ldots, N$ (as a matter of fact, $\gamma_{K_P+1} = \cdots = \gamma_N = 0$). However, the compressed likelihood under H_1 is still given by (46).

The statement of Theorem 3 follows.

Following the lead of previous derivation, it is also straightforward to prove the following result where the rank of $\tilde{\boldsymbol{R}}_{xx}$ is assumed unknown (i.e., less than or equal to N). It will be exploited in the derivation of the second-order detectors for known subspace.

Theorem 6: Let $\boldsymbol{S}_P \in \mathbb{C}^{N\times N}$ be a positive semidefinite matrix and $\boldsymbol{S}_S \in \mathbb{C}^{N\times N}$ a positive definite matrix. Define $\boldsymbol{T}_P = \boldsymbol{S}_S^{-1/2}\boldsymbol{S}_P\boldsymbol{S}^{-1/2} = \boldsymbol{V}\boldsymbol{\Gamma}\boldsymbol{V}^\dagger, \boldsymbol{V}^\dagger\boldsymbol{V} = \boldsymbol{I}_N, \boldsymbol{\Gamma} = \text{diag}\,[\gamma_1, \ldots, \gamma_N], \gamma_1 \geq \ldots \geq \gamma_N$. Then, the function

$$\begin{aligned} h(\boldsymbol{R}, \tilde{\boldsymbol{R}}_{xx}, \gamma) = & -NK_S \log\gamma - K_P \log\det(\tilde{\boldsymbol{R}}_{xx} + \boldsymbol{R}) \\ & - K_S \log\det(\boldsymbol{R}) \\ & - \text{Tr}\left[\left(\tilde{\boldsymbol{R}}_{xx} + \boldsymbol{R}\right)^{-1}\boldsymbol{S}_P\right] - \frac{1}{\gamma}\text{Tr}\left[\boldsymbol{R}^{-1}\boldsymbol{S}_S\right] \end{aligned}$$

for any $\gamma > 0$ is maximized over the set of positive definite matrices $\boldsymbol{R} \in \mathbb{C}^{N\times N}$ and positive semidefinite matrices $\tilde{\boldsymbol{R}}_{xx} \in \mathbb{C}^{N\times N}$ with unknown rank at

$$\widehat{\boldsymbol{R}} = \widehat{\boldsymbol{M}}\widehat{\boldsymbol{M}}^\dagger \quad \text{and} \quad \widehat{\tilde{\boldsymbol{R}}}_{xx} = \widehat{\boldsymbol{M}}\widehat{\boldsymbol{\Lambda}}\widehat{\boldsymbol{M}}^\dagger - \widehat{\boldsymbol{R}}$$

where $\widehat{\boldsymbol{\Lambda}} = \text{diag}\,(\widehat{\lambda}_1, \ldots, \widehat{\lambda}_N) \in \mathbb{R}^{N\times N}$, and

$$\widehat{\lambda}_i(\gamma) = \max\left(\frac{K_S}{K_P}\gamma\gamma_i, 1\right), \quad i = 1, \ldots, N.$$

The matrix $\widehat{\boldsymbol{M}}$ is

$$\widehat{\boldsymbol{M}} = \boldsymbol{K}\widehat{\boldsymbol{D}}^{-1}\widehat{\boldsymbol{\Lambda}}^{-1/2}$$

where $\boldsymbol{K} = \boldsymbol{S}_S^{1/2}\boldsymbol{V} \in \mathbb{C}^{N\times N}$ and $\widehat{\boldsymbol{D}}^2 = \text{diag}(\widehat{d}_1^2, \ldots, \widehat{d}_N^2) \in \mathbb{R}^{N\times N}$,

$$\widehat{d}_i^2(\gamma) = \frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)}, \quad i = 1, \ldots, N.$$

The maximum of h is

$$\begin{aligned} h(\widehat{\boldsymbol{R}}, \widehat{\tilde{\boldsymbol{R}}}_{xx}, \gamma) = & -NK_S \log\gamma - 2K \log|\det(\boldsymbol{K})| - NK \\ & + K\sum_{i=1}^{N}\log\frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)} + K_S\sum_{i=1}^{N}\log\widehat{\lambda}_i(\gamma) \end{aligned} \tag{47}$$

with $K = K_P + K_S$.

Appendix B
Proof of Theorem 4

Define

$$g(\gamma) = -NK_S \log\gamma + g_1(\gamma)$$

with

$$g_1(\gamma) = \begin{cases} \sum_{i=1}^{r-1} K\log\frac{\gamma K}{\gamma\gamma_i + \widehat{\lambda}_i(\gamma)} + \sum_{i=1}^{r-1} K_S \log\widehat{\lambda}_i(\gamma) \\ + \sum_{i=r}^{N} K\log\frac{\gamma K}{\gamma\gamma_i + 1}, \\ \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_r} \\ \sum_{i=1}^{r}\left[K\log\frac{K_P}{\gamma_i} + K_S\log\frac{K_S\gamma\gamma_i}{K_P}\right] \\ + \sum_{i=r+1}^{N} K\log\frac{\gamma K}{\gamma\gamma_i + 1}, \\ \gamma \geq \frac{K_P}{K_S}\frac{1}{\gamma_r} \end{cases}.$$

More specifically, we have that

$$g_1(\gamma) = \begin{cases} K\sum_{i=1}^{N}\log\frac{\gamma K}{\gamma\gamma_i + 1}, \\ \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_1} \\ \sum_{i=1}^{j-1}\left[K\log\frac{K_P}{\gamma_i} + K_S\log\frac{K_S\gamma\gamma_i}{K_P}\right] \\ + K\sum_{i=j}^{N}\log\frac{\gamma K}{\gamma\gamma_i + 1}, \\ \frac{K_P}{K_S}\frac{1}{\gamma_{j-1}} \leq \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_j}, j = 2, \ldots, r \\ \sum_{i=1}^{r}\left[K\log\frac{K_P}{\gamma_i} + K_S\log\frac{K_S\gamma\gamma_i}{K_P}\right] \\ + \sum_{i=r+1}^{N} K\log\frac{\gamma K}{\gamma\gamma_i + 1}, \\ \gamma \geq \frac{K_P}{K_S}\frac{1}{\gamma_r} \end{cases}. \tag{48}$$

Then, notice that

$$\lim_{\gamma\to 0} g(\gamma) = \lim_{\gamma\to 0}\left[-NK_S\log\gamma + K\sum_{i=1}^{N}\log\frac{\gamma K}{\gamma\gamma_i + 1}\right] = -\infty$$

and

$$\begin{aligned} \lim_{\gamma\to+\infty} g(\gamma) &= \lim_{\gamma\to+\infty}\left[-NK_S\log\gamma + \sum_{i=1}^{r} K_S\log\frac{K_S\gamma\gamma_i}{K_P}\right. \\ &\quad \left. + \sum_{i=r+1}^{N} K\log\frac{\gamma K}{\gamma\gamma_i + 1}\right] + \sum_{i=1}^{r} K\log\frac{K_P}{\gamma_i} \\ &= \lim_{\gamma\to+\infty}\left[(r-N)K_S\log\gamma + (N-r)K\log\gamma\right. \\ &\quad \left. + \sum_{i=r+1}^{N} K\log\frac{K}{\gamma\gamma_i + 1}\right] + \sum_{i=1}^{r} K\log\frac{K_P}{\gamma_i} \\ &\quad + \sum_{i=1}^{r} K_S\log\frac{K_S\gamma_i}{K_P} \\ &= \lim_{\gamma\to+\infty}\left[K_P(N-r)\log\gamma - K\sum_{i=r+1}^{K_P}\log(\gamma\gamma_i + 1)\right] \end{aligned}$$

$$+\sum_{i=1}^{r} K\log\frac{K_P}{\gamma_i}+\sum_{i=1}^{r} K_S\log\frac{K_S\gamma_i}{K_P}$$
$$+\sum_{i=r+1}^{N} K\log K=-\infty$$

exploiting $\gamma_i\neq 0, i=r+1,\ldots,K_P$, and provided that $(K_P-r)K>(N-r)K_P$ or, equivalently, $(K_P-r)K_S>(N-r)K_P-(K_P-r)K_P=(N-K_P)K_P$.

Thus, the maximum corresponds to a stationary point. To compute the stationary points we observe that

$$\frac{dg_1}{d\gamma}(\gamma)=\begin{cases}\begin{cases}\sum_{i=1}^{N}\frac{K}{\gamma(\gamma\gamma_i+1)},\\ \gamma<\frac{K_P}{K_S}\frac{1}{\gamma_1}\end{cases}\\ \begin{cases}\sum_{i=1}^{j-1}\frac{K_S}{\gamma}+\sum_{i=j}^{N}\frac{K}{\gamma(\gamma\gamma_i+1)},\\ \frac{K_P}{K_S}\frac{1}{\gamma_{j-1}}\leq\gamma<\frac{K_P}{K_S}\frac{1}{\gamma_j}, j=2,\ldots,r\end{cases}\\ \begin{cases}\frac{rK_S}{\gamma}+\sum_{i=r+1}^{N}\frac{K}{\gamma(\gamma\gamma_i+1)},\\ \gamma\geq\frac{K_P}{K_S}\frac{1}{\gamma_r}\end{cases}\end{cases}.$$

It is easy to check that

$$\frac{dg}{d\gamma}(\gamma)=-\frac{NK_S}{\gamma}+\frac{dg_1}{d\gamma}(\gamma)>0,\quad \gamma<\frac{K_P}{K_S}\frac{1}{\gamma_1}.$$

In fact, $\gamma\gamma_i<\frac{K_P}{K_S}$ implies

$$\frac{dg_1}{d\gamma}(\gamma)=\sum_{i=1}^{N}\frac{K}{\gamma(\gamma\gamma_i+1)}>\frac{1}{\gamma}\sum_{i=1}^{N}K_S=\frac{NK_S}{\gamma}.$$

Similarly, $\gamma\gamma_i<\frac{K_P}{K_S}$, $i=j,\ldots,N$, implies that

$$\sum_{i=j}^{N}\frac{K}{\gamma(\gamma\gamma_i+1)}>\frac{1}{\gamma}\sum_{i=j}^{N}K_S=\frac{1}{\gamma}K_S(N-j+1)$$

and eventually

$$\frac{dg}{d\gamma}(\gamma)>0,\quad \frac{K_P}{K_S}\frac{1}{\gamma_{j-1}}\leq\gamma<\frac{K_P}{K_S}\frac{1}{\gamma_j}, j=2,\ldots,r.$$

Moreover, for $\gamma\geq\frac{K_P}{K_S}\frac{1}{\gamma_r}$,

$$\frac{dg}{d\gamma}(\gamma)=-\frac{NK_S}{\gamma}+\frac{rK_S}{\gamma}+\sum_{i=r+1}^{N}\frac{K}{\gamma(\gamma\gamma_i+1)}$$
$$=-\frac{(N-r)K_S}{\gamma}+\sum_{i=r+1}^{K_P}\frac{K}{\gamma(\gamma\gamma_i+1)}+\frac{(N-K_P)K}{\gamma}$$
$$\geq-\frac{(N-r)K_S}{\gamma}+\frac{K(K_P-r)}{\gamma(\gamma\gamma_r+1)}+\frac{(N-K_P)K}{\gamma}.$$

In particular, $\frac{dg}{d\gamma}(\gamma)$ is non-negative at $\gamma=\frac{K_P}{K_S}\frac{1}{\gamma_r}$. Thus, since

$$\lim_{\gamma\to+\infty}\gamma\frac{dg}{d\gamma}(\gamma)=-(N-r)K_S+(N-K_P)K$$
$$=-(N-r)K_S+(N-K_P)K_S+(N-K_P)K_P$$
$$=K_S(r-K_P)+(N-K_P)K_P<0,$$

it follows that the maximum is attained at the unique solution of the equation[4]

$$\sum_{i=r+1}^{K_P}\frac{K}{(\gamma\gamma_i+1)}=-(N-K_P)K+(N-r)K_S$$
$$=(K_P-r)K_S-(N-K_P)K_P.$$

Appendix C
Proof of Theorem 5

Observe that $K_P\geq r$ implies that the rank of the matrix $\tilde{\boldsymbol{S}}_S^{-1/2}\tilde{\boldsymbol{S}}_P\tilde{\boldsymbol{S}}_S^{-1/2}\in\mathbb{C}^{r\times r}$ is r, and hence that $\gamma_r\neq 0$ (wp 1).[5] Define

$$f(\gamma)=-rK_S\log\gamma+f_1(\gamma)+f_2(\gamma)$$

where

$$f_1(\gamma)=K\sum_{i=1}^{r}\log\frac{\gamma K}{\gamma\gamma_i+\widehat{\lambda}_i(\gamma)}+K_S\sum_{i=1}^{r}\log\widehat{\lambda}_i(\gamma)$$

and

$$f_2(\gamma)=-K\sum_{i=1}^{N-r}\log\left(\frac{1}{\gamma}+\delta_i\right)-(N-r)K_S\log\gamma.$$

Thus, supposing that $K_P\geq r$, we have that

$$f_1(\gamma)=\begin{cases}\begin{cases}K\sum_{i=1}^{r}\log\frac{\gamma K}{\gamma\gamma_i+1},\\ \gamma<\frac{K_P}{K_S}\frac{1}{\gamma_1}\end{cases}\\ \begin{cases}K\sum_{i=1}^{j}\log\frac{K_P}{\gamma_i}+K\sum_{i=j+1}^{r}\log\frac{\gamma K}{\gamma\gamma_i+1}\\ +K_S\sum_{i=1}^{j}\log\left(\frac{K_S}{K_P}\gamma\gamma_i\right),\\ \frac{K_P}{K_S}\frac{1}{\gamma_j}\leq\gamma<\frac{K_P}{K_S}\frac{1}{\gamma_{j+1}}, 1\leq j\leq r-1\end{cases}\\ \begin{cases}K\sum_{i=1}^{r}\log\frac{K_P}{\gamma_i}+K_S\sum_{i=1}^{r}\log\left(\frac{K_S}{K_P}\gamma\gamma_i\right),\\ \gamma\geq\frac{K_P}{K_S}\frac{1}{\gamma_r}\end{cases}\end{cases}.$$

Write $f_2(\gamma)$ as

$$f_2(\gamma)=-K\sum_{i=1}^{\min(K_P,N-r)}\log\left(\frac{1+\gamma\delta_i}{\gamma}\right)$$
$$+[K(N-r-\min(K_P,N-r))-(N-r)K_S]\log\gamma.$$

Thus, it is easy to check that $\lim_{\gamma\to 0}f(\gamma)=-\infty$ and $\lim_{\gamma\to+\infty}f(\gamma)=-\infty$. It follows that the global maximum is achieved at a stationary point.

[4]The solution is apparently unique since the left-hand side of the equation is a stricly decreasing function of $\gamma>0$.

[5]First observe that $\tilde{\boldsymbol{S}}_P=\boldsymbol{B}\tilde{\boldsymbol{Z}}_P\tilde{\boldsymbol{Z}}_P^{\dagger}\boldsymbol{B}^{\dagger}$ with $\tilde{\boldsymbol{Z}}_P=[\tilde{\boldsymbol{z}}_1\cdots\tilde{\boldsymbol{z}}_{K_P}]$. Moreover, $\tilde{\boldsymbol{z}}_k\sim\mathcal{CN}_N(\boldsymbol{0}_{N,1},\boldsymbol{C})$ with $\boldsymbol{C}\in\mathbb{C}^{N\times N}$ a positive definite matrix. It follows $\boldsymbol{B}\tilde{\boldsymbol{z}}_k\sim\mathcal{CN}_r(\boldsymbol{0}_{r,1},\boldsymbol{D})$ with $\boldsymbol{D}=\boldsymbol{B}\boldsymbol{C}\boldsymbol{B}^{\dagger}\in\mathbb{C}^{r\times r}$ a positive definite matrix (in fact, $\mathrm{rk}\{\boldsymbol{B}\}=r$ and hence $\mathrm{rk}\{\boldsymbol{B}\boldsymbol{C}^{1/2}\}=r$, exploiting property 4.3.1 (8) in [31]). The result follows from the fact that the columns of $\tilde{\boldsymbol{Z}}_P$ are independent vectors.

Moreover, the derivative of $f_1'(\gamma) = -rK_S \log\gamma + f_1(\gamma)$ with respect to γ is given by

$$\frac{df_1'(\gamma)}{d\gamma} = \begin{cases} -\frac{rK_S}{\gamma} + \frac{K}{\gamma}\sum_{i=1}^{r}\frac{1}{\gamma\gamma_i+1}, \\ \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_1} \\ -\frac{(r-j)K_S}{\gamma} + \frac{K}{\gamma}\sum_{i=j+1}^{r}\frac{1}{\gamma\gamma_i+1}, \\ \frac{K_P}{K_S}\frac{1}{\gamma_j} \le \gamma < \frac{K_P}{K_S}\frac{1}{\gamma_{j+1}}, 1 \le j \le r-1 \\ 0, \\ \gamma \ge \frac{K_P}{K_S}\frac{1}{\gamma_r} \end{cases}$$

while the derivative of $f_2(\gamma)$ can be written as

$$\frac{df_2(\gamma)}{d\gamma} = \frac{1}{\gamma}\left[K\sum_{i=1}^{\min(K_P,N-r)}\frac{1}{1+\gamma\delta_i} + (K(N-r-\min(K_P,N-r)) - (N-r)K_S)\right].$$

Notice that $\frac{df_1'(\gamma)}{d\gamma}$ is positive if $\gamma < \frac{K_P}{K_S}\frac{1}{\gamma_r}$ and is equal to zero if $\gamma \ge \frac{K_P}{K_S}\frac{1}{\gamma_r}$. It is apparent that the derivative of $f_2(\gamma)$ is positive and strictly decreasing up to the unique value of γ, say γ^*, that solves the equation

$$\sum_{i=1}^{\min(K_P,N-r)}\frac{1}{1+\gamma\delta_i} + (N-r-\min(K_P,N-r)) - \frac{(N-r)K_S}{K} = 0,$$

while it is negative if $\gamma > \gamma^*$. In fact, the function $\sum_{i=1}^{\min(K_P,N-r)}\frac{1}{1+\gamma\delta_i}$ is positive and strictly decreasing. Moreover,

$$\lim_{\gamma\to 0} K\sum_{i=1}^{\min(K_P,N-r)}\frac{1}{1+\gamma\delta_i} = K\min(K_P,N-r)$$

and

$$\min(K_P,N-r) + \frac{(N-r)K_P}{K} - \min(K_P,N-r) > 0,$$

but

$$\lim_{\gamma\to+\infty} K\sum_{i=1}^{\min(K_P,N-r)}\frac{1}{1+\gamma\delta_i} = 0$$

and $K(N-r-\min(K_P,N-r)) - (N-r)K_S < 0$ (when $K_P \le N-r$ recall that $K_S \ge N$).

As a consequence we conclude that if $\gamma^* \ge \frac{K_P}{K_S}\frac{1}{\gamma_r}$ the derivative of $f(\gamma)$ has a unique zero at γ^*; otherwise, the stationary points of f and hence its global maximum belong to the interval $(\gamma^*, \frac{K_P}{K_S}\frac{1}{\gamma_r})$. In fact,

$$\frac{df}{d\gamma}(\gamma^*) = \frac{df_1'}{d\gamma}(\gamma^*) + \frac{df_2}{d\gamma}(\gamma^*) > 0$$

and

$$\frac{df}{d\gamma}\left(\frac{K_P}{K_S}\frac{1}{\gamma_r}\right) = \frac{df_1'}{d\gamma}\left(\frac{K_P}{K_S}\frac{1}{\gamma_r}\right) + \frac{df_2}{d\gamma}\left(\frac{K_P}{K_S}\frac{1}{\gamma_r}\right) < 0.$$

REFERENCES

[1] H. L. Van Trees, *Optimum Array Processing* (Detection, Estimation, and Modulation Theory, Part IV). Hoboken, NJ, USA: Wiley, 2002.

[2] R. Klemm, *Principles of Space-Time Adaptive Processing.* (IEE Radar, Sonar, Navigation and Avionics Series 12), London, U.K.: Inst. Elec. Eng., 2002.

[3] M. I. Skolnik, *Radar Handbook* (*Ser. Electronic Eng.*). New York, NY, USA: McGraw-Hill, 1990.

[4] F. Bandiera, D. Orlando, and G. Ricci, "Advanced radar detection schemes under mismatched signal models," in *Synthesis Lectures on Signal Processing*, San Rafael, California, CA, USA: Morgan & Claypool Publishers, 2009.

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

[6] E. J. Kelly and K. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal models," Lincoln Lab, MIT, Lexington, US, Tech. Rep. 848, 1989.

[7] W.-S. Chen and I. S. Reed, "A new CFAR detection test for radar," *Digit. Signal Process.*, vol. 1, no. 4, pp. 198–214, 1991.

[8] F. Robey, D. Fuhrmann, E. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[9] E. Conte, M. Lops, and G. Ricci, "Asymptotically optimum radar detection in compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 617–625, Apr. 1995.

[10] L. Scharf and L. T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proc. Conf. Rec. 30th Asilomar Conf. Signals, Syst. Compu.*, 1996, vol. 2, pp. 1114–1117.

[11] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Boston, MA, USA: Addison-Wesley Publishing Company, 1991.

[12] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

[13] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2538–2541, 1999.

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

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

[16] K. Gerlach, M. Steiner, and F. C. Lin, "Detection of a spatially distributed target in white noise," *IEEE Signal Process. Lett.*, vol. 4, no. 7, pp. 198–200, Jul. 1997.

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

[18] F. Gini and A. Farina, "Vector subspace detection in compound-Gaussian clutter. Part I: Survey and new results," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1295–1311, Oct. 2002.

[19] 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. Signal Process.*, vol. 55, no. 4, pp. 1223–1237, Apr. 2007.

[20] F. Bandiera, O. Besson, D. Orlando, G. Ricci, and L. L. Scharf, "GLRT-based direction detectors in homogeneous noise and subspace interference," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2386–2394, Jun. 2007.

[21] D. Ciuonzo, A. De Maio, and D. Orlando, "A unifying framework for adaptive radar detection in homogeneous plus structured interference–Part I: On the maximal invariant statistic," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2894–2906, Jun. 2016.

[22] D. Ciuonzo, A. De Maio, and D. Orlando, "A unifying framework for adaptive radar detection in homogeneous plus structured interference–Part II: Detectors design," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2907–2919, Jun. 2016.

[23] A. Coluccia, A. Fascista, and G. Ricci, "A novel approach to robust radar detection of range-spread targets," *Signal Process.*, vol. 166, 2020, Art. no. 107223.

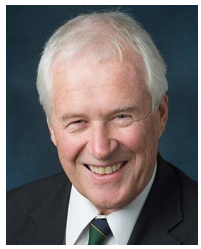
[24] G. Ricci and L. L. Scharf, "Adaptive radar detection of extended Gaussian targets," in *Proc. 12th Annu. Adaptive Sensor Array Process. Workshop*, Lexington, MA, USA, Mar. 16–18, 2004, vol. 1.

[25] O. Besson, A. Coluccia, E. Chaumette, G. Ricci, and F. Vincent, "Generalized likelihood ratio test for detection of Gaussian rank-one signals in Gaussian noise with unknown statistics," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 1082–1092, Feb. 2017.
[26] A. Antman and A. Leshem, "Radio transient detection in radio astronomical arrays," *IEEE Trans. Signal Process.*, vol. 68, pp. 5648–5663, 2020.
[27] H. L. Van Trees, *Detection, Estimation, and Modulation Theory* (*Part I: Detection, Estimation, and Linear Modulation Theory*). Hoboken, NJ, USA: Wiley, 2004.
[28] Y. Jin and B. Friedlander, "A CFAR adaptive subspace detector for second-order Gaussian signals," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 871–884, Mar. 2005.
[29] M. A. Richards, W. A. Holm, and J. Scheer, *Principles of Modern Radar: Basic Principles (Electromagnetics and Radar Series)*. London, U.K.: Institution of Engineering and Technology, 2010.
[30] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. Hoboken, NJ, USA: Wiley, 1982.
[31] H. Lütkepohl, *Handbook of Matrices*. Hoboken, NJ, USA: Wiley, 1996.
[32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1985.
[33] P. Stoica and Y. Selen, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: A refresher," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 112–114, Jan. 2004.
[34] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.
[35] L. Mirsky, "On the trace of matrix products," *Mathematische Nachrichten*, vol. 20, pp. 171–174, 1959.

Danilo Orlando (Senior Member, IEEE) was born in Gagliano del Capo, Italy, on August 9, 1978. He received the Dr. Eng. degree (with Hons.) in computer engineering and the Ph.D. degree (with max. score) in information engineering from the University of Salento (formerly University of Lecce), Lecce, Italy, in 2004 and 2008, respectively. From July 2007 to July 2010, he was with the University of Cassino, Cassino, Italy, engaged in a Research Project on algorithms for track-before-detect of multiple targets in uncertain scenarios. From September to November 2009, he was a Visiting Scientist with the NATO Undersea Research Centre, La Spezia, Italy. From September 2011 to April 2015, he was with Elettronica S.p.A. engaged as System Analyst with the field of electronic warfare. In May 2015, he joined Università degli Studi Niccolò Cusano, Rome, Italy, where he is currently an Associate Professor. In 2007, he has held Visiting Positions with the Department of Avionics and Systems, ENSICA (now Institut Supérieur de l'Aéronautique et de l'Espace, ISAE), Toulouse, France, and from 2017 to 2019, with Chinese Academy of Science, Beijing, China. He is the author or coauthor of more than 150 scientific publications in international journals, conferences, and books. His main research interests include statistical signal processing with more emphasis on adaptive detection and tracking of multiple targets in multisensor scenarios. He was the Senior Area Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the Associate Editor for the IEEE OPEN JOURNAL ON SIGNAL PROCESSING, *EURASIP Journal on Advances in Signal Processing, and MDPI Remote Sensing*.

Giuseppe Ricci (Senior Member, IEEE) was born in Naples, Italy, on February 15, 1964. He received the Dr. and Ph.D. degrees in electronic engineering from the University of Naples "Federico II", Naples, Italy, in 1990 and 1994, respectively. Since 1995, he has been with the University of Salento (formerly University of Lecce), Lecce, Italy, as an Assistant Professor of telecommunications and since 2002, as a Professor. During 1997–1998 and in April/May 2001, he has held Visiting Positions with the University of Colorado, Boulder, CO, USA. In July/September 2003, March 2005, September 2009, and March 2011, respectively, he has held Visiting Positions with Colorado State University, Fort Collins, CO, USA. In March 2006, he has a visiting position with Ensica, Toulouse, France. In September 2008, he has held a Visiting Position with the University of Connecticut, Storrs, CT, USA, and with Arizona State University, Tempe, AZ, USA, in July/August 2019. He is the Member of the Multisensor Target Tracking Working Group, operating under the auspices of the International Society of Information Fusion, from 2012 till 2016, and of the Special Area Team *Eurasip Signal Processing for Multisensor Systems* from 2016 till 2021. He is also the tutorial Co-Chair of the 2020 IEEE Radar Conference, Florence, Italy, 21–25 September, 2020, and the Rector Delegate for teaching activities and continuing education from February 2016 and June 2016, respectively, till April 2017.

His research interests are in the field of statistical signal processing with emphasis on radar processing, localization algorithms, and CDMA systems.



Louis L. Scharf (Life Fellow, IEEE) received the Ph.D. degree from University of Washington, Seattle, WA, USA, in 1969. He is currently Research Professor of Mathematics and Emeritus Professor of Electrical and Computer Engineering at Colorado State University, Fort Collins, CO, USA. He holds a courtesy appointment in Statistics. He has coauthored the books, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, 1991, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*, Cambridge University Press, 2010. His coauthored book, *A First Course in Electrical and Computer Engineering*, Addison-Wesley, Reading, MA, 1990, was republished by Connexions in 2008. The coauthored book, *Coherence in Signal Processing and Machine Learning*, Springer, will appear in Fall 2022. His research interests include statistical signal processing and machine learning as they apply to space-time adaptive processing for radar, sonar, and communication, modal analysis for electric power system monitoring, spectrum analysis for nonstationary times series modeling, hyperspectral imaging, and image processing for group-invariant classification and registration. Prof. Scharf was the recipient of several awards for his contributions to statistical signal processing, including the Technical Achievement and Society Awards from the IEEE Signal Processing Society, Donald W. Tufts Award for Underwater Acoustic Signal Processing, Diamond Award from the University of Washington, and 2016 IEEE Jack S. Kilby Medal for Signal Processing. In 2021, he received the Education Award from the IEEE Signal Processing Society.

Open Access provided by 'Università del Salento' within the CRUI CARE Agreement