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 on 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], [2], [3], [4], [5], [6], [7], [8], [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 Test (CUT) as well as the secondary or training data, selected through the (leading and lagging) reference window, share the same statistical characterization of the interference [10], [11], [12], [13], [14], [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 on 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], [20], [21], [22], [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

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.
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 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 on 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 potential targets [2], [30], [31], and references therein, whereas detectors of multiple point-like targets typically assume that clutter returns share the same covariance matrix [33], [34], [35], [36], [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 on 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 modifications 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 on 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

Let $z_1, \ldots, z_K$ 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 $\Omega = \{1, \ldots, K\}$ 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 $\Omega_l = \{i_1,1, \ldots, i_1K_l\}$, where $K_l > N$, $l \in L = \{1, \ldots, L\}$, 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 $\Omega_l$, although homogeneous from the standpoint of clutter statistical characterization, are heterogeneous due to target components affecting $z_{i,m}$ for some $m \in \{1, \ldots, K_l\}$. Summarizing, if data indexed by the generic $\Omega_l$ are free of target components, then $Z_l = [z_{i_1,1}, \ldots, z_{i_1K_l}] ~ CN_{K_l}(0, M_l, I)$; when there exists $\Omega^l = \{j_{l1}, \ldots, j_{lT_l}\} \subseteq \Omega_l$ indexing vectors that contain target components with $T_l$ the number of targets within $\Omega_l$, then $\forall h = 1, \ldots, T_l$,

$$z_{j_{l,h}} \sim CN_N(\alpha_{j_{l,h}}v, M_l)$$

(1)

for deterministic targets and

$$z_{j_{l,h}} \sim CN_N(0, M_l + \sigma^2_{j_{l,h}} v v^\dagger)$$

(2)

for fluctuating targets. In both cases, the $z_{j_{l,h}}$ are statistically independent of $z_{i_1,1} \sim CN_N(0, M_l)$, $\forall i_1, m \neq j_{l,h}, m = 1, \ldots, K_l$, $h = 1, \ldots, T_l$. In (1) and (2), $\alpha_{j_{l,h}} \in \mathbb{C}$ and $\sigma^2_{j_{l,h}} > 0$ 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 $v$ [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

$$H_0: \quad z_i \sim CN_N(0, M_l), \quad i \in \Omega_l, \quad l = 1, \ldots, L,$$

(3)

the alternative hypothesis for nonfluctuating targets is

$$H_{1,1}: \quad \left\{ \begin{array}{l}
\left. z_i \sim CN_N(0, M_l), \quad i \in \Omega_l \setminus \Omega_l^l, \right.
\left. z_i \sim CN_N(\alpha_i v, M_l), \quad i \in \Omega_l^l, \right.
\end{array} \right\}, \quad l = 1, \ldots, L,$$

(4)

and the alternative hypothesis for fluctuating targets can be expressed as

$$H_{1,2}: \quad \left\{ \begin{array}{l}
\left. z_i \sim CN_N(0, M_l), \quad i \in \Omega_l \setminus \Omega_l^l, \right.
\left. z_i \sim CN_N(0, M_l + \sigma^2 v v^\dagger), \quad i \in \Omega_l^l, \right.
\end{array} \right\}, \quad l = 1, \ldots, L.$$

(5)

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

$$H_{1,3}: \quad \left\{ \begin{array}{l}
\left. z_i \sim CN_N(0, M_l), \quad i \in \Omega_l \setminus \Omega_l^l, \right.
\left. z_i \sim CN_N(0, M_l + R_l), \quad i \in \Omega_l^l, \right.
\end{array} \right\}, \quad l = 1, \ldots, L,$$

(6)

where $R_l \in \mathbb{C}^{N \times N}$ is a rank one matrix representative of a component 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 \( i \in \Omega_i^f \). 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 \( \mathcal{P}_0 = \{\Omega_1, M_1 : l \in \mathcal{L}\} \), \( \mathcal{P}_{1,1} = \{\Omega_1, \Omega_1', \alpha_1, M_1 : i \in \Omega_i^f, l \in \mathcal{L}\} \), \( \mathcal{P}_{1,2} = \{\Omega_1, \Omega_1', \sigma_2^2, M_1 : i \in \Omega_i^f, l \in \mathcal{L}\} \), and \( \mathcal{P}_{1,3} = \{\Omega_1, \Omega_1', M_1, R_l : l \in \mathcal{L}\} \) the sets of unknown parameters under \( H_0, H_{1,1}, H_{1,2}, \) and \( H_{1,3} \), respectively. The joint PDF of \( \mathbf{Z} = [\mathbf{Z}_1, \ldots, \mathbf{Z}_K] \) under \( H_0 \) is

\[
\begin{align*}
    f_0(\mathbf{Z}; \mathcal{P}_0) = \frac{1}{\pi^N} \prod_{l=1}^L \exp\left(-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_l \mathbf{z}_l^\dagger]\right) \frac{\det(\mathbf{M}_l)}{};
\end{align*}
\]

under \( H_{1,1} \) it can be written as

\[
\begin{align*}
    f_1(\mathbf{Z}; \mathcal{P}_{1,1}) = \frac{1}{\pi^N} \prod_{l=1}^L \left( \prod_{i \in \Omega_i^f} \exp\left(-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_l \mathbf{z}_l^\dagger]\right) \right) \frac{\det(\mathbf{M}_l)}{};
\end{align*}
\]

under \( H_{1,2} \), it is given by

\[
\begin{align*}
    f_1(\mathbf{Z}; \mathcal{P}_{1,2}) = \frac{1}{\pi^N} \prod_{l=1}^L \left( \prod_{i \in \Omega_i^f} \exp\left(-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_l \mathbf{z}_l^\dagger]\right) \right) \frac{\det(\mathbf{M}_l)}{};
\end{align*}
\]

finally, that related to \( H_{1,3} \) is

\[
\begin{align*}
    f_1(\mathbf{Z}; \mathcal{P}_{1,3}) = \frac{1}{\pi^N} \prod_{l=1}^L \left( \prod_{i \in \Omega_i^f} \exp\left(-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_l \mathbf{z}_l^\dagger]\right) \right) \frac{\det(\mathbf{M}_l)}{};
\end{align*}
\]

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 \( \hat{R}_l, \hat{\Omega}_l, \hat{\Omega}_l', \hat{M}_l, l = 1, \ldots, L \), along with \( \hat{\sigma}_m^2, m \in \Omega_i^f \), the estimates of the respective unknown parameters, the GLRT-like rule for testing \( H_0 \) against \( H_{1,i} \) has the following form

\[
\frac{f_1(\mathbf{Z}; \hat{\mathcal{P}}_{1,i})}{f_0(\mathbf{Z}; \mathcal{P}_0)} \frac{H_{1,i}}{\eta} > \frac{H_0}{\eta},
\]

where \( \hat{\mathcal{P}}_{1,i}, i = 1, 2, 3 \), and \( \mathcal{P}_0 \) are the sets of estimates under \( H_{1,i}, i = 1, 2, 3 \), and \( H_0 \), respectively; \( \eta \) is the threshold\(^3\) to be set according to the required probability of false alarm (\( P_f \)).

Now, exploiting the MLA to come up with \( \hat{\mathcal{P}}_{1,1} \) and \( \mathcal{P}_0 \) 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, \( c_k = 1, \ldots, K \), \( k \in \{1, \ldots, K\} \). As better explained below, \( L_0 \), 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 \( L_0 + 1 \leq l \leq L_0 + L, \), Eq. 10 is replaced by

\[
\begin{align*}
    f_1(\mathbf{Z}; \hat{\mathcal{P}}_{1,i}) = \frac{1}{\pi^N} \prod_{l=1}^L \left( \prod_{i \in \Omega_i^f} \exp\left(-\text{Tr}[\mathbf{M}_l^{-1} \mathbf{z}_l \mathbf{z}_l^\dagger]\right) \right) \frac{\det(\mathbf{M}_l)}{}
\end{align*}
\]

where we set \( \alpha_{l,k} = \sigma_{l,k}^2 = 0 \). On the other hand, when \( H_0 \) is in force, we assume that \( c_k = l \), \( l \in \mathcal{L} \), implies that \( \frac{z_k}{l} \sim \mathcal{C}_N(0, \hat{R}_l) \). As a consequence, \( \mathcal{L}_c = L \) under \( H_0 \) and \( \mathcal{L}_c = 2L \) under \( H_{1,i}, i = 1, 2, 3 \).\(^4\) The corresponding sets of unknown parameters associated with the distribution of \( z_k \) are \( \mathcal{P}_0 = \mathcal{P}_0 \cup \mathcal{A} = \{\{1, \ldots, p_{L_0}\} \cup \{1, \ldots, p_{L_0}\} \cup \{1, \ldots, p_{L_0}\} \} \cup \{1, \ldots, p_{L_0}\} \) under \( H_{1,1}, \mathcal{P}_{1,2} = \mathcal{A} \cup \{\sigma_{l,k}^2 M_1 : l \in \mathcal{L}\} \) under \( H_{1,2} \), and \( \mathcal{P}_{1,3} = \mathcal{A} \cup \{M_1, R_l : l \in \mathcal{L}\} \) under \( H_{1,3} \).

\(^3\)Hereafter, \( \eta \) denotes the generic detection threshold.

\(^4\)Recall that \( \sum_{s=0}^{L} \sum_{l=1}^{L} p_{L_s+l-1} = 1 \) under \( H_{1,i}, i = 1, 2, 3 \).
With the above remarks in mind, we rewrite the PDF of $z_k$ under $H_{1,i}$, $i = 1, 2, 3$, as
\[
g_1(z_k; \mathcal{P}_{1,i,k}) = \sum_{s=0}^{L} \sum_{l=1}^{L_s} p_{Ls+l} f(z_k | c_k = Ls + l; \Theta_{s,i,k,l}),
\]
where
\[
\Theta_{s,i,k,l} = \begin{cases} 
\{\alpha_{1,k}, M_l\}, & \text{for } s = 1, i = 1, \\
\{\sigma_{2,k}^2, M_l\}, & \text{for } s = 1, i = 2, \\
\{R_s, M_l\}, & \text{for } s = 1, i = 3, \\
\{M_l\}, & \text{for } s = 0, i = 1, 2, 3.
\end{cases}
\]
As for the PDF of $z_k$ under $H_0$, it can be written as in [2] by setting $s = 0$, $\forall l = 1, \ldots, L$, and, hence, neglecting the outer summation in the previous PDFs, namely,
\[
g_0(z_k; \mathcal{P}_0) = \sum_{l=1}^{L} p_l f(z_k | c_k = l; M_l).
\]
In (13), $f(z_k | c_k = Ls + l; \Theta_{s,1,k,l})$ denotes the PDF of a complex Gaussian vector with covariance matrix $M_l$ and mean $\alpha_{1,k} \mathbf{v}$ if $s = 1$ whereas the mean is zero for $s = 0$; $f(z_k | c_k = Ls + l; \Theta_{s,2,k,l})$ is the PDF of a complex Gaussian vector with zero mean and covariance matrix $M_l + \sigma_{2,k}^2 \mathbf{v} \mathbf{v}^H$ if $s = 1$, when $s = 0$ the covariance matrix is $M_l$; $f(z_k | c_k = Ls + l; \Theta_{s,3,k,l})$ is the PDF of a complex Gaussian vector with zero mean and covariance matrix $M_l + sR_s$; finally, in (15), $f(z_k | c_k = l; M_l)$ is the complex Gaussian PDF with zero mean and covariance matrix $M_l$. Exploiting the above characterizations, given $i = 1, 2, 3$, we can also build up the following alternative test
\[
\prod_{k=1}^{K} g_1(z_k; \mathcal{P}_{1,i,k}) \frac{g_1(z_k; \mathcal{P}_{1,i,k})}{g_0(z_k; \mathcal{P}_0)} \geq \eta.
\]
Notice that 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 on 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 $Z$ for deterministic targets as follows
\[
\mathcal{L}(Z; \mathcal{P}_{1,1}) = \sum_{k=1}^{K} \log g_1(z_k; \mathcal{P}_{1,1,k})
= \sum_{k=1}^{K} \log \left( \sum_{s=0}^{L} \sum_{l=1}^{L_s} p_{Ls+l} f(z_k | c_k = Ls + l; \Theta_{s,1,k,l}) \right),
\]
where $\mathcal{P}_{1,i} = \cup_{k=1}^{K} \mathcal{P}_{1,i,k}$.

The first step of the EM is the E-step that leads to the following update rule
\[
q^{(h-1)}_k(Ls + l) = p(c_k = Ls + l | z_k; \hat{f}^{(h-1)}_{1,1,k})
= f(z_k | c_k = Ls + l; \hat{\Theta}_{s,1,k,l}) \frac{g_1(z_k; \mathcal{P}_{1,1,k})}{g_0(z_k; \mathcal{P}_0)},
\]
\[
\sum_{j=0}^{L} \sum_{l=1}^{L_s} f(z_k | c_k = Lj + l; \hat{\Theta}_{j,1,k,l}) \frac{g_1(z_k; \mathcal{P}_{1,1,k})}{g_0(z_k; \mathcal{P}_0)}
\]
s = 0, 1, where $\hat{\Theta}_{s,1,k,l}$ and $\hat{\mathcal{P}}_{1,1,k,l}$ are the estimates of the unknown parameters at the previous step. 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
\[
f(z_k | c_k = Ls + l; \hat{\Theta}_{s,1,k,l}) e^{-u(s)} \frac{g_1(z_k; \mathcal{P}_{1,1,k})}{g_0(z_k; \mathcal{P}_0)} \approx \frac{L}{L_s} \sum_{j=0}^{L} \sum_{l=1}^{L_s} f(z_k | c_k = Lj + l; \hat{\Theta}_{j,1,k,l}) e^{-u(j)} \frac{g_1(z_k; \mathcal{P}_{1,1,k})}{g_0(z_k; \mathcal{P}_0)}
\]
where, given the PDF of $z_k$, $u(j)$, $j = 0, 1$, 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
\[
\hat{f}_{1,1}^{(h)} = \arg \max_{f_{1,1}} \left\{ \sum_{k=1}^{K} \sum_{s=0}^{L} \sum_{l=1}^{L_s} q^{(h-1)}_k(Ls + l) \log f(z_k | c_k = Ls + l; \Theta_{s,1,k,l}) \right\}
\]
that is tantamount to
\[
\hat{f}_{1,1}^{(h)} = \arg \max_{f_{1,1}} \left\{ \sum_{k=1}^{K} \sum_{s=0}^{L} \sum_{l=1}^{L_s} q^{(h-1)}_k(Ls + l) \log f(z_k | c_k = Ls + l; \Theta_{s,1,k,l}) \right\}
\]
\[
+ \sum_{k=1}^{K} \sum_{s=0}^{L} \sum_{l=1}^{L_s} q^{(h-1)}_k(Ls + l) \log \mathcal{P}_{Ls}.
\]
constraint \( \sum_{s=0}^{L} p_{Ls+l} = 1 \). The final result is

\[
\hat{p}_{Ls+l}^{(k)} = \frac{1}{K} \sum_{k=1}^{K} q_{k}^{(h-1)}(Ls + l), \quad s = 0, 1, l = 1, \ldots, L. \quad (21)
\]

Thus, we can focus on the maximization problem with respect to the \( \alpha_{s,k} \)s and \( M_{l}s \), i.e.,

\[
\max_{\alpha_{s,k}} \max_{M_{l}} K \sum_{k=1}^{1} L \sum_{s=0}^{L} q_{k}^{(h-1)}(Ls + l) \times \log f(z_k | c_k = Ls + l; \Theta_{s,1,k,l}). \quad (22)
\]

Replacing the PDFs with their expressions and ignoring the terms independent of the parameters of interest, the above problem is tantamount to

\[
\min_{\alpha_{s,k}} \min_{M_{l}} K \sum_{k=1}^{L} L \sum_{s=0}^{L} q_{k}^{(h-1)}(Ls + l) \{ \log \det(M_{l}) + \text{Tr} \left[ M_{l}^{-1}(z_k - \alpha_{s,k}v)(z_k - \alpha_{s,k}v)^\dagger \right] \}. \quad (23)
\]

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 \( \alpha_{s,k} \)s are known and estimate \( M_{l} \);
- set the \( M_{l} \)s to the values obtained at the previous step and estimate the \( \alpha_{s,k} \)s.

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

\[
\min_{M_{l}} \left\{ - \log \det(M_{l}) K \sum_{k=1}^{1} L \sum_{s=0}^{L} q_{k}^{(h-1)}(Ls + l) + \text{Tr} \left[ M_{l}^{-1} K \sum_{k=1}^{1} L \sum_{s=0}^{L} q_{k}^{(h-1)}(Ls + l) \times \left( z_k - \tilde{\alpha}_{s,k}^{(h-1),(m-1)} \right) (z_k - \tilde{\alpha}_{s,k}^{(h-1),(m-1)} v) \right] \right\}. \quad (24)
\]

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

\[
\tilde{M}_{l}^{(h-1),(m)} = \frac{1}{K} K \sum_{k=1}^{K} L \sum_{s=0}^{L} q_{k}^{(h-1)}(Ls + l) \times \left( z_k - \tilde{\alpha}_{s,k}^{(h-1),(m-1)} v \right) (z_k - \tilde{\alpha}_{s,k}^{(h-1),(m-1)} v)^\dagger. \quad (25)
\]

Now, assuming that in (23) \( M_{l} = \tilde{M}_{l}^{(h-1),(m)} \), \( l = 1, \ldots, L \), we obtain

\[
\min_{\alpha_{s,k}} \max_{M_{l}} K \sum_{k=1}^{L} L \sum_{s=0}^{L} q_{k}^{(h-1)}(L + l) (z_k - \alpha_{1,k}v) \times \left( \tilde{M}_{l}^{(h-1),(m)} \right)^{-1} (z_k - \alpha_{1,k}v), \quad (26)
\]

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

\[
\tilde{\alpha}_{1,k}^{(h-1),(m)} = \frac{L}{L \sum_{l=1}^{L} q_{k}^{(h-1)}(L + l)v^\dagger \left( \tilde{M}_{l}^{(h-1),(m)} \right)^{-1} v}. \quad (27)
\]

This inner procedure continues until a convergence criterion is not satisfied. The final estimates of \( \alpha_{1,k} \) and \( M_{l} \) are denoted by \( \tilde{\alpha}_{1,k}^{(h)} \) and \( \tilde{M}_{l}^{(h)} \), 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 \( z_k \). More precisely, replacing \( \tilde{\Theta}_{s,1,k,l}^{(h-1)} \) in (18) with \( \tilde{\Theta}_{s,2,k,l}^{(h-1)} \) for model (5) and with \( \tilde{\Theta}_{s,3,k,l}^{(h-1)} \) for model (6), we come up with the respective E-steps. Let us recall here that \( \tilde{\Theta}_{s,2,k,l}^{(h-1)} \) and \( \tilde{\Theta}_{s,3,k,l}^{(h-1)} \) 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 (notice that (21) is still valid also in these cases).

1) Fluctuating Targets Model (5):

The problem to be solved has the following expression

\[
\min_{\alpha_{s,k}^{2,k}} \min_{M_{l}} \sum_{k=1}^{K} L \sum_{s=0}^{L} q_{k}^{(h-1)}(Ls + l) \times \left\{ \log \det(M_{l} + \sigma_{s,k}^{2} V) + \text{Tr} \left[ (M_{l} + \sigma_{s,k}^{2} V)^{-1} S_{k} \right] \right\}, \quad (28)
\]

where \( V = v^\dagger v \) and \( S_{k} = z_k z_k^\dagger \). The joint maximization with respect to the \( \sigma_{s,k}^{2} \)s and the \( M_{l} \)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 \( x(P_{1,2}) \) with \( P_{1,2} = \cup_{k=1}^{K} P_{1,2,k} \) and observe that it can be recast as \( x(P_{1,2}) = x_{1}(M_{1}, \ldots, M_{L}) + x_{2}(P_{1,2}) \), where

\[
x_{1}(M_{1}, \ldots, M_{L}) = \sum_{k=1}^{K} L \sum_{l=1}^{L} q_{k}^{(h-1)}(l) \{ \log \det(M_{l}) + \text{Tr} \left[ M_{l}^{-1} S_{k} \right] \}, \quad (29)
\]

\[
x_{2}(P_{1,2}) = x_{1}(M_{1}, \ldots, M_{L}) + x_{2}(P_{1,2}) = x_{1}(M_{1}, \ldots, M_{L}) + x_{2}(P_{1,2}) = \sum_{k=1}^{K} L \sum_{l=1}^{L} q_{k}^{(h-1)}(l) \{ \log \det(M_{l}) + \text{Tr} \left[ M_{l}^{-1} S_{k} \right] \}.
\]
\[ x_2(p_{1,2}) = \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \left\{ \log \det(M_l) + \sigma_{1,k}^2 V \right\} + \logdet(M_l + \sigma_{1,k}^2 V) + \text{Tr} \left[ \left( M_l + \sigma_{1,k}^2 V \right)^{-1} S_k \right]. \]  

(30)

Since \( x_1(\cdot) \) depends on the covariance matrices only, we first estimate \( M_l, l = 1, \ldots, L \), by exploiting \( x_1(\cdot) \), then replace such estimates in \( x_2(\cdot) \), and minimize it with respect to \( \sigma_{1,k}^2, k = 1, \ldots, K \). Thus, the estimates of \( M_l, l = 1, \ldots, L \), based on \( x_1(\cdot) \) are given by

\[ \tilde{M}_l^{(h)} = \left\{ \sum_{k=1}^{K} q_k^{(h-1)}(l) \right\}^{-1} \sum_{k=1}^{K} q_k^{(h-1)}(l) S_k. \]  

(31)

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

\[
\min_{\tilde{M}_l, R_l} \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \left\{ \log \det \left( \tilde{M}_l^{(h)} + \sigma_{1,k}^2 V \right) + \text{Tr} \left[ \left( \tilde{M}_l^{(h)} + \sigma_{1,k}^2 V \right)^{-1} S_k \right] \right\}.
\]

(32)

In order to solve (32), we consider a given \( k \) and the related objective function

\[
h_k(\sigma_{1,k}^2) = \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \left\{ \log \det \left( \tilde{M}_l^{(h)} + \sigma_{1,k}^2 V \right) + \text{Tr} \left[ \left( \tilde{M}_l^{(h)} + \sigma_{1,k}^2 V \right)^{-1} S_k \right] \right\} \\
= \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \log \left( 1 + \sigma_{1,k}^2 v^H \tilde{M}_l^{(h)} v \right) + \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \log \det \tilde{M}_l^{(h)} \\
- \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \sigma_{1,k}^2 v^H \tilde{M}_l^{(h)} v \left| z_k^H \tilde{M}_l^{(h)} z_k \right|^2 \\
+ \sum_{l=1}^{L} q_k^{(h-1)}(L + l) z_k^H \tilde{M}_l^{(h)} z_k.
\]

(33)

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

\[
\sum_{l=1}^{L} q_k^{(h-1)}(L + l) v^H \tilde{M}_l^{(h)} v \left( 1 + \sigma_{1,k}^2 v^H \tilde{M}_l^{(h)} v \right)^{-1} \\
+ \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \sigma_{1,k}^2 v^H \tilde{M}_l^{(h)} v \left| z_k^H \tilde{M}_l^{(h)} z_k \right|^2 \\
+ \sum_{l=1}^{L} q_k^{(h-1)}(L + l) z_k^H \tilde{M}_l^{(h)} z_k.
\]

\[
- \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \left| z_k^H \tilde{M}_l^{(h)} z_k \right|^2 \\
+ \sum_{l=1}^{L} \sigma_{1,k}^2 q_k^{(h-1)}(L + l) \left| z_k^H \tilde{M}_l^{(h)} z_k \right|^2 \\
+ \sum_{l=1}^{L} \left( 1 + \sigma_{1,k}^2 v^H \tilde{M}_l^{(h)} v \right)^{-1} \left( 1 + \sigma_{1,k}^2 v^H \tilde{M}_l^{(h)} v \right)^{-1} v^H \tilde{M}_l^{(h)} v = 0
\]

(34)

namely

\[
\sum_{l=1}^{L} q_k^{(h-1)}(L + l) \left( \sigma_{1,k}^2 q_k^{(h)} a_k^2 + a_k^2 - b_k^2 \right) = 0,
\]

(35)

where \( a_k^2 = v^H \tilde{M}_l^{(h)} v \) and \( b_k^2 = \left| z_k^H \tilde{M}_l^{(h)} z_k \right|^2 \). 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 (33).

2) Fluctuating Targets Model (6): Let us assume that \( H_{1,3} \) (see (6)) is true and recall that \( R_k \) is positive semidefinite of rank 1. The maximization step consists in solving

\[
\min_{M_l, R_l} \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \left\{ \log \det(M_l + s R_l) + \text{Tr} \left[ (M_l + s R_l)^{-1} z_k z_k^H \right] \right\}.
\]

(36)

Notice that the objective function can be further rewritten as

\[
\sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(l) \left\{ \log \det(M_l) + \text{Tr} \left[ (M_l)^{-1} z_k z_k^H \right] \right\} + \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(L + l) \left\{ \log \det(M_l + R_l) + \text{Tr} \left[ (M_l + R_l)^{-1} z_k z_k^H \right] \right\}.
\]

(37)

Thus, in order to minimize the above function with respect to \( M_l \) and \( R_l \), let us consider the generic index \( l \) and the corresponding term

\[
q_l \log \det(M_l) + \text{Tr} \left[ (M_l)^{-1} S_l \right] \\
+ q_l + L \log \det(M_l + R_l) + \text{Tr} \left[ (M_l + R_l)^{-1} S_{l+L} \right],
\]

(38)

where

\[
q_l = \sum_{k=1}^{K} q_k^{(h-1)}(l), \quad S_l = \sum_{k=1}^{K} q_k^{(h-1)}(l) z_k z_k^H, \quad q_{l+L} = \sum_{k=1}^{K} q_k^{(h-1)}(l + L) z_k z_k^H.
\]

(39)

Then, define \( A_l = M_l^{1/2} U_l \) with \( U_l \in \mathbb{C}^{N \times N} \) the unitary matrix whose columns are the eigenvectors of \( M_l^{-1/2}(M_l + \)
\( R_t \) \( M_t^{-1/2} \) and \( \Lambda_t = \text{diag} (\lambda_{t,1}, \ldots, 1) \), \( \lambda_{t,1} > 1 \), the diagonal matrix containing the eigenvalues of \( M_t^{-1/2} (M_t + R_t) M_t^{-1/2} \). It clearly follows that \( M_t = A_t \Lambda_t A_t^\dagger \), \( M_t + R_t = A_t \Lambda_t A_t^\dagger \), and (38) can be written as

\[
2(q_l + q_l + L) \log |\det(A_l)| + \text{Tr} [A_l^{-1} S_l A_l^{-\dagger}] \\
+ q_l + L \log \det(A_l) + \text{Tr} [A_l^{-1} A_l S_l + L A_l^{-\dagger}].
\] (41)

Now, let us exploit the eigendecomposition of \( S_t^{-1/2} S_{t+L} S_t^{-1/2} = V_t \Gamma_t V_t^\dagger \) where \( \Gamma_t = \text{diag} (\gamma_{t,1}, \ldots, \gamma_{t,N}) \), \( \gamma_{t,1} \geq \ldots \geq \gamma_{t,N} \geq 0 \), is the diagonal matrix of the eigenvalues and \( V_t \in \mathbb{C}^{N \times N} \) is the unitary matrix of the corresponding eigenvectors. Denoting by \( B_l = S_l^{1/2} V_l \), we obtain that \( S_l = B_l B_l^\dagger \), \( S_{t+L} = B_l \Gamma_l B_l^\dagger \), and (41) can be recast as

\[
c_l \left[ \log |\det(B_l)| - \frac{1}{2} \log \det(A_l) - \log |\det(X_l)| \right] \\
+ \text{Tr} [X_l^\dagger \Lambda_l X_l] + q_l + L \log \det(A_l) + \text{Tr} [X_l^\dagger \Lambda_l X_l].
\] (42)

where \( c_l = 2(q_l + q_l + L) \) and \( X_l = \Lambda_l^{-2/3} A_l^{-1} B_l \). Exploiting the singular value decomposition of \( X_l \), given by \( X_l = T_l D_l Q_l \), \( D_l = \text{diag} (d_{l,1}, \ldots, d_{l,N}) \) and \( d_{l,1} \leq \ldots \leq d_{l,N}, \) the last equation becomes

\[
c_l \left[ \log |\det(B_l)| - \frac{1}{2} \log \det(A_l) - \log |\det(D_l)| \right] \\
+ \text{Tr} [D_l^2 T_l^\dagger A_l T_l] + q_l + L \log \det(A_l) + \text{Tr} [D_l^2 Q_l \Gamma_l Q_l].
\] (43)

By Theorem 1 of [41], the minimization with respect to \( T_l \) and \( Q_l \) leads to

\[
c_l \left[ \log |\det(B_l)| - \frac{1}{2} \log \det(A_l) - \log |\det(D_l)| \right] \\
+ \text{Tr} [D_l^2 T_l^\dagger A_l T_l] + q_l + L \log \det(A_l) + \text{Tr} [D_l^2 T_l^\dagger T_l] \\
= c_l \log |\det(B_l)| + a_l \log \lambda_{l,1} + d_{l,1}^2 (\lambda_{l,1} + \gamma_{l,1}) \\
+ \sum_{i=2}^N d_{l,i}^2 (1 + \gamma_{l,i}) - \frac{c_l}{2} \sum_{i=1}^N \log d_{l,i}^2,
\] (44)

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

\[
\hat{d}_{l,i}^2 = \begin{cases} 
\frac{c_l}{N \lambda_{l,1} + \gamma_{l,1}} & \text{if } i = 1, \\
\frac{c_l}{\lambda_{l,i} + \gamma_{l,i}} & \text{if } i > 1.
\end{cases}
\] (45)

Finally, replacing \( \hat{d}_{l,i}^2 \) with \( \hat{d}_{l,i}^2 \) in (44) yields

\[
c_l \log |\det(B_l)| + a_l \log \lambda_{l,1} + N \frac{c_l}{2} \\
- \frac{c_l}{2} \log \lambda_{l,1} + \gamma_{l,1} - \frac{c_l}{2} \sum_{i=2}^N \log \frac{c_l}{1 + \gamma_{l,i}}.
\] (46)

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

\[
\hat{\lambda}_{l,1} = \max \left\{ -2a_l \gamma_{l,1}, 2a_l + c_l \right\}
\] (47)

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 \( \forall k = 1, \ldots, K \), we proceed as follows

\[
\begin{cases}
\mathbb{C}^N \left( 0, \hat{\mathcal{M}}_{k_{\hat{k}}}^{(\max)} \right), & 1 \leq \hat{k} \leq L, \\
\mathbb{C}^N (\hat{\sigma}_{k}^{(\max)} v, \hat{\mathcal{M}}_{k_{\hat{k}}-L}^{(\max)}), & L + 1 \leq \hat{k} \leq 2L,
\end{cases}
\] (48)

under \( H_{1,1} \),

\[
\begin{cases}
\mathbb{C}^N \left( 0, \hat{\mathcal{M}}_{k_{\hat{k}}}^{(\max)} \right), & 1 \leq \hat{k} \leq L, \\
\mathbb{C}^N (0, \hat{\mathcal{M}}_{k_{\hat{k}}-L}^{(\max)} + \hat{\sigma}_{k}^{(\max)} v v^\dagger), & L + 1 \leq \hat{k} \leq 2L,
\end{cases}
\] (49)

under \( H_{1,2} \), and

\[
\begin{cases}
\mathbb{C}^N \left( 0, \hat{\mathcal{M}}_{k_{\hat{k}}-L}^{(\max)} \right), & 1 \leq \hat{k} \leq L, \\
\mathbb{C}^N (0, \hat{\mathcal{M}}_{k_{\hat{k}}-L}^{(\max)} + \hat{R}_{k_{\hat{k}}-L}^{(\max)}) , & L + 1 \leq \hat{k} \leq 2L,
\end{cases}
\] (50)

under \( H_{1,3} \), where \( \hat{\sigma}_{k}^{(\max)} \) is the maximum number of EM iterations and

\[
\hat{k} = \arg \max_{l = 1, \ldots, 2L} q_{k_{l}}^{(\max)} (l).
\] (51)

When the proposed decision schemes decide for \( H_{1,1} \), \( i = 1, 2, 3 \), 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.

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 \( C_l = \sigma_{c,l}^2 M_c, l = 1, \ldots, L \), 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), while the \((i,j)\)-th element of \( M_c \) is equal to \( \rho_{l}^{(i,j)\rightarrow(i,j)} \) with \( \rho_{l} = 0.9 \). 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 \( \text{SINR} = |\sigma_{k}^{(\max)} v^\dagger M_{c}^{-1} v | \), whereas the SINR for fluctuating targets is \( \text{SINR} = \sigma_{k}^{(\max)} / (\sigma_{c}^{2} + \sigma_{n}^{2}) \), where \( M_c = C_l + \sigma_{n}^2 I \) with
Algorithm 1: Workflow of the Proposed Cognitive Procedures.

**Input:** $L$, $Z$, $v$

**Output:** $\Omega_l$, $\Omega_l^i$, $l = 1, \ldots, L$, $\hat{P}_{1,i,k}$, $i = 1, 2, 3$, $k = 1, \ldots, K$, detected hypothesis

1. **Latent Variable Model**: introduce the hidden random variables $\xi_l$ with unknown PMF, which represent not only the clutter types but also the presence of a possible target within $z_k$, $k = 1, \ldots, K$;

2. **Initialization**: initialize the parameters of the recursive procedures, namely $\hat{P}_{1,i,k}$, $i = 1, 2, 3$, $k = 1, \ldots, K$; as well as those of the stopping criterion, such as $\max_m$ and $\max_{m_{max}}$;

3. **E-step**: compute the conditional expectation of the log-likelihood of $z_k$ at the $(h-1)$th iteration, obtaining the estimation update of the posterior probability $q_k^{(h-1)}(Ls + l)$;

4. **M-step**: maximize the log-likelihood and provide closed-form updates for $\hat{P}_{1,i,k}$, $i = 1, 2, 3$, $k = 1, \ldots, K$;

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 $q_k^{(h_{max})}(l)$, $l = 1, \ldots, L$, $k = 1, \ldots, K$; 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 $\Omega_l$ and $\Omega_l^i$, $l = 1, \ldots, L$;

7. **Detection**: the fully adaptive detector is obtained by replacing $\Omega_l$, $\Omega_l^i$, $\hat{P}_{1,i,k}$, $l = 1, \ldots, L$, $k = 1, \ldots, K$, in (11) and (16) under the signal-plus-noise hypothesis and similarly under the noise-only hypothesis.

\[
\sigma_n^2 = 1 \text{ the thermal noise power, } \alpha_{k,l} \text{ and } \sigma_{k,l}^2 \text{ are the amplitude and power associated with a target in the } k \text{th range bin of the } l \text{th region, the nominal steering vector is computed assuming that the AoAs are zero in all scenarios. The penalty term in (18) is set as } u(s) = (N^2 + k_{1,l}s)(1 + \rho)/4 \text{ with } \rho = 3, \text{ and } k_{1,1} = 2 \text{ (under } H_{1,1}), k_{1,2} = 1 \text{ (under } H_{1,2}), \text{ and } k_{1,3} = N \text{ (under } H_{1,3}), \text{ that is borrowed from the generalized information criterion with the parameter equal to } 3 \text{ [39].}
\]

A. Operating Scenario With Two Clutter Regions

Let us first 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 $M_1$ and $M_2$ with CNR$_1 = 20$ dB, 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 $\hat{P}_{1,i}^{(0)} = 1/Lc_i$, $l = 1, \ldots, L_c$; the initial value of $M$, namely $\hat{M}^{(0)}_1$, is generated in the same way as in Section IV-A of [2]. A possible choice for $\hat{a}_s^{(0)}$, $s = 0, 1, k = 1, \ldots, K$, under $H_{1,1}$ is

\[
\hat{a}_{s,k}^{(0)} = \begin{cases} 0, & s = 0, \\
\max_{l=1,\ldots,L} \left( \frac{|v^1(M_l^{(0)})^{-1}z_k|}{v^1(M_l^{(0)})^{-1}v} \right), & s = 1. \end{cases}
\]

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

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 $\Delta L_i(h) = [\mathcal{L}(Z; \hat{P}_{1,i}^{(h)}) - \mathcal{L}(Z; \hat{P}_{1,i}^{(h-1)})] / \mathcal{L}(Z; \hat{P}_{1,i}^{(h)})$, $i = 1, 2, 3$. Then, the inner cyclic estimation procedure under $H_{1,1}$ terminates when the convergence criterion $\Omega(m) = \left[ \hat{M}_l^{(h-1)} - \hat{M}_l^{(h-1)-(m-1)} \right] / ||\hat{M}_l^{(h-1)} - \hat{M}_l^{(h-1)-(m-1)}|| < \delta$ 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 convergence and computational load. In Fig. 1, where we plot the mean curves (over 1000 independent trials) of $\Delta L_1$ (namely, under $H_{1,1}$) versus $h$ for different values of $m_{max}$ and assuming SINR = 20 dB, confirms that $m_{max} = 5$ returns a relative variation of $\Delta L$ less than $10^{-4}$ when $h > 5$. In Fig. 2(a), the mean values of $\Delta L_i(h)$, $i = 1, 2, 3$, are plotted resorting to 1000 independent runs for 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 $\Delta L_i(h)$, $i = 1, 2, 3$, versus $h$ (number of iterations of the EM procedure) for SINR = 20 dB, (b) Average values of $\Delta L_i(15)$, $i = 1, 2, 3$, versus SINR (two targets, two clutter regions).

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 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 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 $\hat{\mathbf{r}}$ and $\mathbf{r}$. Precisely, $\hat{\mathbf{r}}$ is a vector of size $K$ that contains 0 except for the range bins classified as target, whereas $\mathbf{r}$ 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.

In Fig. 4, we evaluate the Probability of Detection ($P_d$) returned by detectors\textsuperscript{7} (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 those 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 $H_{1,1}$ (11) because the former exploits 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

\begin{table}[h]
\centering
\caption{RMSCE values under $H_{1,i}$, $i = 1, 2, 3$, for different SINRs (two targets, two clutter regions).}
\begin{tabular}{|c|c|c|c|}
\hline
SINR & $H_{1,1}$ & $H_{1,2}$ & $H_{1,3}$ \\
\hline
10 dB & 4.5560 & 4.7304 & 4.7452 \\
15 dB & 4.1490 & 4.3380 & 4.3906 \\
25 dB & 3.7114 & 3.0842 & 2.9630 \\
35 dB & 3.4347 & 2.7841 & 2.6029 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{7}Notice 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.
increases, the classification capabilities due to the (four targets, two clutter regions). To illustrate this, the Hausdorff distance is shown in Fig. 5. In this figure, the RMSCE values with respect to the classification loss are reported in Table II for different SINRs.

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

B. Operating Scenario With Three Clutter Regions

In this subsection, 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 CNR$>_1 = 20$ dB, CNR$>_2 = 30$ dB, and CNR$>_3 = 40$ dB. Moreover, we include four targets in the 16th, 36th, 75th, and 85th range 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.

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 $H_{1,1}$ is 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. In this scenario, the number of considered classes becomes 6, namely,

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

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.

V. CONCLUSION

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 on different signal models. Their common denominator is the EM algorithm that has been suitably modified and/or approximated in order to come up with a rather likely picture of the proposed architectures can provide a rather likely picture of the

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.

REFERENCES

[1] A. Coluccia, D. Orlando, and G. Ricci, “A GLRT-like CFAR detector for heterogeneous environments,” Signal Process., vol. 194, 2022, Art. no. 108401.

[2] P. Addabbo, S. Han, D. Orlando, and G. Ricci, “Learning strategies for Radar Clutter Classification,” IEEE Trans. Signal Process., 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 Trans. Signal Process., vol. 50, no. 8, pp. 1908–1915, Aug. 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] G. R. Ricci, M. Lops, and G. Ricci, “Asymptotically optimum radar detection in compound-Gaussian clutter,” IEEE Trans. Aerosp. Electron. Syst., vol. 31, no. 2, pp. 617–625, Apr. 1995.

[6] E. Conte, A. De Maio, and G. Ricci, “Covariance matrix estimation for adaptive CFAR detection in compound-gaussian clutter,” IEEE Trans. Aerosp. Electron. Syst., vol. 38, no. 2, pp. 415–426, Apr. 2002.

[7] C. Hao, D. Orlando, J. Liu, and C. Yin, Advances in Adaptive Radar Detection and Range Estimation. Berlin, Germany:Springer, 2021.

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

[9] F. Gini and M. Greco, “Suboptimum approach to adaptive coherent radar detection in compound-gaussian clutter,” IEEE Trans. Aerosp. Electron. Syst., vol. 35, no. 3, pp. 1095–1104, Jul. 1999.

[10] M. A. Richards, J. A. Scheer, and W. A. Holm, Principles of Modern Radar: Basic Principles. Raleigh, NC, USA: SciTech Publishing, 2010.

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

[12] P. Addabbo, D. Orlando, and G. Ricci, “Advanced Radar Detection Schemes Under Mismatched Signal Models,” in Synthesis Lectures on Signal Processing. San Rafael, CA, USA: Morgan & Claypool Publishers, 2009.

[13] F. Bandiera, D. Orlando, and G. Ricci, “Adaptive CFAR matched filter detector,” IEEE Trans. Aerosp. Electron. Syst., vol. 28, no. 1, pp. 208–216, Jan. 1992.

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

[15] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, “A CFAR adaptive matched filter detector,” IEEE Trans. Aerosp. Electron. Syst., vol. 47, no. 1, pp. 208–216, Jan. 1992.

[16] F. Bandiera, D. Orlando, and G. Ricci, “Advanced Radar Detection Schemes Under Mismatched Signal Models,” in Synthesis Lectures on Signal Processing. San Rafael, CA, USA: Morgan & Claypool Publishers, 2009.

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

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

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

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

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

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

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

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

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

[26] F. Gini and A. Farina, “Vector subspace detection in compound-gaussian clutter Part I: Survey and new results,” IEEE Trans. Aerosp. Electron. Syst., vol. 38, no. 4, pp. 1295–1311, Oct. 2002.
Linjie Yan received the B.E. degree in communication engineering from the Shandong University of Science and Technology, Qingdao, China, and the Ph.D. degree in signal and information processing from the Faculty of Information Technologies, Czech Technical University in Prague, Prague, Czech Republic. She is currently an Assistant Professor with the College of Information Science and Technology, China University of Petroleum, Qingdao, China. Her research interests include statistical signal processing, radar signal processing, and machine learning.

Sudan Han received the B.S. degree in information engineering and the M.S. and Ph.D. degrees in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2012, 2015, and 2019, respectively. She is currently an Assistant Researcher with the Beijing Institute of Electronics, Beijing, China. Her research interests include statistical signal processing, radar signal processing, and machine learning.

Chengpeng Hao (Senior Member, IEEE) received the B.S. and M.S. degrees in electronic engineering from Beijing Broadcasting Institute, Beijing, China, in 1998 and 2001, respectively, and the Ph.D. degree in signal and information processing from the Institute of Acoustics, Chinese Academy of Sciences (IACAS), Beijing, China, in 2004. He is currently a Professor with IACAS. From 2013 to 2014, he has held a visiting position with the Department of Electrical and Computer Engineering, Queens University, Kingston, ON, Canada. He has authored or coauthored more than 100 journal and conference papers. His research interests include statistical signal processing with more emphasis on adaptive sonar and radar signal processing. He is currently an Associate Editor for several international journals, including the IEEE ACCESS, Signal, Image and Video Processing, Springer, and the EURASIP Journal on Advances in Signal Processing, Springer.

Danilo Orlando (Senior Member, IEEE) was born in Gagliano del Capo, Italy, in 1978. He received the Dr. Eng. degree (with Honors) in computer engineering and the Ph.D. degree (with maximum score) in information engineering from the University of Salento (formerly University of Lecce), Lecce, Italy, in 2004 and 2008, respectively. From July 2007 to July 2010, he was with the University of Cassino, Cassino, Italy, engaged in a research project on algorithms for track-before-detect of multiple targets in uncertain scenarios. From September to November 2009, he was a Visiting Scientist with the NATO Undersea Research Centre, La Spezia, Italy. From September 2011 to April 2015, he was with Elettronica S.p.A., engaged as System Analyst in the field of electronic warfare. In May 2015, he joined the University of Salento, Lecce, Italy, where he is currently an Associate Professor. In 2007, he has held visiting positions with the Department of Avionics and Systems, ENSICA (now Institut Supérieur de l'Aéronautique et de l'Espace, ISAE), Toulouse, France, and from 2017 to 2019, with the Chinese Academy of Science, Beijing, China. He has authored or coauthored more than 170 scientific publications in international journals, conferences, and books. His main research interests include statistical signal processing with more emphasis on adaptive detection and tracking of multiple targets in multitemporal scenarios. He was the Senior Area Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and currently an Associate Editor for the IEEE OPEN JOURNAL ON SIGNAL PROCESSING, EURASIP JOURNAL ON ADVANCES IN SIGNAL PROCESSING, AND REMOTE SENSING.

Giuseppe Ricci (Senior Member, IEEE) was born in Naples, Italy, in 1964. He received the Dr. and Ph.D. degrees in electronic engineering from the University of Naples Federico II, Naples, Italy, in 1990 and 1994, respectively. Since 1995, he has been with the University of Salento (former University of Lecce), Lecce, Italy, first as an Assistant Professor of telecommunications and since 2002 as a professor. He has held visiting positions with the University of Colorado, Boulder, CO, USA, from 1997 to 1998 and in April/May 2001, with Colorado State University, Fort Collins, CO, USA, in July/September 2003, March 2005, September 2009, and March 2011, with Ensica, Toulouse, France, in March 2006, with the University of Connecticut, Storrs, CT, USA, in September 2008, and with the Arizona State University, Tempe, AZ, USA, from July to August 2019. His research interests include statistical signal processing with more emphasis on radar processing, localization algorithms, and CDMA systems. From 2012 until 2016, he was a Member of the Multisensor Target Tracking Working Group, operating under the auspices of the International Society for Information Fusion, and from 2016 until 2021, he has been a Member of the Technical Area Committee in Signal Processing for Multisensor Systems of Eurosip. From February 2016 and June 2016, until April 2017, he was also a Rector Delegate for teaching activities and continuing education.