On the Statistical Invariance for Adaptive Radar Detection in Partially-homogeneous Disturbance plus Structured Interference

D. Ciuonzo, *Senior Member, IEEE,* A. De Maio, *Fellow, IEEE,* and D. Orlando, *Senior Member, IEEE*

***Abstract*—This paper deals with the problem of adaptive vector subspace signal detection in partially-homogeneous Gaussian disturbance and structured (unknown) deterministic interference within the framework of invariance theory. It is first proved that the Maximal Invariant Statistic (MIS) for the problem at hand is scalar-valued and coincides with the well-known Adaptive Normalized Matched Filter (ANMF) evaluated after data projection in the complementary subspace of the interfering signal. Secondly, the statistical characterization of the MIS under both hypotheses is derived. Then, it is shown the statistical equivalence of (Two-step) Generalized-Likelihood Ratio test, Rao and Wald tests, as well as the more recently considered Durbin and Gradient test, to the above statistic. Finally, simulation results are provided to confirm our findings and analyze the performance trend of the MIS with the relevant parameters.**

***Index Terms*—Adaptive Radar Detection, Constant False-Alarm Rate (CFAR), Invariance Theory, Maximal Invariants, Vector subspace Model, Partially-homogeneous interference, Coherent Interference.**

I. Introduction

A. Motivation and Related Works

THE PROBLEM of adaptive detection has been object of enormous interest in the last decades. Many excellent works appeared in the open literature, dealing with the design and performance analysis of suitable detectors under several specific settings (see for instance [1] and references therein).

Most of the proposed solutions assume a Homogeneous Environment (HE), wherein a set of secondary data (free of signal) components, but sharing the same spectral properties of the disturbance in the cells under test (primary data), is available [2], [3]. The HE often leads to elegant closed-form solutions which provide satisfactory performance in many cases [4], [5]. Unfortunately, the HE might not be met in some realistic situations: see, for example [6]–[8].

Manuscript received 5th February 2016; revised 27th July 2016; accepted 9th October 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Gustau Camps-Valls.

This work was partially sponsored by the Italian Ministry of Education, University and Research (MIUR) within the Project “Tecnologie abilitanti e sistemi innovativi a scansione elettronica del fascio in banda millimetrica e centimetrica per applicazioni radar a bordo di velivoli” (TELEMACO).
D. Ciuonzo and A. De Maio are with University of Naples "Federico II", DIETI, Via Claudio 21, 80125 Naples, Italy. E-mail: domenico.ciuonzo@ieee.org; ademaio@unina.it.
D. Orlando is with Università degli Studi “Niccolò Cusano”, Via Don Carlo Gnocchi 3, 00166 Roma, Italy. E-mail: danilo.orlando@unicusano.it.

Indeed, relevant scenarios are often non-homogeneous due to environmental factors and system considerations [6]–[8]. Among the frequently used assumptions to depict a non-homogeneous scenario there is the Partially Homogeneous Environment (PHE), i.e., both the test data and secondary data share the same disturbance covariance matrix structure up to an unknown scaling factor. Though keeping a relative tractability of the considered model, the present assumption provides increased robustness by allowing the (expected) power level of the disturbance to differ between the test data and the set of training data, which may appear in practice due to variations in terrain as well as the presence of guard cells [6], [8], [9]. Additionally, the PHE subsumes the HE as a special case. It is worth noticing that in more challenging scenarios, such as the naval context with high resolution Radars, more advanced disturbance models, such as spherically-invariant random processes, should be considered in order to account for clutter heterogeneity, as in [10].

The well-known Adaptive Normalized Matched Filter (ANMF) [11], [12], also known as Adaptive Coherence Estimator (ACE) [13] is the most common detector employed in PHE. In fact, it was proved to be the Generalized Likelihood Ratio Test (GLRT) for the aforementioned model in [14]. In the same paper, it was also shown that the same test could be obtained as the result of a two-step GLRT (2S-GLRT) design procedure, that is, devising the GLRT under known covariance of the disturbance and then making it adaptive via its substitution with the sample covariance matrix of secondary data. Later, the strong appeal of the ACE was confirmed in [15], where it was shown that other two theoretically-founded detection approaches, such as the Rao and Wald tests, both lead to the ANMF. It is worth remarking that the equivalence of GLRT and 2S-GLRT in PHE also holds in the case of a target return belonging to a multi-rank subspace, as proved in [16]. Such assumption is very important as it it represents an effective analytic tool to incorporate the partial knowledge of the target response in the detector design and, hence, to mitigate the performance degradation due to steering vector errors [17]–[19]. In fact, by constraining the target steering vector to lie in a suitable subspace of the observation space, it is possible to capture the energy of the potentially distorted wavefront of a mainlobe target. As a result, design architectures based on this model have the potential of declaring the presence of targets whose signature differs from the nominal one. Evidently, the subspace idea can be similarly adopted to model coherent interfering signals impinging on the radar antenna, whose directions of arrival have been estimated within some uncertainty. For instance, in [4], [20] the authors devise adaptive decision schemes to reveal extended targets when the interference comprises a random contribution (referring to clutter and thermal noise) and a structured unwanted component. In this respect, the present work similarly exploits the subspace approach for modeling both the signal and coherent interference signatures.

Other remarkable studies concerning detectors design in PHE appeared in the literature in the last years. A multi-dimensional analysis was conducted in [21], where an invariant approach was proposed to derive the MIS and a two-step GLRT with distributed targets and rank-one signals. Remarkably, the latter test was shown to have a Constant False-Alarm Rate (CFAR). On the other hand, Rao and Wald tests were developed in [22] to detect distributed (range-spread) targets with perfect knowledge of the target steering vector. More recently, Liu et al. treated adaptive detection of multidimensional (subspace) signals through the derivation of several well-founded architectures, such as the (2S-) GLRT, the Rao test and the Wald test (as well as other heuristic statistics) [23].

Last but not least, the PHE has been employed in conjunction with a-priori knowledge of the covariance disturbance, thus giving rise to knowledge-aided detectors. For example, an interesting work analyzed a "generalized" PHE, that is, a partially-specified a-priori distribution of the disturbance covariance matrix was assumed under a Bayesian approach [24]. In the aforementioned study, it was shown that a closed form of the GLRT exists and still coincides with the standard ACE. Similarly, a knowledge-aided version of the ACE via the Bayesian approach was also proposed in [25], where the a-priori distribution of the covariance disturbance was completely specified. The peculiarity of the PHE in the design of adaptive detectors was exploited along with the assumption of a per-symmetric covariance structure for devising a plain GLRT in [26]. Later, Rao and Wald tests were derived in the same context in [27]. Also, a per-symmetric ACE was obtained in [28], as the result of a 2S-GLRT technique in a PHE. Finally, adaptive detectors with range estimation capabilities (and possibly in the case of oversampling) were also developed in the PHE in the recent works [29] and [30], respectively.

We point out that all the aforementioned works (with the sole exception of [21]) differ from the literature [18], [19], [31], [32], where the adaptive (subspace) detection problem has been handled by resorting to the so-called *Principle of Invariance* [33], [34]. Indeed, the aforementioned principle, when exploited at the design stage, allows to focus on decision rules enjoying some desirable practical features. The preliminary step consists in identifying a suitable group of transformations which leaves the formal structure of the hypothesis testing problem unaltered. Of course, the group invariance requirement leads to a (lossy) data reduction. The least compression of the original data ensuring the desired invariance is represented by the *Maximal Invariant Statistic* (MIS), which organizes the original data into equivalence classes. Therefore, every invariant test can be expressed in terms of the MIS [34]. Accordingly, the parameter space is usually compressed after reduction by invariance and the dependence on the original set of parameters is mapped into a maximal invariant in the parameter space (the *induced maximal invariant*) [34]. When referring to radar adaptive detection, the mentioned principle represents an effective tool for obtaining a statistic which is invariant with respect to the set of nuisance parameters, therefore constituting the natural enabler for CFAR rules. With specific reference to the PHE and subspace signals (that is, in *absence of structured interference*), the most relevant works in this context are represented by [35], [36], where it was demonstrated that the MIS is the ACE. Also, in [36] it was claimed that the corresponding likelihood ratio is a monotone function of its argument. Such result implies that the ACE is also the Uniformly Most Powerful Invariant (UMPI) test.

Differently, the adaptive subspace detection in the joint presence of random and subspace structured interference in a PHE under the realm of invariance, to the best of our knowledge, has not been considered yet. In this context, we remark that a recent study, based on invariance theory, appeared in [37], dealing with structured interference but focusing however on the HE. In this respect, the aim of this paper is to build upon the results of [37] and fill this gap. To this end, this work is focused on adaptive detection of a subspace signal competing with two interference sources. The former is a completely random term, modeled as a Gaussian vector with unknown covariance matrix, and represents the returns from clutter and thermal noise. The latter is a subspace structured signal (with unknown location parameters) and accounts for the presence of (possible) multiple pulsed coherent jammers impinging on the radar antenna from some directions. Also, as anticipated, the non-homogeneity between the primary and the secondary data is modeled through the use of the PHE. Hence, the practical importance of the resulting group action is explained as a way to impose the CFAR property with respect to: (i) the clutter plus noise covariance matrix, (ii) the jammer location parameters and (iii) the difference in the expected power of the random disturbance among test and secondary data.

B. Summary of the Contributions

The main contributions of the present work can be summarized as follows:

- The considered problem is analytically formulated as a binary hypothesis test and the Principle of Invariance is exploited at the design stage to concentrate the attention on radar detectors enjoying some desirable practical features. We start from the more intuitive *canonical form* representation for the problem, developed in [37]. Such representation usually helps obtaining the maximal invariant statistics and gaining insights on the problem under investigation; The group of transformations which leaves the problem invariant is identified, thus allowing the search for a MIS.
- Given the aforementioned group of transformations, we are able to derive the explicit expression of the MIS, which *coincides* with the ANMF after projecting out the jammer interference. Such result encompasses that in [36], obtained in the absence of deterministic structured interference.

- For the model under investigation, we obtain the closed-form expressions for the (2S-) GLRT, the Rao test and the Wald test [38], as well as the Gradient test [39] and the Durbin test [40]. We then demonstrate that they are *all statistically equivalent to the MIS*, thus extending the result in [15], which does not consider the presence of deterministic interference.
- A theoretical performance analysis of the MIS is obtained, in terms of its distribution under both hypotheses. This also allows to obtain its false-alarm and detection probabilities in explicit form. The problem of synthesizing the Most Powerful Invariant (MPI) detector [34] is also addressed and a discussion on the existence of the UMPI test is provided.

C. Paper Organization and Manuscript Notation

The remainder of the paper is organized as follows: in Sec. II, we formulate the problem under investigation; in Sec. III, we obtain the MIS for the problem at hand and provide its statistical characterization; in Sec. IV, we derive all the aforementioned theoretically-founded detectors and verify their statistical equivalence to the MIS; finally, in Sec. V we draw some simulation results and in Sec. VI we provide some concluding remark. Proofs and derivations are confined to the Appendices.

Notation - Lower-case (resp. Upper-case) bold letters denote vectors (resp. matrices), with a_n (resp. $A_{n,m}$) representing the n-th (resp. the (n,m)-th) element of the vector $\boldsymbol{a}$ (resp. matrix $\boldsymbol{A}$); $\mathbb{R}^N$, $\mathbb{C}^N$, and $\mathbb{H}^{N\times N}$ are the sets of N-dimensional vectors of real numbers, of complex numbers, and of $N\times N$ Hermitian matrices, respectively, while $\mathbb{R}^+$ denotes the set of positive-valued real numbers; $\mathbb{E}\{\cdot\}$, $(\cdot)^T$, $(\cdot)^\dagger$, $\mathrm{Tr}\left[\cdot\right]$, $\|\cdot\|$, $\Re\{\cdot\}$ and $\Im\{\cdot\}$, denote expectation, transpose, Hermitian, matrix trace, Euclidean norm, real part, and imaginary part operators, respectively; $\boldsymbol{0}_{N\times M}$ (resp. $\boldsymbol{I}_N$) denotes the $N\times M$ null (resp. identity) matrix; $\boldsymbol{0}_N$ (resp. $\boldsymbol{1}_N$) denotes the null (resp. ones) column vector of length N; $\det(\boldsymbol{A})$ denotes the determinant of matrix $\boldsymbol{A}$; $\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}$ denotes the gradient of scalar valued function $f(\boldsymbol{x})$ w.r.t. vector $\boldsymbol{x}$ arranged in a column vector, while $\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}^T}$ its transpose (i.e., a row vector); the symbol "$\sim$" means "distributed as"; $\boldsymbol{x}\sim\mathcal{CN}_N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes a complex (proper) Gaussian-distributed vector $\boldsymbol{x}$ with mean vector $\boldsymbol{\mu}\in\mathbb{C}^{N\times 1}$ and covariance matrix $\boldsymbol{\Sigma}\in\mathbb{H}^{N\times N}$; $b\sim\mathcal{C}\beta_{M,N}$ (resp. $b\sim\mathcal{C}\beta_{M,N}(\delta)$) denotes a random variable distributed according to a complex central (resp. a complex noncentral) Beta distribution with (M,N) complex degrees of freedom (resp. with (M,N) complex degrees of freedom and noncentrality parameter δ); $f\sim\mathcal{CF}_{N,M}$ (resp. $f\sim\mathcal{CF}_{N,M}(\delta)$) denotes a random variable distributed according to a complex central (resp. a complex noncentral) F-distribution with (N,M) complex degrees of freedom (resp. with (N,M) complex degrees of freedom and noncentrality parameter δ); $\boldsymbol{P}_{\boldsymbol{A}}$ denotes the orthogonal projection of the full column rank matrix $\boldsymbol{A}$, that is, $\boldsymbol{P}_{\boldsymbol{A}}\triangleq[\boldsymbol{A}(\boldsymbol{A}^\dagger\boldsymbol{A})^{-1}\boldsymbol{A}^\dagger]$, while $\boldsymbol{P}_{\boldsymbol{A}}^\perp$ its complement, that is, $\boldsymbol{P}_{\boldsymbol{A}}^\perp\triangleq(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{A}})$.

II. Problem Formulation

In this section, we describe the detection problem at hand and recall its canonical form representation. Assume that a sensing system collects data from $N>1$ channels (spatial and/or temporal). The returns from the cell under test, after pre-processing, are properly sampled and organized to form a N-dimensional vector, denoted with $\boldsymbol{r}$. We want to test whether $\boldsymbol{r}$ contains useful target echoes or not, assuming the presence of an additional interfering signal. The target signature is modeled as a vector in a known subspace, spanned by $\boldsymbol{H}\in\mathbb{C}^{N\times r}$, where $r\geq 1$ and $\boldsymbol{H}$ is assumed a full column rank matrix. Therefore, the useful echoes are modeled in a non-redundant form as $\boldsymbol{H}\boldsymbol{p}$ with $\boldsymbol{p}\in\mathbb{C}^{r\times 1}$ (i.e., as a linear combination of the columns of $\boldsymbol{H}$). On the other hand, the interference component consists of *two contributions*. The former is representative of the combined (random) effect of clutter echoes and thermal noise, while the latter accounts for possible coherent sources impinging on the receive antenna from directions different to that where the radar system is steered[1]. More specifically, the structured interfering signal is assumed to belong to a known subspace, spanned by $\boldsymbol{J}\in\mathbb{C}^{N\times t}$, where $t\geq 1$ and $\boldsymbol{J}$ is assumed a full column rank matrix; hence, the interference can be expressed in a non-redundant form as $\boldsymbol{J}\boldsymbol{q}$, where $\boldsymbol{q}\in\mathbb{C}^{t\times 1}$ (i.e., as a linear combination of the columns of $\boldsymbol{J}$). The attractiveness of the model under investigation consists in effectively dealing with scenarios where the presence of one or multiple coherent pulsed jammers from standoff platforms attempt to protect a target located in the mainlobe of the radar antenna (with reference to Doppler processing). Finally, we assume that also the matrix $\begin{bmatrix}\boldsymbol{J} & \boldsymbol{H}\end{bmatrix}\in\mathbb{C}^{N\times J}$, where $J\triangleq(t+r)<N$, is *full column rank*, namely that the columns of $\boldsymbol{J}$ are linearly independent of those of $\boldsymbol{H}$. Such assumption implies non-overlapping uncertainty regions for the target and (coherent) interference steering vectors.

In summary, the decision problem at hand can be formulated in terms of the following binary hypothesis test

$$\begin{cases}\mathcal{H}_0: & \begin{cases}\boldsymbol{r}=\boldsymbol{J}\boldsymbol{q}+\boldsymbol{n}_0,\\ \boldsymbol{r}_k=\boldsymbol{n}_{0k}, \quad k=1,\ldots,K\end{cases}\\ \mathcal{H}_1: & \begin{cases}\boldsymbol{r}=\boldsymbol{H}\boldsymbol{p}+\boldsymbol{J}\boldsymbol{q}+\boldsymbol{n}_0\\ \boldsymbol{r}_k=\boldsymbol{n}_{0k}, \quad k=1,\ldots,K\end{cases}\end{cases} \tag{1}$$

where

- $\boldsymbol{H}\boldsymbol{p}\in\mathbb{C}^{N\times 1}$and $\boldsymbol{J}\boldsymbol{q}\in\mathbb{C}^{N\times 1}$ are the target and interferences signatures, respectively, with $\boldsymbol{p}$ and $\boldsymbol{q}$ deterministic and unknown vectors;
- $\boldsymbol{n}_0\sim\mathcal{CN}_N(\boldsymbol{0}_N,\boldsymbol{M}_0)$ and $\boldsymbol{n}_{0k}\sim\mathcal{CN}_N(\boldsymbol{0}_N,\gamma\boldsymbol{M}_0)$, $k=1,\ldots,K$, where the positive definite covariance

[1] Indeed, in order to face with deceptive signals, a certain knowledge of their constitutive parameters is required, which can be estimated by means of an Electronic Support Measure (ESM) system. However, the estimates depend on several factors which increase the uncertainty tied to the estimation procedure. For instance, the direction of arrival estimate of a jamming signal usually possesses a non-zero root mean square error which determines an uncertain angular sector $[\theta_1,\theta_2]$ (which also depends on the operating frequency). As a consequence, the spatial signature of the coherent interference is not perfectly known. Therefore, at the design stage, the partial knowledge of such parameter can be accommodated by modeling the structured interferer as a vector which belongs to a subspace covering the whole angular sector.

matrix $\boldsymbol{M}_0 \in \mathbb{H}^{N\times N}$ and the scaling factor $\gamma \in \mathbb{R}^+$ are both unknown deterministic quantities (such assumptions determine a PHE).

The model in Eq. (1) can be recast in the more convenient *canonical form*, as shown in [37]. Indeed, we first consider the QR decomposition [41] of the partitioned matrix $\begin{bmatrix}\boldsymbol{J} & \boldsymbol{H}\end{bmatrix} = \boldsymbol{Q}\boldsymbol{R}$, where $\boldsymbol{Q} \in \mathbb{C}^{N\times J}$ is a slice of a unitary matrix and $\boldsymbol{R} \in \mathbb{C}^{J\times J}$ is a non-singular upper triangular matrix. Then, we define a unitary matrix $\boldsymbol{U} \in \mathbb{C}^{N\times N}$ whose first J columns collectively equal $\boldsymbol{Q}$. Then the product $\boldsymbol{U}^\dagger\boldsymbol{Q}$ rotates the columns of $\boldsymbol{Q}$ onto the first J elementary vectors of the standard basis of $\mathbb{C}^{N\times 1}$, i.e., $\boldsymbol{U}^\dagger\boldsymbol{Q} = \begin{bmatrix}\boldsymbol{E}_t & \boldsymbol{E}_r\end{bmatrix}$ where we have denoted $\boldsymbol{E}_t \triangleq \begin{bmatrix}\boldsymbol{I}_t & \boldsymbol{0}_{t\times(N-t)}\end{bmatrix}^T$ and $\boldsymbol{E}_r \triangleq \begin{bmatrix}\boldsymbol{0}_{r\times t} & \boldsymbol{I}_r & \boldsymbol{0}_{r\times(N-J)}\end{bmatrix}^T$, respectively. Thus, without loss of generality, we can cast the problem in the equivalent form:

$$\begin{cases}\mathcal{H}_0: & \begin{cases}\boldsymbol{z} = \boldsymbol{E}_t\,\boldsymbol{\theta}_{10} + \boldsymbol{n}, \\ \boldsymbol{z}_k = \boldsymbol{n}_k, \quad k = 1,\ldots,K\end{cases} \\ \mathcal{H}_1: & \begin{cases}\boldsymbol{z} = \underbrace{\boldsymbol{E}_t\,\boldsymbol{\theta}_{11} + \boldsymbol{E}_r\,\boldsymbol{\theta}_2}_{\boldsymbol{A}\boldsymbol{\theta}_e} + \boldsymbol{n} \\ \boldsymbol{z}_k = \boldsymbol{n}_k, \quad k = 1,\ldots,K\end{cases}\end{cases}, \tag{2}$$

where we have used the notation $\boldsymbol{z} \triangleq \boldsymbol{U}^\dagger\boldsymbol{r} \in \mathbb{C}^{N\times 1}$ and $\boldsymbol{z}_k \triangleq \boldsymbol{U}^\dagger\boldsymbol{r}_k \in \mathbb{C}^{N\times 1}$, for the transformed primary and secondary data, respectively. Additionally, $\boldsymbol{\theta}_{1i} \in \mathbb{C}^{t\times 1}$ and $\boldsymbol{\theta}_2 \in \mathbb{C}^{r\times 1}$ have been used to denote the unknown deterministic vectors accounting for the interference and the useful signal, respectively. Then, aiming at a compact notation, we have defined $\boldsymbol{A} \triangleq \begin{bmatrix}\boldsymbol{E}_t & \boldsymbol{E}_r\end{bmatrix}$ and $\boldsymbol{\theta}_e \triangleq \begin{bmatrix}\boldsymbol{\theta}_{11}^T & \boldsymbol{\theta}_2^T\end{bmatrix}^T$ in Eq. (2). With reference to the disturbance, we have defined $\boldsymbol{n} \sim \mathcal{CN}_N(\boldsymbol{0}_N, \boldsymbol{M})$ and $\boldsymbol{n}_k \sim \mathcal{CN}_N(\boldsymbol{0}_N, \gamma\boldsymbol{M})$, respectively, with $\boldsymbol{M} \triangleq \boldsymbol{U}^\dagger\boldsymbol{M}_0\boldsymbol{U}$ representing the transformed covariance matrix. Finally, for the sake of notational convenience, we denote the matrix collecting primary and secondary data as $\boldsymbol{Z} \triangleq \begin{bmatrix}\boldsymbol{z} & \boldsymbol{z}_1 & \cdots & \boldsymbol{z}_K\end{bmatrix}$.

The probability density function (pdf) of the transformed data, when the hypothesis $\mathcal{H}_1$ is in force, is denoted by $f_1(\cdot\,;\,\cdot)$:

$$f_1(\boldsymbol{Z};\boldsymbol{\theta}_e,\boldsymbol{M}) = \pi^{-N(K+1)}\det(\boldsymbol{M})^{-(K+1)}\,\gamma^{-NK} \times \exp\left(-\mathrm{Tr}\left[\boldsymbol{M}^{-1}((\boldsymbol{z}-\boldsymbol{A}\boldsymbol{\theta}_e)(\boldsymbol{z}-\boldsymbol{A}\boldsymbol{\theta}_e)^\dagger + \gamma^{-1}\boldsymbol{S})\right]\right), \tag{3}$$

where we have defined $\boldsymbol{S} \triangleq \sum_{k=1}^K \boldsymbol{z}_k\boldsymbol{z}_k^\dagger$, while the corresponding pdf under $\mathcal{H}_0$, denoted in the following with $f_0(\cdot\,;\,\cdot)$, can be obtained replacing $\boldsymbol{A}\boldsymbol{\theta}_e$ with $\boldsymbol{E}_t\boldsymbol{\theta}_{10}$ in Eq. (3).

In the present manuscript we will consider decision rules which declare $\mathcal{H}_1$ (resp. $\mathcal{H}_0$) if $\Phi(\boldsymbol{Z}) \geq \eta$ (resp. $\Phi(\boldsymbol{Z}) < \eta$), where $\Phi(\cdot) : \mathbb{C}^{N\times(K+1)} \to \mathbb{R}$ indicates the generic form of a decision statistic based on $\boldsymbol{Z}$ and η denotes the threshold set to guarantee a predetermined probability of false alarm (P_{fa}). Finally, for notational convenience, we define the matrices $\boldsymbol{A}_1 \triangleq \boldsymbol{S}^{-1/2}\boldsymbol{A}$ and $\boldsymbol{A}_0 \triangleq \boldsymbol{S}^{-1/2}\boldsymbol{E}_t$, respectively, which will be thoroughly exploited in the remainder of the manuscript.

III. Maximal Invariant Statistic

In what follows, we will search for functions of data sharing invariance with respect to those parameters (namely, the nuisance parameters $\boldsymbol{M}$, $\boldsymbol{\theta}_{1i}$ and γ) which are irrelevant for the specific decision problem. To this end, we resort to the so-called "Principle of Invariance" [34], whose main idea consists in finding transformations that properly cluster data without altering: (i) the formal structure of the hypothesis testing problem given by $\mathcal{H}_0 : \|\boldsymbol{\theta}_2\| = 0$, $\mathcal{H}_1 : \|\boldsymbol{\theta}_2\| > 0$; ($ii$) the Gaussian assumption for the received data matrix under each hypothesis; (iii) the subspace containing the useful signal components. The next subsection is devoted to the definition of a suitable group which fulfills the above requirements.

A. Desired invariance properties

First, let us consider the sufficient statistic[2] $\{\boldsymbol{z}, \boldsymbol{S}\}$. Now, denote by $\mathcal{GL}(N)$ the linear group of $N \times N$ non-singular matrices and introduce the sets

$$\mathcal{G} \triangleq \left\{\boldsymbol{G} \triangleq \begin{bmatrix}\boldsymbol{G}_{11} & \boldsymbol{G}_{12} & \boldsymbol{G}_{13} \\ \boldsymbol{0}_{r\times t} & \boldsymbol{G}_{22} & \boldsymbol{G}_{23} \\ \boldsymbol{0}_{(N-J)\times t} & \boldsymbol{0}_{(N-J)\times r} & \boldsymbol{G}_{33}\end{bmatrix} \in \mathcal{GL}(N) \right. \\ \left. : \boldsymbol{G}_{11} \in \mathcal{GL}(t),\, \boldsymbol{G}_{22} \in \mathcal{GL}(r),\, \boldsymbol{G}_{33} \in \mathcal{GL}(N-J)\right\}; \tag{4}$$

and

$$\mathcal{F} \triangleq \left\{\boldsymbol{f} \triangleq \begin{bmatrix}\boldsymbol{f}_1 \\ \boldsymbol{0}_{N-t}\end{bmatrix} \in \mathbb{C}^{N\times 1} : \boldsymbol{f}_1 \in \mathbb{C}^{t\times 1}\right\}, \tag{5}$$

along with the composition operator "$\circ$", defined as:

$$(\boldsymbol{G}_a, \boldsymbol{f}_a, \gamma_a) \circ (\boldsymbol{G}_b, \boldsymbol{f}_b, \gamma_b) = (\boldsymbol{G}_b\boldsymbol{G}_a, \boldsymbol{G}_b\boldsymbol{f}_a + \boldsymbol{f}_b, \gamma_a\gamma_b). \tag{6}$$

The sets and the composition operator are here represented compactly as $\mathcal{L} \triangleq (\mathcal{G} \times \mathcal{F} \times \mathbb{R}^+, \circ)$. Then, it is not difficult to show that $\mathcal{L}$ constitutes a *group*[3]. Also, the aforementioned group leaves the hypothesis testing problem in Eq. (2) invariant under the action $\ell(\cdot,\cdot)$ defined by:

$$\ell(\boldsymbol{z}, \boldsymbol{S}) = \left(\boldsymbol{G}\boldsymbol{z} + \boldsymbol{f}, \gamma\,\boldsymbol{G}\,\boldsymbol{S}\,\boldsymbol{G}^\dagger\right) \quad \forall(\boldsymbol{G}, \boldsymbol{f}, \gamma) \in \mathcal{L}. \tag{7}$$

The proof of the aforementioned statement is omitted for the sake of brevity and can be obtained following similar steps as in [37]. Moreover, it is important to point out that invariance of the hypothesis testing problem w.r.t. $\mathcal{L}$ implies that the latter group preserves the family of distributions (that is, $\boldsymbol{G}\boldsymbol{z}+\boldsymbol{f}$ and $\boldsymbol{G}\,\boldsymbol{S}\,\boldsymbol{G}^\dagger$ are Gaussian- and Wishart-distributed, respectively), while not altering the peculiar structure of the hypothesis testing under investigation (including the subspaces $\boldsymbol{A}$ and $\boldsymbol{E}_t$). At the same time, $\mathcal{L}$ determines those transformations (through the action $\ell(\cdot)$ defined in (7)) which are relevant from a practical point of view, as they allow claiming the CFAR property (with respect to $\boldsymbol{M}$, $\boldsymbol{\theta}_{1i}$ and γ) as a consequence of the invariance.

[2]Indeed, Fisher-Neyman factorization theorem ensures that the optimal decision from $\{\boldsymbol{z}, \boldsymbol{S}\}$ is tantamount to deciding from raw data $\boldsymbol{Z}$ [42].

[3]Indeed $\mathcal{L}$ satisfies the following elementary axioms: (i) it is *closed* with respect to the operation "$\circ$", (ii) it satisfies the associative property and (iii) there exist both the identity and the inverse elements.

B. Derivation of the MIS

In Sec. III-A we have identified a group $\mathcal{L}$ which leaves the problem under investigation unaltered. It is thus reasonable finding decision rules that are invariant under $\mathcal{L}$. In order to accomplish this objective, we invoke the Principle of Invariance because it allows to construct statistics that organize data into distinguishable equivalence classes. Such functions are referred to as Maximal Invariant Statistics and, given the group of transformations, every invariant test can be written as a function of the maximal invariant [33]. Therefore, MIS represents the *least compression* of raw data $\{\boldsymbol{z}, \boldsymbol{S}\}$ providing invariance under $\mathcal{L}$.

Before presenting the explicit expression of the MIS, we give the following preliminary definitions based on the partitioning of $\boldsymbol{z}$ and $\boldsymbol{S}$:

$$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \boldsymbol{z}_3 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} & \boldsymbol{S}_{13} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} & \boldsymbol{S}_{23} \\ \boldsymbol{S}_{31} & \boldsymbol{S}_{32} & \boldsymbol{S}_{33} \end{bmatrix}, \tag{8}$$

where $\boldsymbol{z}_1 \in \mathbb{C}^{t\times 1}$, $\boldsymbol{z}_2 \in \mathbb{C}^{r\times 1}$, and $\boldsymbol{z}_3 \in \mathbb{C}^{(N-J)\times 1}$, respectively; $\boldsymbol{S}_{ij}$, $(i,j) \in \{1,2,3\} \times \{1,2,3\}$, is a sub-matrix whose size can be obtained replacing 1, 2, and 3 with t, r, and $(N-J)$, respectively. We are thus ready to present the proposition providing the expression of a MIS for the problem at hand.

Proposition 1. *A MIS with respect to $\mathcal{L}$ for the problem in Eq. (2) is given by:*

$$t(\boldsymbol{z}, \boldsymbol{S}) \triangleq \frac{t_a}{t_b} = \frac{\boldsymbol{z}_{2.3}^{\dagger}\, \boldsymbol{S}_{2.3}^{-1} \boldsymbol{z}_{2.3}}{\boldsymbol{z}_3^{\dagger} \boldsymbol{S}_{33}^{-1} \boldsymbol{z}_3}, \tag{9}$$

where $\boldsymbol{z}_{2.3} \triangleq (\boldsymbol{z}_2 - \boldsymbol{S}_{23}\boldsymbol{S}_{33}^{-1}\boldsymbol{z}_3)$ and $\boldsymbol{S}_{2.3} \triangleq (\boldsymbol{S}_{22} - \boldsymbol{S}_{23}\boldsymbol{S}_{33}^{-1}\boldsymbol{S}_{32})$, respectively.

Proof: The proof is given in Appendix A. ∎

Some important remarks are now in order.

- The MIS is simply given by the ratio of t_a and t_b, where the second component (t_b) represents an *ancillary part,* that is, its distribution does not depend on the hypothesis in force; such result generalizes that obtained [36] for the case of no subspace interference.
- Exploiting [34, Thm. 6.2.1], every invariant test may be written as a function of Eq. (9). Therefore, it naturally follows that every invariant test is CFAR.
- It is useful observing that the numerator and the denominator of the MIS can be also re-written as (the proof is omitted for the sake of brevity as similar steps can be found in [4]):

$$t_a = \boldsymbol{z}_{w1}^{\dagger}\,(\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp} - \boldsymbol{P}_{\boldsymbol{A}_1}^{\perp})\,\boldsymbol{z}_{w1}, \tag{10}$$

$$t_b = \boldsymbol{z}_{w1}^{\dagger}\,\boldsymbol{P}_{\boldsymbol{A}_1}^{\perp}\,\boldsymbol{z}_{w1}, \tag{11}$$

  where we have denoted $\boldsymbol{z}_{w1} \triangleq \boldsymbol{S}^{-1/2}\boldsymbol{z}$ and we recall that $\boldsymbol{A}_1 = \boldsymbol{S}^{-1/2}\boldsymbol{A}$ and $\boldsymbol{A}_0 = \boldsymbol{S}^{-1/2}\boldsymbol{E}_t$, respectively. We recall that these quantities relate to whitened data and subspaces based on the sole signal-free samples. Sec. IV will heavily rely on these results to establish statistical equivalence of some decision statistics and the MIS.
- Since all the maximal invariant statistics are related by one-to-one transformations, the following statistic:

$$t_k \triangleq \frac{t(\boldsymbol{z}, \boldsymbol{S})}{1 + t(\boldsymbol{z}, \boldsymbol{S})} = \frac{t_a}{t_a + t_b} = \frac{\boldsymbol{z}_{w1}^{\dagger}(\boldsymbol{P}_{\boldsymbol{A}_1} - \boldsymbol{P}_{\boldsymbol{A}_0})\,\boldsymbol{z}_{w1}}{\boldsymbol{z}_{w1}^{\dagger}\,\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\,\boldsymbol{z}_{w1}} \tag{12}$$

  is also a MIS.
- Finally, it can be shown that the computational complexity of (9) is dominated by the term $\mathcal{O}(KN^2)$, where $\mathcal{O}(\cdot)$ denotes the usual Landau notation. Indeed, such term arises from the computation of the unscaled covariance $\boldsymbol{S} \triangleq \sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k^{\dagger}$ required for evaluation of partioning elements defined in (8). Those are in fact needed in the inverse matrices employed at numerator and denominator.

C. Statistical distribution of the MIS

This subsection is devoted to the statistical characterization of the MIS under both hypotheses. The obtained results are based mainly on previous findings in [37] and thus only the main steps will be underlined in what follows. First, we notice that MIS $t(\boldsymbol{z}, \boldsymbol{S})$ in Eq. (9) can be rewritten as:

$$t(\boldsymbol{z}, \boldsymbol{S}) = \frac{\tau_k}{1-\beta}, \tag{13}$$

where we have denoted $\tau_k \triangleq \frac{t_a}{1+t_b}$ and $\beta \triangleq \frac{1}{1+t_b}$, respectively.

Secondly, it can be shown that

$$\beta \sim \mathcal{C}\beta_{K-(N-J)+1,\,N-J}, \tag{14}$$

as β is a function of the sole ancillary part of the MIS. Therefore, its distribution does not depend on the hypothesis being in force[4].

On the other hand, given β the random variable τ_k is distributed as

$$\tau_k|\beta, \mathcal{H}_1 \sim \mathcal{CF}_{r,\,K-(N-t)+1}(\delta) \tag{15}$$

when $\mathcal{H}_1$ is in force and

$$\tau_k|\beta, \mathcal{H}_0 \sim \mathcal{CF}_{r,\,K-(N-t)+1} \tag{16}$$

when $\mathcal{H}_0$ holds, where we have denoted $\delta^2 \triangleq \mathrm{SINR} \times \beta$, with $\mathrm{SINR} \triangleq (\boldsymbol{\theta}_2^{\dagger}\,\boldsymbol{M}_{2.3}^{-1}\,\boldsymbol{\theta}_2)$ being the Signal-to-Interference-plus-Noise-Ratio (SINR). We underline that $\boldsymbol{M}_{2.3}$ is defined as $\boldsymbol{M}_{2.3} \triangleq (\boldsymbol{M}_{22} - \boldsymbol{M}_{23}\boldsymbol{M}_{33}^{-1}\boldsymbol{M}_{32})$, where an analogous partitioning as that defined for $\boldsymbol{S}$ in Eq. (8) has been exploited.

Therefore, collecting the aforementioned results, the pdfs of $t(\boldsymbol{z}, \boldsymbol{S})$ under $\mathcal{H}_1$ and $\mathcal{H}_0$ can be expressed (by means of marginalization) as:

$$f_t(x|\mathcal{H}_1) = \int_0^1 (1-y)\, f_{\tau_k}(x\,(1-y)|\,\beta = y;\,\delta)\, f_\beta(y)\, dy,$$

$$f_t(x|\mathcal{H}_0) = \int_0^1 (1-y)\, f_{\tau_k}(x\,(1-y)|\,\beta = y; \delta = 0)\, f_\beta(y)\, dy, \tag{17}$$

[4]We recall that the present analysis is based on perfect matching conditions, that is, the nominal steering vectors of the target and the interference equal the actual ones. In the case of mismatch, the analysis could be generalized following [43].

respectively.

Finally, we conclude the section with a discussion on the induced maximal invariant in the parameter space [34]. The induced maximal invariant represents the reduced set of unknown parameters on which the hypothesis testing in the invariant domain depends. It can be readily shown that for our model this corresponds to SINR $\triangleq (\boldsymbol{\theta}_2^\dagger \boldsymbol{M}_{2.3}^{-1} \boldsymbol{\theta}_2)$. As a result, when the hypothesis $\mathcal{H}_0$ is in force, the SINR equals zero and thus $f_t(x|\mathcal{H}_0)$ does not depend on any unknown parameter. Therefore every function of the MIS satisfies the CFAR property.

IV. Detectors design

In this section we consider several decision statistics designed according to well-founded design criteria. Initially, we concentrate on the derivation of the well-known GLRT (including its two-step version) [14], [36], Rao and Wald tests [38]. Then, we devise the explicit form of less commonly used detection statistics, such as the Gradient (Terrell) test [39] and the Durbin (naive) test [40], which have been shown to be asymptotically distributed as the three aforementioned detectors (under very mild technical conditions).

A. Preliminary definitions

As a preliminary step towards the derivation of suitable detectors for the problem at hand, we give the following auxiliary definitions:

- $\boldsymbol{\theta}_r \triangleq \begin{bmatrix} \Re\{\boldsymbol{\theta}_2\}^T & \Im\{\boldsymbol{\theta}_2\}^T \end{bmatrix}^T \in \mathbb{R}^{2r\times 1}$ is the vector collecting the parameters of interest;
- $\boldsymbol{\theta}_s \triangleq \begin{bmatrix}\boldsymbol{\theta}_{s,a}^T & \boldsymbol{\theta}_{s,b}^T\end{bmatrix}^T \in \mathbb{R}^{(2t+N^2+1)\times 1}$ is the vector of nuisance parameters containing: (a) $\boldsymbol{\theta}_{s,a} \triangleq \begin{bmatrix} \Re\{\boldsymbol{\theta}_{1i}\}^T & \Im\{\boldsymbol{\theta}_{1i}\}^T \end{bmatrix}^T \in \mathbb{R}^{2t\times 1}$; (b) $\boldsymbol{\theta}_{s,b} \triangleq \begin{bmatrix}\gamma & \boldsymbol{\Xi}(\boldsymbol{M})^T\end{bmatrix} \in \mathbb{R}^{(N^2+1)\times 1}$ where $\boldsymbol{\Xi}(\cdot)$ is a real-valued column vector, mapping $\boldsymbol{M}$ to its (equivalent) minimal description in terms of N^2 independent variables;
- $\boldsymbol{\theta} \triangleq \begin{bmatrix} \boldsymbol{\theta}_r^T & \boldsymbol{\theta}_s^T \end{bmatrix}^T \in \mathbb{R}^{(2J+N^2+1)\times 1}$ is the overall unknown parameter vector;
- $\widehat{\boldsymbol{\theta}}_0 \triangleq \begin{bmatrix} \boldsymbol{\theta}_{r,0}^T & \widehat{\boldsymbol{\theta}}_{s,0}^T \end{bmatrix}^T$, with $\boldsymbol{\theta}_{r,0} = \boldsymbol{0}_{2r}$ (that is, the true value of $\boldsymbol{\theta}_r$ under $\mathcal{H}_0$) and $\widehat{\boldsymbol{\theta}}_{s,0}$ denoting the Maximum Likelihood (ML) estimate of $\boldsymbol{\theta}_s$ under $\mathcal{H}_0$;
- $\widehat{\boldsymbol{\theta}}_1 \triangleq \begin{bmatrix} \widehat{\boldsymbol{\theta}}_{r,1}^T & \widehat{\boldsymbol{\theta}}_{s,1}^T \end{bmatrix}^T$, with $\widehat{\boldsymbol{\theta}}_{r,1}$ and $\widehat{\boldsymbol{\theta}}_{s,1}$ denoting the ML estimates of $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_s$, respectively, under $\mathcal{H}_1$.

B. GLR

The generic form of the GLR is given by [38]:

$$\frac{\max_{\{\boldsymbol{\theta}_e,\boldsymbol{M},\gamma\}} f_1(\boldsymbol{Z};\,\boldsymbol{\theta}_e,\boldsymbol{M},\gamma)}{\max_{\{\boldsymbol{\theta}_{10},\boldsymbol{M},\gamma\}} f_0(\boldsymbol{Z};\,\boldsymbol{\theta}_{10},\boldsymbol{M},\gamma)}\,. \tag{18}$$

A detailed derivation is here skipped as it can be easily obtained by generalizing [15] to the case of additional deterministic interference. Therefore, we will only highlight the main steps in what follows. First, it can be shown that:

$$\widehat{\boldsymbol{\theta}}_e = (\boldsymbol{A}^\dagger \boldsymbol{S}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}\,, \tag{19}$$
$$\widehat{\boldsymbol{\theta}}_{10} = (\boldsymbol{E}_t^\dagger \boldsymbol{S}^{-1}\boldsymbol{E}_t)^{-1}\boldsymbol{E}_t^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}\,, \tag{20}$$

under $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively, while the corresponding ML estimates for the scale parameter γ are:

$$\widehat{\gamma}_1 = \frac{N}{K+1-N}\,\frac{1}{\boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\boldsymbol{A}_1}^\perp \boldsymbol{z}_{w1}} \tag{21}$$

under $\mathcal{H}_1$, and

$$\widehat{\gamma}_0 = \frac{N}{K+1-N}\,\frac{1}{\boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{z}_{w1}} \tag{22}$$

under $\mathcal{H}_0$. Finally, the ML estimates of the covariance matrix $\boldsymbol{M}$ (under $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively) are:

$$\widehat{\boldsymbol{M}}_1 = \frac{1}{K+1}\left[\frac{1}{\widehat{\gamma}_1}\boldsymbol{S} + (\boldsymbol{S}^{1/2}\boldsymbol{P}_{\boldsymbol{A}_1}^\perp \boldsymbol{z}_{w1}) \times (\boldsymbol{S}^{1/2}\boldsymbol{P}_{\boldsymbol{A}_1}^\perp \boldsymbol{z}_{w1})^\dagger\right], \tag{23}$$
$$\widehat{\boldsymbol{M}}_0 = \frac{1}{K+1}\left[\frac{1}{\widehat{\gamma}_0}\boldsymbol{S} + (\boldsymbol{S}^{1/2}\boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{z}_{w1}) \times (\boldsymbol{S}^{1/2}\boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{z}_{w1})^\dagger\right]. \tag{24}$$

Therefore, after substitution, the final form of the N-th root of GLR is given by:

$$t_{\mathrm{glr}} \triangleq \frac{\boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{z}_{w1}}{\boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\boldsymbol{A}_1}^\perp \boldsymbol{z}_{w1}} \tag{25}$$
$$= 1 + \frac{t_a}{t_b} = 1 + t(\boldsymbol{z},\boldsymbol{S})\,, \tag{26}$$

where Eq. (26) directly follows from considerations in Sec. III-B, thus proving *statistical equivalence* of GLR to the MIS.

Finally, before proceeding further, we state some useful properties of ML covariance estimates (later exploited in this paper) in the form of the following lemma.

Lemma 2. *The ML estimates of $\boldsymbol{M}$ under $\mathcal{H}_1$ and $\mathcal{H}_0$ satisfy the following equalities:*

$$\hat{\boldsymbol{M}}_1^{-1}\boldsymbol{A} = (K+1)\,\widehat{\gamma}_1\,\boldsymbol{S}^{-1}\boldsymbol{A}\,, \tag{27}$$
$$\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{E}_t = (K+1)\,\widehat{\gamma}_0\,\boldsymbol{S}^{-1}\boldsymbol{E}_t\,. \tag{28}$$

Proof: The proof is analogous to that presented in [4] and thus not reported here for the sake of brevity. ■

C. Two-step GLR (2S-GLR)

The 2S-GLRT design procedure consists in evaluating the GLR statistic under the assumption that $\boldsymbol{R} \triangleq \boldsymbol{M}\gamma$ is known and then plugging-in a reasonable estimate of $\boldsymbol{R}$ obtained from secondary data. Of course this implies the estimation of the inverse scale parameter $\nu \triangleq \gamma^{-1}$ from the primary data. As for the (one-step) GLR, we only underline the main steps and a detailed derivation is omitted for brevity. The GLR statistic for known $\boldsymbol{R}$ can be expressed in implicit form as [17]:

$$\frac{\max_{\{\boldsymbol{\theta}_e,\,\nu\}}\ f_1(\boldsymbol{z};\boldsymbol{\theta}_e,\nu,\boldsymbol{R})}{\max_{\{\boldsymbol{\theta}_{10},\,\nu\}}\ f_0(\boldsymbol{z};\boldsymbol{\theta}_{10},\nu,\boldsymbol{R})}\,. \tag{29}$$

The ML estimates of $\boldsymbol{\theta}_e$ (under $\mathcal{H}_1$) and $\boldsymbol{\theta}_{10}$ (under $\mathcal{H}_0$) are equal to

$$\widehat{\boldsymbol{\theta}}_e(\boldsymbol{R}) = (\boldsymbol{A}^\dagger \boldsymbol{R}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^\dagger \boldsymbol{R}^{-1}\boldsymbol{z}\,, \tag{30}$$
$$\widehat{\boldsymbol{\theta}}_{10}(\boldsymbol{R}) = (\boldsymbol{E}_t^\dagger \boldsymbol{R}^{-1}\boldsymbol{E}_t)^{-1}\boldsymbol{E}_t^\dagger \boldsymbol{R}^{-1}\boldsymbol{z}\,, \tag{31}$$

respectively, while the ML estimates for the (inverse) scale parameter ν are:

$$\hat{\nu}_1(\boldsymbol{R}) = \frac{1}{N}\left(\boldsymbol{z}^\dagger\, \boldsymbol{R}^{-1/2}\, \boldsymbol{P}^{\perp}_{\breve{\boldsymbol{A}}_1}\, \boldsymbol{R}^{-1/2}\, \boldsymbol{z}\right), \tag{32}$$

$$\hat{\nu}_0(\boldsymbol{R}) = \frac{1}{N}\left(\boldsymbol{z}^\dagger\, \boldsymbol{R}^{-1/2}\, \boldsymbol{P}^{\perp}_{\breve{\boldsymbol{A}}_0}\, \boldsymbol{R}^{-1/2}\, \boldsymbol{z}\right), \tag{33}$$

when $\mathcal{H}_1$ and $\mathcal{H}_0$ are in force, respectively. In Eqs. (32) and (33), we have also exploited the definitions $\breve{\boldsymbol{A}}_1 \triangleq (\boldsymbol{R}^{-1/2}\boldsymbol{A})$ and $\breve{\boldsymbol{A}}_0 \triangleq (\boldsymbol{R}^{-1/2}\boldsymbol{E}_t)$, respectively. Therefore, after substitution, the following statistic is obtained (as the N-th root of Eq. (29)):

$$\frac{\boldsymbol{z}^\dagger\, \boldsymbol{R}^{-1/2}\, \boldsymbol{P}^{\perp}_{\breve{\boldsymbol{A}}_0}\, \boldsymbol{R}^{-1/2}\, \boldsymbol{z}}{\boldsymbol{z}^\dagger\, \boldsymbol{R}^{-1/2}\, \boldsymbol{P}^{\perp}_{\breve{\boldsymbol{A}}_1}\, \boldsymbol{R}^{-1/2}\, \boldsymbol{z}}\,. \tag{34}$$

We recall that the expression in Eq. (34) depends on $\boldsymbol{R}$, the latter being unknown. Therefore, we now turn our attention on finding an estimate for the covariance $\boldsymbol{R}$. Clearly, in order to obtain a meaningful estimate to be plugged in both the numerator and denominator of Eq. (29), such estimate should be based only on signal-free (secondary) data.

It is not difficult to show that such estimate is given by the sample covariance based only on secondary data, that is, $\widehat{\boldsymbol{R}}_{\rm sd} = K^{-1}\boldsymbol{S} = K^{-1}\sum_{k=1}^{K}\boldsymbol{z}_k\boldsymbol{z}_k^\dagger$. Then, substitution of $\widehat{\boldsymbol{R}}_{\rm sd}$ into Eq. (34) leads to the final form of 2S-GLR:

$$t_{\rm 2s-glr} \triangleq \frac{\boldsymbol{z}_{w1}^\dagger\, \boldsymbol{P}^{\perp}_{\boldsymbol{A}_0}\, \boldsymbol{z}_{w1}}{\boldsymbol{z}_{w1}^\dagger\, \boldsymbol{P}^{\perp}_{\boldsymbol{A}_1}\, \boldsymbol{z}_{w1}}\,. \tag{35}$$

From direct comparison of Eq. (35) with Eq. (25) the following important proposition can be immediately stated.

Proposition 3. *The 2S-GLR statistic $t_{\rm 2s-glr}$ in Eq. (35) coincides with the one-step GLR statistic in Eq. (25). Thus it is statistically equivalent to MIS obtained in Eq. (9).*

D. Rao statistic

The generic form for the Rao statistic is given by [38]:

$$\left.\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_r^T}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} [\boldsymbol{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} \left.\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_r}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} \tag{36}$$

where

$$\boldsymbol{I}(\boldsymbol{\theta}) \triangleq \mathbb{E}\left\{\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}\right\}, \tag{37}$$

denotes the Fisher Information Matrix (FIM) and $[\boldsymbol{I}^{-1}(\boldsymbol{\theta})]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}$ indicates the sub-matrix obtained by selecting from the FIM inverse only the elements corresponding to the vector of signal parameter $\boldsymbol{\theta}_r$.

In order to obtain the explicit form of Rao statistic, we first observe that:

$$\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2^*} = \underbrace{\boldsymbol{E}_r^\dagger\, \boldsymbol{M}^{-1}(\boldsymbol{z} - \boldsymbol{E}_t\boldsymbol{\theta}_{11} - \boldsymbol{E}_r\boldsymbol{\theta}_2)}_{\triangleq\, \boldsymbol{g}_r}, \tag{38}$$

where $\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2^*}$ denotes the complex gradient of a real function [44], whose explicit expression is

$$\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2} = \frac{1}{2}\left[\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\, \Re\{\boldsymbol{\theta}_2\}} - j\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\, \Im\{\boldsymbol{\theta}_2\}}\right] \tag{39}$$

and also satisfies

$$\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2^*} = \left(\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2}\right)^*, \tag{40}$$

from which it is readily inferred that:

$$\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\, \Re\{\boldsymbol{\theta}_2\}} = 2\,\Re\{\boldsymbol{g}_r\}, \quad \frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\, \Im\{\boldsymbol{\theta}_2\}} = 2\,\Im\{\boldsymbol{g}_r\}\,. \tag{41}$$

Therefore, collecting the above results we get:

$$\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_r} = \begin{bmatrix} 2\,\Re\{\boldsymbol{g}_r\}^T & 2\,\Im\{\boldsymbol{g}_r\}^T \end{bmatrix}^T. \tag{42}$$

We now turn our attention to the evaluation of $(\boldsymbol{\theta}_r,\boldsymbol{\theta}_r)$ block of FIM inverse. Such evaluation can be greatly simplified by exploiting the following equality:

$$[\boldsymbol{I}^{-1}(\boldsymbol{\theta})]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} = [\boldsymbol{I}_a^{-1}(\boldsymbol{\theta})]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}, \tag{43}$$

where

$$\boldsymbol{I}_a(\boldsymbol{\theta}) \triangleq \mathbb{E}\left\{\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\begin{bmatrix}\boldsymbol{\theta}_r^T & \boldsymbol{\theta}_{s,a}^T\end{bmatrix}^T}\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\begin{bmatrix}\boldsymbol{\theta}_r^T & \boldsymbol{\theta}_{s,a}^T\end{bmatrix}}\right\}, \tag{44}$$

as the block terms corresponding to $(\boldsymbol{\theta}_r,\gamma)$, $(\boldsymbol{\theta}_{s,a},\gamma)$, $(\boldsymbol{\theta}_r,\boldsymbol{\Xi})$ and $(\boldsymbol{\theta}_{s,a},\boldsymbol{\Xi})$ in the FIM $\boldsymbol{I}(\boldsymbol{\theta})$ are *all null*. Evaluation of $\boldsymbol{I}_a(\boldsymbol{\theta})$ in Eq. (44) is obtained starting from the following observation:

$$\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\begin{bmatrix}\boldsymbol{\theta}_r^T & \boldsymbol{\theta}_{s,a}^T\end{bmatrix}^T} = \boldsymbol{P}\,\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\begin{bmatrix}\Re\{\boldsymbol{\theta}_e\}^T & \Im\{\boldsymbol{\theta}_e\}^T\end{bmatrix}^T}, \tag{45}$$

where $\boldsymbol{P}$ is a suitably defined permutation matrix[5]:

$$\boldsymbol{P} \triangleq \begin{bmatrix} \boldsymbol{0}_{r\times t} & \boldsymbol{I}_r & \boldsymbol{0}_{r\times t} & \boldsymbol{0}_{r\times r} \\ \boldsymbol{0}_{r\times t} & \boldsymbol{0}_{r\times r} & \boldsymbol{0}_{r\times t} & \boldsymbol{I}_r \\ \boldsymbol{I}_t & \boldsymbol{0}_{t\times r} & \boldsymbol{0}_{t\times t} & \boldsymbol{0}_{t\times r} \\ \boldsymbol{0}_{t\times t} & \boldsymbol{0}_{t\times r} & \boldsymbol{I}_t & \boldsymbol{0}_{t\times r} \end{bmatrix}. \tag{46}$$

Then exploiting Eq. (45), we obtain:

$$\boldsymbol{I}_a(\boldsymbol{\theta}) = \tag{47}$$

$$\boldsymbol{P}\,\mathbb{E}\left\{\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\begin{bmatrix}\Re\{\boldsymbol{\theta}_e\}^T & \Im\{\boldsymbol{\theta}_e\}^T\end{bmatrix}^T}\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\begin{bmatrix}\Re\{\boldsymbol{\theta}_e\}^T & \Im\{\boldsymbol{\theta}_e\}^T\end{bmatrix}}\right\}\boldsymbol{P}^T$$

Using an analogous rationale as in Eq. (42), it can be proved that:

$$\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\begin{bmatrix}\Re\{\boldsymbol{\theta}_e\}^T & \Im\{\boldsymbol{\theta}_e\}^T\end{bmatrix}} = \begin{bmatrix} 2\,\Re\{\boldsymbol{g}_e\}^T & 2\,\Im\{\boldsymbol{g}_e\}^T \end{bmatrix}, \tag{48}$$

where we have similarly defined $\boldsymbol{g}_e \triangleq \boldsymbol{A}^\dagger\,\boldsymbol{M}^{-1}(\boldsymbol{z} - \boldsymbol{E}_t\,\boldsymbol{\theta}_{11} - \boldsymbol{E}_r\,\boldsymbol{\theta}_2)$. The explicit expression of the inner matrix in Eq. (47) is obtained exploiting the result in (48) and is given by:

$$\begin{bmatrix} 2\,\Re\{\boldsymbol{K}\} & -2\,\Im\{\boldsymbol{K}\} \\ 2\,\Im\{\boldsymbol{K}\} & 2\,\Re\{\boldsymbol{K}\} \end{bmatrix}, \tag{49}$$

where $\boldsymbol{K} \triangleq (\boldsymbol{A}^\dagger\boldsymbol{M}^{-1}\boldsymbol{A})$. Tedious mathematical steps reveal that:

$$\boldsymbol{I}_a^{-1}(\boldsymbol{\theta}) = \boldsymbol{P}\begin{bmatrix} \frac{1}{2}\Re\{\boldsymbol{K}^{-1}\} & -\frac{1}{2}\Im\{\boldsymbol{K}^{-1}\} \\ \frac{1}{2}\Im\{\boldsymbol{K}^{-1}\} & \frac{1}{2}\Re\{\boldsymbol{K}^{-1}\} \end{bmatrix}\boldsymbol{P}^T, \tag{50}$$

[5]Recall that every permutation matrix is a special orthogonal matrix, that is $\boldsymbol{P}^{-1} = \boldsymbol{P}^T$.

since the block structure in Eq. (49) is closed under inversion[6]. Then, extracting the $(\boldsymbol{\theta}_r, \boldsymbol{\theta}_r)$ block provides:

$$[\boldsymbol{I}_a^{-1}(\boldsymbol{\theta})]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} = \begin{bmatrix} \frac{1}{2}\Re\{\boldsymbol{B}\} & -\frac{1}{2}\Im\{\boldsymbol{B}\} \\ \frac{1}{2}\Im\{\boldsymbol{B}\} & \frac{1}{2}\Re\{\boldsymbol{B}\} \end{bmatrix}, \tag{51}$$

where we have defined:

$$\boldsymbol{B} \triangleq \Big[\boldsymbol{E}_r^\dagger \boldsymbol{M}^{-1}\boldsymbol{E}_r - (\boldsymbol{E}_r^\dagger \boldsymbol{M}^{-1}\boldsymbol{E}_t)(\boldsymbol{E}_t^\dagger \boldsymbol{M}^{-1}\boldsymbol{E}_t)^{-1}\boldsymbol{E}_t^\dagger \boldsymbol{M}^{-1}\boldsymbol{E}_r\Big]^{-1}. \tag{52}$$

Given the specific structure of Eqs. (51) and (42), it can be recognized the proportionality (by a factor $1/2$) of Eq. (36) to an equivalent Hermitian quadratic form:

$$t_{\text{rao}} \triangleq \left(\boldsymbol{g}_r^\dagger \boldsymbol{B}\,\boldsymbol{g}_r\right)\big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0}. \tag{53}$$

The substitution $\boldsymbol{\theta} \to \widehat{\boldsymbol{\theta}}_0$ only requires replacing the true $\boldsymbol{M}$ with $\hat{\boldsymbol{M}}_0$ (given by Eq. (23)) in $\boldsymbol{B}$. Differently, $\boldsymbol{\theta} \to \widehat{\boldsymbol{\theta}}_0$ on $\boldsymbol{g}_r$ implies substitution of $\boldsymbol{\theta}_{11}$, $\boldsymbol{\theta}_2$, and $\boldsymbol{M}$ with $\widehat{\boldsymbol{\theta}}_{10}$ (given by Eq. (20)), $\boldsymbol{0}_r$ and $\hat{\boldsymbol{M}}_0$ (given by Eq. (23)), respectively. Simple manipulations[7] then give:

$$t_{\text{rao}} = \boldsymbol{z}_{w0}^\dagger \left(\boldsymbol{P}_{\bar{\boldsymbol{A}}_1} - \boldsymbol{P}_{\bar{\boldsymbol{A}}_0}\right)\boldsymbol{z}_{w0}, \tag{54}$$

where we have denoted $\bar{\boldsymbol{A}}_0 \triangleq \hat{\boldsymbol{M}}_0^{-1/2}\boldsymbol{E}_t$, $\bar{\boldsymbol{A}}_1 \triangleq \hat{\boldsymbol{M}}_0^{-1/2}\boldsymbol{A}$ and $\boldsymbol{z}_{w0} \triangleq \hat{\boldsymbol{M}}_0^{-1/2}\boldsymbol{z}$, respectively. We recall that these quantities relate to whitened data and subspaces based on ML covariance estimate under $\mathcal{H}_0$.

Proposition 4. *The Rao statistic t_{rao} in Eq. (54) can be expressed in the alternative form:*

$$t_{\text{rao}} = \frac{N\,t_k}{\frac{K+1}{K+1-N} + \frac{N}{K+1-N}\,t_k}, \tag{55}$$

which thus proves that Rao statistic is statistically equivalent to MIS obtained in Eq. (12), since Eq. (55) is a monotone function of the MIS t_k.

Proof: The proof is given in Appendix B. ∎

E. Wald statistic

The generic form for the Wald statistic is given by [38]:

$$(\hat{\boldsymbol{\theta}}_{r,1} - \boldsymbol{\theta}_{r,0})^T \{[\boldsymbol{I}^{-1}(\widehat{\boldsymbol{\theta}}_1)]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}\}^{-1}(\hat{\boldsymbol{\theta}}_{r,1} - \boldsymbol{\theta}_{r,0}). \tag{56}$$

First, we notice that $\boldsymbol{\theta}_{r,0} = \boldsymbol{0}_{2r}$. Then, we observe that $\hat{\boldsymbol{\theta}}_{r,1}$ can be obtained via the following steps. We recall that $\hat{\boldsymbol{\theta}}_{r,1} = \left[\Re\{\hat{\boldsymbol{\theta}}_2\}^T \quad \Im\{\hat{\boldsymbol{\theta}}_2\}^T\right]^T$, where $\hat{\boldsymbol{\theta}}_2 \in \mathbb{C}^{r\times 1}$ is simply a sub-vector of $\hat{\boldsymbol{\theta}}_e \in \mathbb{C}^{J\times 1}$ (obtained by taking its last r elements), whose closed-form expression is given in Eq. (19), thus leading to:

$$\hat{\boldsymbol{\theta}}_2 = \boldsymbol{\Lambda}_{21}\left(\boldsymbol{E}_t^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}\right) + \boldsymbol{\Lambda}_{22}\left(\boldsymbol{E}_r^\dagger \boldsymbol{S}^{-1}\boldsymbol{z}\right), \tag{57}$$

where we have partitioned the matrix $\boldsymbol{\Lambda} \triangleq (\boldsymbol{A}^\dagger \boldsymbol{S}^{-1}\boldsymbol{A})^{-1}$ as

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12} \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} \end{bmatrix}, \tag{58}$$

where $\boldsymbol{\Lambda}_{ij}$, $(i,j) \in \{1,2\}\times\{1,2\}$, denotes a sub-matrix whose size can be obtained replacing 1 and 2 with t and r, respectively. Since it holds $\boldsymbol{\Lambda}_{21} = -\boldsymbol{\Lambda}_{22}\left(\boldsymbol{E}_r^\dagger \boldsymbol{S}^{-1}\boldsymbol{E}_t\right)\left(\boldsymbol{E}_t^\dagger \boldsymbol{S}^{-1}\boldsymbol{E}_t\right)^{-1}$ and by virtue of Lemma 2, we obtain:

$$\hat{\boldsymbol{\theta}}_2 = (K+1)\,\hat{\gamma}_1\,\boldsymbol{\Gamma}_{22}\,\boldsymbol{E}_r^\dagger \boldsymbol{S}^{-1/2}\boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{z}_{w1}, \tag{59}$$

where we have partitioned the matrix $\boldsymbol{\Gamma} \triangleq (\boldsymbol{A}^\dagger \hat{\boldsymbol{M}}_1^{-1}\boldsymbol{A})^{-1}$ analogously as $\boldsymbol{\Lambda}$ in Eq. (58). We now focus on evaluation of the matrix in Eq. (56), which is achieved by an analogous procedure as the case of FIM block needed for Rao test evaluation in Sec. IV-D. Indeed, starting from Eq. (51), it can be proved that:

$$\{[\boldsymbol{I}^{-1}(\widehat{\boldsymbol{\theta}}_1)]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}\}^{-1} = \begin{bmatrix} \frac{1}{2}\Re\{\boldsymbol{\Gamma}_{22}^{-1}\} & -\frac{1}{2}\Im\{\boldsymbol{\Gamma}_{22}^{-1}\} \\ \frac{1}{2}\Im\{\boldsymbol{\Gamma}_{22}^{-1}\} & \frac{1}{2}\Re\{\boldsymbol{\Gamma}_{22}^{-1}\} \end{bmatrix}, \tag{60}$$

as the block structure is closed under inversion. Finally, exploiting the well-known equivalence between a real-valued Hermitian quadratic form and its real symmetric quadratic counterpart, we demonstrate proportionality (by a factor 2) of Eq. (56) with:

$$\hat{\boldsymbol{\theta}}_2^\dagger \boldsymbol{\Gamma}_{22}^{-1}\hat{\boldsymbol{\theta}}_2 = \boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\boldsymbol{A}_0}^\perp (\boldsymbol{S}^{-1/2}\boldsymbol{E}_r \boldsymbol{\Gamma}_{22}\boldsymbol{E}_r^\dagger \boldsymbol{S}^{-1/2}) \times \boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{z}_{w1}(K+1)^2\hat{\gamma}_1^2 \triangleq t_{\text{wald}}. \tag{61}$$

Tedious manipulations then give the alternative (and more) compact form:

$$t_{\text{wald}} = (K+1)\,\hat{\gamma}_1\,\boldsymbol{z}_{w1}^\dagger\left(\boldsymbol{P}_{\boldsymbol{A}_1} - \boldsymbol{P}_{\boldsymbol{A}_0}\right)\boldsymbol{z}_{w1}. \tag{62}$$

The technical details are not reported here, since a similar proof can be found in [4]. We only underline that this is achieved by proving the matrix equivalence:

$$\boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{S}^{-1/2}\boldsymbol{E}_r\boldsymbol{\Gamma}_{22}\boldsymbol{E}_r^\dagger \boldsymbol{S}^{-1/2}\boldsymbol{P}_{\boldsymbol{A}_0}^\perp = \frac{1}{(K+1)\hat{\gamma}_1}\left(\boldsymbol{P}_{\boldsymbol{A}_1} - \boldsymbol{P}_{\boldsymbol{A}_0}\right). \tag{63}$$

Based on the compact form of Wald statistic in Eq. (62), we can now claim the following important result.

Proposition 5. *The Wald statistic t_{wald} in Eq. (62) can be expressed in the alternative form:*

$$t_{\text{wald}} = \frac{(K+1)N}{K+1-N}\,t(\boldsymbol{z},\boldsymbol{S}), \tag{64}$$

which thus proves that Wald test is statistically equivalent to MIS obtained in Eq. (9).

Proof: The proof is obtained substituting Eq. (21) into (62) and from direct comparison with Eq. (9) (with the help of the alternative expressions provided in Eqs. (10) and (11), respectively). ∎

[6] Indeed $\boldsymbol{K}\boldsymbol{K}^{-1} = \boldsymbol{I}_J$ implies both the equivalences $\Re\{\boldsymbol{K}\}\Re\{\boldsymbol{K}^{-1}\} - \Im\{\boldsymbol{K}\}\Im\{\boldsymbol{K}^{-1}\} = \boldsymbol{I}_J$ and $\Re\{\boldsymbol{K}\}\Im\{\boldsymbol{K}^{-1}\} + \Im\{\boldsymbol{K}\}\Re\{\boldsymbol{K}^{-1}\} = \boldsymbol{0}_{J\times J}$. A similar pair of equalities can be also drawn from $\boldsymbol{K}^{-1}\boldsymbol{K} = \boldsymbol{I}_J$. Therefore, the aforementioned set of conditions ensures that the inner matrix in Eq. (50) corresponds to the inverse of Eq. (49).

[7] The steps are very tedious and a similar proof can be found in [4].

F. Gradient (Terrell) statistic

The Gradient (Terrell) test requires the evaluation of the following statistic [39], [45]:

$$\left.\frac{\partial \ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_r^T}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_0} (\hat{\boldsymbol{\theta}}_{r,1}-\boldsymbol{\theta}_{r,0})\,. \tag{65}$$

The appeal of Eq. (65) arises from the fact that it does not require neither to invert the FIM nor to evaluate a compressed likelihood function under both hypotheses (as opposed to GLR, Wald, and Rao statistics). As a consequence, this structural simplicity can make the Gradient statistic computation easier. Moreover, under some mild technical conditions, such test is asymptotically equivalent to the GLR, Rao and Wald statistics [39].

In order to obtain the explicit form of Terrell statistic, we first recall that $\boldsymbol{\theta}_{r,0}=\boldsymbol{0}_{2r}$. Also, we observe that $\hat{\boldsymbol{\theta}}_{r,1}$ can be similarly obtained as in the case of Wald test, that is, $\hat{\boldsymbol{\theta}}_{r,1}=\left[\Re\{\hat{\boldsymbol{\theta}}_2\}^T \quad \Im\{\hat{\boldsymbol{\theta}}_2\}^T\right]^T$, where closed form of $\hat{\boldsymbol{\theta}}_2$ has been reported in Eq. (59). Similarly, the log-likelihood gradient evaluated $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_0$ in Eq. (65) is obtained similarly as in the case of Rao test, and it is equal to Eq. (42), where $\boldsymbol{g}_{r,0}\triangleq \boldsymbol{E}_r^\dagger \hat{\boldsymbol{M}}_0^{-1}(\boldsymbol{z}-\boldsymbol{E}_t\hat{\boldsymbol{\theta}}_{10})$. Therefore, combining the aforementioned results, Eq. (65) is proved to be equal to

$$\begin{aligned} t_{\mathrm{grad}} &\triangleq \Re\{\boldsymbol{g}_{r,0}^\dagger\,\hat{\boldsymbol{\theta}}_{r,1}\}=\Re\{\boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\boldsymbol{A}_0}^\perp \boldsymbol{S}^{-1/2}\boldsymbol{E}_r \\ &\times\hat{\gamma}_1(K+1)\boldsymbol{\Gamma}_{22}\,\boldsymbol{E}_r^\dagger\hat{\boldsymbol{M}}_0^{-1}(\boldsymbol{z}-\boldsymbol{E}_t\hat{\boldsymbol{\theta}}_{10})\}\,. \end{aligned} \tag{66}$$

After some manipulations, we obtain the following alternative (compact) form of Gradient statistic:

$$t_{\mathrm{grad}}=\Re\{\boldsymbol{z}_{w1}^\dagger\,(\boldsymbol{P}_{\boldsymbol{A}_1}-\boldsymbol{P}_{\boldsymbol{A}_0})\,\boldsymbol{S}^{1/2}\,\hat{\boldsymbol{M}}_0^{-1/2}\boldsymbol{z}_{w0}\}\,. \tag{67}$$

As done previously, the detailed steps are not reported here, since a similar proof can be found in [4]. Although seemingly different from Eq. (12), the following statistical equivalence result is proved hereinafter.

Proposition 6. *The Gradient (Terrell) statistic t_{grad} in Eq. (67) can be expressed in the alternative form:*

$$t_{\mathrm{grad}}=N\,t_k\,, \tag{68}$$

which thus proves that Gradient (Terrell) test is statistically equivalent to MIS obtained in Eq. (12).

Proof: The proof is given in Appendix C. ■

G. Durbin (Naive) statistic

The Durbin test (also referred to as "Naive test") consists in the evaluation of the following decision statistic [40]:

$$\begin{aligned}(\hat{\boldsymbol{\theta}}_{r,01}-\boldsymbol{\theta}_{r,0})^T\Big\{\Big[\boldsymbol{I}\left(\hat{\boldsymbol{\theta}}_0\right)\Big]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}\Big[\boldsymbol{I}^{-1}\left(\hat{\boldsymbol{\theta}}_0\right)\Big]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}\times \\ \Big[\boldsymbol{I}\left(\hat{\boldsymbol{\theta}}_0\right)\Big]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}\Big\}\,(\hat{\boldsymbol{\theta}}_{r,01}-\boldsymbol{\theta}_{r,0})\,,\end{aligned} \tag{69}$$

where the estimate $\hat{\boldsymbol{\theta}}_{r,01}$ is defined as:

$$\hat{\boldsymbol{\theta}}_{r,01}\triangleq\arg\max_{\boldsymbol{\theta}_r}\,f_1(\boldsymbol{Z};\boldsymbol{\theta}_r,\hat{\boldsymbol{\theta}}_{s,0})\,. \tag{70}$$

In general, the Durbin statistic is asymptotically equivalent to GLR, Rao, and Wald statistics, as shown in [40]. However, for the considered problem, a stronger result holds, as stated by the following theorem.

Proposition 7. *The Durbin statistic for the hypothesis testing model considered in Eq. (2) is coincident with the Rao statistic. Therefore, the test is statistically equivalent to the MIS.*

Proof: The proof is given in Appendix D. ■

H. Table summary and related discussion

In Tab. I, we gather all the results obtained in the previous sub-sections. First, the most important claim is that all the considered tests are statistically equivalent to the MIS, which is recognized as an ANMF (ACE) in the complementary subspace of the deterministic interference. Hereinafter we discuss such result in relationship to the existing literature.

We recall that statistical equivalence of the GLRT to the MIS in the case of no structured interference was first proved in [14], [36]. The same equivalence result was later shown to hold for Rao and Wald tests in [15]. The findings in the present manuscript confirm that statistical equivalence of the three well-known statistics to the MIS *still holds* in presence of an additional structured interference.

Additionally, in Sec. IV-G we have proved that Durbin test coincides with the Rao test in the finite sample case. Such result not only extends the findings obtained for quite general signal models in a homogeneous scenario [4], [46], but also further confirms the effectiveness of the MIS as a theoretically-founded decision statistic, that is, also arising from the Durbin test construction.

Similar conclusions apply to the case of Gradient (Terrell) test, whose statistical equivalence to the MIS (and, as an immediate consequence, to all the other tests) represents a result whose peculiarity only pertains to the PHE, since in the homogeneous case only statistical equivalence to (one-step) GLRT holds [4].

V. Simulation results

In the previous section we have shown that all the considered statistics are statistically equivalent to the MIS. For this reason, in the following, we concentrate on its performance assessment. To this end, we first provide the false alarm and detection probabilities of the MIS in a (convenient) one-dimensional integral form, as reported in Eqs. (73) and (74), respectively (at the top of next page). The detailed derivation is not reported here, as it can be obtained by generalizing the results contained in [43]. Indeed, the two probabilities are calculated by noticing that:

$$P_{fa}^{\mathrm{mis}}=\mathbb{E}_\beta\{\Pr(\tau_k\geq\eta\,(1-\beta)|\beta,\mathcal{H}_0)\}\,, \tag{71}$$

$$P_d^{\mathrm{mis}}=1-\mathbb{E}_\beta\{\Pr(\tau_k\leq\eta\,(1-\beta)|\beta,\mathcal{H}_1)\}\,. \tag{72}$$

The probability term within the expectation (depending on β) in the above equations corresponds to the (complementary) cumulative distribution function of a complex F-distribution, whose explicit expression is available [43]. Differently, the outer expectation cannot be evaluated in closed form and a numerical integration is usually performed.

Table I
DETECTORS COMPARISON AND THEIR FUNCTIONAL DEPENDENCE OF THE MIS (VIZ. CFARNESS).

Detector	*Standard Expression*	*MIS function*
GLR/2S-GLR	$\frac{\boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\bar{\boldsymbol{A}}_0}^\perp \boldsymbol{z}_{w1}}{\boldsymbol{z}_{w1}^\dagger \boldsymbol{P}_{\bar{\boldsymbol{A}}_1}^\perp \boldsymbol{z}_{w1}}$	$1+t(\boldsymbol{z},\boldsymbol{S})$
Rao/Durbin	$\boldsymbol{z}_{w0}^\dagger (\boldsymbol{P}_{\bar{\boldsymbol{A}}_1} - \boldsymbol{P}_{\bar{\boldsymbol{A}}_0})\, \boldsymbol{z}_{w0}$	$N\, t_k / \left(\frac{(K+1)}{K+1-N} + \frac{N}{K+1-N}\, t_k \right)$
Wald	$\boldsymbol{z}_{w1}^\dagger (\boldsymbol{P}_{\boldsymbol{A}_1} - \boldsymbol{P}_{\boldsymbol{A}_0})\, \boldsymbol{z}_{w1}\, (K+1)\hat{\gamma}_1$	$\frac{(K+1)\,N}{K+1-N}\, t(\boldsymbol{z},\boldsymbol{S})$
Gradient	$\Re\{\boldsymbol{z}_{w1}^\dagger (\boldsymbol{P}_{\boldsymbol{A}_1} - \boldsymbol{P}_{\boldsymbol{A}_0})\, \boldsymbol{S}^{1/2}\, \hat{\boldsymbol{M}}_0^{-1/2} \boldsymbol{z}_{w0}\}$	$N\, t_k$

$$P_{fa}^{\mathrm{mis}} = \frac{K!}{(N-J-1)!\,(K-N+J)!} \sum_{\ell=0}^{r-1} \binom{K-N+J}{\ell} \int_0^1 (1-x)^{N-J-1+\ell}\, \eta^\ell \left[\frac{x}{1+\eta(1-x)}\right]^{K-N+J} dx \tag{73}$$

$$P_d^{\mathrm{mis}} = 1 - \frac{K!}{(N-J-1)!\,(K-N+J)!} \int_0^1 x^{K-N+J} \frac{(1-x)^{N-J-1}[\eta(1-x)]^r}{[1+\eta(1-x)]^{K-N+J}} \sum_{k=0}^{K-N+t} \binom{K-N+J}{r+k} [\eta(1-x)]^k \times \exp\left(-\frac{\mathrm{SINR}\,x}{1+\eta(1-x)}\right) \sum_{i=0}^{k} \left(\frac{\mathrm{SINR}\,x}{1+\eta(1-x)}\right)^i \frac{1}{i!}\, dx \tag{74}$$

Also, for the sake of completeness, we will also consider in our analysis the MPI detector, i.e., the clairvoyant (since its implementation requires the SINR to be known in advance) test based on the likelihood ratio of the MIS, whose expression is given by:

$$t_{\mathrm{mpi}} \triangleq \frac{f_t(x|\mathcal{H}_1)}{f_t(x|\mathcal{H}_0)}\,, \tag{75}$$

where $f_t(x|\mathcal{H}_i)$, $i \in \{0,1\}$, is obtained from Eq. (17). In what follows we compare the tests based on the two statistics via simulation results.

In Fig. 1, we report the probability of detection P_d vs. the SINR (assuming $P_{fa} = 10^{-3}$) for MIS and MPI detectors. We recall that the $\mathrm{SINR} = (\boldsymbol{\theta}_2^\dagger \boldsymbol{M}_{2.3}^{-1} \boldsymbol{\theta}_2)$ corresponds to the induced maximal invariant (cf. Sec. III-C). Two cases have been considered: scenario (a), corresponding to $r=2$ and $t=4$ (small uncertainty on the target steering vector, high uncertainty on the interferer steering vector) and scenario (b), corresponding to $r=4$ and $t=2$ (higher uncertainty on the target steering vector, smaller uncertainty on the interferer steering vector). As to the disturbance, we model it as an exponentially-correlated Gaussian vector with covariance matrix (in canonical space) $\boldsymbol{M} = \sigma_n^2 \boldsymbol{I}_N + \sigma_c^2 \boldsymbol{M}_c$, where $\sigma_n^2 > 0$ is the thermal noise power, $\sigma_c^2 > 0$ denotes the clutter power, and the (i,j)-th element of $\boldsymbol{M}_c$ is given by $0.95^{|i-j|}$. The clutter-to-noise ratio σ_c^2/σ_n^2 is set here to $30\,\mathrm{dB}$, with $\sigma_n^2 = 1$.

All the curves have been obtained via standard Monte Carlo (MC) counting techniques, except for those describing MIS performance evaluated through the theoretical expressions in Eqs. (73) and (74), respectively (solid and dashed lines for $K=16$ and $K=24$). More specifically, with reference to MC techniques, the thresholds necessary to ensure a preassigned value of P_{fa} have been evaluated exploiting $(2 \cdot 10^3)/P_{fa}$ independent trials, while the P_d values are estimated over $3 \cdot 10^4$ independent trials.

We point out that the specific value of the deterministic interference $\boldsymbol{\theta}_{1i}$, as well as the scale parameter γ, does not need to be specified at each trial considered (for both P_{fa} and P_d evaluation); the reason is that the performance of each detector depends is independent on nuisance parameters under $\mathcal{H}_0$, as a consequence of the invariance property. Differently, under $\mathcal{H}_1$ their performance depends on the unknown parameters solely through the induced maximal invariant, which is *independent* on $\boldsymbol{\theta}_{1i}$ and γ (cf. Sec. III-C).

In order to average the performance of P_d with respect to $\boldsymbol{\theta}_2$, for each independent trial we generate[8] the signal vector as $\boldsymbol{\theta}_2 = \alpha_B \boldsymbol{\theta}_g$, where $\boldsymbol{\theta}_g \sim \mathcal{CN}(\boldsymbol{0}_r, \boldsymbol{I}_r)$ and $\alpha_B \in \mathbb{R}$. The latter coefficient is a scaling factor used to enforce a deterministic SINR value, that is, $\alpha_B \triangleq \sqrt{\mathrm{SINR} / \left(\boldsymbol{\theta}_g^\dagger \boldsymbol{M}_{2.3}^{-1} \boldsymbol{\theta}_g\right)}$.

From inspection of the figure, the following considerations can be drawn. First of all, as expected MIS curves obtained with MC simulations *coincide* with the theoretical ones provided by Eqs. (73) and (74), respectively. Secondly, the ACE (viz. MIS) and the MPI detector performance *seem to coincide* for all the considered scenarios. This numerical evidence allows us to conjecture that also in the case of additional structured interference the ACE may correspond to the UMPI test, thus constituting a possible extension of the result claimed in [36]. Thirdly, it is evident that increasing the number of secondary data K leads to improved performance of both detectors. Finally, we notice that performance in scenario (b) are worse than those corresponding to the scenario (a). This is simply explained as the increased rank of the target return subspace means increased uncertainty in the knowledge of the true steering vector of the target.

[8]We point out that different sampling procedures for $\boldsymbol{\theta}_2$ would lead to the same performance, as long as the value α_B ensures the SINR to assume a deterministic value.

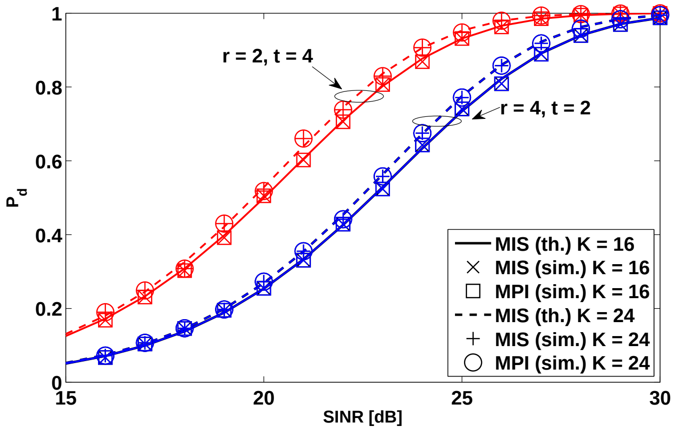

Figure 1. P_d vs. SINR for MIS (theoretical vs. simulated: solid and dashed lines vs. "×" and "+" markers) and MPI ("□" and "○" markers) detectors; setup parameters: $N = 8$, $K \in \{16, 24\}$. Scenario (*a*) (red plots) $r = 2$, $t = 4$; Scenario (*b*) (blue plots) $r = 4$, $t = 2$.

VI. Conclusions

In this paper we provided a complete study for the problem of adaptive detection of a vector subspace target in the presence of vector subspace interference for PHE. The study has been conducted with the help of the statistical theory of invariance. We first obtained the group of transformations leaving the hypothesis testing problem invariant, thus allowing the identification of transformations which enforce the CFAR property. Then, a MIS was derived for the aforementioned group. It was found that the MIS for the problem at hand coincides with the ACE (ANMF) in the complementary subspace of the structured interference. Furthermore, a statistical characterization of the obtained MIS under both hypotheses was obtained.

Then, we focused on the derivation of GLR, 2S-GLR, Rao, Wald, Durbin, and Gradient statistics for the considered problem. Remarkably, all the aforementioned statistics have been shown to be statistically equivalent to the MIS. The statistical characterization of the MIS was exploited for the derivation of explicit forms of detection and false alarm probabilities in one-dimensional integral form. Finally, simulation results were provided with the intent of investigating the performance of the aforementioned test in comparison to the (clairvoyant) MPI detector. Numerical evidence has shown that the test based on the MIS may represent the UMPI test also in a scenario with additional structured interference, thus extending the claim of [36].

Appendix A
Proof of Proposition 1

The proof is obtained by noticing that the action $\ell(\cdot,\cdot)$ (cf. Eq. (7)) can be re-interpreted as the sequential application of the following sub-actions:

$$\ell_1(\boldsymbol{z},\boldsymbol{S}) = (\boldsymbol{G}\,\boldsymbol{z}+\boldsymbol{f},\boldsymbol{G}\,\boldsymbol{S}\,\boldsymbol{G}^\dagger) \quad \forall(\boldsymbol{G},\boldsymbol{f})\in\mathcal{L}_1$$
$$\ell_2(\boldsymbol{z},\boldsymbol{S}) = (\boldsymbol{z},\gamma\boldsymbol{S}) \quad \forall\gamma\in\mathcal{L}_2, \tag{76}$$

where $\mathcal{L}_1 \triangleq \{\mathcal{G}\times\mathcal{F},"\circ"\}$ and $\mathcal{L}_2 \triangleq \{\mathbb{R}^+,"\times"\}$ (i.e., the composition operator for $\mathcal{L}_2$ simply corresponds to the product). Then, it is recognized that the MIS for the sub-action $\ell_1(\cdot,\cdot)$ has been already obtained in [37], as the former represents the relevant action enforcing desirable invariances in a homogeneous background (viz. $\gamma = 1$). Such statistic is two-dimensional and given by $\boldsymbol{t}(\boldsymbol{z},\boldsymbol{S}) = \begin{bmatrix} t_a & t_b\end{bmatrix}^T$, where $t_a \triangleq \boldsymbol{z}_{2.3}^\dagger\,\boldsymbol{S}_{2.3}^{-1}\boldsymbol{z}_{2.3}$ and $t_b \triangleq \boldsymbol{z}_3^\dagger\boldsymbol{S}_{33}^{-1}\boldsymbol{z}_3$, respectively. Now, define the action $\ell_2^\star(\cdot,\cdot)$ acting on the couple of positive-valued scalars (a_1,a_2) (which correspond to t_a and t_b, respectively) as:

$$\ell_2^\star(a_1,\,a_2) = \left(\gamma^{-1}a_1,\gamma^{-1}a_2\right) \quad \forall\,\gamma\in\mathcal{L}_2. \tag{77}$$

It is not difficult to show that a MIS for the elementary operation $\ell_2^\star(\cdot,\cdot)$ in Eq. (77) is given by $t_2(a_1,a_2) \triangleq \frac{a_1}{a_2}$. This is clearly achieved by verifying that both *invariance* and *maximality* properties [34] hold for $t_2(\cdot,\cdot)$. Indeed invariance follows from $t_2(\gamma^{-1}a_1,\gamma^{-1}a_2) = \frac{\gamma^{-1}a_1}{\gamma^{-1}a_2} = \frac{a_1}{a_2} = t_2(a_1,a_2)$, while *maximality* can be proved as follows. Suppose that $t_2(a_1,a_2) = t_2(\bar{a}_1,\bar{a}_2)$, which implies:

$$\bar{a}_2 = \frac{\bar{a}_1}{a_1}a_2\,. \tag{78}$$

Then there exists a $\gamma\in\mathcal{L}_2$, equal to $\gamma = \frac{a_1}{\bar{a}_1}$, which ensures $\left(\gamma^{-1}a_1,\gamma^{-1}a_2\right) = (\bar{a}_1,\bar{a}_2)$. This demonstrates that $t_2(a_1,a_2)$ is a MIS for $\ell_2^\star(\cdot,\cdot)$. Additionally, we notice that

$$\boldsymbol{t}(\bar{\boldsymbol{z}},\bar{\boldsymbol{S}}) = \boldsymbol{t}(\boldsymbol{z},\boldsymbol{S}) \Rightarrow$$
$$\boldsymbol{t}(\bar{\boldsymbol{z}},\gamma\bar{\boldsymbol{S}}) = \boldsymbol{t}(\boldsymbol{z}\,,\gamma\boldsymbol{S})\,,\;\forall\gamma\in\mathcal{L}_2, \tag{79}$$

since $\boldsymbol{t}(\boldsymbol{z},\gamma\boldsymbol{S}) = \frac{1}{\gamma}\boldsymbol{t}(\boldsymbol{z},\boldsymbol{S})$ holds. Therefore, exploiting [34, p. 217, Thm. 6.2.2], it follows that a MIS for the action $\ell(\cdot,\cdot)$ is the composite function $t(\boldsymbol{z},\boldsymbol{S}) \triangleq t_2(\boldsymbol{t}(\boldsymbol{z},\boldsymbol{S})) = \frac{t_a}{t_b} = \frac{\boldsymbol{z}_{2.3}^\dagger\,\boldsymbol{S}_{2.3}^{-1}\boldsymbol{z}_{2.3}}{\boldsymbol{z}_3^\dagger\boldsymbol{S}_{33}^{-1}\boldsymbol{z}_3}$.

Appendix B
Proof of Proposition 4

In this Appendix we derive the alternative form of Rao statistic in Eq. (55). To this end, we consider each term in the difference of Eq. (54) separately. We start from expression in Eq. (24) and notice that $\hat{\boldsymbol{M}}_0$ can be rewritten as $\hat{\boldsymbol{M}}_0 = \bar{\boldsymbol{M}}_0 + \boldsymbol{v}_0\boldsymbol{v}_0^\dagger$, where $\bar{\boldsymbol{M}}_0 \triangleq [\hat{\gamma}_0(K+1)]^{-1}\boldsymbol{S}$ and $\boldsymbol{v}_0 \triangleq (K+1)^{-1/2}\boldsymbol{S}^{1/2}\boldsymbol{P}_{\boldsymbol{A}_0}^\perp\boldsymbol{z}_{w1}$, respectively. Therefore, after (repeated) application of matrix inversion lemma [41], it readily follows that:

$$\hat{\boldsymbol{M}}_0^{-1} = \bar{\boldsymbol{M}}_0^{-1} - \frac{\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0\boldsymbol{v}_0^\dagger\,\bar{\boldsymbol{M}}_0^{-1}}{1+\boldsymbol{v}_0^\dagger\,\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0}\,, \tag{80}$$

$$(\boldsymbol{A}^\dagger\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{A})^{-1} = \left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{A}\right)^{-1} - \frac{\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0\right)\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0\right)^\dagger\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{A}\right)^{-1}}{1+\boldsymbol{v}_0^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0+\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0\right)^\dagger\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0\right)}. \tag{81}$$

Then, exploiting Eq. (80), we get:

$$\boldsymbol{A}^\dagger\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{z} = \boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1/2}\left[\bar{\boldsymbol{M}}_0^{-1/2}\boldsymbol{z} - \frac{\bar{\boldsymbol{M}}_0^{-1/2}\boldsymbol{v}_0(\boldsymbol{v}_0^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{z})}{1+\boldsymbol{v}_0^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0}\right]. \tag{82}$$

The term within square brackets in Eq. (82) also equals

$$\left[\frac{(K+1-N)}{(K+1)}\boldsymbol{I}_N + \frac{N}{(K+1)}\boldsymbol{P}_{\boldsymbol{A}_0}\right]\bar{\boldsymbol{M}}_0^{-1/2}\boldsymbol{z}\,. \tag{83}$$

Such result is drawn noticing that the following equality holds

$$\frac{(\boldsymbol{v}_0^\dagger\,\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{z})}{1+\boldsymbol{v}_0^\dagger\,\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0} = \frac{N}{\sqrt{K+1}}\,, \tag{84}$$

and observing that $\boldsymbol{v}_0 = (K+1)^{-1/2}\left(\bar{\boldsymbol{M}}_0^{1/2}\boldsymbol{P}_{\bar{\boldsymbol{A}}_0}^{\perp}\bar{\boldsymbol{M}}_0^{-1/2}\boldsymbol{z}\right)$. Also, exploiting Eq. (81), it is proved:

$$\begin{aligned}&\bar{\boldsymbol{M}}_0^{-1/2}\boldsymbol{A}(\boldsymbol{A}^\dagger\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1/2}\\ &= \boldsymbol{P}_{\boldsymbol{A}_1}\left[\boldsymbol{I}_N - \frac{\bar{\boldsymbol{M}}_0^{-1/2}\boldsymbol{v}_0\boldsymbol{v}_0^\dagger\bar{\boldsymbol{M}}_0^{-1/2}}{x_0}\right]\boldsymbol{P}_{\boldsymbol{A}_1}\end{aligned} \tag{85}$$

where we have used the projector equivalence $\boldsymbol{P}_{\bar{\boldsymbol{M}}_0^{-1/2}\boldsymbol{A}} = \boldsymbol{P}_{\boldsymbol{A}_1}$ and we have also defined

$$\begin{aligned}x_0 \triangleq\ & 1+\boldsymbol{v}_0^\dagger\,\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0\\ &+\boldsymbol{v}_0^\dagger\,\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{A}\left(\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^\dagger\bar{\boldsymbol{M}}_0^{-1}\boldsymbol{v}_0\,.\end{aligned} \tag{86}$$

Then, it provides:

$$\begin{aligned}\boldsymbol{z}_{w0}^\dagger\,\boldsymbol{P}_{\bar{\boldsymbol{A}}_1}\boldsymbol{z}_{w0} \;=\; \boldsymbol{z}_{w1}^\dagger&\left[\frac{N\left[(K+1-N)\boldsymbol{P}_{\boldsymbol{A}_1}+N\boldsymbol{P}_{\boldsymbol{A}_0}\right]^2\boldsymbol{z}_{w1}}{(K+1-N)\boldsymbol{z}_{w1}^\dagger\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}}\right.\\ &\left.-\frac{N^2}{K+1}\frac{t_k^2}{x_0}\right]\end{aligned} \tag{87}$$

Moreover, we recall that the second term in (54) can be rewritten as:

$$\boldsymbol{z}_{w0}^\dagger\,\boldsymbol{P}_{\bar{\boldsymbol{A}}_0}\,\boldsymbol{z}_{w0} = \frac{(K+1)\,N}{(K+1-N)}\,\frac{\boldsymbol{z}_{w1}^\dagger\boldsymbol{P}_{\boldsymbol{A}_0}\boldsymbol{z}_{w1}}{\boldsymbol{z}_{w1}^\dagger\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}}\,, \tag{88}$$

where we have exploited Lemma 2. Combining Eqs. (87) and (88), we obtain:

$$t_{\mathrm{rao}} = \frac{N(K+1-N)}{K+1}t_k - \frac{N^2}{K+1}\frac{t_k^2}{x_0}\,. \tag{89}$$

Finally, noticing that $x_0 = 1+\frac{N}{K+1-N}+\frac{N}{K+1-N}\,t_k$, we prove the claimed result.

Appendix C
Proof of Proposition 6

Hereinafter we show that the gradient statistic is statistically equivalent to the MIS, by proving Eq. (68). With this intent, we analyze each term in the difference of Eq. (67) separately.

As in Appendix B, we rewrite the estimated covariance (under $\mathcal{H}_0$) $\hat{\boldsymbol{M}}_0$ as $\hat{\boldsymbol{M}}_0 = \bar{\boldsymbol{M}}_0 + \boldsymbol{v}_0\,\boldsymbol{v}_0^\dagger$, where $\bar{\boldsymbol{M}}_0 \triangleq [\hat{\gamma}_0(K+1)]^{-1}\,\boldsymbol{S}$ and $\boldsymbol{v}_0 \triangleq (K+1)^{-1/2}\boldsymbol{S}^{1/2}\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}$, respectively. Therefore, after application of matrix inversion lemma (as in Eq. (80)), it readily follows that:

$$\begin{aligned}&\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{z} = (K+1)\,\hat{\gamma}_0\\ &\times\left[\frac{\boldsymbol{S}^{-1}\boldsymbol{z}+\hat{\gamma}_0\left(\boldsymbol{z}_{w1}^\dagger\,\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}\right)\left(\boldsymbol{S}^{-1/2}\,\boldsymbol{P}_{\boldsymbol{A}_0}\boldsymbol{z}_{w1}\right)}{1+\hat{\gamma}_0\,\boldsymbol{z}_{w1}^\dagger\,\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}}\right].\end{aligned} \tag{90}$$

Then, exploiting the result in Eq. (90), it follows:

$$\begin{aligned}&(\boldsymbol{P}_{\boldsymbol{A}_1}-\boldsymbol{P}_{\boldsymbol{A}_0})\,\boldsymbol{S}^{1/2}\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{z}\\ &= \frac{(K+1)\hat{\gamma}_0(\boldsymbol{P}_{\boldsymbol{A}_1}-\boldsymbol{P}_{\boldsymbol{A}_0})\boldsymbol{z}_{w1}}{1+\hat{\gamma}_0\,\boldsymbol{z}_{w1}^\dagger\,\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}} && (91)\\ &= N\,\frac{(\boldsymbol{P}_{\boldsymbol{A}_1}-\boldsymbol{P}_{\boldsymbol{A}_0})\boldsymbol{z}_{w1}}{\boldsymbol{z}_{w1}^\dagger\,\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}} && (92)\end{aligned}$$

where second line arises from $(\boldsymbol{P}_{\boldsymbol{A}_1}-\boldsymbol{P}_{\boldsymbol{A}_0})\boldsymbol{P}_{\boldsymbol{A}_0} = \boldsymbol{0}_{N\times N}$ and the final expression is obtained by exploiting the closed form of $\hat{\gamma}_0$, given in Eq. (22). Therefore, substitution of Eq. (92) into (67) provides

$$\begin{aligned}&\Re\left\{\boldsymbol{z}_{w1}^\dagger(\boldsymbol{P}_{\boldsymbol{A}_1}-\boldsymbol{P}_{\boldsymbol{A}_0})\boldsymbol{S}^{1/2}\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{z}\right\}\\ &= N\,\frac{\boldsymbol{z}_{w1}^\dagger\,(\boldsymbol{P}_{\boldsymbol{A}_1}-\boldsymbol{P}_{\boldsymbol{A}_0})\,\boldsymbol{z}_{w1}}{\boldsymbol{z}_{w1}^\dagger\,\boldsymbol{P}_{\boldsymbol{A}_0}^{\perp}\boldsymbol{z}_{w1}}\,,\end{aligned} \tag{93}$$

where $\Re\{\cdot\}$ has been dropped, since Eq. (93) contains only Hermitian (real-valued) quadratic forms. This concludes the proof.

Appendix D
Proof of Proposition 7

In order to prove coincidence between Rao and Durbin tests, it suffices to show the equivalence (cf. Eqs. (36) and (69))

$$\left.\frac{\partial\ln f_1(\boldsymbol{Z};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_r}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} = \left[\boldsymbol{I}\left(\widehat{\boldsymbol{\theta}}_0\right)\right]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r}(\hat{\boldsymbol{\theta}}_{r,01}-\boldsymbol{\theta}_{r,0})\,, \tag{94}$$

for the model under investigation. Following Eq. (42), we know that left hand side in Eq. (94) equals $\left[2\Re\{\boldsymbol{g}_{r,0}\}^T \quad 2\,\Im\{\boldsymbol{g}_{r,0}\}^T\right]^T$, where

$$\boldsymbol{g}_{r,0} = \boldsymbol{E}_r^\dagger\hat{\boldsymbol{M}}_0^{-1/2}\,\boldsymbol{P}_{\bar{\boldsymbol{A}}_0}^{\perp}\,\boldsymbol{z}_{w0}\,. \tag{95}$$

Differently the right hand side of Eq. (94) can be calculated as explained hereinafter. Indeed, we first notice that $\boldsymbol{\theta}_{r,0} = \boldsymbol{0}_{2r}$. Then, we observe that $\hat{\boldsymbol{\theta}}_{r,01}$ can be shown to be equal to

$$\hat{\boldsymbol{\theta}}_{r,01} = \left[\Re\{\boldsymbol{\psi}_0\}^T \quad \Im\{\boldsymbol{\psi}_0\}^T\right]^T\,, \tag{96}$$

where:

$$\begin{aligned}\boldsymbol{\psi}_0 &\triangleq (\boldsymbol{E}_r^\dagger\,\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{E}_r)^{-1}\boldsymbol{E}_r^\dagger\,\hat{\boldsymbol{M}}_0^{-1/2}(\boldsymbol{z}-\boldsymbol{E}_t\hat{\boldsymbol{\theta}}_{10}) && (97)\\ &= (\boldsymbol{E}_r^\dagger\,\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{E}_r)^{-1}\boldsymbol{E}_r^\dagger\,\hat{\boldsymbol{M}}_0^{-1/2}\boldsymbol{P}_{\bar{\boldsymbol{A}}_0}^{\perp}\boldsymbol{z}_{w0}. && (98)\end{aligned}$$

Clearly, Eq. (98) is obtained by plugging in the closed form expression of $\hat{\boldsymbol{\theta}}_{10}$ (given by Eq. (20)) in Eq. (97). Furthermore, the block $(\boldsymbol{\theta}_r,\boldsymbol{\theta}_r)$ of the FIM of interest (evaluated in $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_0$) equals:

$$\left[\boldsymbol{I}\left(\widehat{\boldsymbol{\theta}}_0\right)\right]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} = \begin{bmatrix}2\,\Re\{\boldsymbol{\Psi}_0\} & -2\,\Im\{\boldsymbol{\Psi}_0\}\\ 2\,\Im\{\boldsymbol{\Psi}_0\} & 2\,\Re\{\boldsymbol{\Psi}_0\}\end{bmatrix}\,, \tag{99}$$

where we have defined $\boldsymbol{\Psi}_0 \triangleq (\boldsymbol{E}_r^\dagger\,\hat{\boldsymbol{M}}_0^{-1}\boldsymbol{E}_r)$. Combining Eqs. (96) and (99), and exploiting the well-known equivalence:

$$\begin{bmatrix}\Re\{\boldsymbol{D}\} & -\Im\{\boldsymbol{D}\}\\ \Im\{\boldsymbol{D}\} & \Re\{\boldsymbol{D}\}\end{bmatrix}\begin{bmatrix}\Re\{\boldsymbol{\delta}\}\\ \Im\{\boldsymbol{\delta}\}\end{bmatrix} = \begin{bmatrix}\Re\{\boldsymbol{D}\boldsymbol{\delta}\}\\ \Im\{\boldsymbol{D}\boldsymbol{\delta}\}\end{bmatrix}\,, \tag{100}$$

where $\boldsymbol{D}$ and $\boldsymbol{\delta}$ denote a generic matrix and a generic vector, respectively, it holds:

$$\left[\boldsymbol{I}\left(\widehat{\boldsymbol{\theta}}_0\right)\right]_{\boldsymbol{\theta}_r,\boldsymbol{\theta}_r} \hat{\boldsymbol{\theta}}_{r,01} = 2 \begin{bmatrix} \Re\{\boldsymbol{\Psi}_0 \boldsymbol{\psi}_0\} \\ \Im\{\boldsymbol{\Psi}_0 \boldsymbol{\psi}_0\} \end{bmatrix}, \qquad (101)$$

which thus proves Eq. (94) and, consequently, the claimed proposition.

REFERENCES

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

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

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

[4] 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, 2016.

[5] W. Liu, W. Xie, J. Liu, and Y. Wang, "Adaptive double subspace signal detection in Gaussian background, Part I: Homogeneous environments," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2345–2357, May 2014.

[6] K. F. McDonald and R. S. Blum, "Exact performance of STAP algorithms with mismatched steering and clutter statistics," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2750–2763, Oct. 2000.

[7] W. L. Melvin, "Space-time adaptive radar performance in heterogeneous clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 621–633, Apr. 2000.

[8] J. Ward, "Space-time adaptive processing for airborne radar," MIT Lincoln Lab., Tech. Rep. 1015, 1994.

[9] M. C. Wicks, M. Rangaswamy, R. Adve, and T. B. Hale, "Space-time adaptive processing: a knowledge-based perspective for airborne radar," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 51–65, Jan. 2006.

[10] N. Bon, A. Khenchaf, and R. Garello, "GLRT subspace detection for range and Doppler distributed targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 2, pp. 678–696, 2008.

[11] E. Conte, M. Lops, and G. Ricci, "Adaptive radar detection in compound-Gaussian clutter," in *Proc. of the European Signal processing Conference (EUSIPCO)*, 1994, pp. 526–529.

[12] ——, "Adaptive matched filter detection in spherically invariant noise," *IEEE Signal Process. Lett.*, vol. 3, no. 8, pp. 248–250, Aug. 1996.

[13] L. L. Scharf and T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proc. of 30th Asilomar Conference on Signals, Systems and Computers*, 1996, pp. 1114–1117.

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

[15] A. De Maio and S. Iommelli, "Coincidence of the Rao test, Wald test, and GLRT in partially homogeneous environment," *IEEE Signal Process. Lett.*, vol. 15, pp. 385–388, 2008.

[16] J. Liu, Z.-J. Zhang, Y. Yang, and H. Liu, "A CFAR adaptive subspace detector for first-order or second-order Gaussian signals based on a single observation," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5126–5140, Nov. 2011.

[17] E. J. Kelly and K. M. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal models," Massachussets Inst. of Tech. Lexington Lincoln Lab, Tech. Rep. No. TR-848., 1989.

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

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

[20] 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.

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

[22] C. Hao, X. Ma, X. Shang, and L. Cai, "Adaptive detection of distributed targets in partially homogeneous environment with Rao and Wald tests," *Signal Processing*, vol. 92, no. 4, pp. 926–930, Apr. 2012.

[23] W. Liu, W. Xie, J. Liu, and Y. Wang, "Adaptive double subspace signal detection in Gaussian background-Part II: Partially homogeneous environments," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2358–2369, May 2014.

[24] S. Bidon, O. Besson, and J.-Y. Tourneret, "The adaptive coherence estimator is the generalized likelihood ratio test for a class of heterogeneous environments," *IEEE Signal Process. Lett.*, vol. 15, pp. 281–284, 2008.

[25] P. Wang, Z. Sahinoglu, M.-O. Pun, H. Li, and B. Himed, "Knowledge-aided adaptive coherence estimator in stochastic partially homogeneous environments," *IEEE Signal Process. Lett.*, vol. 18, no. 3, pp. 193–196, Mar. 2011.

[26] M. Casillo, A. De Maio, S. Iommelli, and L. Landi, "A persymmetric GLRT for adaptive detection in partially-homogeneous environment," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 1016–1019, 2007.

[27] C. Hao, D. Orlando, X. Ma, and C. Hou, "Persymmetric Rao and Wald tests for partially homogeneous environment," *IEEE Signal Process. Lett.*, vol. 19, no. 9, pp. 587–590, Sep. 2012.

[28] Y. Gao, G. Liao, S. Zhu, X. Zhang, and D. Yang, "Persymmetric adaptive detectors in homogeneous and partially homogeneous environments," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 331–342, Jan. 2014.

[29] A. De Maio, C. Hao, and D. Orlando, "An adaptive detector with range estimation capabilities for partially homogeneous environment," *IEEE Signal Process. Lett.*, vol. 21, no. 3, pp. 325–329, Mar. 2014.

[30] C. Hao, D. Orlando, G. Foglia, X. Ma, and C. Hou, "Adaptive radar detection and range estimation with oversampled data for partially homogeneous environment," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1359–1363, Sep. 2015.

[31] S. Bose and A. Steinhardt, "Adaptive array detection of uncertain rank one waveforms," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2801–2809, Nov. 1996.

[32] 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.

[33] L. L. Scharf, *Statistical signal processing*. Addison-Wesley Reading, MA, 1991, vol. 98.

[34] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*. Springer Science & Business Media, 2006.

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

[36] 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.

[37] A. De Maio and D. Orlando, "Adaptive detection of a subspace signal embedded in subspace structured plus Gaussian interference via invariance," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2156–2167, 2015.

[38] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Prentice Hall PTR, Jan. 1998.

[39] G. R. Terrell, "The gradient statistic," *Computing Science and Statistics*, vol. 34, pp. 206–215, 2002.

[40] J. Durbin, "Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables," *Econometrica: Journal of the Econometric Society*, pp. 410–421, 1970.

[41] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 2012.

[42] R. J. Muirhead, *Aspects of multivariate statistical theory*. John Wiley & Sons, 2009, vol. 197.

[43] F. Bandiera, D. Orlando, and G. Ricci, "Advanced radar detection schemes under mismatched signal models," *Synthesis lectures on signal processing*, vol. 4, no. 1, pp. 1–105, 2009.

[44] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[45] A. J. Lemonte and S. L. P. Ferrari, "The local power of the gradient test," *Annals of the Institute of Statistical Mathematics*, vol. 64, no. 2, pp. 373–381, 2012.

[46] A. De Maio and E. Conte, "Adaptive detection in Gaussian interference with unknown covariance after reduction by invariance," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 2925–2934, Jun. 2010.