

Adaptive Radar Detection and Classification Algorithms for Multiple Coherent Signals

Sudan Han, Linjie Yan, Yuxuan Zhang, Pia Addabbo, *Senior Member, IEEE*,
Chengpeng Hao, *Senior Member, IEEE*, and Danilo Orlando, *Senior Member, IEEE*

Abstract—In this paper, we address the problem of target detection in the presence of coherent (or fully correlated) signals, which can be due to multipath propagation effects or electronic attacks by smart jammers. To this end, we formulate the problem at hand as a multiple-hypothesis test that, besides the conventional radar alternative hypothesis, contains additional hypotheses accounting for the presence of an unknown number of interfering signals. In this context and leveraging the classification capabilities of the Model Order Selection rules, we devise penalized likelihood-ratio-based detection architectures that can establish, as a byproduct, which hypothesis is in force. Moreover, we propose a suboptimum procedure to estimate the angles of arrival of multiple coherent signals ensuring (at least for the considered parameters) almost the same performance as the exhaustive search. Finally, the performance assessment, conducted over simulated data and in comparison with conventional radar detectors, highlights that the proposed architectures can provide satisfactory performance in terms of probability of detection and correct classification.

Index Terms—Adaptive Radar Detection, Classification, Electronic Counter-Countermeasures, Fully Coherent Signals, Generalized Likelihood Ratio Test, Model Order Selection Rules, Multipath, Radar, Smart Jammer.

I. INTRODUCTION

In recent years, radar systems have become ubiquitous in real life due to the advances in digital architectures and miniaturization technologies. More importantly, the huge amount of computational power has paved the way for sophisticated processing algorithms fed by digital samples and leading to new architectures where the presence of analog hardware resources devoted to a specific task is very limited [1]–[3]. As a consequence, modern radar systems are extremely flexible and can incorporate different functions without additional hardware components.

In system design, the turning point is represented by the knocking down of high-frequency sampler and processing board costs that have allowed the development of fully-digital architectures. In this evolving scenario, radar research community continues to devise algorithms with increased complexity

that take full advantage of the potential provided by digital architectures. The *quid pro quo* of the increased computational power is represented by the enhanced performance as corroborated, for instance, by the space and/or time adaptive algorithms developed in the open literature. As a matter of fact, focusing on the radar detection task and starting from the seminal works by Kelly [4], [5], a plethora of decision schemes have been proposed by enriching the design assumptions in order to account for a priori information and/or specific aspects of the application/system under consideration [4], [6]–[20].

It is also important to underline that the benefits coming from the aforementioned technology advances have been also exploited by the Electronic Warfare (EW) systems, which have adapted themselves to the more and more reliable capabilities of radar systems leading to a more effective class of electronic countermeasures referred to as smart jammers [3], [21]–[25]. For instance, modern noise-like jamming systems are capable of transmitting narrow-band interfering signals which are concurrent with the radar pulses according to the radar pulse repetition interval after having estimated it (EW integrated systems). Moreover, they can use an intrapulse modulation to generate a noise bandwidth at radio frequency, while maintaining phase coherence over a group of successive jamming pulses [25]. Remarkably, this kind of jammers can generate signals that are coherent (i.e., fully correlated) with the desired signal even though the former impinge on the radar from different directions [26], [27]. An analogous situation occurs in scenarios where, due to multipath propagation, replicas of the original signal come back to the receiver with a sufficiently small delay difference [27], [28].

The main drawback caused by the presence of multiple coherent signals is that they can completely destroy the performance of the most common high-resolution direction finding algorithms for adaptive array systems [29], [30]. As a matter of fact, coherent signals appear as a single signal impinging on the array of sensors and that, more importantly, arrives from a direction which is quite different from that of the sought signals. As a consequence, high-resolution eigenstructure-based techniques as, for instance, MULTIPLE SIGNAL CLASSIFICATION algorithm [29], fail to correctly resolve the signals jeopardizing the Angle of Arrival (AoA) estimation [26]. Similar remarks also hold for another important approach to AoA estimation, namely the CLEAN algorithm [31] which consists in iterative cancellations of strong signals under the assumption that the spatial covariance matrix results from the sum of contributions associated with uncorrelated sources. Remarkably, it can be used for the detection of weak targets embedded in strong

Sudan Han is with the National Innovation Institute of Defense Technology, Beijing, China. E-mail: xiaoxiaosu0626@163.com.

Pia Addabbo is with Università degli studi “Giustino Fortunato”, Benevento, Italy. E-mail: p.addabbo@unifortunato.eu.

Linjie Yan, Yuxuan Zhang, and Chengpeng Hao are with Institute of Acoustics, Chinese Academy of Sciences, Beijing, China. E-mail: yanlinjie16@163.com; zhangyuxuan@mail.ioa.ac.cn; haochengp@mail.ioa.ac.cn.

Daniilo Orlando is with the Engineering Faculty of Università degli Studi “Niccolò Cusano”, via Don Carlo Gnocchi 3, 00166 Roma, Italy. E-mail: danilo.orlando@unicusano.it.

interference signals [32]–[34].

In order to mitigate the effects of coherent signals, data feeding the direction finding algorithm can be suitably preprocessed in order to decrease the correlation of the impinging signals. In this respect, spatial smoothing technique [30] combines data obtained from synthetic subarrays whose size is lower than that of the original array with the drawback of a reduced angular resolution. In the context of subspace-based algorithms, in [35] a method based upon the weighted subspace fitting criterion is proposed to jointly detect and estimate the number and the related parameters of the coherent signals. However, this method does not allow for the control of the probability of false alarm (P_{fa}) which is of primary concern in radar. Alternative approaches against the coherent signals problem can rely on the Maximum Likelihood (ML) estimation [36] which, however, is computationally intensive and requires the knowledge of the number of coherent signals, or the compressed sensing paradigm where a further stage for estimate fusion is required [37], [38].

Thus, in order to save computational resources (due to the activation of additional processing stages also in situations where they are not required) and to take advantage of the full angular resolution of the system, a detection stage capable of deciding for the presence of a target and possible coherent signals by estimating, as a byproduct, their number along with other side information is highly desirable. In fact, when this stage declares the presence of noncoherent signals, direction finding algorithms can be applied without losses in resolution. On the other hand, when coherent signals are present, the rough estimates of the corresponding parameters provided by this stage can be used to drive ML-based direction finding algorithms to reduce their computational load. Remarkably, such stage can be viewed as an Electronic Counter-CounterMeasure.

With the above remarks in mind, in this paper we devise a detection architecture accounting for the presence of coherent signals and that can provide an estimate of their number as well as the AoAs. To this end, at the design stage, we formulate the problem at hand in terms of a multiple hypothesis test, comprising the usual null (or interference-only) hypothesis, the conventional signal-plus-interference hypothesis, and multiple alternative hypotheses that differ in the number of coherent echoes. These signals are assumed to follow Swerling II Target Model, which assumes that the Radar Cross Section (RCS) of the target obeys the chi-squared distribution with two degrees of freedom [1]. Then, assuming an upper bound on the number of impinging signals (that can be dictated by system parameters), we conceive likelihood-ratio-based decision rules which exploit suitable penalty terms borrowed from the Model Order Selection (MOS) rules [39], [40]. Therefore, the proposed architectures can provide an estimate of the actual number of coherent signals. As for the AoA estimation, the angular sector under consideration is sampled to form a discrete set of angular positions that are used at the design stage. Besides, since the exhaustive search can become very time demanding for high numbers of signals, we design a suboptimum iterative procedure providing satisfactory performance at least for the considered numerical examples. Finally, the performance assessment is conducted over synthetic data in comparison

with classical detection architectures and highlights that the proposed schemes represent an effective means to face with the problem of target detection in the presence of coherent signals.

The remainder of the paper is organized as follows. Section II is devoted to problem formulation and the definitions used in the next derivations, while the design of the detection and estimation architectures is described in Section III. Section IV shows the effectiveness of the proposed strategies through numerical examples on simulated data. Finally, Section V contains concluding remarks and charts a course for future works.

A. Notation and Acronyms

In the sequel, vectors and matrices are denoted by bold-face lower-case and upper-case letters, respectively. Symbols $\det(\cdot)$, $\text{Tr}(\cdot)$, $(\cdot)^T$, and $(\cdot)^\dagger$ denote the determinant, trace, transpose, and complex conjugate transpose, respectively. If A and B are two generic sets, then $A \times B$ denotes the Cartesian product between A and B . As to the numerical sets, \mathbb{C} is the set of complex numbers, and $\mathbb{C}^{N \times M}$ is the Euclidean space of $(N \times M)$ -dimensional complex-valued matrices (or vectors if $M = 1$). The imaginary unit is indicated by j . The Euclidean norm of a generic vector \mathbf{x} is denoted by $\|\mathbf{x}\|$ whereas the modulus of a complex number x is denoted by $|x|$. The symbol $\mathbb{E}[\cdot]$ denotes statistical expectation while $\mathbf{0}$ and \mathbf{I} are the null vector/matrix and the identity matrix, respectively, of proper size. Given two events A and B , the conditional probability of A given B is denoted by $P(A|B)$. The acronyms PDF and IID mean Probability Density Function and Independent and Identically Distributed, respectively. For a given matrix \mathbf{A} , $\lambda_{\max}\{\mathbf{A}\}$ denotes the maximum eigenvalue of \mathbf{A} . Finally, we write $\mathbf{x} \sim \mathcal{CN}_N(\mathbf{m}, \mathbf{M})$ if \mathbf{x} is a complex circular N -dimensional normal vector with mean \mathbf{m} and positive definite covariance matrix \mathbf{M} .

II. PROBLEM FORMULATION

Let us assume that a search radar system is equipped with a uniform linear array with N antennas and transmits L pulses in the nominal beam position during one scan cycle. Then, each antenna collects L samples from the cell under test. Denote by $z_l \in \mathbb{C}^{N \times 1}$, $l = 1, \dots, L$, the vector whose entries are the l th returns from each antenna and by $\mathbf{Z}_L = [z_1, \dots, z_L] \in \mathbb{C}^{N \times L}$ the overall data matrix, the classical radar detection problem consists in deciding whether or not \mathbf{Z}_L contains the target of interest, namely, a component whose signature coincides with the nominal steering vector. **Before proceeding with the problem formulation, it is important to state here that the columns of \mathbf{Z}_L are assumed statistically independent and, hence, temporally noncoherent, whereas each column is representative of both spatially correlated clutter and spatially coherent useful signal components (a point better explained below).**

As customary, we assume that a set of K ($K \geq N$) secondary data, $z_k \in \mathbb{C}^{N \times 1}$, $k = 1, \dots, K$, free of target components and sharing the same statistical properties of the interference in the cell under test, is available [4], [7],

[12], [41]. Under the above assumptions, this problem can be formulated in terms of a conventional binary hypothesis test whose expression is (see [4], [7], [12] and references therein)

$$\begin{cases} H_0 : \begin{cases} \mathbf{z}_l = \mathbf{n}_l, & l = 1, \dots, L, \\ \mathbf{z}_k = \mathbf{n}_k, & k = 1, \dots, K, \end{cases} \\ H_1 : \begin{cases} \mathbf{z}_l = \alpha_l \mathbf{v}(\theta_t) + \mathbf{n}_l, & l = 1, \dots, L, \\ \mathbf{z}_k = \mathbf{n}_k, & k = 1, \dots, K, \end{cases} \end{cases} \quad (1)$$

where

- $\mathbf{n}_l, \mathbf{n}_k \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M})$, $l = 1, \dots, L$, $k = 1, \dots, K$, are IID with unknown positive definite covariance matrix $\mathbf{M} \in \mathbb{C}^{N \times N}$ representative of the thermal noise plus clutter; \mathbf{M} is referred to in the following as Interference Covariance Matrix (ICM);
- $\alpha_1, \dots, \alpha_L \sim \mathcal{CN}_1(0, \sigma_\alpha^2)$ with $\sigma_\alpha^2 > 0$ are IID random variables accounting for both the target response (RCS) and channel effects (Swerling II target model [1]);
- $\mathbf{v}(\theta_t) \in \mathbb{C}^{N \times 1}$ is the nominal (spatial) steering vector whose expression is

$$\mathbf{v}(\theta_t) = [1, \exp\{j\pi \sin(\theta_t)\}, \dots, \exp\{j\pi(N-1) \sin(\theta_t)\}]^T,$$

where θ_t is the (known) AoA of the target measured with respect to the array broadside and we have assumed that the inter-element spacing is half the operating wavelength in order to avoid the aliasing of the spatial frequency.

However, as stated in Section I, in practical applications the generic \mathbf{z}_l might not only contain the direct echo but also returns from other directions and that, more importantly, are coherent (i.e., fully correlated) with the former due, for instance, to the effects of multipath propagation and/or the presence of smart jammers [26]–[28]. As a consequence, the conventional radar detection problem (1) consisting of two hypotheses may be no longer representative of the actual operating scenario. In fact, in such case, the alternative hypothesis can be replaced by the following

$$H_{1,M} : \begin{cases} \mathbf{z}_l = \alpha_l \mathbf{v}(\theta_t) + \sum_{i=1}^M \beta_{l,i} \mathbf{v}(\theta_i) + \mathbf{n}_l, & l = 1, \dots, L, \\ \mathbf{z}_k = \mathbf{n}_k, & k = 1, \dots, K, \end{cases} \quad (2)$$

where M is the actual number of coherent signals, which is unknown. However, notice that the uncertainty on M naturally leads to multiple alternative hypotheses, $H_{1,i}$ say, which differ in the number of coherent signals. Therefore, assuming an upper bound, M_c say, on the latter unknown parameter, a possible approach to account for the presence of possible fully-correlated signals consists in considering a multiple hypothesis test where besides the conventional H_0 and H_1 hypotheses, $M_c < N$ additional hypotheses¹, corresponding to scenarios

containing a different number of coherent signals, appear. Therefore, problem (1) can be generalized as follows

$$\begin{cases} H_0 : \begin{cases} \mathbf{z}_l = \mathbf{n}_l, & l = 1, \dots, L, \\ \mathbf{z}_k = \mathbf{n}_k, & k = 1, \dots, K, \end{cases} \\ H_{1,0} : \begin{cases} \mathbf{z}_l = \alpha_l \mathbf{v}(\theta_t) + \mathbf{n}_l, & l = 1, \dots, L, \\ \mathbf{z}_k = \mathbf{n}_k, & k = 1, \dots, K, \end{cases} \\ H_{1,i} : \begin{cases} \mathbf{z}_l = \alpha_l \mathbf{v}(\theta_t) + \sum_{k=1}^i \beta_{l,k} \mathbf{v}(\theta_k) + \mathbf{n}_l, & l = 1, \dots, L, \\ \mathbf{z}_k = \mathbf{n}_k, & k = 1, \dots, K, \end{cases} \end{cases} \quad (3)$$

$i = 1, \dots, M_c$, where the α_l s, the \mathbf{n}_l s, and the \mathbf{n}_k s have been already defined for problem (1) and

- $M_c \geq M$ is an upper bound on the number of coherent signals²;
- given $i \in \{1, \dots, M_c\}$ and $k \in \{1, \dots, i\}$, the coefficients $\beta_{1,k}, \dots, \beta_{L,k} \in \mathbb{C}$ are IID complex Gaussian random variables with zero mean and variance $\sigma_k^2 > 0$; they represent the complex amplitudes of the k th coherent signal over the time³;
- given $i \in \{1, \dots, M_c\}$, $\theta_1, \dots, \theta_i$ are the unknown AoAs of the additional coherent signals.

Finally, we assume that, $\forall i = 1, \dots, M_c$, the random variables $\alpha_l, \beta_{l,1}, \dots, \beta_{l,i}$ are fully correlated. Otherwise stated, given $l \in \{1, \dots, L\}$ and $i \in \{1, \dots, M_c\}$, $\forall k, h \in \{1, \dots, i\}$ and $k \neq h$ the following equalities hold

$$\mathbb{E}[\alpha_l \beta_{l,k}] = \sigma_\alpha \sigma_k \quad \text{and} \quad \mathbb{E}[\beta_{l,k} \beta_{l,h}] = \sigma_k \sigma_h. \quad (4)$$

In order to unburden the notation, hereafter we omit the dependence of the steering vectors on the AoAs and denote by \mathbf{v}_t the nominal steering vector whereas we use \mathbf{p}_i to indicate the steering vector associated with the coherent signal whose AoA is θ_i .

Before concluding this section, we provide some definitions that will come in handy for the next developments. More precisely, the PDF of \mathbf{Z}_L under H_0 has the following expression

$$f_0(\mathbf{Z}_L; \mathbf{M}) = \left[\frac{1}{\pi^N \det(\mathbf{M})} \right]^L \exp \left[-\text{Tr} \left(\mathbf{M}^{-1} \mathbf{Z}_L \mathbf{Z}_L^\dagger \right) \right], \quad (5)$$

whereas its PDF under the generic $H_{1,i}$, $i = 0, \dots, M_c$, is given by

$$f_{1,i}(\mathbf{Z}_L; \mathbf{M}, \mathbf{R}_i, \mathbf{V}_i) = \left[\frac{1}{\pi^N \det(\mathbf{M} + \mathbf{V}_i \mathbf{R}_i \mathbf{V}_i^\dagger)} \right]^L \times \exp \left\{ -\text{Tr} \left[\left(\mathbf{M} + \mathbf{V}_i \mathbf{R}_i \mathbf{V}_i^\dagger \right)^{-1} \mathbf{Z}_L \mathbf{Z}_L^\dagger \right] \right\}, \quad (6)$$

where $\mathbf{V}_i = [\mathbf{v}_t, \mathbf{p}_1, \dots, \mathbf{p}_i] \in \mathbb{C}^{N \times (i+1)}$ and $\mathbf{R}_i = \mathbb{E}[\boldsymbol{\nu}_{l,i} \boldsymbol{\nu}_{l,i}^\dagger] \in \mathbb{C}^{(i+1) \times (i+1)}$, $l = 1, \dots, L$, with $\boldsymbol{\nu}_{l,i} = [\alpha_l, \beta_{l,1}, \dots, \beta_{l,i}]^T \in \mathbb{C}^{(i+1) \times 1}$. It is understood that when $i = 0$, \mathbf{V}_0 and $\boldsymbol{\nu}_{l,0}$ coincide with \mathbf{v}_t and α_l , respectively.

¹The constraint $M_c < N$ is due to both the fact that N degrees of freedom are required to estimate N AoAs also making matrix \mathbf{V}_i , defined in (6), full-column rank.

²Note that problem (3) reduces to (1) when $M_c = 0$.

³It is worth noticing that Swerling II target model is also adopted for the coherent signals.

III. DETECTION ARCHITECTURE DESIGNS

In this section, we devise a detection architecture for problem (3) that relies on a *penalized* log-likelihood ratio test [29], [42] whose generic structure is given by

$$\left[\widehat{\Lambda}_i(\mathbf{Z}_L) - c \cdot h(\hat{i}) \right] \underset{H_0}{\overset{H_{1,i}}{>}} \eta, \quad (7)$$

where

$$\hat{i} = \arg \max_{i=0, \dots, M_c} \left\{ \widehat{\Lambda}_i(\mathbf{Z}_L) - c \cdot h(i) \right\}, \quad (8)$$

η is the detection threshold to be set according to a desired⁴ P_{fa} ,

$$\widehat{\Lambda}_i(\mathbf{Z}_L) = \log \frac{\max_{\mathbf{R}_i} \max_{\theta_1, \dots, \theta_i} f_{1,i}(\mathbf{Z}_L; \widehat{\mathbf{M}}, \mathbf{R}_i, \mathbf{V}_i)}{f_0(\mathbf{Z}_L; \widehat{\mathbf{M}})} \quad (9)$$

with⁵ $\widehat{\mathbf{M}}$ a suitable estimate of \mathbf{M} , and $c \cdot h(i)$ is a penalty term borrowed from the MOS rules [39] with $h(i)$ the number of unknown real-valued parameters under the $H_{1,i}$ hypothesis and c a suitable scaling factor. Specifically, c can be set according to the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Generalized Information Criterion (GIC), namely

$$c = \begin{cases} 1, & \text{for AIC-based Detector (AIC-D),} \\ \log(2LN)/2, & \text{for BIC-based Detector (BIC-D),} \\ (1 + \rho)/2, \rho > 1, & \text{for GIC-based Detector (GIC-D).} \end{cases} \quad (10)$$

Two important remarks are now in order. First, notice that $\widehat{\Lambda}_i(\mathbf{Z}_L)$ is the logarithm of a decision statistic obtained by resorting to a modified Generalized Likelihood Ratio Test (GLRT)-based design procedure [7], where the ICM is firstly assumed known and then replaced by an estimate. In the specific case, we use the ML estimate based upon the training samples only (namely, the sample covariance matrix that is invertible with probability 1 since $K \geq N$), whose expression is⁶

$$\widehat{\mathbf{M}} = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^\dagger. \quad (11)$$

Second, the correlation of the signals makes the columns of the positive semidefinite matrix \mathbf{R}_i proportional. Therefore, the resulting rank is 1 and it can be factorized as

$$\mathbf{R}_i = \mathbf{r}_i \mathbf{r}_i^\dagger, \quad (12)$$

where $\mathbf{r}_i \in \mathbb{C}^{(i+1) \times 1}$. As a consequence, the number of unknown real-valued parameters under $H_{1,i}$ can be written

⁴Hereafter, we denote by η the generic detection threshold.

⁵Notice that for $i = 0$, the maximization at the numerator of (9) is with respect to σ_α^2 only.

⁶This design choice is dictated by the fact that deriving closed-form expression for the plain GLRT is not mathematically tractable at least to the best of authors' knowledge.

as $h(i) = N^2 + 1 + 2i$. With the above remarks in mind, let us proceed by deriving $\widehat{\Lambda}_0(\mathbf{Z}_L)$, namely

$$\begin{aligned} \widehat{\Lambda}_0(\mathbf{Z}_L) &= \log \frac{\max_{\sigma_\alpha^2} f_{1,0}(\mathbf{Z}_L; \widehat{\mathbf{M}}, \sigma_\alpha^2, \mathbf{v}_t)}{f_0(\mathbf{Z}_L; \widehat{\mathbf{M}})} \\ &= \max_{\sigma_\alpha^2} \left\{ -L \log \det(\widehat{\mathbf{M}} + \sigma_\alpha^2 \mathbf{v}_t \mathbf{v}_t^\dagger) - \text{Tr} \left[\left(\widehat{\mathbf{M}} + \sigma_\alpha^2 \mathbf{v}_t \mathbf{v}_t^\dagger \right)^{-1} \right. \right. \\ &\quad \left. \left. \times \mathbf{Z}_L \mathbf{Z}_L^\dagger \right] \right\} + L \log \det(\widehat{\mathbf{M}}) + \text{Tr} \left[\widehat{\mathbf{M}}^{-1} \mathbf{Z}_L \mathbf{Z}_L^\dagger \right]. \end{aligned} \quad (13)$$

Applying the Woodbury identity [43], we come up with the following expression

$$\widehat{\Lambda}_0(\mathbf{Z}_L) = \max_{\sigma_\alpha^2} \left[-L \log(1 + \sigma_\alpha^2 \|\mathbf{v}_w\|^2) + \frac{\sigma_\alpha^2 \mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w}{1 + \sigma_\alpha^2 \|\mathbf{v}_w\|^2} \right], \quad (14)$$

where $\mathbf{v}_w = \widehat{\mathbf{M}}^{-1/2} \mathbf{v}_t$ and $\mathbf{S}_w = \widehat{\mathbf{M}}^{-1/2} \mathbf{Z}_L \mathbf{Z}_L^\dagger \widehat{\mathbf{M}}^{-1/2}$. Let us define the function

$$g_0(\sigma_\alpha^2) = -L \log(1 + \sigma_\alpha^2 \|\mathbf{v}_w\|^2) + \frac{\sigma_\alpha^2 \mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w}{1 + \sigma_\alpha^2 \|\mathbf{v}_w\|^2} \quad (15)$$

and observe that $\lim_{\sigma_\alpha^2 \rightarrow 0} g_0(\sigma_\alpha^2) = 0$ and $\lim_{\sigma_\alpha^2 \rightarrow +\infty} g_0(\sigma_\alpha^2) = -\infty$. Thus, in order to maximize $g_0(\sigma_\alpha^2)$, we find the zeros of its first derivative with respect to σ_α^2 to obtain

$$\widehat{\sigma}_\alpha^2 = \begin{cases} \frac{\mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w - L \|\mathbf{v}_w\|^2}{L \|\mathbf{v}_w\|^4}, & \text{if } \mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w > L \|\mathbf{v}_w\|^2, \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

Moreover, $g_0(\sigma_\alpha^2)$ is monotonic increasing when $\sigma_\alpha^2 < (\mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w - L \|\mathbf{v}_w\|^2) / (L \|\mathbf{v}_w\|^4)$ and decreasing for $\sigma_\alpha^2 > (\mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w - L \|\mathbf{v}_w\|^2) / (L \|\mathbf{v}_w\|^4)$ (provided that $\mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w > L \|\mathbf{v}_w\|^2$). As a consequence, (16) represents a maximum point of $g_0(\sigma_\alpha^2)$. It follows that (14) can be recast as

$$\begin{aligned} \widehat{\Lambda}_0(\mathbf{Z}_L) &= \\ &\begin{cases} -L \log \left(\frac{\mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w}{L \|\mathbf{v}_w\|^2} \right) + \frac{\mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w}{\|\mathbf{v}_w\|^2} - L, & \text{if } \mathbf{v}_w^\dagger \mathbf{S}_w \mathbf{v}_w > L \|\mathbf{v}_w\|^2, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (17)$$

The next step towards the computation of (7) consists in deriving $\widehat{\Lambda}_i(\mathbf{Z}_L)$ for $i = 1, \dots, M_c$. To this end, observe that

$$\begin{aligned} \widehat{\Lambda}_i(\mathbf{Z}_L) &= \log \frac{\max_{\mathbf{R}_i} \max_{\theta_1, \dots, \theta_i} f_{1,i}(\mathbf{Z}_L; \widehat{\mathbf{M}}, \mathbf{R}_i, \mathbf{V}_i)}{f_0(\mathbf{Z}_L; \widehat{\mathbf{M}})} \\ &= \max_{\theta_1, \dots, \theta_i} \left\{ -L \log \det(\widehat{\mathbf{M}} + \mathbf{V}_i \mathbf{R}_i \mathbf{V}_i^\dagger) - \text{Tr} \left[\left(\widehat{\mathbf{M}} + \mathbf{V}_i \mathbf{R}_i \mathbf{V}_i^\dagger \right)^{-1} \right. \right. \\ &\quad \left. \left. \times \mathbf{Z}_L \mathbf{Z}_L^\dagger \right] \right\} + L \log \det(\widehat{\mathbf{M}}) + \text{Tr} \left(\widehat{\mathbf{M}}^{-1} \mathbf{Z}_L \mathbf{Z}_L^\dagger \right). \end{aligned} \quad (18)$$

Replacing (12) into the most right-hand side of the above equation and applying the Woodbury identity leads to

$$\begin{aligned} \widehat{\Lambda}_i(\mathbf{Z}_L) &= \max_{\mathbf{r}_i} \max_{\theta_1, \dots, \theta_i} \left\{ -L \log(1 + \mathbf{r}_i^\dagger \mathbf{V}_{w,i}^\dagger \mathbf{V}_{w,i} \mathbf{r}_i) \right. \\ &\quad \left. + \frac{\mathbf{r}_i^\dagger \mathbf{V}_{w,i}^\dagger \mathbf{S}_w \mathbf{V}_{w,i} \mathbf{r}_i}{1 + \mathbf{r}_i^\dagger \mathbf{V}_{w,i}^\dagger \mathbf{V}_{w,i} \mathbf{r}_i} \right\}, \end{aligned} \quad (19)$$

where $\mathbf{V}_{w,i} = \widehat{\mathbf{M}}^{-1/2} \mathbf{V}_i$ and \mathbf{S}_w has been previously defined. Since \mathbf{V}_i has a Vandermonde structure and $\widehat{\mathbf{M}}$ is nonsingular, also matrices $\mathbf{A}_i = \mathbf{V}_{w,i}^\dagger \mathbf{V}_{w,i}$ and $\mathbf{B}_i = \mathbf{V}_{w,i}^\dagger \mathbf{S}_w \mathbf{V}_{w,i}$ are nonsingular and (19) can be expressed in terms of these matrices as

$$\widehat{\Lambda}_i(\mathbf{Z}_L) = \max_{\mathbf{r}_i} \max_{\theta_1, \dots, \theta_i} \left\{ -L \log(1 + \mathbf{r}_i^\dagger \mathbf{A}_i \mathbf{r}_i) + \frac{\mathbf{r}_i^\dagger \mathbf{B}_i \mathbf{r}_i}{1 + \mathbf{r}_i^\dagger \mathbf{A}_i \mathbf{r}_i} \right\}. \quad (20)$$

Exploiting the Cholesky decomposition of \mathbf{A}_i , given by

$$\mathbf{A}_i = \mathbf{C}_i \mathbf{C}_i^\dagger, \quad (21)$$

where $\mathbf{C}_i \in \mathbb{C}^{(i+1) \times (i+1)}$ is a lower triangular matrix, the right-hand side of (20) can be recast as

$$\begin{aligned} & \max_{\mathbf{r}_i, \theta_1, \dots, \theta_i} \left\{ -L \log(1 + \mathbf{r}_i^\dagger \mathbf{C}_i \mathbf{C}_i^\dagger \mathbf{r}_i) + \frac{\mathbf{r}_i^\dagger \mathbf{C}_i \mathbf{C}_i^{-1} \mathbf{B}_i (\mathbf{C}_i^\dagger)^{-1} \mathbf{C}_i^\dagger \mathbf{r}_i}{1 + \mathbf{r}_i^\dagger \mathbf{C}_i \mathbf{C}_i^\dagger \mathbf{r}_i} \right\} \\ & = \max_{\mathbf{a}_i} \max_{\theta_1, \dots, \theta_i} \left\{ -L \log(1 + \|\mathbf{a}_i\|^2) + \frac{\mathbf{a}_i^\dagger \mathbf{D}_i \mathbf{a}_i}{1 + \|\mathbf{a}_i\|^2} \right\}, \end{aligned} \quad (22)$$

where

$$\mathbf{a}_i = \mathbf{C}_i^\dagger \mathbf{r}_i, \quad \mathbf{D}_i = \mathbf{C}_i^{-1} \mathbf{B}_i (\mathbf{C}_i^\dagger)^{-1}. \quad (23)$$

As further step towards the solution of the above problem, we decompose \mathbf{a}_i as the product of a positive scalar $\sqrt{a_i}$, with $a_i > 0$, times a unit-norm vector $\mathbf{u}_{a,i}$. Thus, problem (22) is tantamount to

$$\max_{a_i > 0} \max_{\substack{\mathbf{u}_{a,i} \\ \|\mathbf{u}_{a,i}\|=1}} \max_{\theta_1, \dots, \theta_i} \left\{ -L \log(1 + a_i) + \frac{a_i \mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i}}{1 + a_i} \right\}. \quad (24)$$

Denoting by

$$g(a_i) = -L \log(1 + a_i) + \frac{a_i}{1 + a_i} \mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i}, \quad (25)$$

the maximization with respect to a_i can be performed by investigating the behavior of $g(a_i)$ over the interval $(0, +\infty)$. Specifically, since the following equalities hold

$$\lim_{a_i \rightarrow 0} g(a_i) = 0, \quad \lim_{a_i \rightarrow +\infty} g(a_i) = -\infty, \quad (26)$$

we can search for the stationary points of $g(a_i)$ in the interior of its domain by setting to zero its first derivative with respect to a_i to obtain

$$\widehat{a}_i = \begin{cases} \frac{\mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i} - L}{L}, & \text{if } \mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i} > L, \\ 0, & \text{otherwise.} \end{cases} \quad (27)$$

$$\widehat{\Lambda}_i(\mathbf{Z}_L) = \begin{cases} \max_{\substack{\mathbf{u}_{a,i} \\ \|\mathbf{u}_{a,i}\|=1}} \max_{\theta_1, \dots, \theta_i} -L \log(\mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i}) + \mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i} + L \log L - L, & \text{if } \mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i} > L, \\ 0, & \text{otherwise.} \end{cases} \quad (29)$$

$$\widehat{\Lambda}_i(\mathbf{Z}_L) = \begin{cases} \max_{\theta_1, \dots, \theta_i} -L \log \lambda_{\max}\{\mathbf{D}_i\} + \lambda_{\max}\{\mathbf{D}_i\} + L \log L - L, & \text{if } \lambda_{\max}\{\mathbf{D}_i\} > L, \\ 0, & \text{otherwise.} \end{cases} \quad (30)$$

Moreover, it is not difficult to show that $g(a_i)$ is increasing when $a_i < (\mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i} - L)/L$ and decreasing in the opposite case (provided that $\mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i} > L$). As a consequence, (27) represents a maximum point. Now, plugging (27) into (24), the maximization problem can be recast as (29) at the bottom of this page.

The maximization over $\mathbf{u}_{a,i}$ can be conducted by defining the function $f(x) = x - L \log x + L \log L - L$, which is monotonic increasing for $x \geq L$. In fact, its first derivative is $\frac{df(x)}{dx} = 1 - \frac{L}{x} \geq 0, \forall x \geq L$. Therefore, the maximum of $f(\mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i})$ with respect to $\mathbf{u}_{a,i}$ can be obtained by directly optimizing argument, i.e., $\mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i}$. Exploiting the Rayleigh-Ritz Theorem [43], we obtain that

$$\max_{\substack{\mathbf{u}_{a,i} \\ \|\mathbf{u}_{a,i}\|=1}} \mathbf{u}_{a,i}^\dagger \mathbf{D}_i \mathbf{u}_{a,i} = \lambda_{\max}\{\mathbf{D}_i\}, \quad (28)$$

and the maximum is attained when $\mathbf{u}_{a,i}$ is the normalized eigenvector of \mathbf{D}_i corresponding to its maximum eigenvalue. Gathering the above results, we obtain (30) at the bottom of this page.

The last optimization problem to be solved is with respect to the AoAs of the coherent signals. In this respect, since deriving closed-form expressions for the ML estimates of $\theta_1, \dots, \theta_i$ represents a difficult task (at least to the best of authors' knowledge), we devise grid-search methods. More precisely, the angular sector under surveillance is sampled to form a discrete set of angular positions which is denoted by $\Omega = \{\omega_1, \dots, \omega_S\}$ (see Figure 1). The cardinality of Ω can be chosen according to both the computational power available at the receiver and the system requirements in terms of reactivity time intervals. It is clear that a dense sampling of the angular sector increases the processing time but can provide better estimation results with respect to the case where the sampling interval is large. However, in the latter case, the computational load is lower than in the former case. Nevertheless, an iterative approach can be pursued for the maximization under the generic $H_{1,i}$ hypothesis and for each $i \in \{1, \dots, M_c\}$. Specifically, given $i \in \{1, \dots, M_c\}$ and assuming that $H_{1,i}$ holds true, we start by roughly sampling the angular sector of interest and use this search grid to come up with preliminary estimates of the assumed i AoAs. Then, we use these estimates to identify a restricted angular sector (resulting from the union of partially overlapped narrow sectors) comprising suitable guard bands.

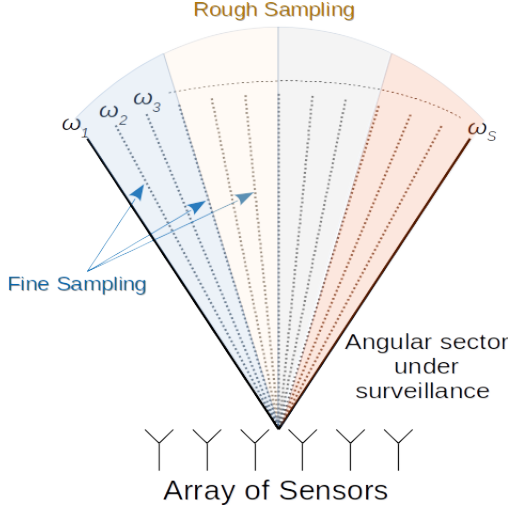


Figure 1: Sampling of the angular sector under surveillance.

The new angular sector of interest is sampled through a shorter sampling interval and the maximization under $H_{1,i}$ is repeated over the new search grid. As a result, the grid points can be very close to the actual angular positions (at least for high signal power values) improving the estimation quality. Finally, notice that this procedure is applied for each $i = \{1, \dots, M_c\}$ before forming the decision statistic and, in a single application, the number of supposed interferers is maintained constant. As for the single search procedure under the generic $H_{1,i}$ hypothesis with $i \geq 1$, observe that in situations where, according to the system geometry and antenna beamwidth, M_c is enough low, an exhaustive grid-based search can be conducted over the set

$$\underbrace{\Omega \times \Omega \times \dots \times \Omega}_i = \Omega^i, \quad \forall i \in \{1, \dots, M_c\}, \quad (31)$$

with the constraint that admissible candidates are such that $\omega_1 \neq \omega_2 \neq \dots \neq \omega_i$. On the other hand, exhaustive grid search can be prohibitive for large M_c values. In this case, in order to limit the number of operations, we design a suboptimum search method relying on the cyclic optimization paradigm [44]. Specifically, let us assume that $H_{1,i}$ is true and that $i-1$ estimates, $\hat{\Theta}_{1,i}^{(0)} = \{\hat{\theta}_2^{(0)}, \dots, \hat{\theta}_i^{(0)}\} \subset \Omega$ say, are available, then we select $\theta_1 \in \Omega$ exploiting the following criterion

$$\hat{\theta}_1^{(1)} = \arg \max_{\theta_1 \in \Omega, \theta_1 \notin \hat{\Theta}_{1,i}^{(0)}} \left\{ -L \log \lambda_{\max} \{ \mathbf{D}_i(\theta_1, \hat{\Theta}_{1,i}^{(0)}) \} + \lambda_{\max} \{ \mathbf{D}_i(\theta_1, \hat{\Theta}_{1,i}^{(0)}) \} \right\}, \quad (32)$$

where $\mathbf{D}_i(\theta_1, \hat{\Theta}_{1,i}^{(0)})$ is computed as \mathbf{D}_i in (23) with the difference that \mathbf{V}_i is replaced by the following matrix

$$\mathbf{V}_i(\theta_1, \hat{\Theta}_{1,i}^{(0)}) = [\mathbf{v}_t, \mathbf{v}(\theta_1), \mathbf{v}(\hat{\theta}_2^{(0)}), \dots, \mathbf{v}(\hat{\theta}_i^{(0)})]. \quad (33)$$

The new estimate $\hat{\theta}_1^{(1)}$ is used to form $\hat{\Theta}_{2,i}^{(0)} =$

$\{\hat{\theta}_1^{(1)}, \hat{\theta}_3^{(0)}, \dots, \hat{\theta}_i^{(0)}\}$ and the update of $\hat{\theta}_2$ is obtained as

$$\hat{\theta}_2^{(1)} = \arg \max_{\theta_2 \in \Omega, \theta_2 \notin \hat{\Theta}_{2,i}^{(0)}} \left\{ -L \log \lambda_{\max} \{ \mathbf{D}_i(\theta_2, \hat{\Theta}_{2,i}^{(0)}) \} + \lambda_{\max} \{ \mathbf{D}_i(\theta_2, \hat{\Theta}_{2,i}^{(0)}) \} \right\}, \quad (34)$$

where $\mathbf{D}_i(\theta_2, \hat{\Theta}_{2,i}^{(0)})$ is built up using

$$\mathbf{V}_i(\theta_2, \hat{\Theta}_{2,i}^{(0)}) = [\mathbf{v}_t, \mathbf{v}(\hat{\theta}_1^{(1)}), \mathbf{v}(\theta_2), \mathbf{v}(\hat{\theta}_3^{(0)}), \dots, \mathbf{v}(\hat{\theta}_i^{(0)})]. \quad (35)$$

The above steps continue until the i th update occurs to obtain $\hat{\Theta}_{1,i}^{(1)} = \{\hat{\theta}_2^{(1)}, \dots, \hat{\theta}_i^{(1)}\}$, which can be used to repeat the entire refinement procedure.

Summarizing, the update of the k th AoA within the n th ($n \geq 1$) procedure cycle has the following expression

$$\hat{\theta}_k^{(n)} = \arg \max_{\theta_k \in \Omega, \theta_k \notin \hat{\Theta}_{k,i}^{(n-1)}} \left\{ -L \log \lambda_{\max} \{ \mathbf{D}_i(\theta_k, \hat{\Theta}_{k,i}^{(n-1)}) \} + \lambda_{\max} \{ \mathbf{D}_i(\theta_k, \hat{\Theta}_{k,i}^{(n-1)}) \} \right\}, \quad (36)$$

$1 \leq k \leq i$, where

$$\hat{\Theta}_{k,i}^{(n-1)} = \{\hat{\theta}_1^{(n)}, \dots, \hat{\theta}_{k-1}^{(n)}, \hat{\theta}_{k+1}^{(n-1)}, \dots, \hat{\theta}_i^{(n-1)}\} \quad (37)$$

and $\mathbf{D}_i(\theta_k, \hat{\Theta}_{k,i}^{(n-1)})$ is defined as in (23) with $\mathbf{V}_i(\theta_k, \hat{\Theta}_{k,i}^{(n-1)}) = [\mathbf{v}_t, \mathbf{v}(\hat{\theta}_1^{(n)}), \dots, \mathbf{v}(\hat{\theta}_{k-1}^{(n)}), \mathbf{v}(\theta_k), \mathbf{v}(\hat{\theta}_{k+1}^{(n-1)}), \dots, \mathbf{v}(\hat{\theta}_i^{(n-1)})]$. It is important to observe that the above procedure leads to a nondecreasing sequence of log-likelihood function values, namely

$$\begin{aligned} \mathcal{L}_i(\hat{\theta}_1^{(0)}, \hat{\Theta}_{1,i}^{(0)}) &\leq \mathcal{L}_i(\hat{\theta}_1^{(1)}, \hat{\Theta}_{1,i}^{(0)}) = \mathcal{L}_i(\hat{\theta}_2^{(0)}, \hat{\Theta}_{2,i}^{(0)}) \leq \\ \mathcal{L}_i(\hat{\theta}_2^{(1)}, \hat{\Theta}_{2,i}^{(0)}) &\leq \dots \leq \mathcal{L}_i(\hat{\theta}_k^{(n)}, \hat{\Theta}_{k,i}^{(n-1)}) \leq \dots \end{aligned} \quad (38)$$

where

$$\begin{aligned} \mathcal{L}_i(\hat{\theta}_k^{(n)}, \hat{\Theta}_{k,i}^{(m)}) &= -L \log \lambda_{\max} \{ \mathbf{D}_i(\hat{\theta}_k^{(n)}, \hat{\Theta}_{k,i}^{(m)}) \} \\ &+ \lambda_{\max} \{ \mathbf{D}_i(\hat{\theta}_k^{(n)}, \hat{\Theta}_{k,i}^{(m)}) \} + L \log L - L. \end{aligned} \quad (39)$$

The entire procedure may terminate when

$$\Delta \mathcal{L}_i(n) = \frac{|\mathcal{L}_i(\hat{\theta}_1^{(n+1)}, \hat{\Theta}_{1,i}^{(n+1)}) - \mathcal{L}_i(\hat{\theta}_1^{(n)}, \hat{\Theta}_{1,i}^{(n)})|}{|\mathcal{L}_i(\hat{\theta}_1^{(n+1)}, \hat{\Theta}_{1,i}^{(n+1)})|} < \epsilon, \quad (40)$$

where $\epsilon > 0$, or $n \geq N_{\max}$ where N_{\max} is the maximum allowable number of iterations.

Finally, as for the initialization of the procedure, given the received vectors, we proceed by first computing

$$r_1 = \left| \frac{1}{L} \sum_{l=1}^L \mathbf{z}_l^\dagger \mathbf{v}(\omega_1) \right|^2, \dots, r_S = \left| \frac{1}{L} \sum_{l=1}^L \mathbf{z}_l^\dagger \mathbf{v}(\omega_S) \right|^2, \quad (41)$$

then, sort the above values in decreasing order, namely

$$r_{k_1} \geq r_{k_2} \geq \dots \geq r_{k_S}, \quad (42)$$

and select $\omega_{k_1}, \dots, \omega_{k_i}$. The behavior of the proposed architectures coupled with the above estimation procedure is assessed in the next section by means of numerical examples.

IV. ILLUSTRATIVE EXAMPLES AND DISCUSSION

In this section, the performance of the proposed detection architectures is investigated drawing upon synthetic data and considering three operating scenarios. In the first case, only the target of interest is present, whereas in the second scenario, an additional coherent signal is considered. Finally, the third scenario contains the signal of interest along with two additional coherent signals. The performance metrics are

- the Probability of Detection under $H_{1,i}$, $i \geq 0$, ($P_{d,i}$) defined as the probability of rejecting H_0 when the latter is false and given a preassigned $P_{fa} = P(\text{reject } H_0 | H_0 \text{ is true})$;
- the Probability of Correct Classification (P_{cc}), namely the probability of deciding⁷ for $H_{1,i}$ under $H_{1,i}$;
- the Root Mean Square Error (RMSE) in angle whose expression is

$$\sqrt{\mathbb{E} \left[\frac{1}{M} \sum_{m=1}^M \min_{k=1, \dots, \hat{i}} \{(\theta_m - \hat{\theta}_k)^2\} \right]}, \quad (43)$$

where $\hat{\theta}_k$, $k = 1, \dots, \hat{i}$, are the AoA estimates of the coherent signals when $H_{1,\hat{i}}$ is declared. Since deriving closed-form expressions for the above quantities represents a mathematically intractable task (at least to the best of authors' knowledge), we estimate them resorting to the Monte Carlo counting techniques (also replacing the statistical expectation of the RMSE with the sample mean over the Monte Carlo trials). More precisely, we exploit 10^3 independent trials to estimate the considered performance metrics, whereas the detection thresholds are set over $100/P_{fa}$, with $P_{fa} = 10^{-3}$.

As stated before, the analysis starts from the conventional case where only the target echoes impinge on the radar and proceeds with two more difficult cases that assume the presence of two and three coherent signals (including that of interest), respectively. As for the operating setup, the coherent signals share the same power as the target signal, whereas the angular sector under surveillance ranges from -25 to 25 degrees and is sampled at 1 degree. The Signal-to-Interference-plus-Noise Ratio (SINR) is defined as $\text{SINR} = \sigma_\alpha^2 \mathbf{v}_t^\dagger \mathbf{M}^{-1} \mathbf{v}_t$, where $\mathbf{M} = \sigma_n^2 \mathbf{I} + \text{CNR} \mathbf{M}_c$, $\text{CNR} = 20$ dB is the Clutter-to-Noise Ratio, $\sigma_n^2 = 1$ is the noise power, and \mathbf{M}_c is the clutter covariance matrix whose (i, j) th entry is defined as $\mathbf{M}_c(i, j) = \rho_c^{|i-j|}$ with $\rho_c = 0.9$ the one-lag correlation coefficient. Moreover, all the numerical examples assume $N = 16$, $L = 32$, $\theta_t = 0^\circ$, and $M_c = 4$.

As for the angular positions of the coherent signals, we begin the analysis assuming that they belong to the search grid. In this case, the obtained curves represent an upper bound on the performance that can be attained exploiting fine and fine search grids (possibly considering the previously described iterative maximization procedure). Then, we investigate the behavior of the proposed architectures when the actual positions of the coherent signals are in between the points of the search grid. Specifically, we generate them as uniformly

distributed in intervals of different sizes and centered on the nominal search grid points (a point better explained below).

Finally, we compare the proposed architectures with well-known decision schemes, namely the Generalized Adaptive Matched Filter (GAMF) and the Generalized Adaptive Subspace Detector (GASD) given by [41]

$$\sum_{l=1}^L \frac{|\mathbf{v}_t^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{z}_l|^2}{\mathbf{v}_t^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{v}_t} \underset{H_0}{\overset{H_{1,0}}{>}} \eta \quad \text{and} \quad (44)$$

$$\sum_{l=1}^L \frac{|\mathbf{v}_t^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{z}_l|^2}{\mathbf{v}_t^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{v}_t \sum_{m=1}^L \mathbf{z}_m^\dagger \widehat{\mathbf{M}}^{-1} \mathbf{z}_m} \underset{H_0}{\overset{H_{1,0}}{>}} \eta, \quad (45)$$

respectively.

Before presenting the detection and classification results for each of the aforementioned cases, it is important to assess the convergence rate of the estimation procedure under each hypothesis as well as the sensitivity of the P_{fa} with respect to ρ_c and CNR. To this end, in Figure 2, we plot the root mean square values for $\Delta \mathcal{L}_i(n)$, $i = 2, 3, 4$, versus n under H_0 , $H_{1,0}$, $H_{1,1}$, and $H_{1,2}$; as for $H_{1,1}$ and $H_{1,2}$, we set $\theta_1 = 10^\circ$ and $\theta_2 = 18^\circ$. Inspection of the figure highlights that a number of iterations $n = 5$ is enough to guarantee $\epsilon \leq 10^{-3}$ under each hypothesis. As for the P_{fa} behavior, in Figure 3, we estimate it for different values of ρ_c given CNR (subplot (a)) and for different values of CNR given ρ_c (subplot (b)) when the thresholds are computed assuming the nominal values for these parameters (namely, $\rho_c = 0.9$ and $\text{CNR} = 20$ dB) and $P_{fa} = 10^{-3}$. It turns out that all the considered architectures can guarantee P_{fa} values contained within the interval $[0.0007, 0.0011]$ providing a rather robust behavior to the considered parameter variations.

In the next three figures, we investigate the detection and classification performance when the conventional alternative hypothesis $H_{1,0}$, which contemplates the presence of the signal of interest only (first scenario), is in force. Specifically, in Figure 4, we show the $P_{d,0}$ curves as functions of the SINR assuming $N = 16$, $L = 32$, $K = 32$, and different values for ρ . This preliminary analysis allows us to quantify the sensitivity of GIC-D with respect to its tuning parameter in comparison with the other considered architectures. It turns out that increasing ρ leads to improved detection performances for GIC-D, whose loss with respect to both the GAMF and GASD, which overcome the other architectures, ranges from about 1.9 dB for $\rho = 20$ to about 0.5 dB for $\rho = 80$ at $P_{d,0} = 0.9$. On the other hand, AIC-D and BIC-D experience a loss of about 2.2 dB with respect to GAMF and GASD at $P_{d,0} = 0.9$. In Figure 5, we analyze the effects of K on the performance by doubling the value used in Figure 4. Moreover, in order to quantify the loss due to the estimation of \mathbf{M} we also report the curves of the clairvoyant detectors which assume that \mathbf{M} is known. The most evident change with respect to the previous figure is that all the detection curves clearly move towards the left part of the plot, namely a significant improvement in performance occurs. In addition, GIC-D with $\rho = 80$, GAMF, and GASD almost share the same performance while AIC-D and BIC-D continue to exhibit a loss of about 2 dB at $P_{d,0} = 0.9$

⁷The detection threshold is set according to the preassigned P_{fa} also in this case.

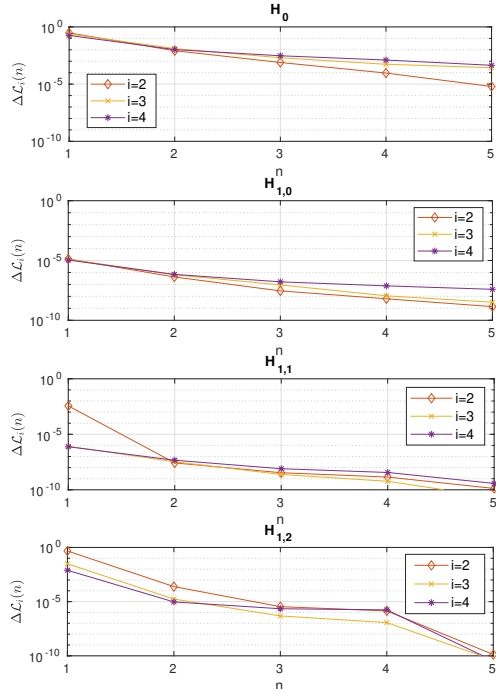


Figure 2: Convergence curves under H_0 , $H_{1,0}$, $H_{1,1}$, and $H_{1,2}$ assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, and $\theta_2 = 18^\circ$.

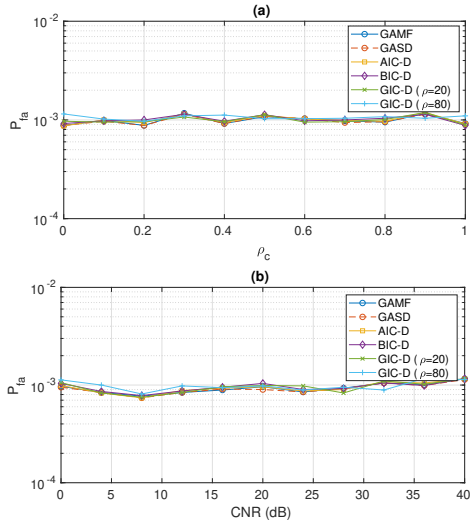


Figure 3: P_{fa} versus ρ_c (subplot (a)) and CNR (subplot (b)) for the GAMF, GASD, AIC-D, BIC-D, and GIC-D with different values of ρ assuming $N = 16$, $L = 32$, and $K = 32$.

with respect to the former. The figure also highlights that the loss associated to the estimation of M is of about 3.5 dB (at $P_{d,0} = 0.9$). In the last figure of this first study case (Figure 6), we analyze the classification capabilities of the proposed architectures through the probabilities of classification (P_c) and the P_{cc} estimated at two different SINR values. The figure clearly highlights the inclination of AIC-D to overestimate the hypothesis model, whereas the penalty term of BIC-D allows to mitigate this effect even though the P_{cc} is less than 0.4. On the other hand, GIC-D can be suitably tuned in order

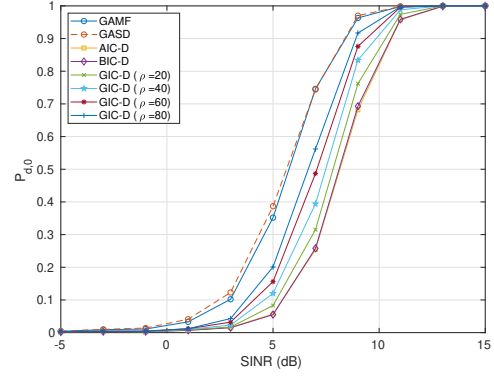


Figure 4: $P_{d,0}$ versus SINR for the GAMF, GASD, AIC-D, BIC-D, and GIC-D with different values of ρ assuming $N = 16$, $L = 32$, and $K = 32$.

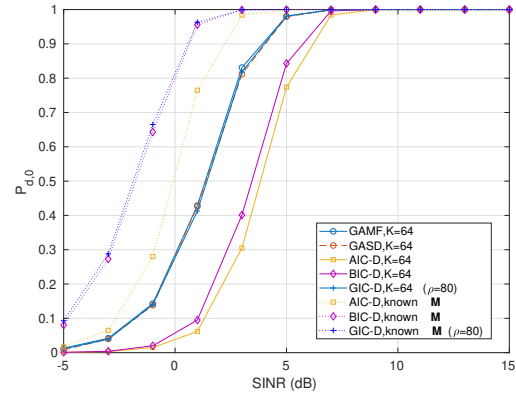


Figure 5: $P_{d,0}$ versus SINR for the GAMF, GASD, AIC-D, BIC-D, GIC-D with $\rho = 80$, AIC-D with known M , BIC-D with known M , and GIC-D with known M and $\rho = 80$ assuming $N = 16$, $L = 32$, and $K = 64$.

to attain satisfactory classification performance. Specifically, when $\rho = 80$, GIC-D correctly classifies the environment with a probability very close to 1.

Now, we address the intermediate scenario, where besides the signal of interest, a coherent signal enters the antenna from an angular direction θ_1 . The hypothesis corresponding to this scenario is $H_{1,1}$. In Figure 7, we plot $P_{d,1}$ versus SINR for $N = 16$, $L = 32$, $K = 32$, and $\theta_1 = 10^\circ$; moreover, several values of ρ are considered. Notice that in the presence of the additional coherent signal, GAMF and GASD are no longer at the top of the performance ranking with the GASD not capable of providing a $P_{d,1}$ above 0.1 due to its selective behavior [12], at least for the considered parameter values. On the other hand, the remaining detectors share about the same performance with a gain of about 10 dB with respect to the GAMF; observe also that GIC-D slightly improves its performance as ρ grows. In the next figure, we assess the behavior of the considered decision schemes when θ_1 takes on different values. Specifically, we set $\theta_1 = 15^\circ, 20^\circ$, whereas the remaining parameters are the same as in the previous figure except for $\rho = 80$. The figure shows that the performance of

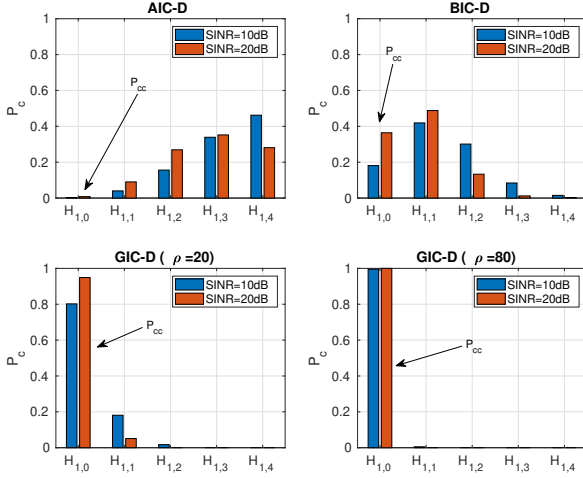


Figure 6: Classification probabilities for the AIC-D, BIC-D, GIC-D with $\rho = 20$, and GIC-D with $\rho = 80$ under $H_{1,0}$ assuming $N = 16$, $L = 32$, and $K = 32$.

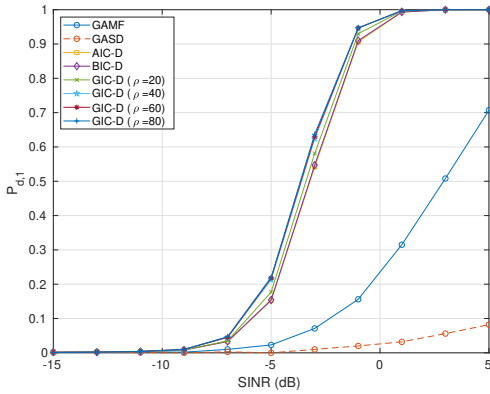


Figure 7: $P_{d,1}$ versus SINR for the GAMF, GASD, AIC-D, BIC-D, and GIC-D with different values of ρ assuming $N = 16$, $L = 32$, $K = 32$, and $\theta_1 = 10^\circ$.

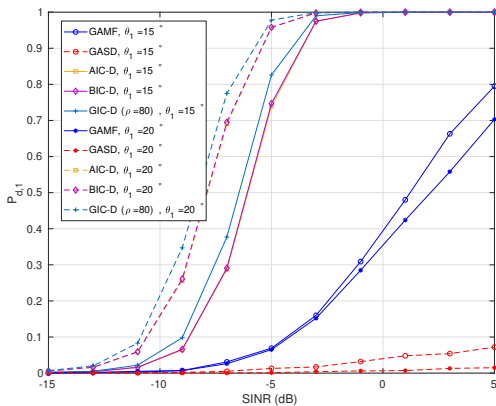


Figure 8: $P_{d,1}$ versus SINR for the GAMF, GASD, AIC-D, BIC-D, and GIC-D assuming different values of θ_1 , $N = 16$, $L = 32$, and $K = 32$.

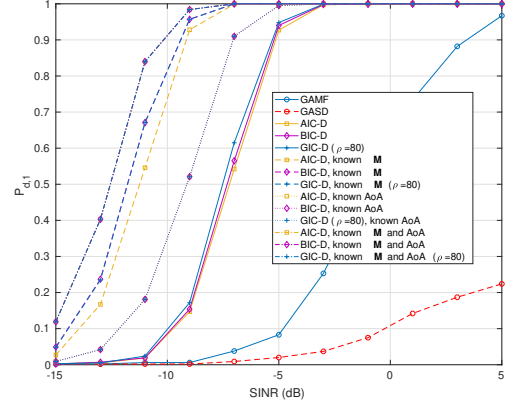


Figure 9: $P_{d,1}$ versus SINR for the GAMF, GASD, AIC-D, AIC-D with known parameters, AIC-D with known M only, AIC-D with known AoA only, BIC-D, BIC-D with known parameters, BIC-D with known M only, BIC-D with known AoA only, GIC-D with $\rho = 80$, GIC-D with $\rho = 80$ and known parameters, GIC-D with $\rho = 80$ and known M only, and GIC-D with $\rho = 80$ and known AoA only, assuming $N = 16$, $L = 32$, $K = 64$, and $\theta_1 = 10^\circ$.

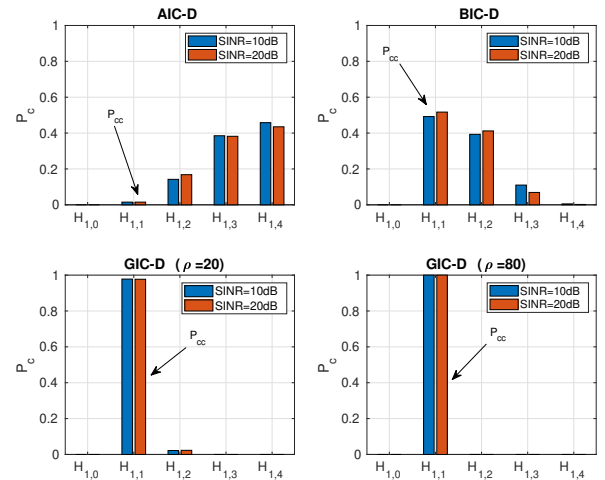


Figure 10: Classification probabilities for the AIC-D, BIC-D, and GIC-D with $\rho = 20, 80$ under $H_{1,1}$ assuming $N = 16$, $L = 32$, $K = 32$, and $\theta_1 = 10^\circ$.

the proposed architectures exhibits a gain of about 2 dB (at $P_{d,1} = 0.9$) when the angular separation between the coherent signals increases with the GIC-D slightly outperforming the other schemes, whereas the performance of the GAMF and GASD degrades due to the fact that the composition of the coherent signals results in a direction that moves away from the nominal steering angle. In Figure 9, we account for a different value of K leaving unaltered the parameters N , L and also quantify the loss due to the estimation of the AoA and/or M . As expected, increasing K improves the performance of all detectors. Moreover, it turns out that estimating the AoA is not as crucial as the estimation of the covariance matrix since

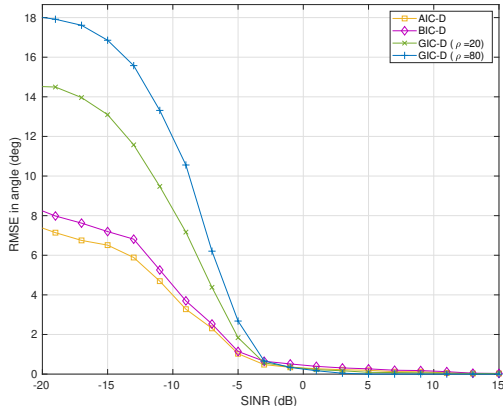


Figure 11: RMSE in angle versus SINR for the AIC-D, BIC-D, and GIC-D with $\rho = 20, 80$ under $H_{1,1}$ assuming $N = 16$, $L = 32$, $K = 32$, and $\theta_1 = 10^\circ$.

the loss associated with the estimation of AoA is of about 1 dB at $P_{d,1} = 0.9$ and with respect to the architectures where all the parameters are known, by contrast that associated with the estimation of M is approximately 3 dB. Again, GIC-D with $\rho = 80$ slightly overcomes the other proposed schemes. In the last two figures related to this intermediate scenario, we show the classification and estimation performance. Specifically, Figure 10 contains the classification histograms, whereas the curves of RMSE in angle versus the SINR are presented in Figure 11. The former figure confirms the behavior observed in Figure 6 and, hence, the superior performance of GIC-D with respect to AIC-D and BIC-D that are inclined to overestimate the number of coherent signals. As for the RMSE, all the proposed architectures share almost the same performance for $\text{SINR} \geq -5$ dB. The main differences occur for low SINR values where, based upon (43), the overestimation of the number of coherent signals for the AIC-D and BIC-D leads to lower RMSE values than GIC-D. As a matter of fact, given (43) and for low SINR values, it is more likely to draw an estimate from the set provided by AIC-D/BIC-D that is in a narrower neighborhood of a true angular position with respect to all the estimates provided by GIC-D. As a consequence, the resulting error for AIC-D/BIC-D takes on lower values than the error for GIC-D.

Finally, the last scenario is the most challenging since it encompasses the presence of two coherent signals with nominal positions $\theta_1 = 10^\circ$ and $\theta_2 = 18^\circ$ in addition to the signal of interest. Besides, we consider two situations that differ in the actual positions of the coherent signals. In the first situation, they are exactly located at θ_1 and θ_2 , whereas in the second case, the positions of the two coherent signals are uniformly generated in the intervals $[\theta_1 - \Delta\theta, \theta_1 + \Delta\theta]$ and $[\theta_2 - \Delta\theta, \theta_2 + \Delta\theta]$, where $\Delta\theta \in \{0.3^\circ, 0.5^\circ\}$. The detection performance for matched signals is shown in Figure 12, where we also plot the architectures that assume partial/full knowledge of the parameter values. Moreover, we compare the results obtained through the exhaustive search for the AoA estimation (left subplot) with those provided by the proposed

suboptimum procedure (right subplot). The figure highlights that the search procedures return detection curves that are very close to each other. In addition, the same remarks for Figure 9 also hold in this case. The classification performance and the RMSE curves for both the exhaustive and suboptimum search procedures are shown in Figure 13-15. Inspection of these figures confirms the excellent classification capabilities of GIC-D, whereas AIC-D and BIC-D can provide better AoA estimates than GIC-D for low SINR values (≤ -10 dB). In addition, it is important to underline that there does not exist a valuable difference in performance between the exhaustive search and the suboptimum search algorithm. Finally, the last four figures assess the behavior of the considered architectures (coupled with the suboptimum search procedure) when the AoAs of the coherent signals are uniformly generated at each Monte Carlo trial in an interval of length $2\Delta\theta$ and centered around the nominal positions. Figures 16 and 17 assume $\Delta\theta = 0.3$, whereas in Figures 18 and 19, we set $\Delta\theta = 0.5$. These numerical examples highlight that, from the detection point of view, the mismatch between the search grid points and the actual positions of the coherent signals leads to a slight performance deterioration, that is more noticeable for $\Delta\theta = 0.5$, while leaving the previously observed hierarchy unaltered. Note that we do not assume any mismatch related to the signal of interest since it requires a different performance analysis that accounts for the mismatch degree and is out of the scope of the present work. As for the classification performance, AIC-D and BIC-D continue to exhibit a clear inclination to overestimate the number of signals, while GIC-D with $\rho = 20$ returns a probability of correct classification lower than that for matched signals. Finally, the classification performance of GIC-D with $\rho = 80$ is not degraded by the coherent signal mismatches.

Summarizing, the above analysis has singled out GIC-D with $\rho = 80$ as the recommended detection architecture capable of providing a satisfactory detection and classification performance in the presence of coherent signals at least for the considered scenarios.

V. CONCLUSIONS

In this paper, we focused on the adaptive radar detection in the presence of fully correlated signals besides that of interest. Such additional signals may be due to multipath propagation effects or to the action of malicious platforms (smart jammers). In order to account for different operating scenarios, at the design stage, we have considered a multiple-hypothesis test that also includes the classical radar signal-plus-interference hypothesis and devised likelihood-ratio-based decision schemes whose statistics under a specific hypothesis depend on a suitable penalty factor tuned according to the number of unknown parameters under that hypothesis (leveraging the approach of the MOS rules). As a result, such architectures are provided with classification capabilities returning, as a byproduct, an estimate of the number of coherent signals impinging on the radar system. The performance analysis has been carried out resorting to simulated data considering three different scenarios with an increasing number of coherent signals. Moreover,

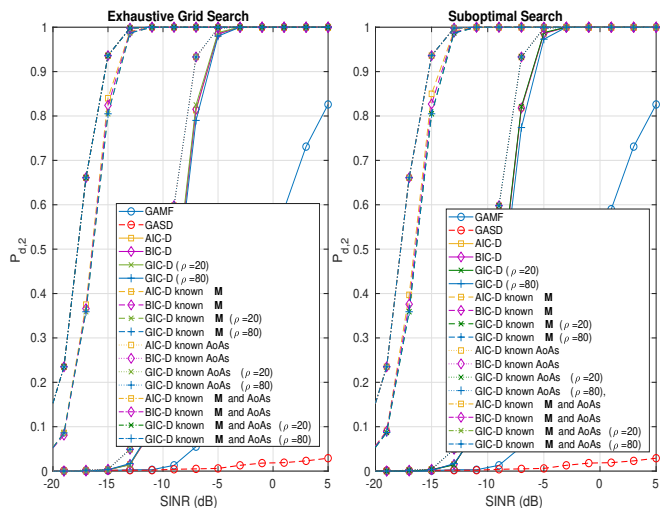


Figure 12: $P_{d,2}$ versus SINR for the GAMF, GASD, AIC-D, AIC-D with known parameters, AIC-D with known M only, AIC-D with known AoA only, BIC-D, BIC-D with known parameters, BIC-D with known M only, BIC-D with known AoA only, GIC-D with $\rho = 20, 80$, GIC-D with $\rho = 20, 80$ and known parameters, GIC-D with $\rho = 20, 80$ and known M only, and GIC-D with $\rho = 20, 80$ and known AoA only, exhaustive grid search (left subplot) and suboptimum search (right subplot) assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, and $\theta_2 = 18^\circ$.

for comparison purposes, the curves for the GAMF and GASD have been also reported. The analysis has singled out the GIC-D with $\rho = 80$ as the recommended detection architecture since it overcomes the remaining proposed decision schemes in terms of both detection and classification performance.

Future research tracks may include the design of (possibly space-time) processing architectures that account for coherent signals spread along the range dimension or aimed at operating in scenarios where multiple coherent and/or uncorrelated signals are present.

REFERENCES

- [1] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar: Basic Principles*. Raleigh, NC: Scitech Publishing, 2010.
- [2] W. L. Melvin and J. A. Scheer, *Principles of Modern Radar: Advanced Techniques*, S. Publishing, Ed., Edison, NJ, 2013.
- [3] M. A. Richards, W. L. Melvin, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar: Radar Applications, Volume 3*, ser. Electromagnetics and Radar. Institution of Engineering and Technology, 2013.
- [4] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 2, pp. 115–127, 1986.
- [5] E. J. Kelly and K. Forsythe, "Adaptive Detection and Parameter Estimation for Multidimensional Signal Models," Lincoln Lab, MIT, Lexington, US, Technical Report 848, 1989.
- [6] W. L. Melvin, "Space-time Adaptive Radar Performance in Heterogeneous Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 621–633, 2000.
- [7] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 208–216, 1992.
- [8] F. Gini and A. Farina, "Vector Subspace Detection in Compound-Gaussian Clutter Part I: Survey and New Results," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 4, pp. 1295–1311, 2002.
- [9] D. Orlando and G. Ricci, "A Rao Test With Enhanced Selectivity Properties in Homogeneous Scenarios," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5385–5390, 2010.
- [10] W. Liu, W. Xie, and Y. Wang, "Rao and Wald Tests for Distributed Targets Detection With Unknown Signal Steering," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1086–1089, 2013.
- [11] Y. I. Abramovich and B. A. Johnson, "GLRT-based detection-estimation for undersampled training conditions," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3600–3612, 2008.
- [12] F. Bandiera, D. Orlando, and G. Ricci, *Advanced Radar Detection Schemes Under Mismatched Signal Models*. San Rafael, US: Synthesis Lectures on Signal Processing No. 8, Morgan & Claypool Publishers, 2009.
- [13] W. Liu, W. Xie, J. Liu, and Y. Wang, "Adaptive Double Subspace Signal Detection in Gaussian Background—Part I: Homogeneous Environments," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2345–2357, 2014.
- [14] W. Liu, J. Liu, Y. Gao, G. Wang, and Y. L. Wang, "Multichannel signal detection in interference and noise when signal mismatch happens,"

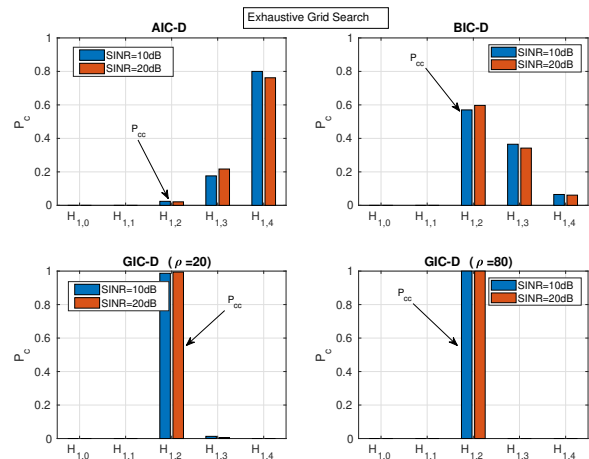


Figure 13: Classification probabilities for the AIC-D, BIC-D, and GIC-D with $\rho = 20, 80$ under $H_{1,2}$ assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, $\theta_2 = 18^\circ$, and the exhaustive grid search.

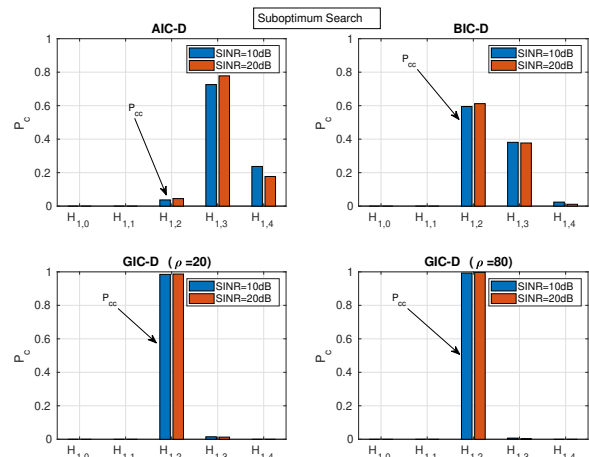


Figure 14: Classification probabilities for the AIC-D, BIC-D, and GIC-D with $\rho = 20, 80$ under $H_{1,2}$ assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, $\theta_2 = 18^\circ$, and the suboptimum search.

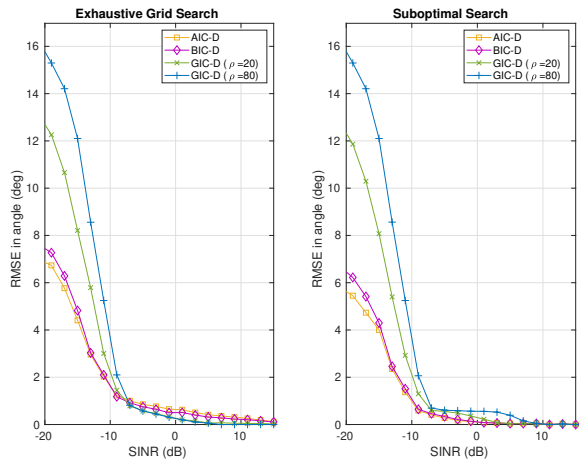


Figure 15: RMSE in angle versus SINR for the AIC-D, BIC-D, and GIC-D with $\rho = 20, 80$ under $H_{1,2}$ assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, $\theta_2 = 18^\circ$, exhaustive search grid (left subplot), and suboptimum search (right subplot).

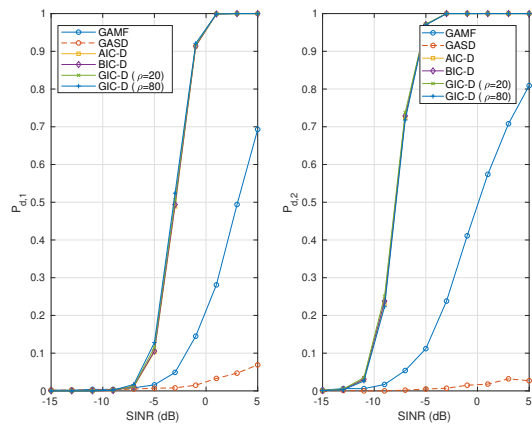


Figure 16: $P_{d,1}$ versus SINR (left subplot) and $P_{d,2}$ versus SINR (right subplot) for the GAMF, GASD, AIC-D, BIC-D, and GIC-D with $\rho = 20, 80$ when the actual AoAs of the coherent signals are uniformly generated in between $\theta_i - 0.3^\circ$ and $\theta_i + 0.3^\circ$, $i = 1, 2$, assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, $\theta_2 = 18^\circ$, and the suboptimum search.

Signal Processing, vol. 166, p. 107268, 2020.

- [15] J. Liu, S. Sun, and W. Liu, "One-step persymmetric GLRT for subspace signals," *IEEE Transaction on Signal Processing*, vol. 14, no. 67, pp. 3639–3648, July 15 2019.
- [16] C. Hao, S. Gazor, G. Foglia, B. Liu, and C. Hou, "Persymmetric adaptive detection and range estimation of a small target," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 4, pp. 2590–2604, 2015.
- [17] G. Foglia, C. Hao, A. Farina, G. Giunta, D. Orlando, and C. Hou, "Adaptive Detection of Point-Like Targets in Partially Homogeneous Clutter With Symmetric Spectrum," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 4, pp. 2110–2119, 2017.
- [18] C. Hao, D. Orlando, G. Foglia, and G. Giunta, "Knowledge-Based Adaptive Detection: Joint Exploitation of Clutter and System Symmetry Properties," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1489–1493, October 2016.
- [19] G. Foglia, C. Hao, G. Giunta, and D. Orlando, "Knowledge-aided adaptive detection in partially homogeneous clutter: Joint exploitation of persymmetry and symmetric spectrum," *Digital Signal Processing*,

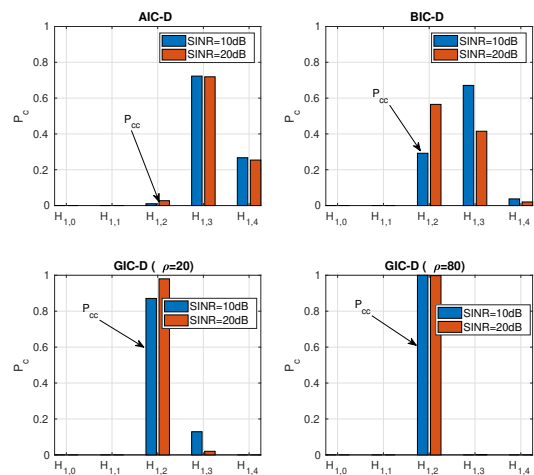


Figure 17: Classification probabilities for AIC-D, BIC-D, GIC-D with $\rho = 20, 80$ under $H_{1,2}$ when the actual AoAs of the coherent signals are uniformly generated in between $\theta_i - 0.3^\circ$ and $\theta_i + 0.3^\circ$, $i = 1, 2$, assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, $\theta_2 = 18^\circ$, and the suboptimum search.

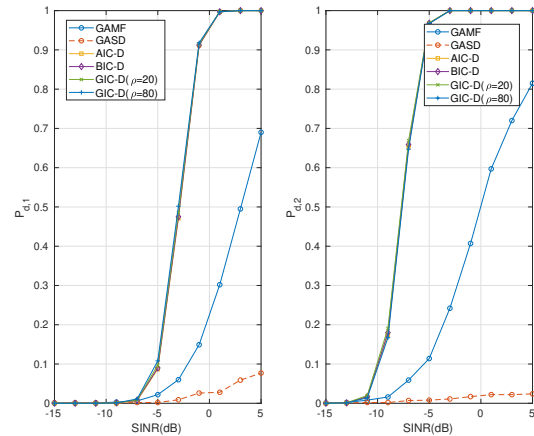


Figure 18: $P_{d,1}$ versus SINR (left subplot) and $P_{d,2}$ versus SINR (right subplot) for the GAMF, GASD, AIC-D, BIC-D, and GIC-D with $\rho = 20, 80$ when the actual AoAs of the coherent signals are uniformly generated in between $\theta_i - 0.5^\circ$ and $\theta_i + 0.5^\circ$, $i = 1, 2$, assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, $\theta_2 = 18^\circ$, and the suboptimum search.

vol. 67, no. Supplement C, pp. 131 – 138, 2017.

- [20] P. Wang, Z. Sahinoglu, M. Pun, and H. Li, "Persymmetric Parametric Adaptive Matched Filter for Multichannel Adaptive Signal Detection," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3322–3328, 2012.
- [21] D. K. Barton, *Radar Equations for Modern Radar*, ser. Artech House radar library. Artech House, 2013.
- [22] D. J. Bachmann, R. J. Evans, and B. Moran, "Game Theoretic Analysis of Adaptive Radar Jamming," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1081–1100, 2011.
- [23] H. Gao, J. Wang, C. Jiang, and X. Zhang, "Equilibrium between a statistical MIMO radar and a jammer," in *2015 IEEE Radar Conference (RadarCon)*, 2015, pp. 0461–0466.
- [24] D. C. Schleher, *Electronic Warfare in the Information Age*, ser. Artech

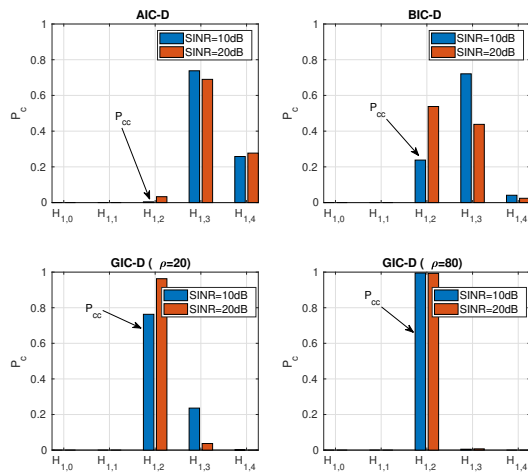


Figure 19: Classification probabilities for AIC-D, BIC-D, GIC-D with $\rho = 20, 80$ under $H_{1,2}$ when the actual AoAs of the coherent signals are uniformly generated in between $\theta_i - 0.5^\circ$ and $\theta_i + 0.5^\circ$, $i = 1, 2$, assuming $N = 16$, $L = 32$, $K = 32$, $\theta_1 = 10^\circ$, $\theta_2 = 18^\circ$, and the suboptimum search.

House radar library. Artech House, 1999.

- [25] F. Neri, *Introduction to Electronic Defense Systems*. Norwood, MA: Artech House, 2018.
- [26] N. Yuen and B. Friedlander, "DOA estimation in multipath: an approach using fourth-order cumulants," *IEEE Transactions on Signal Processing*, vol. 45, no. 5, pp. 1253–1263, 1997.
- [27] A. J. Weiss and B. Friedlander, "Performance analysis of spatial smoothing with interpolated arrays," *IEEE Transactions on Signal Processing*, vol. 41, no. 5, pp. 1881–1892, 1993.
- [28] Tie-Jun Shan and T. Kailath, "Adaptive beamforming for coherent signals and interference," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 3, pp. 527–536, 1985.
- [38] L. Yan, P. Addabbo, C. Hao, D. Orlando, and A. Farina, "New ECCM Techniques Against Noiselike and/or Coherent Interferers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 2, pp. 1172–1188, 2020.
- [29] H. L. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. John Wiley & Sons, 2002.
- [30] Tie-Jun Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, 1985.
- [31] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Pearson Prentice Hall, 2005.
- [32] F. Ansari, M. Taban, and S. Gazor, "A novel sequential algorithm for clutter and direct signal cancellation in passive bistatic radars," *EURASIP Journal on Advances in Signal Processing*, no. 134, 2016.
- [33] B. Feng, T. Wang, C. Liu, C. Chen, and W. Chen, "An effective CLEAN algorithm for interference cancellation and weak target detection in passive radar," in *Conference Proceedings of 2013 Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, 2013, pp. 160–163.
- [34] K. Kulpa, "The CLEAN type algorithms for radar signal processing," in *2008 Microwaves, Radar and Remote Sensing Symposium*, 2008, pp. 152–157.
- [35] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 11, pp. 2436–2449, 1991.
- [36] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 10, pp. 1553–1560, 1988.
- [37] X. Tan, W. Roberts, J. Li, and P. Stoica, "Sparse Learning via Iterative Minimization With Application to MIMO Radar Imaging," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1088–1101, March 2011.
- [39] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004.
- [40] P. Stoica and P. Babu, "On the Proper Forms of BIC for Model Order Selection," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4956–4961, September 2012.
- [41] E. Conte, A. De Maio, and G. Ricci, "GLRT-based adaptive detection algorithms for range-spread targets," *IEEE Transactions on Signal Processing*, vol. 49, no. 7, pp. 1336–1348, July 2001.
- [42] P. Addabbo, D. Orlando, and G. Ricci, "Adaptive Radar Detection of Dim Moving Targets in Presence of Range Migration," *IEEE Signal Processing Letters*, vol. 26, no. 10, pp. 1461–1465, Oct 2019.
- [43] R. A. Horn and C. R. Johnson, *Matrix Analysis*, C. U. Press, Ed., 1985.
- [44] P. Stoica and Y. Selen, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 112–114, 2004.