

Innovative Cognitive Approaches for Joint Radar Clutter Classification and Multiple Target Detection in Heterogeneous Environments

Linjie Yan, Sudan Han, Chengpeng Hao, *Senior Member, IEEE*, Danilo Orlando, *Senior Member, IEEE*, and Giuseppe Ricci, *Senior Member, IEEE*

Abstract—The joint adaptive detection of multiple point-like targets in scenarios characterized by different clutter types is still an open problem in the radar community. In this paper, we provide a solution to this problem by devising detection architectures capable of classifying the range bins according to their clutter properties and detecting possible multiple targets whose positions and number are unknown. Remarkably, the information provided by the proposed architectures makes the system aware of the surrounding environment and can be exploited to enhance the entire detection and estimation performance of the system. At the design stage, we assume three different signal models and apply the latent variable model in conjunction with estimation procedures based upon the expectation-maximization algorithm. In addition, for some models, the maximization step cannot be computed in closed-form (at least to the best of authors' knowledge) and, hence, suitable approximations are pursued, whereas, in other cases, the maximization is exact. The performance of the proposed architectures is assessed over synthetic data and shows that they can be effective in heterogeneous scenarios providing an initial snapshot of the radar operating scenario.

Index Terms—Cognitive systems, Clutter, Expectation-Maximization, Heterogeneous Environments, Multiple targets, Interference Classification, Radar.

I. INTRODUCTION

The design of adaptive detectors in heterogeneous environments continues to attract the attention of the radar community. As a matter of fact, it represents a compelling task and several interesting solutions have been proposed over the years [1]–[9, and references therein]. The complexity of such solutions and of the considered scenarios is *somehow directly proportional* to the advances in electronic technology and digitalization allowing for miniaturized high performance sensors and processing boards.

There exist different models to address heterogeneity in radar data that generalize the well-known homogeneous environment. This model assumes that the range Cell Under

(Corresponding author: Danilo Orlando.)

Linjie Yan and Chengpeng Hao are with the Institute of Acoustics, Chinese Academy of Sciences, No. 21, North 4th Ring Road West, Haidian District, 100864 Beijing, China. E-mail: yanlinjie16@163.com; haochengp@mail.ioa.ac.cn.

Sudan Han is with the National Innovation Institute of Defense Technology, Yard 53, East Street, Fengtai District, 100010 Beijing, China. E-mail: xiaoxiaosu0626@163.com.

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

Giuseppe Ricci is with the Dipartimento di Ingegneria dell’Innovazione, Università del Salento, Via Monteroni, 73100 Lecce, Italy. E-mail: giuseppe.ricci@unisalento.it.

Test (CUT) as well as the secondary data, selected through the (leading and lagging) reference window, share the same statistical characterization of the interference [10]–[15, and references therein].¹ The idea is that the interference in the CUT can be equalized by exploiting the estimates of its distribution parameters obtained through the secondary data. The most important aspect of this approach is that the corresponding decision rules possess the Constant False Alarm Rate (CFAR) property which allows radar engineers to set the detection threshold regardless the interference spectral properties. When the reference window includes range bins characterized by different statistical properties (e.g., different reflectivity coefficients, different grazing angles, clutter edges, unwanted intentional/unintentional targets, etc.), the quality of the estimates based upon secondary data might considerably deplete leading to two main effects [16], [17]: a detection performance degradation, due to the attenuation of target components, and an increase of the false alarm rate, due to undernullled clutter.

An alternative to the homogeneous environment that still allows for mathematical tractability is represented by the so-called *partially-homogeneous environment*, where interference in the secondary data shares the same structure of the covariance matrix as in the CUT but a different power level [18]. It is important to notice that even though in this case CUT and secondary data are not homogeneous, the statistical characterization of the training samples is assumed to be the invariant over the range. In scenarios of practical value, this behavior is not always valid as corroborated by different experimental measurements [19]–[23]. Such results indicate that an effective model for clutter in high-resolution radars or at low grazing angles is the so-called compound-Gaussian [24], [25] formed by the product of a speckle and a texture component. For short time intervals, the texture component can be approximated as a constant leading to a *fully-heterogeneous environment* where each range bin is characterized by its own power level [1], [4].

Another approach to deal with heterogeneous interference consists in the classification of the region under surveillance in terms of clutter properties. Then, only the range bins belonging

¹In what follows, we use the term interference to denote the overall disturbance affecting radar data that, generally speaking, is the sum of three components: thermal noise, clutter, and possible intentional interference. When we use the term clutter to denote the interference, it is understood that the clutter component is the most significant.

to these homogeneous regions come into play to estimate the related distribution parameters. Thus, for each region a set of estimates is available to make adaptive the decision schemes used to detect targets in that region. In this respect, the knowledge-aided paradigm is an effective means to guide the system towards reliable clutter parameter estimates [26], [27]. It consists in accounting for all the available *a priori* information about the region of interest to exclude inhomogeneities from the computation of the Sample Covariance Matrix (SCM) [16]. More recently, classification strategies based upon the latent variable model [28] and the Expectation-Maximization (EM) algorithm [29] have been proposed in [2] for different models of the disturbance covariance matrix. Such methods are conceived to partition the range bins within the region of interest into homogeneous (not necessarily contiguous) clusters that can be used for training data selection according to the position of the cell under test. However, the design assumptions of these algorithms do not contemplate the presence of prospective targets and/or point-like inhomogeneities within the region of interest. As better explained below, when training data are contaminated by outliers, conventional detectors might experience a performance degradation and an increase of the probability of false alarm. The same kind of clustering but with the constraint of regions formed by contiguous range bins is developed in [30], [31] where a sliding window moves over the entire radar window and for each position a test on the presence of a clutter edge is performed. A fusion strategy is also provided to merge the returned estimates of the clutter edge positions.

However, it is important to observe that modern radar systems can operate in target-rich scenarios where the risk of incorporating target components into the covariance matrix estimate becomes high. More importantly, such an estimate is commonly used to whiten data and, hence, the presence of target components would dramatically attenuate the energy of target echoes with a reduction of the receiver sensitivity. A tangible example of this problem is provided by the so-called Adaptive Matched Filter with De-emphasis (AMFD) with the de-emphasis parameter equal to 1 [32] that is equivalent to the Adaptive Matched Filter (AMF) derived in [12] where the SCM over the secondary data is updated with the CUT. In fact, the AMFD returns poor detection performance with respect to the AMF and Kelly's detector [13]. All these issues point out the need of developing classification schemes not only capable of labeling the range bins according to the statistical properties of the clutter but also of identifying the range bins containing structured signals in order to somehow exclude them from clutter properties estimation. Remarkably, this information can be used to build up architectures that, unlike classical detectors, jointly process the entire radar window to detect multiple point-like targets within each clutter region (and, hence, taking advantage of the entire energy present in the collected data).

A. Contributions and Organization of this Work

Inspired by the above remarks, in the present work we further extend the problem solved in [2] by accounting for the

presence of possible multiple targets and/or outliers. Otherwise stated, the detection/classification problem considered here is given by the union of the two main challenging problems related to radar: estimation of heterogeneous clutter properties and detection of an unknown number of multiple targets. In literature, such problems have been separately addressed due to the high number of unknown parameters that can give rise to possible solutions characterized by a huge computational cost.

More specifically, the main contributions are the following.

- We come up with a suitable general model for joint classification and detection of multiple clutter edges and point-like targets (or outliers).
- We devise radar architectures that jointly classify the clutter returns in terms of their covariance matrix (by partitioning the region of interest on the basis of clutter properties) as in [2] but with the additional capability of detecting possible multiple point-like targets. It is important to highlight that the design of this kind of architectures represents the main technical novelty of this paper. As a matter of fact, existing clutter covariance classifiers do not account for the possible inhomogeneity due to structured signals backscattered from possible targets [2], [30], [31, and references therein], whereas detectors of multiple point-like targets assume that clutter returns share the same covariance matrix [33]–[37, and references therein].
- At the design stage, we consider three target-plus-noise hypotheses that account for different target models comprising deterministic and fluctuating responses. For each hypothesis, we conceive estimation procedures exploiting hidden random variables representative of different situations that can occur at range bin level. To be more definite, the bin under consideration might contain clutter only characterized by a specific class of distribution parameters (that identify a clutter region) or, in addition, target components. Thus, unlike [2], the set of classes considered here encompass the presence of a target in each range bin and for each clutter type. This modification makes more compelling and difficult the problem at hand with respect to [2] and, as shown in the next sections, we obtain estimate updates based upon suboptimal approximations of the objective functions to be optimized. Three important remarks are in order. First, the classification of a range bin as occupied by a target is tantamount to detecting that target; second, notice that we do not establish any assumption about the position and the number of targets; finally, the proposed architectures take a unique decision by processing all the range bins belonging to the region of interest unlike the conventional detectors that work on a single range bin basis.
- Since applying the Maximum Likelihood Approach (MLA) to come up with the estimates of the unknown parameters leads to intractable mathematics, in the developed framework, we resort to the EM algorithm [29] (or heuristic modification of it) in conjunction with cyclic optimization procedures [38]. Borrowing the idea of the

Generalized Likelihood Ratio Test (GLRT), the estimates returned by these procedures are used to build up two decision statistics based upon the likelihood ratio.

The numerical examples are obtained over synthetic data and highlight that the proposed architectures represent an effective means to detect multiple point-like targets in heterogeneous scenarios returning reliable detection and classification performance. More importantly, detection and classification information provided by the proposed algorithms can be suitably exploited by *ad hoc* (possibly conventional) decision schemes and estimation procedures to enhance the entire performance of the system.

The remainder of the paper is organized as follows. The next section contains the problem formulation along with useful preliminary definitions, whereas Section III is devoted to the design of the classification architectures. The illustrative examples and the discussion about the detection and classification performance are provided in Section IV. Finally, in Section V, we draw the conclusions and lay down possible future research lines.

II. PROBLEM FORMULATION AND PRELIMINARY DEFINITIONS

Let² us denote by $\mathbf{z}_1, \dots, \mathbf{z}_K$ the N -dimensional vectors representing the returns from K range bins forming the region of interest illuminated by the radar system. The size N of such vectors represents the number of space, time, or space-time channels. An important assumption of practical value is that the statistical characterization of the clutter component is range-dependent [10] and, hence, the corresponding distribution parameters might change over the range (e.g., due to the presence of clutter boundaries). Thus, the set $\Omega = \{1, \dots, K\}$ of indices can be partitioned into a given number, L say, of subsets whose elements index data vectors sharing the same statistical characterization of the clutter; the l th subset is denoted by $\Omega_l = \{i_{l,1}, \dots, i_{l,K_l}\}$, where $K_l > N$, $l \in \mathcal{L} = \{1, \dots, L\}$, denotes its cardinality. The value of L is assumed known and can be set exploiting the a priori information about the terrain types composing the region of interest. However, it is also likely that multiple point-like targets are present within the region of interest with the consequence that data vectors indexed by Ω_l , although homogeneous from the standpoint of clutter statistical

²*Notation:* vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. Symbols $\det(\cdot)$, $\text{Tr}(\cdot)$, $\|\cdot\|$, and $(\cdot)^\dagger$ denote the determinant, trace, Euclidean norm, and conjugate transpose, respectively. As to 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 matrices (or vectors if $M = 1$). The modulus of $x \in \mathbb{C}$ is denoted by $|x|$. \mathbf{I} and $\mathbf{0}$ stand for the identity matrix and the null vector or matrix of proper size. Given $a_1, \dots, a_N \in \mathbb{C}$, $\text{diag}(a_1, \dots, a_N) \in \mathbb{C}^{N \times N}$ indicates the diagonal matrix whose i th diagonal element is a_i . The acronyms PDF and PMF stand for Probability Density Function and Probability Mass Function, respectively, whereas the conditional PDF of a random variable x given another random variable y is denoted by $f(x|y)$. 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} while given a matrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M] \in \mathbb{C}^{N \times M}$, writing $\mathbf{X} \sim \mathcal{CN}_{N,M}(\mathbf{m}, \mathbf{M}, \mathbf{I})$ means that $\mathbf{x}_i \sim \mathcal{CN}_N(\mathbf{m}, \mathbf{M})$, $i = 1, \dots, M$, and the \mathbf{x}_i s are statistically independent.

characterization, are heterogeneous due to target components affecting $\mathbf{z}_{i_{l,m}}$ for some $m \in \{1, \dots, K_l\}$. Summarizing, if data indexed by the generic Ω_l are free of target components, then $\mathbf{Z}_l = [\mathbf{z}_{i_{l,1}}, \dots, \mathbf{z}_{i_{l,K_l}}] \sim \mathcal{CN}_{N,K_l}(\mathbf{0}, \mathbf{M}_l, \mathbf{I})$; when there exists $\Omega_l^t = \{j_{l,1}, \dots, j_{l,T_l}\} \subseteq \Omega_l$ indexing vectors that contain target components with T_l the number of targets within Ω_l , then $\forall h = 1, \dots, T_l$,

$$\mathbf{z}_{j_{l,h}} \sim \mathcal{CN}_N(\alpha_{j_{l,h}} \mathbf{v}, \mathbf{M}_l) \quad (1)$$

for deterministic targets and

$$\mathbf{z}_{j_{l,h}} \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l + \sigma_{j_{l,h}}^2 \mathbf{v} \mathbf{v}^\dagger) \quad (2)$$

for fluctuating targets. In both cases, the $\mathbf{z}_{j_{l,h}}$ s are statistically independent of $\mathbf{z}_{i_{l,m}} \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l)$, $\forall i_{l,m} \neq j_{l,h}$, $m = 1, \dots, K_l$, $h = 1, \dots, T_l$. In (1) and (2), $\alpha_{j_{l,h}} \in \mathbb{C}$ and $\sigma_{j_{l,h}}^2 > 0$ are the target response and the target power from the $j_{l,h}$ th range bin and along the nominal (space, time, or space-time) steering vector \mathbf{v} [13], respectively.

From the above aspects, it turns out that the problem at hand has a twofold nature, namely it can be viewed as

- a classification problem of clutter returns;
- a detection problem of multiple targets in heterogeneous clutter.

Thus, assuming that the clutter distribution parameters are unknown and, hence, must be estimated from data, we can formulate such a problem in terms of a binary hypothesis test where the null hypothesis is given by

$$H_0 : \mathbf{z}_i \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l), \quad i \in \Omega_l, \quad l = 1, \dots, L, \quad (3)$$

the alternative hypothesis for nonfluctuating targets is

$$H_{1,1} : \begin{cases} \mathbf{z}_i \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l), & i \in \Omega_l \setminus \Omega_l^t, \\ \mathbf{z}_i \sim \mathcal{CN}_N(\alpha_i \mathbf{v}, \mathbf{M}_l), & i \in \Omega_l^t, \end{cases} \quad l = 1, \dots, L, \quad (4)$$

and the alternative hypothesis for fluctuating targets can be expressed as

$$H_{1,2} : \begin{cases} \mathbf{z}_i \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l), & i \in \Omega_l \setminus \Omega_l^t, \\ \mathbf{z}_i \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l + \sigma_i^2 \mathbf{v} \mathbf{v}^\dagger), & i \in \Omega_l^t, \end{cases} \quad l = 1, \dots, L. \quad (5)$$

If we forget the structure of the target component, namely $\sigma_i^2 \mathbf{v} \mathbf{v}^\dagger$, and assume that possible swarms of targets are present in the operating scenario, the alternative hypothesis for random targets can be further recast as

$$H_{1,3} : \begin{cases} \mathbf{z}_i \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l), & i \in \Omega_l \setminus \Omega_l^t, \\ \mathbf{z}_i \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l + \mathbf{R}_l), & i \in \Omega_l^t, \end{cases} \quad l = 1, \dots, L, \quad (6)$$

where $\mathbf{R}_l \in \mathbb{C}^{N \times N}$ is a rank one matrix representative of the specific target swarm. It is understood that each element of the swarm shares the same velocity and motion direction. As for the power, actually, each target should have its own power level and, hence, the index of the matrix representative of the target should be $i \in \Omega_l^t$. However, such a model leads to intractable mathematics in the subsequent developments. The same effect is also observed if we assume a common mean power for all targets, namely, only one target covariance matrix \mathbf{R} . For the above reasons, we follow an alternative route where a swarm

power class corresponds to a clutter covariance structure. This line of reasoning is motivated by the fact that when the swarm is located in a clutter region close to the radar, its power would likely be higher than that of an analogous fleet of targets in a clutter region far from the radar. As a consequence, it is possible to associate a swarm power to each clutter covariance class. Finally, estimating the unknown quantities under (4), (5), and (6) allows us to accomplish the classification task.

Before concluding this section, we provide the expressions of the PDF that will be used in what follows for the design of the detection architectures. Let us start by defining $\mathcal{P}'_0 = \{\Omega_l, \mathbf{M}_l : l \in \mathcal{L}\}$, $\mathcal{P}'_{1,1} = \{\Omega_l, \alpha_i, \mathbf{M}_l : i \in \Omega_l^t, l \in \mathcal{L}\}$, $\mathcal{P}'_{1,2} = \{\Omega_l, \sigma_i^2, \mathbf{M}_l : i \in \Omega_l^t, l \in \mathcal{L}\}$, and $\mathcal{P}'_{1,3} = \{\Omega_l, \Omega_l^t, \mathbf{M}_l, \mathbf{R}_l : l \in \mathcal{L}\}$ the sets of unknown parameters under H_0 , $H_{1,1}$, $H_{1,2}$, and $H_{1,3}$, respectively. The joint PDF of $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_K]$ under H_0 is

$$f_0(\mathbf{Z}; \mathcal{P}'_0) = \prod_{l=1}^L \prod_{i \in \Omega_l} \frac{\exp\{-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_i \mathbf{z}_i^\dagger]\}}{\pi^N \det(\mathbf{M}_l)}; \quad (7)$$

under $H_{1,1}$ it can be written as

$$f_1(\mathbf{Z}; \mathcal{P}'_{1,1}) = \prod_{l=1}^L \left[\prod_{i \in \Omega_l \setminus \Omega_l^t} \frac{\exp\{-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_i \mathbf{z}_i^\dagger]\}}{\pi^N \det(\mathbf{M}_l)} \times \prod_{i \in \Omega_l^t} \frac{\exp\{-\text{Tr}[\mathbf{M}_l^{-1} (\mathbf{z}_i - \alpha_i \mathbf{v})(\mathbf{z}_i - \alpha_i \mathbf{v})^\dagger]\}}{\pi^N \det(\mathbf{M}_l)} \right]; \quad (8)$$

under $H_{1,2}$, it is given by

$$f_1(\mathbf{Z}; \mathcal{P}'_{1,2}) = \prod_{l=1}^L \left[\prod_{i \in \Omega_l \setminus \Omega_l^t} \frac{\exp\{-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_i \mathbf{z}_i^\dagger]\}}{\pi^N \det(\mathbf{M}_l)} \times \prod_{i \in \Omega_l^t} \frac{\exp\{-\text{Tr}[(\mathbf{M}_l + \sigma_i^2 \mathbf{v} \mathbf{v}^\dagger)^{-1} \mathbf{z}_i \mathbf{z}_i^\dagger]\}}{\pi^N \det(\mathbf{M}_l + \sigma_i^2 \mathbf{v} \mathbf{v}^\dagger)} \right]; \quad (9)$$

finally, that related to $H_{1,3}$ is

$$f_1(\mathbf{Z}; \mathcal{P}'_{1,3}) = \prod_{l=1}^L \left[\prod_{i \in \Omega_l \setminus \Omega_l^t} \frac{\exp\{-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_i \mathbf{z}_i^\dagger]\}}{\pi^N \det(\mathbf{M}_l)} \times \prod_{i \in \Omega_l^t} \frac{\exp\{-\text{Tr}[(\mathbf{M}_l + \mathbf{R}_l)^{-1} \mathbf{z}_i \mathbf{z}_i^\dagger]\}}{\pi^N \det(\mathbf{M}_l + \mathbf{R}_l)} \right]. \quad (10)$$

III. ARCHITECTURE DESIGNS AND ESTIMATION PROCEDURES

The architectures devised in this section are grounded on the main idea behind the GLRT with the difference that, instead of the MLA, suitable procedures are used to come up with the parameter estimates. Thus, denoting by $\widehat{\mathbf{R}}_l$, $\widehat{\Omega}_l$, $\widehat{\Omega}_l^t$, $\widehat{\mathbf{M}}_l$, $l = 1, \dots, L$, along with $\widehat{\sigma}_m^2$ and $\widehat{\alpha}_m$, $m \in \widehat{\Omega}_l^t$, the estimates of the respective unknown parameters, the GLRT-like rule for testing H_0 against $H_{1,i}$ has the following form

$$\frac{f_1(\mathbf{Z}; \widehat{\mathcal{P}}'_{1,i})}{f_0(\mathbf{Z}; \widehat{\mathcal{P}}'_0)} \underset{H_0}{\overset{H_{1,i}}{>}} \eta, \quad (11)$$

where $\widehat{\mathcal{P}}'_{1,i}$, $i = 1, 2, 3$, and $\widehat{\mathcal{P}}'_0$ are the sets of estimates under $H_{1,i}$, $i = 1, 2, 3$, and H_0 , respectively; η is the threshold³ to be set according to the required probability of false alarm (P_{fa}). Now, exploiting the MLA to come up with $\widehat{\mathcal{P}}'_{1,i}$ and $\widehat{\mathcal{P}}'_0$ does not seem viable from the standpoints of mathematical tractability and computational load. In fact, this approach requires to account for all the range bin configurations in terms of presence of targets and/or specific clutter covariance components see also Section III of [2], [36]. For this reason, we resort to the EM algorithm and the latent variable model. Specifically, we extend the methodology proposed in [2] to the case contemplating the presence of targets and introduce hidden discrete random variables that are representative of the covariance class associated with a given range bin as well as of the presence of a target with a given Angle of Arrival (AoA) and/or normalized Doppler frequency in that range bin.

Under the alternative hypotheses, let us assume that K independent and identically distributed discrete random variables, c_k s say, are associated with the range bins under test and that they are not observable. Such random variables have unknown PMF $P(c_k = l) = p_l$, $k = 1, \dots, K$ and $l \in \{1, \dots, L_c\}$. As better explained below, L_c accounts for the number of clutter covariance classes and the presence of a possible target. Thus, we can use the index l to code different operating situations. Specifically, under the alternative hypotheses, when $c_k = l$ and $L_s + 1 \leq l \leq L_s + L$, $s = 0, 1$, then

$$\mathbf{z}_k \sim \begin{cases} \mathcal{CN}_N(\alpha_{s,k} \mathbf{v}, \mathbf{M}_{l-L_s}), & \text{under } H_{1,1}, \\ \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_{l-L_s} + \sigma_{s,k}^2 \mathbf{v} \mathbf{v}^\dagger), & \text{under } H_{1,2}, \\ \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_{l-L_s} + s \mathbf{R}_{l-L_s}), & \text{under } H_{1,3}, \end{cases} \quad (12)$$

where we set $\alpha_{0,k} = \sigma_{0,k}^2 = 0$. On the other hand, when H_0 is in force, we assume that $c_k = l$, $l \in \mathcal{L}$, implies that $\mathbf{z}_k \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l)$. As a consequence, $L_c = L$ under H_0 and $L_c = 2L$ under $H_{1,i}$, $i = 1, 2, 3$.⁴ The corresponding sets of unknown parameters associated with the distribution of \mathbf{z}_k are $\mathcal{P}_0 = \mathcal{P}'_0 \cup \mathcal{A}$ with $\mathcal{A} = \{p_1, \dots, p_{L_c}\}$ under H_0 , $\mathcal{P}_{1,1,k} = \mathcal{A} \cup \{\alpha_{1,k}, \mathbf{M}_l : l \in \mathcal{L}\}$ under $H_{1,1}$, $\mathcal{P}_{1,2,k} = \mathcal{A} \cup \{\sigma_{1,k}^2, \mathbf{M}_l : l \in \mathcal{L}\}$ under $H_{1,2}$, and $\mathcal{P}_{1,3,k} = \mathcal{A} \cup \{\mathbf{M}_l, \mathbf{R}_l : l \in \mathcal{L}\}$ under $H_{1,3}$.

With the above remarks in mind, we rewrite the PDF of \mathbf{z}_k under $H_{1,i}$, $i = 1, 2, 3$, as

$$g_1(\mathbf{z}_k; \mathcal{P}_{1,i,k}) = \sum_{s=0}^1 \sum_{l=1}^L p_{L_s+l} f(\mathbf{z}_k | c_k = L_s + l; \Theta_{s,i,k,l}), \quad (13)$$

where

$$\Theta_{s,i,k,l} = \begin{cases} \{\alpha_{1,k}, \mathbf{M}_l\}, & \text{for } s = 1, i = 1, \\ \{\sigma_{1,k}^2, \mathbf{M}_l\}, & \text{for } s = 1, i = 2, \\ \{\mathbf{R}_l, \mathbf{M}_l\}, & \text{for } s = 1, i = 3, \\ \{\mathbf{M}_l\}, & \text{for } s = 0, i = 1, 2, 3. \end{cases} \quad (14)$$

As for the PDF of \mathbf{z}_k under H_0 , it can be written as in [2] by setting $s = 0$, $\forall l = 1, \dots, L$, and, hence, neglecting the outer

³Hereafter, η denotes the generic detection threshold.

⁴Recall that $\sum_{s=0}^1 \sum_{l=1}^L p_{L_s+l} = 1$ under $H_{1,i}$, $i = 1, 2, 3$.

summation in the previous PDFs, namely,

$$g_0(\mathbf{z}_k; \mathcal{P}_0) = \sum_{l=1}^L p_l f(\mathbf{z}_k | c_k = l; \mathbf{M}_l). \quad (15)$$

In (13), $f(\mathbf{z}_k | c_k = Ls + l; \Theta_{s,1,k,l})$ denotes the PDF of a complex Gaussian vector with covariance matrix \mathbf{M}_l and mean $\alpha_{1,k} \mathbf{v}$ if $s = 1$ whereas the mean is zero for $s = 0$; $f(\mathbf{z}_k | c_k = Ls + l; \Theta_{s,2,k,l})$ is the PDF of a complex Gaussian vector with zero mean and covariance matrix $\mathbf{M}_l + \sigma_{1,k}^2 \mathbf{v} \mathbf{v}^\dagger$ if $s = 1$, when $s = 0$ the covariance matrix is \mathbf{M}_l ; $f(\mathbf{z}_k | c_k = Ls + l; \Theta_{s,3,k,l})$ is the PDF of a complex Gaussian vector with zero mean and covariance matrix $\mathbf{M}_l + s \mathbf{R}_l$; finally, in (15), $f(\mathbf{z}_k | c_k = l; \mathbf{M}_l)$ is the multivariate complex Gaussian PDF with zero mean and covariance matrix \mathbf{M}_l . Exploiting the above characterizations, given $i = 1, 2, 3$, we can also build up the following alternative test

$$\prod_{k=1}^K \frac{g_1(\mathbf{z}_k; \hat{\mathcal{P}}_{1,i,k})}{g_0(\mathbf{z}_k; \hat{\mathcal{P}}_0)} \underset{H_0}{\overset{H_{1,i}}{>}} \eta. \quad (16)$$

Notice that also in this case, obtaining the maximum likelihood estimates of the unknown parameters is a difficult task. However, the presence of the hidden random variables allows us to solve the estimation problems under all the considered hypotheses by devising iterative procedures based upon the EM algorithm [29]. It is worth noting once again that such estimates can be exploited to obtain data partitions which can also be used in (11).

The estimation procedures under H_0 have been already developed in [2] and, hence, we focus on those under $H_{1,i}$, $i = 1, 2, 3$.

A. Estimation procedures for deterministic targets ($H_{1,1}$)

In order to apply the EM algorithm under $H_{1,1}$, let us exploit (13) with $i = 1$ and write the joint log-likelihood of \mathbf{Z} for deterministic targets as follows

$$\begin{aligned} \mathcal{L}(\mathbf{Z}; \mathcal{P}_{1,1}) &= \sum_{k=1}^K \log g_1(\mathbf{z}_k; \mathcal{P}_{1,1,k}) \\ &= \sum_{k=1}^K \log \left[\sum_{s=0}^1 \sum_{l=1}^L p_{Ls+l} f(\mathbf{z}_k | c_k = Ls + l; \Theta_{s,1,k,l}) \right], \end{aligned}$$

where $\mathcal{P}_{1,1} = \cup_{k=1}^K \mathcal{P}_{1,1,k}$.⁵ The first step of the EM is the E-step that leads to the following update rule

$$\begin{aligned} q_k^{(h-1)}(Ls + l) &= p \left(c_k = Ls + l | \mathbf{z}_k; \hat{\mathcal{P}}_{1,1,k}^{(h-1)} \right) \\ &= \frac{f(\mathbf{z}_k | c_k = Ls + l; \hat{\Theta}_{s,1,k,l}^{(h-1)}) \hat{p}_{Ls+l}^{(h-1)}}{\sum_{j=0}^1 \sum_{i=1}^L f(\mathbf{z}_k | c_k = Lj + i; \hat{\Theta}_{j,1,k,i}^{(h-1)}) \hat{p}_{Lj+i}^{(h-1)}}, \quad (17) \end{aligned}$$

⁵In what follows, we denote by $\mathcal{L}(\mathbf{Z}; \mathcal{P}_{1,i})$ with $\mathcal{P}_{1,i} = \cup_{k=1}^K \mathcal{P}_{1,i,k}$, $i = 1, 2, 3$, the joint log-likelihood function under $H_{1,i}$.

$s = 0, 1$, where $\hat{\Theta}_{s,1,k,l}^{(h-1)}$ and $\hat{p}_{Ls+l}^{(h-1)}$ are the estimates of the unknown parameters at the previous step.⁶ In what follows, in order to mitigate the inclination of the likelihood function to select the hypotheses that include target components, we borrow the likelihood approximations that give rise to the Model Order Selection rules [39] and write (17) as follows

$$(17) \approx \frac{f(\mathbf{z}_k | c_k = Ls + l; \hat{\Theta}_{s,1,k,l}^{(h-1)}) e^{-u(s)} \hat{p}_{Ls+l}^{(h-1)}}{\sum_{j=0}^1 \sum_{i=1}^L f(\mathbf{z}_k | c_k = Lj + i; \hat{\Theta}_{j,1,k,i}^{(h-1)}) e^{-u(j)} \hat{p}_{Lj+i}^{(h-1)}}, \quad (18)$$

where, given the PDF of \mathbf{z}_k , $u(j)$, $j = 0, 1$, is a penalty function depending on the number of the unknown parameters (a point better explained in Section IV).

The second step is the M-step that requires to solve the following problem

$$\begin{aligned} \hat{\mathcal{P}}_{1,1}^{(h)} &= \arg \max_{\mathcal{P}_{1,1}} \left\{ \sum_{k=1}^K \sum_{s=0}^1 \sum_{l=1}^L q_k^{(h-1)}(Ls + l) \right. \\ &\quad \left. \times \log \frac{f(\mathbf{z}_k | c_k = Ls + l; \Theta_{s,1,k,l}) p_{Ls+l}}{q_k^{(h-1)}(Ls + l)} \right\} \quad (19) \end{aligned}$$

that is tantamount to

$$\begin{aligned} \hat{\mathcal{P}}_{1,1}^{(h)} &= \arg \max_{\mathcal{P}_{1,1}} \left\{ \sum_{k=1}^K \sum_{s=0}^1 \sum_{l=1}^L q_k^{(h-1)}(Ls + l) \right. \\ &\quad \times \log f(\mathbf{z}_k | c_k = Ls + l; \Theta_{s,1,k,l}) \\ &\quad \left. + \sum_{k=1}^K \sum_{s=0}^1 \sum_{l=1}^L q_k^{(h-1)}(Ls + l) \log p_{Ls+l} \right\}. \quad (20) \end{aligned}$$

Now, we proceed with the maximization step and observe that the optimization problem with respect to p_{Ls+l} , $s = 0, 1$, $l = 1, \dots, L$, is independent of that over \mathbf{M}_l and $\alpha_{1,k}$. Thus, we can proceed by separately addressing these two problems. Starting from the optimization over \mathcal{A} , observe that it can be solved by using the method of Lagrange multipliers to take into account the constraint $\sum_{s=0}^1 \sum_{l=1}^L p_{Ls+l} = 1$. The final result is

$$\hat{p}_{Ls+l}^{(h)} = \frac{1}{K} \sum_{k=1}^K q_k^{(h-1)}(Ls + l), \quad s = 0, 1, \quad l = 1, \dots, L. \quad (21)$$

Thus, we can focus on the maximization problem with respect to the $\alpha_{1,k,s}$ and \mathbf{M}_l s, i.e.,

$$\begin{aligned} \max_{\substack{\alpha_{1,k} \\ k=1, \dots, K}} \max_{\substack{\mathbf{M}_l \\ l=1, \dots, L}} \sum_{k=1}^K \sum_{s=0}^1 \sum_{l=1}^L q_k^{(h-1)}(Ls + l) \\ \times \log f(\mathbf{z}_k | c_k = Ls + l; \Theta_{s,1,k,l}). \quad (22) \end{aligned}$$

Replacing the PDFs with their expressions and ignoring the terms independent of the parameters of interest, the above

⁶In what follows, the estimate of a parameter vector \mathbf{p} at the h th step is denoted by $\hat{\mathbf{p}}^{(h)}$.

problem is tantamount to

$$\min_{\alpha_{1,k}} \min_{\mathbf{M}_l} \sum_{k=1}^K \sum_{s=0}^1 \sum_{l=1}^L q_k^{(h-1)} (Ls + l) \left\{ \log \det(\mathbf{M}_l) + \text{Tr} \left[\mathbf{M}_l^{-1} (\mathbf{z}_k - \alpha_{s,k} \mathbf{v})(\mathbf{z}_k - \alpha_{s,k} \mathbf{v})^\dagger \right] \right\}. \quad (23)$$

A suboptimum solution to this optimization can be found by resorting to a cyclic procedure that repeats the following steps until a stopping criterion is not satisfied

- assume that the $\alpha_{1,k}$ s are known and estimate \mathbf{M}_l ;
- set the \mathbf{M}_l s to the values obtained at the previous step and estimate the $\alpha_{1,k}$ s.

Thus, denote by $\hat{\alpha}_{1,k}^{(h-1),(m-1)}$, $k = 1, \dots, K$, the estimates of the $\alpha_{1,k}$ s at the $(h-1)$ th EM step and $(m-1)$ th step of this inner procedure (we set $\hat{\alpha}_{0,k}^{(h-1),(m-1)} = 0$), then we solve the following problem

$$\min_{\mathbf{M}_l} \left\{ -\log \det(\mathbf{M}_l^{-1}) \sum_{k=1}^K \sum_{s=0}^1 q_k^{(h-1)} (Ls + l) + \text{Tr} \left[\mathbf{M}_l^{-1} \sum_{k=1}^K \sum_{s=0}^1 q_k^{(h-1)} (Ls + l) \times (\mathbf{z}_k - \hat{\alpha}_{s,k}^{(h-1),(m-1)} \mathbf{v})(\mathbf{z}_k - \hat{\alpha}_{s,k}^{(h-1),(m-1)} \mathbf{v})^\dagger \right] \right\}. \quad (24)$$

The minimizer can be obtained by resorting to the following inequality [40] $\log \det(\mathbf{A}) \leq \text{Tr}[\mathbf{A}] - N$, where \mathbf{A} is any N -dimensional matrix with nonnegative eigenvalues, and, hence, we come up with

$$\widehat{\mathbf{M}}_l^{(h-1),(m)} = \frac{1}{\sum_{s=0}^1 \sum_{k=1}^K q_k^{(h-1)} (Ls + l)} \sum_{s=0}^1 \sum_{k=1}^K q_k^{(h-1)} (Ls + l) \times (\mathbf{z}_k - \hat{\alpha}_{s,k}^{(h-1),(m-1)} \mathbf{v})(\mathbf{z}_k - \hat{\alpha}_{s,k}^{(h-1),(m-1)} \mathbf{v})^\dagger. \quad (25)$$

Now, assuming that in (23) $\mathbf{M}_l = \widehat{\mathbf{M}}_l^{(h-1),(m)}$, $l = 1, \dots, L$, we obtain

$$\min_{\alpha_{1,k}} \sum_{k=1}^K \sum_{l=1}^L q_k^{(h-1)} (L + l) (\mathbf{z}_k - \alpha_{1,k} \mathbf{v})^\dagger \times \left(\widehat{\mathbf{M}}_l^{(h-1),(m)} \right)^{-1} (\mathbf{z}_k - \alpha_{1,k} \mathbf{v}), \quad (26)$$

and setting to zero the first derivative of the objective function with respect to $\alpha_{1,k}$ leads to

$$\hat{\alpha}_{1,k}^{(h-1),(m)} = \frac{\sum_{l=1}^L q_k^{(h-1)} (L + l) \mathbf{v}^\dagger \left(\widehat{\mathbf{M}}_l^{(h-1),(m)} \right)^{-1} \mathbf{z}_k}{\sum_{l=1}^L q_k^{(h-1)} (L + l) \mathbf{v}^\dagger \left(\widehat{\mathbf{M}}_l^{(h-1),(m)} \right)^{-1} \mathbf{v}}. \quad (27)$$

This inner procedure continues until a convergence criterion is not satisfied. The final estimates of $\alpha_{1,k}$ and \mathbf{M}_l are denoted by $\hat{\alpha}_{1,k}^{(h)}$ and $\widehat{\mathbf{M}}_l^{(h)}$, respectively, and are used in the next cycle

of the EM-based procedure. Before moving on, notice that the composition of the EM and this alternating procedure yields a nondecreasing sequence of log-likelihood values.

B. Estimation procedures for fluctuating targets ($H_{1,2}$ and $H_{1,3}$)

In this subsection, we consider the alternative hypotheses defined by (5) and (6) as well as the related PDFs in the steps of the EM algorithm. Starting from the expectation step and following the same line of reasoning as for $H_{1,1}$, the final result maintains the form of (17) and its approximation (18) but for the PDF of \mathbf{z}_k . More precisely, replacing $\hat{\Theta}_{s,1,k,l}^{(h-1)}$ in (18) with $\hat{\Theta}_{s,2,k,l}^{(h-1)}$ for model (5) and with $\hat{\Theta}_{s,3,k,l}^{(h-1)}$ for model (6), we come up with the respective E-steps. Let us recall here that $\hat{\Theta}_{s,2,k,l}^{(h-1)}$ and $\hat{\Theta}_{s,3,k,l}^{(h-1)}$ contain the estimates of the unknown parameters under $H_{1,2}$ and $H_{1,3}$, respectively, at the previous step.

As for the maximization steps, we separately proceed as described in the next subsections.

1) *Fluctuating targets model (5)*: The problem to be solved has the following expression

$$\min_{\sigma_{1,k}^2} \min_{\mathbf{M}_l} \sum_{k=1}^K \sum_{s=0}^1 \sum_{l=1}^L q_k^{(h-1)} (Ls + l) \times \left\{ \log \det(\mathbf{M}_l + \sigma_{s,k}^2 \mathbf{V}) + \text{Tr} [(\mathbf{M}_l + \sigma_{s,k}^2 \mathbf{V})^{-1} \mathbf{S}_k] \right\}, \quad (28)$$

where $\mathbf{V} = \mathbf{v}\mathbf{v}^\dagger$ and $\mathbf{S}_k = \mathbf{z}_k \mathbf{z}_k^\dagger$. The joint maximization with respect to the $\sigma_{s,k}^2$ s and the \mathbf{M}_l s is difficult from the standpoint of mathematics (at least to the best of authors' knowledge). For this reason, we exploit a heuristic approach that leads to suitable solutions. Specifically, let us denote the objective function in (28) by $x(\mathcal{P}_{1,2})$ with $\mathcal{P}_{1,2} = \cup_{k=1}^K \mathcal{P}_{1,2,k}$ and observe that it can be recast as $x(\mathcal{P}_{1,2}) = x_1(\mathbf{M}_1, \dots, \mathbf{M}_L) + x_2(\mathcal{P}_{1,2})$, where $x_1(\mathbf{M}_1, \dots, \mathbf{M}_L) = \sum_{k=1}^K \sum_{l=1}^L q_k^{(h-1)} (l) \{ \log \det(\mathbf{M}_l) + \text{Tr} [\mathbf{M}_l^{-1} \mathbf{S}_k] \}$ and $x_2(\mathcal{P}_{1,2}) = \sum_{k=1}^K \sum_{l=1}^L q_k^{(h-1)} (L + l) \{ \log \det(\mathbf{M}_l + \sigma_{1,k}^2 \mathbf{V}) + \text{Tr} [(\mathbf{M}_l + \sigma_{1,k}^2 \mathbf{V})^{-1} \mathbf{S}_k] \}$. Since $x_1(\cdot)$ depends on the covariance matrices only, we first estimate \mathbf{M}_l , $l = 1, \dots, L$, by exploiting $x_1(\cdot)$, then replace such estimates in $x_2(\cdot)$, and minimize it with respect to $\sigma_{1,k}^2$, $k = 1, \dots, K$. Thus, the estimates of \mathbf{M}_l , $l = 1, \dots, L$, based upon $x_1(\cdot)$ are given by

$$\widehat{\mathbf{M}}_l^{(h)} = \left\{ \sum_{k=1}^K q_k^{(h-1)} (l) \right\}^{-1} \sum_{k=1}^K q_k^{(h-1)} (l) \mathbf{S}_k. \quad (29)$$

Now, we consider $x_2(\cdot)$ and set $\mathbf{M}_l = \widehat{\mathbf{M}}_l^{(h)}$, $l = 1, \dots, L$. Then, the problem at hand becomes

$$\min_{\substack{\sigma_{1,k}^2 \\ k=1, \dots, K}} \sum_{k=1}^K \sum_{l=1}^L q_k^{(h-1)} (L+l) \left\{ \log \det \left(\widehat{\mathbf{M}}_l^{(h)} + \sigma_{1,k}^2 \mathbf{V} \right) + \text{Tr} \left[\left(\widehat{\mathbf{M}}_l^{(h)} + \sigma_{1,k}^2 \mathbf{V} \right)^{-1} \mathbf{S}_k \right] \right\}. \quad (30)$$

In order to solve (30), we consider a given k and the related objective function

$$\begin{aligned} h_k(\sigma_{1,k}^2) &= \sum_{l=1}^L q_k^{(h-1)} (L+l) \\ &\times \left\{ \log \det \left(\widehat{\mathbf{M}}_l^{(h)} + \sigma_{1,k}^2 \mathbf{V} \right) + \text{Tr} \left[\left(\widehat{\mathbf{M}}_l^{(h)} + \sigma_{1,k}^2 \mathbf{V} \right)^{-1} \mathbf{S}_k \right] \right\} \\ &= \sum_{l=1}^L q_k^{(h-1)} (L+l) \log \left[1 + \sigma_{1,k}^2 \mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v} \right] \\ &+ \sum_{l=1}^L q_k^{(h-1)} (L+l) \log \det \left(\widehat{\mathbf{M}}_l^{(h)} \right) \\ &- \sum_{l=1}^L q_k^{(h-1)} (L+l) \sigma_{1,k}^2 \frac{|\mathbf{z}_k^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}|^2}{1 + \sigma_{1,k}^2 \mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}} \\ &+ \sum_{l=1}^L q_k^{(h-1)} (L+l) \mathbf{z}_k^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{z}_k. \end{aligned} \quad (31)$$

Thus, setting to zero the first derivative with respect to $\sigma_{1,k}^2$, we obtain

$$\begin{aligned} &\sum_{l=1}^L \frac{q_k^{(h-1)} (L+l) \mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}}{1 + \sigma_{1,k}^2 \mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}} \\ &- \sum_{l=1}^L \frac{q_k^{(h-1)} (L+l) |\mathbf{z}_k^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}|^2}{1 + \sigma_{1,k}^2 \mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}} \\ &+ \sum_{l=1}^L \frac{\sigma_{1,k}^2 q_k^{(h-1)} (L+l) |\mathbf{z}_k^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}|^2}{\left(1 + \sigma_{1,k}^2 \mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v} \right)^2 \left(\mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v} \right)^{-1}} = 0 \end{aligned} \quad (32)$$

namely

$$\sum_{l=1}^L \frac{q_k^{(h-1)} (L+l) [\sigma_{1,k}^2 (a_l^{(h)})^2 + a_l^{(h)} - b_{k,l}^{(h)}]}{\left(1 + \sigma_{1,k}^2 a_l^{(h)} \right)^2} = 0, \quad (33)$$

where $a_l^{(h)} = \mathbf{v}^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}$ and $b_{k,l}^{(h)} = |\mathbf{z}_k^\dagger (\widehat{\mathbf{M}}_l^{(h)})^{-1} \mathbf{v}|^2$. The solutions of the last equation can be found by means of numerical routines and, then, we choose the positive solution (if any) that minimizes (31).

2) *Fluctuating targets model* (6): Let us assume that $H_{1,3}$ (see equation (6)) is true and recall that \mathbf{R}_k is positive semidefinite of rank 1. The maximization step consists in solving

$$\min_{\substack{\mathbf{M}_l, \mathbf{R}_l \\ l=1, \dots, L}} \sum_{k=1}^K \sum_{s=0}^1 \sum_{l=1}^L q_k^{(h-1)} (Ls+l) \left\{ \log \det(\mathbf{M}_l + s\mathbf{R}_l) + \text{Tr} \left[(\mathbf{M}_l + s\mathbf{R}_l)^{-1} \mathbf{z}_k \mathbf{z}_k^\dagger \right] \right\}. \quad (34)$$

Notice that the objective function can be further rewritten as

$$\begin{aligned} &\sum_{k=1}^K \sum_{l=1}^L q_k^{(h-1)} (l) \left\{ \log \det(\mathbf{M}_l) + \text{Tr} \left[\mathbf{M}_l^{-1} \mathbf{z}_k \mathbf{z}_k^\dagger \right] \right\} \\ &+ \sum_{k=1}^K \sum_{l=1}^L q_k^{(h-1)} (L+l) \left\{ \log \det(\mathbf{M}_l + \mathbf{R}_l) + \text{Tr} \left[(\mathbf{M}_l + \mathbf{R}_l)^{-1} \mathbf{z}_k \mathbf{z}_k^\dagger \right] \right\}. \end{aligned} \quad (35)$$

Thus, in order to minimize the above function with respect to \mathbf{M}_l and \mathbf{R}_l , let us consider the generic index l and the corresponding term

$$\begin{aligned} &q_l \log \det(\mathbf{M}_l) + \text{Tr} \left[\mathbf{M}_l^{-1} \mathbf{S}_l \right] \\ &+ q_{l+L} \log \det(\mathbf{M}_l + \mathbf{R}_l) + \text{Tr} \left[(\mathbf{M}_l + \mathbf{R}_l)^{-1} \mathbf{S}_{l+L} \right], \end{aligned} \quad (36)$$

where $q_l = \sum_{k=1}^K q_k^{(h-1)} (l)$, $\mathbf{S}_l = \sum_{k=1}^K q_k^{(h-1)} (l) \mathbf{z}_k \mathbf{z}_k^\dagger$, $q_{l+L} = \sum_{k=1}^K q_k^{(h-1)} (L+l)$, and $\mathbf{S}_{l+L} = \sum_{k=1}^K q_k^{(h-1)} (L+l) \mathbf{z}_k \mathbf{z}_k^\dagger$. Then, define $\mathbf{A}_l = \mathbf{M}_l^{1/2} \mathbf{U}_l$ with $\mathbf{U}_l \in \mathbb{C}^{N \times N}$ the unitary matrix whose columns are the eigenvectors of $\mathbf{M}_l^{-1/2} (\mathbf{M}_l + \mathbf{R}_l) \mathbf{M}_l^{-1/2}$ and $\mathbf{\Lambda}_l = \text{diag}(\lambda_{l,1}, 1, \dots, 1)$, $\lambda_{l,1} > 1$, the diagonal matrix containing the eigenvalues of $\mathbf{M}_l^{-1/2} (\mathbf{M}_l + \mathbf{R}_l) \mathbf{M}_l^{-1/2}$. It clearly follows that $\mathbf{M}_l = \mathbf{A}_l \mathbf{A}_l^\dagger$, $(\mathbf{M}_l + \mathbf{R}_l) = \mathbf{A}_l \mathbf{\Lambda}_l \mathbf{A}_l^\dagger$, and (36) can be written as

$$\begin{aligned} &2(q_l + q_{l+L}) \log |\det(\mathbf{A}_l)| + \text{Tr} \left[\mathbf{A}_l^{-1} \mathbf{S}_l \mathbf{A}_l^{-\dagger} \right] \\ &+ q_{l+L} \log \det(\mathbf{\Lambda}_l) + \text{Tr} \left[\mathbf{\Lambda}_l^{-1} \mathbf{A}_l^{-1} \mathbf{S}_{l+L} \mathbf{A}_l^{-\dagger} \right]. \end{aligned} \quad (37)$$

Now, let us exploit the eigendecomposition of $\mathbf{S}_l^{-1/2} \mathbf{S}_{l+L} \mathbf{S}_l^{-1/2} = \mathbf{V}_l \mathbf{\Gamma}_l \mathbf{V}_l^\dagger$ where $\mathbf{\Gamma}_l = \text{diag}(\gamma_{l,1}, \dots, \gamma_{l,N})$, $\gamma_{l,1} \geq \dots \geq \gamma_{l,N} \geq 0$, is the diagonal matrix of the eigenvalues and $\mathbf{V}_l \in \mathbb{C}^{N \times N}$ is the unitary matrix of the corresponding eigenvectors. Denoting by $\mathbf{B}_l = \mathbf{S}_l^{1/2} \mathbf{V}_l$, we obtain that $\mathbf{S}_l = \mathbf{B}_l \mathbf{B}_l^\dagger$, $\mathbf{S}_{l+L} = \mathbf{B}_l \mathbf{\Gamma}_l \mathbf{B}_l^\dagger$, and (37) can be recast as

$$\begin{aligned} &c_l [\log |\det(\mathbf{B}_l)| - (1/2) \log \det(\mathbf{\Lambda}_l) - \log |\det(\mathbf{X}_l)|] \\ &+ \text{Tr} \left[\mathbf{X}_l^\dagger \mathbf{\Lambda}_l \mathbf{X}_l \right] + q_{l+L} \log \det(\mathbf{\Lambda}_l) + \text{Tr} \left[\mathbf{X}_l \mathbf{\Gamma}_l \mathbf{X}_l^\dagger \right], \end{aligned} \quad (38)$$

where $c_l = 2(q_l + q_{l+L})$ and $\mathbf{X}_l = \mathbf{\Lambda}_l^{-1/2} \mathbf{A}_l^{-1} \mathbf{B}_l$. Exploiting the singular value decomposition of \mathbf{X}_l given by $\mathbf{X}_l = \mathbf{T}_l \mathbf{D}_l \mathbf{Q}_l$, $\mathbf{D}_l = \text{diag}(d_{l,1}, \dots, d_{l,N})$ and $d_{l,1} \leq \dots \leq d_{l,N}$,

the last equation becomes

$$c_l [\log |\det(\mathbf{B}_l)| - (1/2) \log \det(\mathbf{\Lambda}_l) - \log \det(\mathbf{D}_l)] + \text{Tr} [\mathbf{D}_l^2 \mathbf{T}_l^\dagger \mathbf{\Lambda}_l \mathbf{T}_l] + q_{l+L} \log \det(\mathbf{\Lambda}_l) + \text{Tr} [\mathbf{D}_l^2 \mathbf{Q}_l \mathbf{T}_l \mathbf{Q}_l^\dagger]. \quad (39)$$

By *Theorem 1* of [41], the minimization with respect to \mathbf{T}_l and \mathbf{Q}_l leads to

$$\begin{aligned} & c_l \left[\log |\det(\mathbf{B}_l)| - \frac{1}{2} \log \det(\mathbf{\Lambda}_l) - \log \det(\mathbf{D}_l) \right] \\ & + \text{Tr} [\mathbf{D}_l^2 \mathbf{\Lambda}_l] + q_{l+L} \log \det(\mathbf{\Lambda}_l) + \text{Tr} [\mathbf{D}_l^2 \mathbf{T}_l] \\ & = c_l \log |\det(\mathbf{B}_l)| + a_l \log \lambda_{l,1} + d_{l,1}^2 (\lambda_{l,1} + \gamma_{l,1}) \\ & + \sum_{i=2}^N d_{l,i}^2 (1 + \gamma_{l,i}) - \frac{c_l}{2} \sum_{i=1}^N \log d_{l,i}^2, \end{aligned} \quad (40)$$

where $a_l = -c_l/2 + q_{l+L}$. Setting to zero the first derivative of the above function with respect to $d_{l,i}^2$, we obtain

$$\hat{d}_{l,i}^2 = \begin{cases} \frac{c_l/2}{\lambda_{l,1} + \gamma_{l,1}}, & \text{if } i = 1, \\ \frac{c_l/2}{1 + \gamma_{l,i}}, & \text{if } i > 1. \end{cases} \quad (41)$$

Finally, replacing $d_{l,i}^2$ with $\hat{d}_{l,i}^2$ in (40) yields

$$\begin{aligned} & c_l \log |\det(\mathbf{B}_l)| + a_l \log \lambda_{l,1} + N \frac{c_l}{2} \\ & - \frac{c_l}{2} \log \frac{c_l/2}{\lambda_{l,1} + \gamma_{l,1}} - \frac{c_l}{2} \sum_{i=2}^N \log \frac{c_l/2}{1 + \gamma_{l,i}} \end{aligned} \quad (42)$$

and setting to zero the first derivative of the last equation with respect to $\lambda_{l,1}$, the result is

$$\hat{\lambda}_{l,1} = \max \left\{ \frac{-2a_l \gamma_{l,1}}{2a_l + c_l}, 1 \right\} = \max \left\{ \frac{q_l \gamma_{l,1}}{q_{l+L}}, 1 \right\}. \quad (43)$$

C. Classification rule

Once the unknown quantities have been estimated, data classification can be accomplished by exploiting suitable rules. Under H_0 , we use the same rule as in [2], whereas under the alternative hypotheses $\forall k = 1, \dots, K$, we proceed as follows

$$\mathbf{z}_k \sim \begin{cases} \mathcal{CN}_N(\mathbf{0}, \widehat{\mathbf{M}}_{\hat{l}_k}^{(h_{\max})}), & 1 \leq \hat{l}_k \leq L, \\ \mathcal{CN}_N(\widehat{\boldsymbol{\alpha}}_k^{(h_{\max})} \mathbf{v}, \widehat{\mathbf{M}}_{\hat{l}_k - L}^{(h_{\max})}), & L + 1 \leq \hat{l}_k \leq 2L, \end{cases} \quad (44)$$

under $H_{1,1}$,

$$\mathbf{z}_k \sim \begin{cases} \mathcal{CN}_N(\mathbf{0}, \widehat{\mathbf{M}}_{\hat{l}_k}^{(h_{\max})}), & 1 \leq \hat{l}_k \leq L, \\ \mathcal{CN}_N(\mathbf{0}, \widehat{\mathbf{M}}_{\hat{l}_k - L}^{(h_{\max})} + (\hat{\sigma}_{1,k}^2)^{(h_{\max})} \mathbf{v} \mathbf{v}^\dagger), & L + 1 \leq \hat{l}_k \leq 2L, \end{cases} \quad (45)$$

under $H_{1,2}$, and

$$\mathbf{z}_k \sim \begin{cases} \mathcal{CN}_N(\mathbf{0}, \widehat{\mathbf{M}}_{\hat{l}_k}^{(h_{\max})}), & 1 \leq \hat{l}_k \leq L, \\ \mathcal{CN}_N(\mathbf{0}, \widehat{\mathbf{M}}_{\hat{l}_k - L}^{(h_{\max})} + \mathbf{R}_{\hat{l}_k - L}^{(h_{\max})}), & L + 1 \leq \hat{l}_k \leq 2L, \end{cases} \quad (46)$$

where h_{\max} is the maximum number of EM iterations and

$$\hat{l}_k = \arg \max_{l=1, \dots, 2L} q_k^{(h_{\max})}(l). \quad (47)$$

When the proposed decision schemes decide for $H_{1,i}$, $i = 1, 2, 3$, the classification rule is that associated with this hypothesis; on the contrary, data are partitioned according to the classification results under H_0 . It is important to stress here that the proposed detectors are fed by all the range bins of interest and perform only one test to detect multiple targets deployed in the region of interest. The workflow of the above cognitive schemes is summarized in Algorithm 1.

Algorithm 1 Workflow of the proposed cognitive procedures.

Input: $L, \mathbf{Z}, \mathbf{v}$

Output: $\widehat{\Omega}_l, \widehat{\Omega}_l^t, l = 1, \dots, L, \widehat{\mathcal{P}}_{1,i,k}, i = 1, 2, 3, k = 1, \dots, K$

- 1: **Latent Variable Model:** introduce the hidden random variables c_k s with unknown PMF, which represent not only the clutter types but also the presence of a possible target within $\mathbf{z}_k, k = 1, \dots, K$;
 - 2: **Initialization:** initialize the parameters of the recursive procedures, namely $\widehat{\mathcal{P}}_{1,i,k}^{(0)}, i = 1, 2, 3, k = 1, \dots, K$, as well as those of the stopping criterion, such as h_{\max} and m_{\max} ;
 - 3: **E-step:** compute the conditional expectation of the log-likelihood of \mathbf{z}_k at the $(h-1)$ th iteration, obtaining the estimation update of the posterior probability $q_k^{(h-1)}(L_s + l)$;
 - 4: **M-step:** maximize the log-likelihood and provide closed-form updates for $\widehat{\mathcal{P}}_{1,i,k}^{(h)}, i = 1, 2, 3, k = 1, \dots, K$;
 - 5: **Recursive Estimation:** if the stopping criterion is satisfied go to step 6 else set $h = h + 1$ and go to step 3;
 - 6: **Classification:** select the class corresponding to the maximum value of $q_k^{(h_{\max})}(l), l = 1, \dots, L, k = 1, \dots, K$; this is the cognitive task of the procedure since it allows us to gather information about the clutter regions and the existence of a target in the k th range bin leading to $\widehat{\Omega}_l$ and $\widehat{\Omega}_l^t, l = 1, \dots, L$;
 - 7: **Detection:** the fully adaptive detector is obtained by replacing $\widehat{\Omega}_l, \widehat{\Omega}_l^t, \widehat{\mathcal{P}}_{1,i,k}, l = 1, \dots, L, k = 1, \dots, K$, in (11) and (16) under the signal-plus-noise hypothesis and similarly under the noise-only hypothesis.
-

IV. NUMERICAL EXAMPLES AND DISCUSSION

In this section, we investigate the classification and detection performances of the proposed architectures resorting to standard Monte Carlo counting techniques. Specifically, let us consider $N = 8$ spatial channels and two operating scenarios differing in the number of targets, L , and K . As for the clutter covariance components, we use $\mathbf{C}_l = \sigma_{c,l}^2 \mathbf{M}_c, l = 1, \dots, L$, where $\sigma_{c,l}^2 > 0$ is the clutter power of the l th region set according to the corresponding Clutter-to-Noise Ratio (CNR $_l$), while the (i, j) th element of \mathbf{M}_c is equal to $\rho_c^{|i-j|}$ with $\rho_c = 0.9$. For simplicity, we assume the same Signal-to-Interference-plus-Noise Ratio (SINR) value for all the synthetic targets. More

precisely, in the case of deterministic targets, it is defined as $\text{SINR} = |\alpha_{k,l}|^2 \mathbf{v}^\dagger \mathbf{M}_l^{-1} \mathbf{v}$, whereas the SINR for fluctuating targets is $\text{SINR} = \sigma_{k,l}^2 / (\sigma_{c,l}^2 + \sigma_n^2)$, where $\mathbf{M}_l = \mathbf{C}_l + \sigma_n^2 \mathbf{I}$ with $\sigma_n^2 = 1$ the thermal noise power, $\alpha_{k,l}$ and $\sigma_{k,l}^2$ are the amplitude and power associated with a target in the k th range bin of the l th region, the nominal steering vector is computed assuming that the AoAs are zero in all scenarios. The penalty term in (18) is set as $u(s) = (N^2 + k_{1,1}s)(1 + \rho)/2$ with $\rho = 3$, $k_{1,1} = 2$ (under $H_{1,1}$), $k_{1,2} = 1$ (under $H_{1,2}$), and $k_{1,3} = N$ (under $H_{1,3}$), that is borrowed from the generalized information criterion with the parameter equal to 3 [39].

A. Operating scenario with two clutter regions

Let us firstly focus on the scenario comprising two clutter regions of $K_1 = K_2 = 32$ range bins; as for the number of targets, we start with the case where two point-like targets are present and, then, we double the number of targets. Each region is characterized by \mathbf{M}_1 and \mathbf{M}_2 with $\text{CNR}_1 = 20$ dB, $\text{CNR}_2 = 30$ dB. Two targets appear in the 15th range bin of the first region and in the 6th range bin of the second region. This operating scenario yields the following classes:

- class 1: the generic vector of the first region does not contain any target component;
- class 2: the generic vector of the second region does not contain any target component;
- class 3: the generic vector of the first region contains target components;
- class 4: the generic vector of the second region contains target components.

As for the parameters initialization of the EM iterations under H_0 and $H_{1,i}$, $i = 1, 2, 3$, we set $\hat{p}_l^{(0)} = 1/L_c$, $l = 1, \dots, L_c$; the initial value of \mathbf{M}_l , namely $\widehat{\mathbf{M}}_l^{(0)}$, is generated in the same way as in Section IV.A of [2]. A possible choice for $\hat{\alpha}_{s,k}^{(0)}$, $s = 0, 1$, $k = 1, \dots, K$, under $H_{1,1}$ is

$$\hat{\alpha}_{s,k}^{(0)} = \begin{cases} 0, & s = 0, \\ \max_{l=1, \dots, L} \left(\frac{\mathbf{v}^\dagger \widehat{\mathbf{M}}_l^{(0)} \mathbf{z}_k}{\mathbf{v}^\dagger \widehat{\mathbf{M}}_l^{(0)} \mathbf{v}} \right), & s = 1. \end{cases} \quad (48)$$

Moreover, the initialization of $\sigma_{s,k}^2$ under $H_{1,2}$ is given by $(\hat{\sigma}_{0,k}^2)^{(0)} = 0$ and $(\hat{\sigma}_{1,k}^2)^{(0)} = |\mathbf{z}_k^\dagger \mathbf{v}|^2$, $k = 1, \dots, K$. The initial value of \mathbf{R}_l under $H_{1,3}$ is grounded on a heuristic way: 1) compute $|\mathbf{z}_k|^2$ for each range bin; 2) sort these quantities in descending order, namely $|\mathbf{z}_{i_1}|^2 \geq |\mathbf{z}_{i_2}|^2 \geq \dots \geq |\mathbf{z}_{i_K}|^2$, where i_k is the ordered index; 3) $\mathbf{R}_l^{(0)} = \mathbf{z}_{i_l} \mathbf{z}_{i_l}^\dagger$, $l = 1, \dots, L$.

Under the above assumptions, we first analyze the convergence performance of the EM algorithm under the three target-plus-noise hypotheses. To this end, let us define $\Delta \mathcal{L}_i(h) = |[\mathcal{L}(\mathbf{Z}; \hat{\mathbf{P}}_{1,i}^{(h)}) - \mathcal{L}(\mathbf{Z}; \hat{\mathbf{P}}_{1,i}^{(h-1)})] / \mathcal{L}(\mathbf{Z}; \hat{\mathbf{P}}_{1,i}^{(h)})|$, $i = 1, 2, 3$. Then, the inner cyclic estimation procedure under $H_{1,1}$ terminates when the convergence criterion $\Omega(m) = \|\widehat{\mathbf{M}}_l^{(h-1),(m)} - \widehat{\mathbf{M}}_l^{(h-1),(m-1)}\| / \|\widehat{\mathbf{M}}_l^{(h-1),(m)}\| + |\hat{\alpha}_{1,k}^{(h-1),(m)} - \hat{\alpha}_{1,k}^{(h-1),(m-1)}| / |\hat{\alpha}_{1,k}^{(h-1),(m)}| < \delta$ is satisfied with $\delta = 10^{-4}$ or when $m = m_{\max}$. In the ensuing analysis, we set $m_{\max} = 5$ that ensures a good compromise between

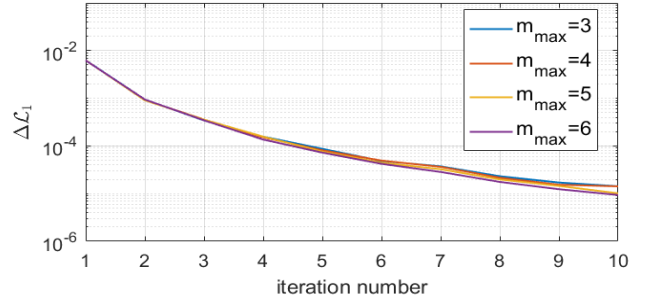


Fig. 1. $\Delta \mathcal{L}_1(h)$ versus h (number of iterations of the EM procedure) for different values of m_{\max} (two targets, two clutter regions).

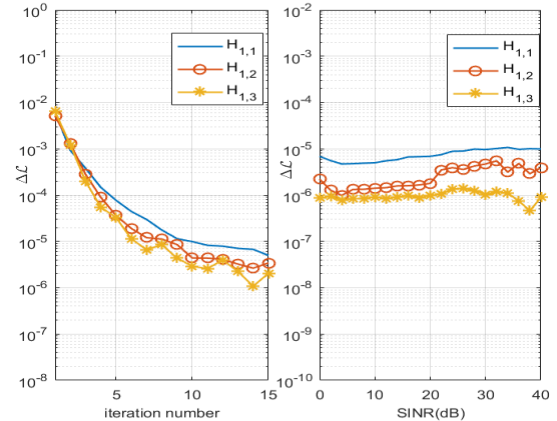


Fig. 2. The convergence performance of the EM procedure: (a) $\Delta \mathcal{L}_i(h)$, $i = 1, 2, 3$, versus h (number of iterations of the EM procedure) for $\text{SINR} = 20$ dB; (b) $\Delta \mathcal{L}_i(15)$, $i = 1, 2, 3$, versus SINR (two targets, two clutter regions).

convergence and computational load. In fact, Fig. 1, where we plot the mean curves (over 1000 independent trials) of $\Delta \mathcal{L}_1$ (namely, under $H_{1,1}$) versus h for different values of m_{\max} and assuming $\text{SINR} = 20$ dB, confirms that $m_{\max} = 5$ returns a relative variation of $\Delta \mathcal{L}$ less than 10^{-4} when $h \geq 5$. In Fig. 2(a), the mean values of $\Delta \mathcal{L}_i(h)$, $i = 1, 2, 3$, are plotted resorting to 1000 independent runs for $\text{SINR} = 20$ dB. The convergence curves show that the log-likelihood variations are roughly lower than 10^{-5} when at least 15 iterations are used for the considered parameters. The mean values of $\Delta \mathcal{L}_i(15)$, $i = 1, 2, 3$, versus SINR are plotted in Fig. 2(b) to verify that the variations maintain low values given $h_{\max} = 15$. Therefore, in the next illustrative examples, we set the maximum number of iteration $h_{\max} = 15$.

A qualitative assessment of the classification results can be found in Fig. 1 of the attached supplemental material where we plot the outcome of one Monte Carlo trial (for two targets) under each alternative hypothesis. The parameters are the same as in the previous figures except for $\text{SINR} = 15, 25, 35$ dB. The inspection of this figure highlights that the three models lead to almost equivalent classification performances. In Table I, we show the Root Mean Square Classification Error (RMSCE) values with respect to the covariance matrix class (its definition can be found in [2] and is omitted here for brevity) that confirms the above conclusion with the estimation

TABLE I
RMSCE VALUES UNDER $H_{1,i}$, $i = 1, 2, 3$, FOR DIFFERENT SINRS (TWO TARGETS, TWO CLUTTER REGIONS)

	$H_{1,1}$	$H_{1,2}$	$H_{1,3}$
SINR = 10 dB	4.5560	4.7304	4.7452
SINR = 15 dB	4.1490	4.3380	4.3906
SINR = 25 dB	3.7114	3.0842	2.9630
SINR = 35 dB	3.4347	2.7841	2.6029

procedure under $H_{1,1}$ ensuring slightly better performance than the other procedures for low SINR values. On the other hand, for high SINR values the procedures under $H_{1,2}$ and $H_{1,3}$ overcome that under $H_{1,1}$. In order to measure the error in target position estimation, we estimate the Root Mean Square (RMS) values of the Hausdorff metric [42], [43] over 1000 trials in Fig. 3 between \hat{r} and r . Precisely, \hat{r} is a vector of size K that contains 0 except for the range bins classified as target, whereas r is an analogous vector containing the true range bin positions of the targets. The figure points out that the classifier under $H_{1,1}$ has lower RMS values than the other two procedures that suffer the existence of more ghosts. Since the effects of the clutter power levels on the classification performance have been assessed in [2], we omit this part of the analysis.

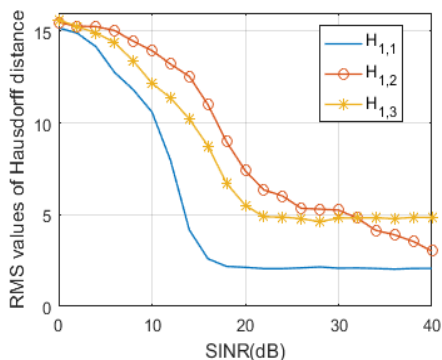


Fig. 3. Hausdorff distance versus SINR under $H_{1,i}$, $i = 1, 2, 3$ (two targets, two clutter regions).

In Fig. 4, we evaluate the Probability of Detection (P_d) returned by detectors⁷ (11) and (16) as a function of the SINR under each hypothesis (and for two targets), respectively. Specifically, the detection thresholds are set over $100/P_{fa}$ independent runs with $P_{fa} = 10^{-3}$, whereas the P_d is computed exploiting 1000 independent trials. The figure points out that the detection performances of detectors (11) and (16) for deterministic targets are roughly better than that for fluctuating targets. This behavior can be explained by the fact that when target Radar Cross Section (RCS) fluctuates, the detection performance impairs with respect to the model where target RCS is steady. Another aspect to be mentioned is that under $H_{1,1}$ (16) outperforms (11) because the former exploits

⁷Notice that the P_d is the probability of declaring the joint presence of multiple point-like targets under $H_{1,i}$, $i = 1, 2, 3$; the information about the number of targets and their positions is obtained from classification results.

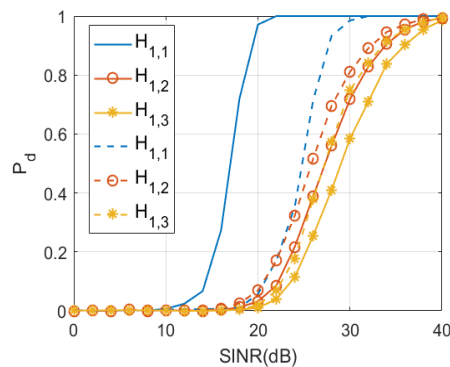


Fig. 4. P_d of the detectors in (11) (dashed line) and (16) (solid line) versus SINR under $H_{1,i}$, $i = 1, 2, 3$, assuming $P_{fa} = 10^{-3}$ (two targets, two clutter regions).

more information than the latter and the estimation procedures are thought for the model associated with (16). However, this trend reverses in the case of random targets that make (16) less numerically stable than (11). Moreover, both detectors for $H_{1,2}$ return slightly better performance than the detectors for $H_{1,3}$ in the medium/high SINR region. Numerical examples not reported here for brevity highlight that when K increases, the classification capabilities (for high SINR values) and the detection performance improve. The drawback is that for low SINR values the Hausdorff distance increases due to the larger number of entries from a wider space.

In the next numerical examples, we consider an operating scenario that comprises four targets. The additional targets occupy the 6th and 55th range bins. The classification results are reported in Fig. 2 of the supplemental material where for SINR = 15 dB, the classifier under $H_{1,1}$ is capable of correctly identifying the range bins with target components whereas the other classifiers miss two targets. When the SINR increases to 25 dB the classifiers under $H_{1,1}$ and $H_{1,2}$ ensure a reliable classification, whereas that for $H_{1,3}$ still misses targets. Finally, when SINR = 35 dB, the performance of all classifiers becomes excellent. As for the Hausdorff metric, in Fig. 5, the classifier for $H_{1,1}$ provides the lowest RMS errors, while that for $H_{1,3}$ is the worst and converges to a larger constant with respect to Fig. 3. In Table II, the RMSCE values with respect to the covariance class are larger than those in Table I due to the classification loss. Finally, the P_d curves are shown in Fig. 6. It turns out that detectors for $H_{1,1}$ and $H_{1,2}$ increase their detection performance with respect to the previous situation with two targets due to the increase of target energy. However, the incorrect classification results of the procedure under $H_{1,3}$ lead to a performance degradation for the related detectors.

B. Operating scenario with three clutter regions

In this subsection, we consider three clutter regions of $K_1 = K_2 = K_3 = 32$ range bins and four point-like targets. The clutter power of these regions is set according to $\text{CNR}_1 = 20$ dB, $\text{CNR}_2 = 30$ dB, and $\text{CNR}_3 = 40$ dB. Moreover, we include four targets in the 16th, 36th, 75th, and 85th range

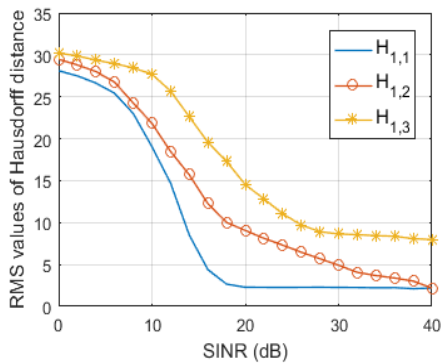


Fig. 5. Hausdorff distance versus SINR under $H_{1,i}$, $i = 1, 2, 3$ (four targets, two clutter regions).

TABLE II

RMSCE VALUES UNDER $H_{1,r}$, $r = 1, 2, 3$, FOR DIFFERENT SINRS (FOUR TARGETS, TWO CLUTTER REGIONS)

	$H_{1,1}$	$H_{1,2}$	$H_{1,3}$
SINR = 10 dB	6.7414	7.6235	7.7879
SINR = 15 dB	5.9711	6.7064	7.4718
SINR = 25 dB	5.4401	4.2819	7.1379
SINR = 35 dB	4.9620	3.8834	6.9627

bin. The values of the other parameters and initialization are the same as those in the Figs. 3-6. In this scenario, the number of considered classes becomes 6, namely,

- class 1: the generic vector of the first region does not contain any target component;
- class 2: the generic vector of the second region does not contain any target component;
- class 3: the generic vector of the third region does not contain any target component;
- class 4: the generic vector of the first region contains target components;
- class 5: the generic vector of the second region contains target components;
- class 6: the generic vector of the third region contains target components.

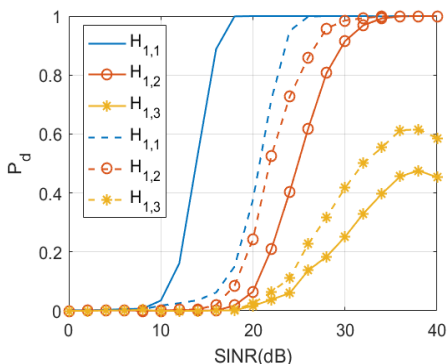


Fig. 6. P_d of the detectors (11) (dashed line) and (16) (solid line) versus SINR under $H_{1,i}$, $i = 1, 2, 3$, assuming $P_{fa} = 10^{-3}$ (four targets, two clutter regions).

TABLE III
RMSCE VALUES UNDER $H_{1,i}$, $i = 1, 2, 3$, FOR DIFFERENT SINRS (FOUR TARGETS, THREE CLUTTER REGIONS)

	$H_{1,1}$	$H_{1,2}$	$H_{1,3}$
SINR = 10 dB	4.4615	5.0104	5.4258
SINR = 15 dB	4.2706	4.5399	4.9164
SINR = 25 dB	3.6266	3.0715	4.1104
SINR = 35 dB	3.4316	2.7658	2.9159

The convergence behavior is almost the same as in the previous subsection and is not reported here for brevity. Thus, also in this case we set $m_{\max} = 5$ and $h_{\max} = 15$.

A qualitative assessment of the classification capabilities can be obtained through Fig. 3 of the supplemental material, where the classification procedure under $H_{1,1}$ turns out to be more robust than the other procedures for SINR = 15 dB. Moreover, as in the previous case, the higher the SINR, the more reliable the classification results. The RMSCE values with respect to the clutter class are reported in Table III and confirm what observed in Fig. 3 of the supplemental material from another perspective. The Hausdorff curves for this scenario are contained in Fig. 7 and exhibit a similar trend as in the previous cases. As a matter of fact, also in this case, the classification procedure under $H_{1,1}$ returns the best performance in terms of target position estimation. Finally, the corresponding P_d curves are confined to Fig. 8 where the highest P_d values are returned by the detectors for $H_{1,1}$. As in Fig. 6, also in this case, the curves of decision schemes for $H_{1,3}$ come after those of the other architectures.

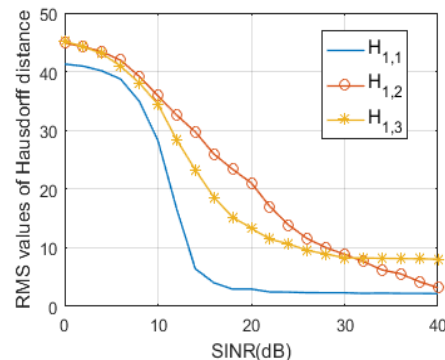


Fig. 7. Hausdorff distance versus SINR under $H_{1,i}$, $i = 1, 2, 3$ (four targets, three clutter regions).

V. CONCLUSIONS

In this paper, we have devised detection architectures dealing with multiple point-like targets in heterogeneous scenarios. At the design stage, neither the number of targets nor their positions have been assumed known as well as the clutter regions within the data window. In this context, we have devised three estimation procedures based upon different signal models. Their common denominator is the EM algorithm that has been suitably modified and/or approximated in order to come up with closed-form updates for the parameter estimates. Then,

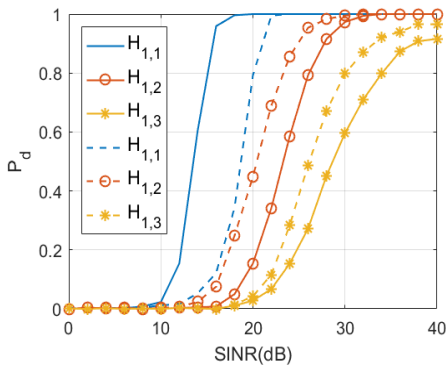


Fig. 8. P_d of the detectors (11) (dashed line) and (16) (solid line) versus SINR under $H_{1,i}$, $i = 1, 2, 3$, assuming $P_{fa} = 10^{-3}$ (four targets, three clutter regions).

such estimates have been used to implement two decision schemes relying on heuristic modifications of the GLRT. The classification and detection capabilities of the proposed architectures have been assessed over synthetic data simulating different operating scenarios with an increasing complexity in terms of the number of clutter regions and targets. This analysis has shown that the proposed architectures can provide a rather likely picture of the entire operating scenarios making the radar system aware of the surrounding environment.

Future research tracks might encompass the design of cognitive schemes that account for more specific covariance structures or that consider the joint presence of point-like as well as range-spread targets.

VI. ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant No. 61571434.

REFERENCES

- [1] A. Coluccia, D. Orlando, and G. Ricci, "A GLRT-like CFAR detector for heterogeneous environments," *Signal Processing*, vol. 194, 108401, 2022.
- [2] P. Addabbo, S. Han, D. Orlando, and G. Ricci, "Learning Strategies for Radar Clutter Classification," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1070–1082, 2021.
- [3] E. Conte, A. De Maio, and G. Ricci, "Recursive Estimation of the Covariance Matrix of a Compound-Gaussian Process and Its Application to Adaptive CFAR Detection," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1908–1915, 2002.
- [4] J. Liu, D. Massaro, D. Orlando, and A. Farina, "Radar Adaptive Detection Architectures for Heterogeneous Environments," *IEEE Trans. Signal Process.*, vol. 68, pp. 4307–4319, 2020.
- [5] E. Conte, M. Lops, and G. Ricci, "Asymptotically optimum radar detection in compound-gaussian clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 2, pp. 617–625, 1995.
- [6] E. Conte, A. De Maio, and G. Ricci, "Covariance matrix estimation for adaptive CFAR detection in compound-Gaussian clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 2, pp. 415–426, 2002.
- [7] C. Hao, D. Orlando, J. Liu, and C. Yin, *Advances in Adaptive Radar Detection and Range Estimation*. Springer, 2021.
- [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] F. Gini and M. Greco, "Suboptimum Approach to Adaptive Coherent Radar Detection in Compound-Gaussian Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 3, pp. 1095–1103, 1999.

- [10] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar: Basic Principles*. Raleigh, NC: Scitech Publishing, 2010.
- [11] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, no. 2, pp. 115–127, 1986.
- [12] 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.
- [13] 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.
- [14] 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, 2019.
- [15] W. Liu, J. Liu, C. Hao, Y. Gao, and Y.-L. Wang, "Multichannel adaptive signal detection: basic theory and literature review," *Science China Information Sciences*, vol. 65, no. 2, p. 121301, 2022.
- [16] J. Guerci, *Cognitive Radar: The Knowledge-aided Fully Adaptive Approach*, ser. Artech House radar library. Artech House, 2010.
- [17] —, *Space-Time Adaptive Processing for Radar, Second Edition*, ser. Artech House radar library. Artech House Publishers, 2014.
- [18] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2538–2541, 1999.
- [19] J. Ward, "Space-time adaptive processing for airborne radar," MIT Lincoln Laboratory, Tech. Rep., 1994.
- [20] K. D. Ward, C. J. Baker, and S. Watts, "Maritime surveillance radar. part 1: Radar scattering from the ocean surface," *Inst. Elect. Eng. Proc. F*, vol. 137, no. 2, pp. 51–62, 1990.
- [21] A. Farina, F. Gini, M. V. Greco, and L. Verrazzani, "High resolution sea clutter data: statistical analysis of recorded live data," *IEE Proc. - Radar, Sonar and Navig.*, vol. 144, no. 3, pp. 121–130, 1997.
- [22] M. Greco, F. Gini, and M. Rangaswamy, "Statistical analysis of measured polarimetric clutter data at different range resolutions," *IEE Proc. - Radar Sonar Navig.*, vol. 153, no. 6, pp. 473–481, 2006.
- [23] K. D. Ward, R. J. A. Tough, and S. Watts, *Sea Clutter, Scattering, the K distribution and radar performance*, 2nd ed. Inst of Engineering & Technology, 2013.
- [24] E. Conte and M. Longo, "Characterisation of radar clutter as a spherically invariant random process," *IEE Proc. F - Commun., Radar and Sig. Process.*, vol. 134, no. 2, pp. 191–197, 1987.
- [25] E. Conte, M. Longo, and M. Lops, "Modelling and simulation of non-rayleigh radar clutter," *IEE Proc. F - Radar and Sig. Process.*, vol. 138, no. 2, pp. 121–130, 1991.
- [26] G. Capraro, A. Farina, H. Griffiths, and M. Wicks, "Knowledge-based radar signal and data processing: a tutorial review," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 18–29, 2006.
- [27] C. T. Capraro, G. T. Capraro, I. Bradaric, D. D. Weiner, M. C. Wicks, and W. J. Baldygo, "Implementing digital terrain data in knowledge-aided space-time adaptive processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 3, pp. 1080–1099, 2006.
- [28] K. Murphy, *Machine Learning: A Probabilistic Perspective*, ser. Adaptive Computation and Machine Learning series. MIT Press, 2012.
- [29] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society (Series B - Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [30] D. Xu, P. Addabbo, C. Hao, J. Liu, D. Orlando, and A. Farina, "Adaptive strategies for clutter edge detection in radar," *Signal Processing*, vol. 186, 108127, 2021.
- [31] T. Wang, D. Xu, C. Hao, P. Addabbo, and D. Orlando, "Clutter Edges Detection Algorithms for Structured Clutter Covariance Matrices," *IEEE Signal Processing Letters*, vol. 29, pp. 642–646, 2022.
- [32] C. Richmond, "The theoretical performance of a class of space-time adaptive detection and training strategies for airborne radar," in *Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284)*, vol. 2, 1998, pp. 1327–1331 vol.2.
- [33] F. Bandiera, D. Orlando, and G. Ricci, "CFAR detection of extended and multiple point-like targets without assignment of secondary data," *IEEE Signal Processing Letters*, vol. 13, no. 4, pp. 240–243, 2006.
- [34] F. Gini, F. Bordononi, and A. Farina, "Multiple radar targets detection by exploiting induced amplitude modulation," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 903–913, 2004.
- [35] U. C. Doyuran and Y. Tanik, "Detection of multiple targets in non-Gaussian clutter," in *2008 IEEE Radar Conference*, 2008, pp. 1–5.

- [36] P. Addabbo, J. Liu, D. Orlando, and G. Ricci, "Novel Parameter Estimation and Radar Detection Approaches for Multiple Point-Like Targets: Designs and Comparisons," *IEEE Signal Processing Letters*, vol. 27, pp. 1789–1793, 2020.
- [37] P. Addabbo, S. Han, F. Biondi, G. Giunta, and D. Orlando, "Adaptive Radar Detection in the Presence of Multiple Alternative Hypotheses Using Kullback-Leibler Information Criterion-Part II: Applications," *IEEE Transactions on Signal Processing*, vol. 69, pp. 3742–3754, 2021.
- [38] 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.
- [39] —, "Model-order selection: A review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004.
- [40] H. Lütkepohl, *Handbook of Matrices*. Wiley, 1997.
- [41] L. Mirsky, "On the trace of matrix products," *Mathematische Nachrichten*, vol. 20, pp. 171–174, 1959.
- [42] D. Schuhmacher, B. Vo, and B. Vo, "A Consistent Metric for Performance Evaluation of Multi-Object Filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, 2008.
- [43] L. Yan, P. Addabbo, C. Hao, D. Orlando, and A. Farina, "New ECCM Techniques Against Noise-like and/or Coherent Interferers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 2, pp. 1172–1188, 2020.