Adaptive Radar Detection in Heterogeneous Clutter Plus Thermal Noise via the Expectation-Maximization Algorithm

ANGELO COLUCCIA, Senior Member, IEEE
ALESSIO FASCISTA, Member, IEEE
Università del Salento, Lecce, Italy

DANILO ORLANDO, Senior Member, IEEE
Università degli Studi "Niccolò Cusano", Roma, Italy

GIUSEPPE RICCI, Senior Member, IEEE
Università del Salento, Lecce, Italy

This article addresses adaptive radar detection of N pulses coherently backscattered by a prospective target in heterogeneous disturbance. As customary $K \geq N$ range cells adjacent to the one under test are used for estimation purposes. The disturbance in each range cell is described by a non-Gaussian model based on a mixture of $L < K$ Gaussian distributions. Gaussian components are characterized by an unknown low-rank matrix plus thermal noise with unknown power level. We first derive a detector inspired by the generalized likelihood ratio test that adaptively estimates the statistical properties of the disturbance from the observed data. To overcome the intractability of the involved maximum-likelihood estimation problem, a suitable approximate strategy based on the expectation-maximization algorithm is developed. This also allows us to classify the cell under test by selecting the "maximum a posteriori Gaussian distribution" for the disturbance (under both hypotheses). Accordingly, a likelihood ratio test is also proposed. An extensive performance analysis, conducted on synthetic data as well as on two different experimental datasets (PhaseOne and IPIX for land and sea radar returns, respectively), shows that the proposed approaches outperform state-of-the-art competitors in terms of both detection capabilities and false alarms control.

Manuscript received 26 April 2023; revised 14 September 2023; accepted 27 September 2023. Date of publication 9 October 2023; date of current version 9 February 2024.

DOI. No. 10.1109/TAES.2023.3322389

Refereeing of this contribution was handled by Igal Bilik.

Authors' addresses: Angelo Coluccia, Alessio Fascista, and Giuseppe Ricci are with the Dipartimento di Ingegneria dell'Innovazione, Università del Salento, 73100 Lecce, Italy, E-mail: (angelo.coluccia@unisalento.it; alessio.fascista@unisalento.it; giuseppe.ricci@unisalento.it); Danilo Orlando is with Università degli Studi "Niccolò Cusano," 00166 Roma, Italy, E-mail: danilo.orlando@unicusano.it. *(Corresponding author: Angelo Coluccia.)*

I. INTRODUCTION

In the last decades, the design of adaptive decision schemes capable of detecting coherent targets buried in Gaussian and non-Gaussian noise has attracted a great interest in the signal processing community. In a seminal article, Kelly [1] used the generalized likelihood ratio test (GLRT) to design an adaptive decision scheme aimed at detecting coherent pulse trains in the presence of Gaussian disturbance with an unknown covariance matrix. Therein, it is assumed that a set of secondary data, free of signal components, but sharing the *same statistical characterization* of the overall interference (i.e., clutter, thermal noise, and possible noise-like jammers) in the cell under test (CUT), is available (*homogeneous* environment). To reasonably meet this condition, secondary data are usually picked from a window of range cells adjacent to the CUT. Building upon such a pioneering work, a plethora of detectors have been designed by following procedures that include statistical tests with modified hypotheses, asymptotic arguments, approximations, and ad hoc strategies [2], [3], [4], [5], [6].

However, experimental data [7], [8], [9], [10], [11] have demonstrated that the Gaussian assumption is not always valid; in particular, in high-resolution radars (especially at low grazing angles), non-Gaussian clutter is generally modeled as a compound-Gaussian process that, when observed on sufficiently short time intervals, degenerates into a spherically invariant random process (SIRP), namely the product of a random variable (RV) (texture) times a Gaussian process [12], [13]. An asymptotically optimum approximation of the GLRT to detect a coherent signal when the disturbance is modeled in terms of an SIRP has been derived in [14]. Therein the covariance matrix of the disturbance is supposed to be known at the design stage and the corresponding detector is commonly referred to as the normalized matched filter (NMF). Remarkably, in clutter-dominated, non-Gaussian environments such a detector possesses the constant false alarm rate (CFAR) property with respect to the probability density function (PDF) of the texture. Adaptive versions of the aforementioned NMF (ANMF) can be obtained when secondary data are available by means of the estimate-and-plug paradigm [15], [16]. Specifically, a sample covariance matrix based on normalized secondary data can be used to estimate the unknown covariance matrix if the secondary data share the same covariance matrix of the disturbance in the CUT up to possibly *different power levels* (the so-called *heterogeneous* environment) [17], [18]. Such a detector inherits the CFAR property with respect to the power levels from the NMF, but it is not CFAR with respect to the structure of the clutter covariance matrix (even in clutter-dominated environments). It is commonly referred to as $\boldsymbol{\Sigma}$-ANMF. Recursive estimators for the structure of the clutter covariance matrix have been proposed in [19], [20], [21], and [22] based on secondary data drawn from a heterogeneous environment; plugging such

estimators into the NMF in place of the unknown covariance matrix can guarantee a distribution of the decision statistic under the noise-only hypothesis independent of both the power levels and the structure of the clutter covariance matrix, provided that the environment is clutter dominated. An approximation of the GLRT for heterogeneous environments has been proposed in [23]. Interestingly, exploiting a specific initialization, it also guarantees the CFAR property in clutter-dominated environments with respect to all of the unknown parameters and may guarantee better performance than the previously proposed estimate-and-plug solutions. Finally, in [24], adaptive detectors for point-like targets in heterogeneous scenarios have been devised by resorting to the so-called directional statistics and the expectation-maximization (EM) algorithm [25], [26].

All the aforementioned contributions share the common assumption that the structure of the disturbance covariance matrix remains unaltered over the entire extent (in space) of the secondary data window. Actually, it depends on the spatial distribution of the clutter point scatterers, and in practice, it is not guaranteed that such a distribution is invariant with respect to the range. In fact, heterogeneous terrains in the form of variations in topology, land cover variations, and land–sea interfaces might lead to different structures for the disturbance covariance matrix [27]. Accounting for these variations at the design stage and exploiting the maximum likelihood (ML) approach yield optimization problems that, from a mathematical point of view, become very difficult and it is not guaranteed to obtain closed-form expressions for their solutions. For this reason, suboptimum approaches have been pursued such as in [28] and [29] where the EM algorithm has been used to address the direction of arrival estimation problem and clutter clusterization in the presence of heterogeneities (i.e., clutter returns characterized by possibly different covariance structures). Finally, the EM algorithm has been employed in [30] to devise detection architectures capable of classifying the range bins according to their clutter properties and detecting possible multiple targets whose positions and number are unknown. Therein, the covariance matrix of each range cell is assumed to belong to a finite set of different classes.

With the aforementioned remarks in mind, in this article, we address the detection of a coherent point-like target under a new heterogeneous scenario where both clutter power levels and covariance structures are modeled as parameters that can vary over the range. Specifically, we formulate the detection problem assuming that in each range cell the disturbance components are non-Gaussian but can be modeled as a mixture of Gaussian terms, each with zero-mean and covariance matrix given by one out of L low-rank covariance matrices, representative of classes of different heterogeneous clutter (e.g., originating from land, sea, etc.), plus thermal noise with unknown power. To the best of authors' knowledge, these more general assumptions have not yet been investigated in the context of target detection in heterogeneous environments. We then provide the following contributions.

1) We formulate the binary hypothesis testing problem using a latent variables model: hidden random variables (weights of the mixture model) are introduced to specify in a probabilistic manner the clutter in each range cell. For this problem, the GLRT-based detector that adaptively estimates the statistical properties of the disturbance is derived.
2) To overcome the intractability of the ML estimation problem stemming from the GLRT approach, we propose a suitable strategy based on the EM algorithm, which provides approximate ML estimates of the unknown parameters under both hypotheses and efficiently solve the detection problem. The outcome of the EM is also used to derive an alternative decision scheme based on the likelihood-ratio test (LRT) fed by the maximum a posteriori (MAP) classification of the disturbance distribution in the CUT, i.e., a detector that decides based on the most probable environment at hand.
3) We conduct an extensive performance analysis to test the effectiveness of the proposed detectors in terms of detection capabilities, false alarms control, and robustness to mismatches on the assumed non-Gaussian heterogeneous disturbance model. The algorithms are validated both on synthetic data as well as on two different experimental radar datasets (PhaseOne and IPIX) including different types of heterogeneous clutter (land and sea, respectively), and also compared against the state-of-the-art detectors. Results show that the proposed strategies guarantee noticeable advantages in terms of detection performance for a preassigned value of the false alarm probability over state-of-the-art methods. Moreover, on real data, the proposed strategies tend to guarantee values of the false alarm rate closer to the nominal one than all the considered competitors except for one of them that, however, experiences an important detection loss with respect to the proposed approaches.

The rest of this article is organized as follows. Section II deals with the problem formulation, while Section III is devoted to the design of the newly proposed decision schemes relying on the EM algorithm. Section IV assesses the performance of the proposed algorithms also in comparison with natural competitors over synthetic and real recorded data. Finally, Section V concludes this article.

Notation: Vectors and matrices are denoted by boldface lower case and upper case letters, respectively. Symbols $\det(\cdot)$, $(\cdot)^{-1}$, $(\cdot)^T$, and $(\cdot)^\dagger$ denote the determinant, inverse, transpose, and conjugate transpose, respectively. As to numerical sets, $\mathbb{C}$ is the set of complex numbers, $\mathbb{C}^{N\times M}$ is the Euclidean space of $(N \times M)$-dimensional complex matrices, and $\mathbb{C}^N$ is the Euclidean space of N-dimensional complex vectors. The nth entry of the vector $\boldsymbol{x}$ is denoted by $[\boldsymbol{x}]_n$. The symbol $\mathbf{0}$ denotes a matrix of zeros of proper dimensions, while the identity matrix of size $N \times N$ is indicated by $\boldsymbol{I}_N$. The acronym RV means random variable, while

IID means independent and identically distributed. The set function P denotes a probability measure. We write $\boldsymbol{x} \sim \mathcal{CN}_N(\boldsymbol{0}, \boldsymbol{M})$ if $\boldsymbol{x} \in \mathbb{C}^N$ is a complex normal (N-dimensional) vector with zero mean and (Hermitian) positive definite covariance matrix $\boldsymbol{M} \in \mathbb{C}^{N\times N}$.

II. PROBLEM FORMULATION AND DEFINITIONS

The general radar detection problem for a coherent point-like target can be formulated as the following hypothesis test:

$$\begin{cases} H_0: & \boldsymbol{z} = \boldsymbol{n}, & \boldsymbol{z}_k = \boldsymbol{n}_k, & k = 1, \ldots, K \\ H_1: & \boldsymbol{z} = \alpha \boldsymbol{v} + \boldsymbol{n}, & \boldsymbol{z}_k = \boldsymbol{n}_k, & k = 1, \ldots, K \end{cases} \tag{1}$$

where $\boldsymbol{z} \in \mathbb{C}^N$ is the vector of samples from the CUT, $\boldsymbol{v} \in \mathbb{C}^N$ is the known (space, time, or space-time) steering vector, $\alpha \in \mathbb{C}$ is an unknown parameter accounting for channel propagation, radar cross section of the target, etc., the $\boldsymbol{z}_k$s are secondary data, $K \geq N$, and $\boldsymbol{n}, \boldsymbol{n}_1, \ldots, \boldsymbol{n}_K$ are the disturbance terms. As customary, secondary data are taken from a window of cells adjacent in range to the CUT, and are supposed free of signal components.

We resort to hidden random variables (latent variable model [26]) to identify the covariance matrix modeling the clutter components in each range cell. As a matter of fact, we introduce the RV c_k for the kth secondary range cell (and c for the CUT) that takes on the value l with probability p_l, $l = 1, \ldots, L$, namely $P(c_k = l) = p_l$ ($P(c = l) = p_l$), and $c_k = l$ ($c = l$) implies that the N-dimensional noise vector $\boldsymbol{n}_k$ ($\boldsymbol{n}$) is a complex normal vector with covariance matrix

$$\boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N \in \mathbb{C}^{N\times N}. \tag{2}$$

In the last equation, $\boldsymbol{R}_l$ is an unknown low-rank[1] positive semidefinite matrix with known rank $r < N$, and $\sigma_w^2 > 0$ is the unknown level of the thermal noise. Due to the inherent complexity of the considered problem, we suppose that L is less than the number K of secondary data. In addition, we suppose that $c, c_1, \ldots, c_K$ are IID RVs.

It follows that $\boldsymbol{n}$ is modeled by a multivariate contaminated normal distribution, i.e., a convex mixture model of Gaussian terms, each conditioned on one of the L classes. Otherwise stated, $\boldsymbol{n}|c = l \sim \mathcal{CN}_N(\boldsymbol{0}, \boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N)$, and hence, the (unconditional) PDF of $\boldsymbol{n}$ is given by

$$f(\boldsymbol{n}; \mathcal{P}_0) = \sum_{l=1}^{L} p_l \frac{1}{\pi^N \det\left(\boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N\right)} \times \exp\left[-\boldsymbol{n}^\dagger \left(\boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N\right)^{-1} \boldsymbol{n}\right] \tag{3}$$

where we recall that p_l is the probability that $c = l$. For future reference, we indicate by $\mathcal{P}_0$ the set of the unknown parameters in the aforementioned PDF associated to the H_0 hypothesis, namely $p_1, \ldots, p_L, \sigma_w^2$ and the elements of the set of the deterministic entries of the positive semidefinite matrices $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_L$, say $\mathcal{R}$. Thus, $\mathcal{P}_0 = \{p_1, \ldots, p_L, \sigma_w^2\} \cup \mathcal{R}$. Notice that some of the entries of the covariance matrices are related by a one-to-one mapping (due to Hermitian symmetry), thus impacting the set of unknowns forming $\mathcal{R}$.

Similarly, the PDF of $\boldsymbol{n}_k$ is given by

$$f(\boldsymbol{n}_k; \mathcal{P}_0) = \sum_{l=1}^{L} p_l \frac{1}{\pi^N \det\left(\boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N\right)} \times \exp\left[-\boldsymbol{n}_k^\dagger \left(\boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N\right)^{-1} \boldsymbol{n}_k\right]. \tag{4}$$

Assuming that $\boldsymbol{n}$ and the $\boldsymbol{n}_k$s are independent random vectors, it follows that the logarithm of the joint PDF of $\boldsymbol{z}$ and $\boldsymbol{Z} = [\boldsymbol{z}_1 \ \cdots \ \boldsymbol{z}_K]$ is given by

$$\mathcal{L}_0(\boldsymbol{z}, \boldsymbol{Z}; \mathcal{P}_0) = \log f(\boldsymbol{z}; \mathcal{P}_0) + \sum_{k=1}^{K} \log f(\boldsymbol{z}_k; \mathcal{P}_0) \tag{5}$$

under H_0 and

$$\mathcal{L}_1(\boldsymbol{z}, \boldsymbol{Z}; \mathcal{P}_1) = \log f(\boldsymbol{z} - \alpha \boldsymbol{v}; \mathcal{P}_0) + \sum_{k=1}^{K} \log f(\boldsymbol{z}_k; \mathcal{P}_0) \tag{6}$$

under H_1 with $\mathcal{P}_1 = \mathcal{P}_0 \cup \{\alpha\}$.

Under these assumptions, we aim at estimating the unknown probabilities $p_l > 0$, $\sum_{l=1}^{L} p_l = 1$, together with all the other unknown parameters from the available data. We remark that the solution of the corresponding hypothesis test via GLRT, for this rather general clutter model, based on a convex mixture of Gaussian terms, is a challenging problem. Thus, to derive an adaptive detector able to estimate the resulting large number of parameters involved in the sets $\mathcal{P}_0$ and $\mathcal{P}_1$, we adopt the EM algorithm, which yields approximate local maxima of the log-likelihood functions ($\mathcal{L}_0$ and $\mathcal{L}_1$) required in the GLRT. Moreover, we present two different ways to solve the detection problem.

The first decision scheme relies on the GLRT where the compressed likelihoods,[2] modeled according to the contaminated normal distribution, are computed by implementing the EM algorithm under both hypotheses. The corresponding detector will be referred to in the following as *EM-based GLRT* (labeled EM-GLRT). The EM algorithm can also be used to construct the MAP estimates of c and the c_ks, given the CUT and the secondary data, and eventually to select a reasonable Gaussian distribution of the disturbance within each cell (a point better clarified in due course). An ad hoc detector, based upon the LRT and implemented by replacing the actual contaminated normal distribution of the CUT with the Gaussian distribution associated to the MAP

[1] Notice that in space-time adaptive processing (STAP), the disturbance covariance matrix exhibits a structure that comprises the sum of (white) noise and clutter covariances, with the clutter component being positive semidefinite and rank deficient [31], [32], [33]. For a uniform array and for fixed pulse repetition frequency, the space-time clutter covariance matrix is essentially low rank due to the inherent oversampling nature of the STAP architecture (see [34] and references therein). In addition, the rank of the clutter covariance matrix is an indicator of both severity of the clutter scenario and the number of degrees of freedom required to equalize the clutter component [35], [36], [37], [38].

[2] Compressed likelihood refers to the maximum of the likelihood function with respect to its unknown parameters.

estimates of c, will also be investigated. Such a detector will be referred to as the *EM-MAP-based LRT* (labeled EM-MAP-LRT).

III. DESIGN OF DETECTORS FOR HETEROGENEOUS ENVIRONMENTS

As already anticipated, to solve the hypothesis testing problem (1) under the modeling assumptions of Section II, we propose two different design approaches. Let us start with the plain GLRT relying on the (approximate) ML estimates of the unknown parameters obtained by the EM algorithm (under both hypotheses). The logarithm of the GLRT is given by

$$\max_{\mathcal{P}_1} \mathcal{L}_1\left(\boldsymbol{z}, \boldsymbol{Z}; \mathcal{P}_1\right) - \max_{\mathcal{P}_0} \mathcal{L}_0\left(\boldsymbol{z}, \boldsymbol{Z}; \mathcal{P}_0\right) \tag{7}$$

from which the EM-GLRT decision scheme is obtained by replacing the ML estimates with their EM counterparts, i.e.,

$$\Lambda_{\text{EM-GLRT}}(\boldsymbol{z}, \boldsymbol{Z}) = \mathcal{L}_1\left(\boldsymbol{z}, \boldsymbol{Z}; \widehat{\mathcal{P}}_1\right) - \mathcal{L}_0\left(\boldsymbol{z}, \boldsymbol{Z}; \widehat{\mathcal{P}}_0\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{8}$$

where η is the threshold to be set according to the desired probability of false alarm (P_{fa}), while $\widehat{\mathcal{P}}_1$ and $\widehat{\mathcal{P}}_0$ denote the estimates of $\mathcal{P}_1$ and $\mathcal{P}_0$, computed by the EM algorithm.

An alternative to the aforementioned approximation of the GLRT can be obtained by considering the MAP estimates of c (i.e., the class of the CUT) obtained starting from the posterior probability

$$P\left(c = l | \boldsymbol{z}; \widehat{\mathcal{P}}_1\right) \tag{9}$$

computed by the EM algorithm under H_1 and the analogous quantity computed under H_0. Accordingly, we can estimate the value c takes on under H_0 and H_1, respectively. As a matter of fact, we compute

$$\hat{l}_1 = \arg\max_l P\left(c = l | \boldsymbol{z}; \widehat{\mathcal{P}}_1\right) \tag{10}$$

under H_1 and

$$\hat{l}_0 = \arg\max_l P\left(c = l | \boldsymbol{z}; \widehat{\mathcal{P}}_0\right) \tag{11}$$

under H_0. It turns out that the EM-MAP-LRT is given by

$$\Lambda_{\text{EM-MAP-LRT}}(\boldsymbol{z}, \boldsymbol{Z}) = \frac{\tilde{f}\left(\boldsymbol{z}; \widehat{\alpha}, \widehat{\boldsymbol{R}}_{\hat{l}_1}, \widehat{\sigma}^2_{w1}\right)}{\tilde{f}\left(\boldsymbol{z}; 0, \widehat{\boldsymbol{R}}_{\hat{l}_0}, \widehat{\sigma}^2_{w0}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{12}$$

where

$$\tilde{f}\left(\boldsymbol{z}; \alpha, \boldsymbol{R}_l, \sigma_w^2\right) = \frac{1}{\det\left(\boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N\right)} \times \exp\left[-(\boldsymbol{z} - \alpha \boldsymbol{v})^{\dagger}\left(\boldsymbol{R}_l + \sigma_w^2 \boldsymbol{I}_N\right)^{-1}(\boldsymbol{z} - \alpha \boldsymbol{v})\right]. \tag{13}$$

$\widehat{\boldsymbol{R}}_{\hat{l}_0}$ and $\widehat{\sigma}^2_{w0}$ denote the estimates of $\boldsymbol{R}_l$ and σ_w^2 under H_0, respectively (again computed by the EM algorithm); similarly, $\widehat{\alpha}$, $\widehat{\boldsymbol{R}}_{\hat{l}_1}$, and $\widehat{\sigma}^2_{w1}$ are the estimates of α, $\boldsymbol{R}_l$, and σ_w^2 under H_1.

We finally derive the EM algorithm for the problem addressed in this work, which is necessary to implement the proposed decision schemes. This approach allows us to recursively estimate the parameters of the sets $\mathcal{P}_i$, $i = 0, 1$. For the sake of clarity, let us denote by[3] Θ_l the set formed by σ_w^2 and the entries of $\boldsymbol{R}_l$, and by $\widehat{\mathcal{P}}_i^{(h-1)}$ and $\widehat{\Theta}_l^{(h-1)}$ the sets of the estimates of the parameters in $\mathcal{P}_i$ and Θ_l, respectively, at the $(h-1)$th iteration of the EM algorithm. The hth iteration computes $\widehat{\mathcal{P}}_i^{(h)}$ starting from $\widehat{\mathcal{P}}_i^{(h-1)}$ to guarantee

$$\mathcal{L}_i(\boldsymbol{z}, \boldsymbol{Z}; \widehat{\mathcal{P}}_i^{(h)}) \geq \mathcal{L}_i(\boldsymbol{z}, \boldsymbol{Z}; \widehat{\mathcal{P}}_i^{(h-1)}). \tag{14}$$

The EM procedure consists of the following two steps, referred to as the E-step and the M-step.

A. E-Step

The E-step is tantamount to implementing the update rule

$$q_k^{(h-1)}(l) = P\left(c_k = l | \boldsymbol{z}_k; \widehat{\mathcal{P}}_0^{(h-1)}\right) = \frac{f\left(\boldsymbol{z}_k | c_k = l; \widehat{\Theta}_l^{(h-1)}\right)\widehat{p}_l^{(h-1)}}{\sum_{i=1}^{L} f\left(\boldsymbol{z}_k | c_k = i; \widehat{\Theta}_i^{(h-1)}\right)\widehat{p}_i^{(h-1)}} \tag{15}$$

under both hypotheses for the secondary data, while for the primary data, we have

$$q^{(h-1)}(l) = P\left(c = l | \boldsymbol{z}; \widehat{\mathcal{P}}_0^{(h-1)}\right) = \frac{f\left(\boldsymbol{z} | c = l; \widehat{\Theta}_l^{(h-1)}\right)\widehat{p}_l^{(h-1)}}{\sum_{i=1}^{L} f\left(\boldsymbol{z} | c = i; \widehat{\Theta}_i^{(h-1)}\right)\widehat{p}_i^{(h-1)}} \tag{16}$$

under H_0 and

$$\tilde{q}^{(h-1)}(l) = P\left(c = l | \boldsymbol{z}; \widehat{\mathcal{P}}_1^{(h-1)}\right) = \frac{f\left(\boldsymbol{z} - \widehat{\alpha}^{(h-1)}\boldsymbol{v} | c = l; \widehat{\Theta}_l^{(h-1)}\right)\widehat{p}_l^{(h-1)}}{\sum_{i=1}^{L} f\left(\boldsymbol{z} - \widehat{\alpha}^{(h-1)}\boldsymbol{v} | c = i; \widehat{\Theta}_i^{(h-1)}\right)\widehat{p}_i^{(h-1)}} \tag{17}$$

under H_1, respectively. In the aforementioned formulas, $f(\cdot | c_k = l; \widehat{\Theta}_l^{(h-1)})$ and $f(\cdot | c = l; \widehat{\Theta}_l^{(h-1)})$ are the estimates of the conditional PDFs of $\boldsymbol{z}_k$ and $\boldsymbol{z}$ under H_0 (or $\boldsymbol{z} - \alpha\boldsymbol{v}$ under H_1) given $c_k = l$ and $c = l$, respectively, at the $(h-1)$th iteration of the EM algorithm, namely the PDF of a complex normal random vector with zero mean and covariance matrix given by $\widehat{\boldsymbol{R}}_l^{(h-1)} + \widehat{\sigma}_w^{2(h-1)}\boldsymbol{I}_N$, with $\widehat{\boldsymbol{R}}_l^{(h-1)}$ and $\widehat{\sigma}_w^{2(h-1)}$ the estimates of $\boldsymbol{R}_l$ and σ_w^2, respectively, at the $(h-1)$th iteration. Similarly, $\widehat{\alpha}^{(h-1)}$ and $\widehat{p}_l^{(h-1)}$ are the estimates of α and p_l, respectively, at the previous step.

B. M-Step

We now focus on the M-step under the H_1 hypothesis, as it provides as special case also the equations for the H_0 hypothesis (for which $\alpha = 0$, hence, does not need to be estimated). We can write that

$$\widehat{\mathcal{P}}_1^{(h)} = \arg\max_{\mathcal{P}_1}\left[g_1\left(p_1, \ldots, p_L, \mathcal{R}, \sigma_w^2\right)\right.$$

[3] For notational convenience, we omit the dependence of the estimates on the hypothesis on which they are obtained.

$$+ g_2\left(p_1, \ldots, p_L, \mathcal{R}, \sigma_w^2, \alpha\right)\Big] \tag{18}$$

with

$$g_1\left(p_1, \ldots, p_L, \mathcal{R}, \sigma_w^2\right) = \sum_{k=1}^{K}\sum_{l=1}^{L} q_k^{(h-1)}(l) \times \log \frac{f\left(z_k | c_k = l; \Theta_l\right) p_l}{q_k^{(h-1)}(l)} \tag{19}$$

and

$$g_2\left(p_1, \ldots, p_L, \mathcal{R}, \sigma_w^2, \alpha\right) = \sum_{l=1}^{L} \tilde{q}^{(h-1)}(l) \times \log \frac{f\left(\boldsymbol{z} - \alpha \boldsymbol{v} | c = l; \Theta_l\right) p_l}{\tilde{q}^{(h-1)}(l)} \tag{20}$$

where $f(\cdot|c_k = l; \Theta_l)$ and $f(\cdot|c = l; \Theta_l)$ are the conditional PDFs of z_k and $\boldsymbol{z}$ under H_0, given $c_k = l$ and $c = l$, respectively. The maximization with respect to the p_ls is tantamount to solving the following optimization problem:

$$\begin{cases} \max\limits_{p_l, l=1,\ldots,L} \sum\limits_{l=1}^{L}\left[\left(\sum\limits_{k=1}^{K} q_k^{(h-1)}(l) + \tilde{q}^{(h-1)}(l)\right)\log p_l\right] \\ \text{s.t.} \sum\limits_{l=1}^{L} p_l = 1 \end{cases}. \tag{21}$$

By applying the method of Lagrange multipliers, it is not difficult to prove (see Appendix A) that

$$\widehat{p}_l^{(h)} = \frac{1}{K+1}\left[\sum_{k=1}^{K} q_k^{(h-1)}(l) + \tilde{q}^{(h-1)}(l)\right]. \tag{22}$$

To the best of authors' knowledge, the problem of estimating the remaining parameters in the right-hand side of (18) cannot be solved in closed form. Thus, we pursue an alternative path and seek for an approximate solution using the following two-step procedure: first we estimate the disturbance-related parameters $\mathcal{R}$ and σ_w^2 by maximizing g_1 only, which is a function of $z_1, \ldots, z_K$, thus obtaining $\widehat{\mathcal{R}}^{(h)}$ and $\widehat{\sigma}_w^{2(h)}$; then, we plug these estimates into g_2, which is a function of $\boldsymbol{z}$, in place of $\mathcal{R}$ and σ_w^2 and maximize it with respect to the remaining target-related parameter α obtaining $\widehat{\alpha}^{(h)}$. Similar suboptimal strategies have been used in the radar signal processing literature to conceive many detectors as, for instance, the well-known adaptive matched filter (AMF) [15].

We start with the maximization of g_1, which can be conducted using the results reported in [29, Proposition 3] that we restate here for the sake of completeness.

PROPOSITION 1 An approximation to the maximizer of the function

$$g_1'(\mathcal{R}, \sigma_w^2) = \sum_{k=1}^{K}\sum_{l=1}^{L} q_k^{(h-1)}(l) \log f\left(z_k | c_k = l; \Theta_l\right) \tag{23}$$

can be obtained as follows:

$$\widehat{\sigma}_w^{2(h)} = \left\{\sum_{l=1}^{L}\sum_{n=r+1}^{N} \gamma_{l,n}^{(h-1)}\right\} \Big/ \left\{\sum_{l=1}^{L}\sum_{k=1}^{K} q_k^{(h-1)}(l)(N-r)\right\} \tag{24}$$

and

$$\widehat{\boldsymbol{R}}_l^{(h)} = \widehat{\boldsymbol{U}}_l^{(h)} \widehat{\boldsymbol{\Lambda}}_l^{(h)} (\widehat{\boldsymbol{U}}_l^{(h)})^{\dagger}, \quad l = 1, \ldots, L \tag{25}$$

with $\widehat{\boldsymbol{U}}_l^{(h)}$ the unitary matrix whose columns are the eigenvectors corresponding to the eigenvalues $\gamma_{l,1}^{(h-1)} \geq \gamma_{l,2}^{(h-1)} \geq \ldots \geq \gamma_{l,N}^{(h-1)}$ of the matrix $\boldsymbol{S}_l^{(h-1)} = \sum_{k=1}^{K} q_k^{(h-1)}(l) \boldsymbol{z}_k \boldsymbol{z}_k^{\dagger}$ and

$$\widehat{\boldsymbol{\Lambda}}_l^{(h)} = \operatorname{diag}\left(\max\left\{\frac{\gamma_{l,1}^{(h-1)}}{\sum_{k=1}^{K} q_k^{(h-1)}(l)} - \widehat{\sigma}_w^{2(h)}, 0\right\}, \ldots, \max\left\{\frac{\gamma_{l,r}^{(h-1)}}{\sum_{k=1}^{K} q_k^{(h-1)}(l)} - \widehat{\sigma}_w^{2(h)}, 0\right\}, 0, \ldots, 0\right). \tag{26}$$

As a final comment, notice that, although the aforementioned proposition assumes that $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_L$ have the same known rank r, the result can be generalized to the case that the matrices have unknown, possibly different, values of the rank [29]. Moreover, it is important to remark that the EM algorithm derived in the present contribution conveniently uses only a part of the results in [29]: in fact, while [29] deals with clutter classification, here we address target detection, and hence, in addition to estimating the unknown disturbance parameters (appearing in g_1), it is also necessary to handle unknown parameters associated to the presence of a target buried in a heterogeneous clutter environment (appearing in the g_2 function).

Maximizing g_2 with respect to α is tantamount to maximizing

$$\begin{aligned} g_2'(\alpha) &= \sum_{l=1}^{L} \tilde{q}^{(h-1)}(l) \log f\left(\boldsymbol{z} - \alpha \boldsymbol{v} | c = l; \widehat{\Theta}_l^{(h)}\right) \\ &= \sum_{l=1}^{L} \tilde{q}^{(h-1)}(l)\Big[-N \log \pi - \log\det\left(\widehat{\boldsymbol{R}}_l^{(h)} + \widehat{\sigma}_w^{2(h)} \boldsymbol{I}_N\right) \\ &\quad - (\boldsymbol{z} - \alpha \boldsymbol{v})^{\dagger}\left(\widehat{\boldsymbol{R}}_l^{(h)} + \widehat{\sigma}_w^{2(h)} \boldsymbol{I}_N\right)^{-1}(\boldsymbol{z} - \alpha \boldsymbol{v})\Big] \end{aligned}$$

where $f(\cdot|c = l; \widehat{\Theta}_l^{(h)})$ is the estimate of the conditional PDF of $\boldsymbol{z} - \alpha \boldsymbol{v}$, under H_1 and given $c = l$, at the hth iteration of the EM algorithm. Moreover, neglecting additive terms independent of α, we introduce the function

$$g_2''(\alpha) = -(\boldsymbol{z} - \alpha \boldsymbol{v})^{\dagger} \boldsymbol{A}^{(h)} (\boldsymbol{z} - \alpha \boldsymbol{v})$$

with

$$\boldsymbol{A}^{(h)} = \left[\sum_{l=1}^{L} \tilde{q}^{(h-1)}(l)\left(\widehat{\boldsymbol{R}}_l^{(h)} + \widehat{\sigma}_w^{2(h)} \boldsymbol{I}_N\right)^{-1}\right].$$

It is not difficult to show that the maximum is achieved at

$$\widehat{\alpha}^{(h)} = \frac{\sum_{l=1}^{L} \tilde{q}^{(h-1)}(l) \boldsymbol{v}_l^{\dagger} \boldsymbol{\zeta}_l}{\sum_{l=1}^{L} \tilde{q}^{(h-1)}(l) \boldsymbol{v}_l^{\dagger} \boldsymbol{v}_l} \tag{27}$$

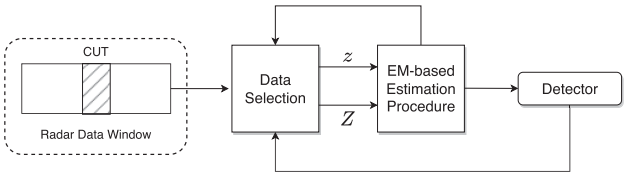

Fig. 1. Possible signal processing unit architecture that includes the proposed approach.

where

$$\boldsymbol{v}_l = \left(\widehat{\boldsymbol{R}}_l^{(h)} + \widehat{\sigma}_w^{2(h)}\boldsymbol{I}_N\right)^{-1/2}\boldsymbol{v} \tag{28}$$

and

$$\boldsymbol{\zeta}_l = \left(\widehat{\boldsymbol{R}}_l^{(h)} + \widehat{\sigma}_w^{2(h)}\boldsymbol{I}_N\right)^{-1/2}\boldsymbol{z}. \tag{29}$$

Estimation of $\widehat{\mathcal{P}}_0$ under H_0 is tantamount to maximizing g in (18) with $\alpha = 0$ and can be conducted exploiting the same proposition used to maximize g_1 under H_1.

Fig. 1 shows a possible architecture that incorporates the proposed approach into the signal processing unit of a radar system.

C. EM Initialization

We now propose how to properly initialize the EM algorithm. In this respect, we assume equiprobable priors for the values of p_l's, namely, $\hat{p}_l^{(0)} = 1/L$, whereas the initial value of the clutter covariance matrices $\hat{\boldsymbol{R}}_l^{(0)}$ is set according to the following strategy:

1) for each range bin, compute the power $g(k) = \frac{1}{N}\boldsymbol{z}_k^\dagger\boldsymbol{z}_k$, $k = 1, \ldots, K$;
2) sort the power values $g(k)$'s in ascending order, namely $\tilde{g}(1) \leq \cdots \leq \tilde{g}(K)$;
3) denote by $\tilde{\boldsymbol{z}}_i$ the vector corresponding to $\tilde{g}(i)$;
4) use the subsets of[4] K/L vectors $\tilde{\boldsymbol{z}}_i$, $i = (l-1)\frac{K}{L} + 1, \ldots, l\frac{K}{L}$, representative of a coarse similarity among the different range bins, to compute the following matrix, that will be used to initialize $\boldsymbol{R}_l$

$$\tilde{\boldsymbol{R}}_l^{(0)} = \frac{L}{K}\sum_{i=(l-1)\frac{K}{L}+1}^{l\frac{K}{L}} \tilde{\boldsymbol{z}}_i\tilde{\boldsymbol{z}}_i^\dagger, \quad l = 1, \ldots, L. \tag{30}$$

To set the initial value of the noise power, we first compute the sample covariance matrix $\boldsymbol{\Sigma} = (1/K)\boldsymbol{Z}\boldsymbol{Z}^\dagger$. Denoting with $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$ the ordered eigenvalues of $\boldsymbol{\Sigma}$, we set

$$\hat{\sigma}_w^{2(0)} = \frac{1}{N-r}\sum_{j=1}^{N-r}\lambda_j. \tag{31}$$

The initial estimate of $\boldsymbol{R}_l$ is computed by removing from (30) the estimated contribution of the noise power (31) as

$$\hat{\boldsymbol{R}}_l^{(0)} = \tilde{\boldsymbol{R}}_l^{(0)} - \hat{\sigma}_w^{2(0)}\boldsymbol{I}_N, \quad l = 1, \ldots, L. \tag{32}$$

[4]For the considered values of K and L, the ratio K/L is an integer.

As to the complex amplitude α, we set its initial value (only required under the H_1 hypothesis) as

$$\hat{\alpha}^{(0)} = \frac{\boldsymbol{v}^\dagger\boldsymbol{z}}{\boldsymbol{v}^\dagger\boldsymbol{v}} \tag{33}$$

that is the ML estimate of α in white Gaussian noise. The discussion about the number of iterations required to obtain $\widehat{\mathcal{P}}_1$ and $\widehat{\mathcal{P}}_0$ is part of the analyses conducted in the next section.

IV. Performance Assessment on Synthetic and Real Data and Results

In this section, we present an extensive performance analysis of the proposed detectors based on both synthetic and real data. We start by determining the number of iterations required to compute $\widehat{\mathcal{P}}_1$ and $\widehat{\mathcal{P}}_0$. The analysis over synthetic data allows us to first evaluate the advantages in terms of nominal detection performance of the proposed approaches against that of the natural competitors. The analysis also tests the robustness of the newly proposed decision schemes with respect to possible mismatches related to the design assumptions. In the second part, the behavior of the proposed detectors is assessed on live data recorded by real radar systems. Such data are characterized by the presence of heterogeneous clutter disturbances in different operational environments.

A. Simulation Setup

The numerical assessment on synthetic data is conducted assuming $N = 8$ and considering two different configurations of K and number of clutter classes L, i.e., $K = 48$, $L = 2$, and $K = 96$, $L = 3$. It is important to stress here that the proposed approach goes beyond the limitations of the conventional radar window used to select secondary data. We use a spatial steering vector $\boldsymbol{v}$ steered at 0°. A desired $P_{\text{fa}} = 10^{-3}$ is assumed and the performance is assessed by Monte Carlo simulations with $100/P_{\text{fa}}$ independent trials to estimate the thresholds; the P_d values are obtained over 10^3 trials. We adopt the general definition for the signal to clutter-plus-noise ratio (SCNR)

$$\text{SCNR} = |\alpha|^2\boldsymbol{v}^\dagger(\boldsymbol{R}^{\text{CUT}} + \sigma_w^2\boldsymbol{I}_N)^{-1}\boldsymbol{v} \tag{34}$$

with $\boldsymbol{R}^{\text{CUT}}$ denoting the covariance matrix associated to the specific clutter class affecting the returns in the CUT (i.e., one among the L possible matrices $\boldsymbol{R}_l$, $l = 1, \ldots, L$). The power of the thermal noise is instead set to $\sigma_w^2 = 0.5$.

1) *Clutter Covariance Matrix:* For simplicity, we consider a spatial-only processing performed through a uniformly spaced linear array of N identical and isotropic (at least in the angular sector of interest) sensors with interelement distance equal to $\lambda/2$, with λ being the wavelength corresponding to the radar carrier frequency. Accordingly, clutter samples can be modeled as the summation of individual patch returns at distinct angles [39], leading to the following covariance structure:

$$\boldsymbol{R}_l = \sigma_{c,l}^2\sum_{\phi_i^l\in\Phi_l}\boldsymbol{v}(\phi_i^l)\boldsymbol{v}(\phi_i^l)^\dagger \tag{35}$$

where

1) $\boldsymbol{v}(\phi_i^l)$ denotes the spatial steering vector having the nth entry equal to $[\boldsymbol{v}(\phi_i^l)]_n = 1/\sqrt{N}e^{j\pi n \sin\phi_i^l}$, $n = 0, \ldots, N-1$;
2) $\Phi_l = \{\phi_1^l, \phi_2^l, \ldots, \phi_{N_c^l}^l\}$; for simplicity, we assume that the number of patch returns is the same for each class, namely $N_c^l = N_c, \forall l$.

We set $N_c = 3$, and hence, $r = 3$ (in fact, N_c is the rank of $\boldsymbol{R}_l$). The specific N_c directions for each class of clutter used in the simulations are $\Phi_1 = \{-5.96°, -1.76°, -2.97°\}$, $\Phi_2 = \{-11.24°, 4.86°, -9.17°\}$, and $\Phi_3 = \{13.95°, -14.17°, 3.13°\}$, respectively.[5] As to $\sigma_{c,l}^2$, it denotes the lth class clutter power and is set as $\sigma_{c,l}^2 = 10l$ dB, $l = 1, \ldots, L$.

2) *State-of-the-Art Competitors:* We compare the performance of the newly proposed detectors against a set of different algorithms designed to operate in heterogeneous environments. More specifically, we consider the ANMF detectors obtained by employing different estimates of the disturbance covariance matrix.

1) the Σ-ANMF in [18] relying on a sample covariance matrix based on normalized secondary data;
2) the R-ANMF detector, relying on the recursive procedure devised in [19];
3) the RP-ANMF detector based on a recursive estimate exploiting the persymmetric structure of the covariance matrix proposed in [21].

In addition, we also consider the approximate GLRT (AGLRT) detector proposed in [23], proven to be a viable approach to detect coherent targets in clutter-dominated heterogeneous environments. The considered competitors represent the state-of-the-art references for radar detection in heterogeneous scenarios.[6] The number of iterations used by the R-ANMF and RP-ANMF detectors to recursively estimate the clutter covariance matrix is set to 3, being this number sufficient to guarantee an acceptable convergence as corroborated by the related literature. On the other hand, the AGLRT uses 20 iterations in its cyclic estimation procedures.

B. Results on Synthetic Data

First, we set $K = 48, L = 2$, and as preliminary step, analyze the requirements of the proposed procedures in terms of number of EM iterations. To this aim, we consider as metric the absolute value of the relative difference between the values assumed by the "compressed" log-likelihood functions, i.e., $\mathcal{L}_0(z, \boldsymbol{Z}; \widehat{\mathcal{P}}_0^{(h)})$ or $\mathcal{L}_1(z, \boldsymbol{Z}; \widehat{\mathcal{P}}_1^{(h)})$, over two successive iterations $h-1$ and h, evaluated under H_0, as a function of the number of iterations h, namely, $i = 0, 1$

$$\Delta\mathcal{L}_i(h) = \left| \frac{\left(\mathcal{L}_i\left(z, \boldsymbol{Z}; \widehat{\mathcal{P}}_i^{(h)}\right) - \mathcal{L}_i\left(z, \boldsymbol{Z}; \widehat{\mathcal{P}}_i^{(h-1)}\right)\right)}{\mathcal{L}_i\left(z, \boldsymbol{Z}; \widehat{\mathcal{P}}_i^{(h)}\right)} \right|. \quad (36)$$

In Fig. 2, we report the average values of $\Delta\mathcal{L}_i(h)$ under both hypotheses as a function of the number of iterations h, computed over 10^5 Monte Carlo trials. As it can be observed, the curves quickly decrease with the iteration number under both hypotheses, and already at $h = 10$ achieve variations lower than 10^{-4}. Notice that such curves represent the log-likelihood variations, and hence, do not have to experience an increasing monotone behavior unlike the sequence of likelihood values. Similar results are obtained also for other parameters setting, hence are omitted for brevity. Therefore, in the following, we will consider for the proposed detectors a number of maximum iterations equal to 10.

The detection performances in terms of P_d as a function of the SCNR are shown in Fig. 3. It can be observed that both the proposed detection schemes significantly outperform the state-of-the-art competitors over the whole range of SCNR, with a gap that in the low SCNR regime is around 4 dB for the EM-MAP-LRT, and increases up to about 5 dB for the EM-GLRT, when compared with the

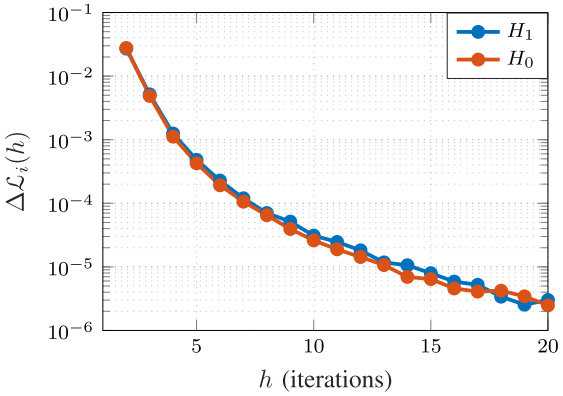

Fig. 2. Average $\Delta\mathcal{L}_i(h)$ as a function of the iteration number h.

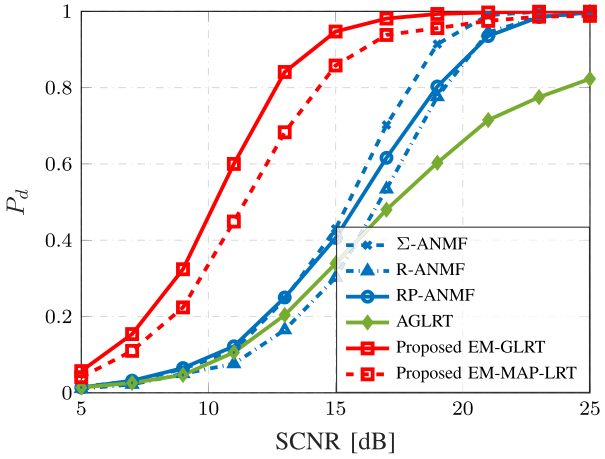

Fig. 3. P_d as a function of the SCNR for the proposed detectors, in comparison with natural competitors, for the case $N = 8$, $K = 48$, and $L = 2$.

[5]Each ϕ_i^l is drawn from a uniform distribution within the first null-to-null beamwidth of the linear array (whose extent is about $4/N$).

[6]Notice that such competitors are the ultimate achievements of different research groups from the radar community, who contributed to their derivation and analysis. In addition to [18], [19], and [21], the interested reader is also referred to, e.g., [2], [7], [9], [10], [20], [22], and [40].

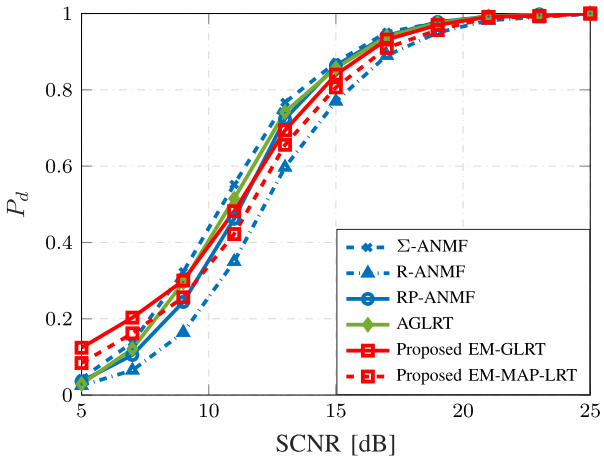

Fig. 4. P_d as a function of the SCNR for the proposed detectors, in comparison with natural competitors, assuming $N = 8$, $K = 48$, and $L = 2$, using a compound-Gaussian model for the clutter.

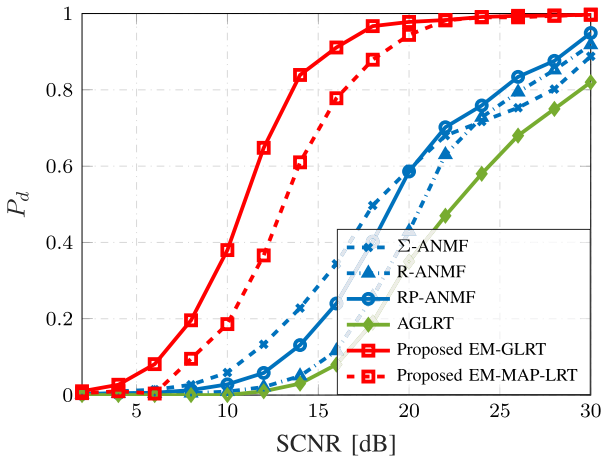

Fig. 5. P_d as a function of the SCNR for the proposed detectors, in comparison with natural competitors, for the case $N = 8$, $K = 96$, and $L = 3$.

Σ-ANMF detector, the latter providing the best P_d's among the considered competitors. These interesting outcomes demonstrate improved performance, i.e., a better capability to exploit the available information in the detection task, despite the higher number of free parameters to be estimated. Differently put, although more parameters need to be estimated to cope with the L clutter classes, the proposed detectors are able to learn the different structures and powers associated with the diverse clutter conditions; consequently, they fully adapt to heterogeneous environments and offer superior performance compared to existing approaches that model clutter heterogeneity only through different power levels. Furthermore, the behavior of the EM-MAP-LRT also reveals that the proposed EM-based recursive estimation scheme is able to correctly infer the specific class of clutter the CUT is embedded in. As a byproduct, it can be used for a coarse clutter classification.

To test the robustness of the proposed approaches to mismatches with respect to the assumed heterogeneous model, we now stick to different scenarios that match the design assumptions of the considered competitors. More specifically, we change the distribution of the disturbance and assume for the clutter contribution a compound-Gaussian model characterized by a texture component distributed as the square root of a Gamma RV with parameters $(\nu, 1/\nu)$ (so that the mean square value is unitary), with $\nu = 0.5$, multiplied by a complex Gaussian vector with zero mean and covariance matrix $\boldsymbol{R}$. As to $\boldsymbol{R}$, we set it equal to the covariance matrix of the class $l = 1$ generated according to (35). In addition, texture values of different range cells (i.e., CUT and secondary data) are supposed to be independent RVs. The disturbance also includes a thermal noise component.

The obtained results in terms of P_d as a function of the SCNR are reported in Fig. 4. Remarkably, the two proposed detectors incur a very limited gap compared to the considered competitors. Overall, we can conclude that the novel detectors offer performance comparable to state-of-the-art competitors in environments where the clutter has the same structure (covariance matrix) and its heterogeneity is related to a varying level of powers in range, whereas significant improvements are achieved when clutter heterogeneity (in range) is described by (a limited number of) different covariance matrices. As better discussed later in Section IV-C, these conditions can be typically found in real experimental data.

To corroborate the aforementioned results, we investigate a second configuration consisting in $K = 96$ secondary data and $L = 3$, which is representative of a processing performed over a larger portion of the monitored area including an additional class of clutter. In Fig. 5, we report the curves of the P_d as a function of the SCNR for the proposed detectors and considered competitors. The obtained performance confirms the superiority of the proposed detection schemes over the state-of-the-art. Remarkably, the proposed EM-GLRT guarantees an advantage that reaches more than 10 dB for $P_d = 0.8$ compared to the Σ-ANMF and RP-ANMF, and even higher with respect to the R-ANMF and the AGLRT. The EM-MAP-LRT also provides significantly higher P_d values than the ANMF and AGLRT detectors, though with a slightly smaller performance gap (about 8 dB at $P_d = 0.8$) compared to the EM-GLRT algorithm. On the other hand, it is also interesting to observe the worse performance of the ANMF detectors: this behavior is related to the increased level of heterogeneity of the clutter returns along the different range bins, and reveals the inability of ANMF detectors to adapt to this type of heterogeneity. Similarly, the AGLRT detector is inherently penalized by the underlying assumption that the structure of the clutter covariance matrix remains the same across the L clutter classes, and that the heterogeneity of the clutter can be sufficiently captured by just considering different power levels in each range bin. Finally, the superior performance of the EM-GLRT with respect to the EM-MAP-LRT can be explained by the fact that the former exploits all the available information, while the decision statistic of the EM-MAP-LRT considers only the class returned by the MAP classifier. As a consequence,

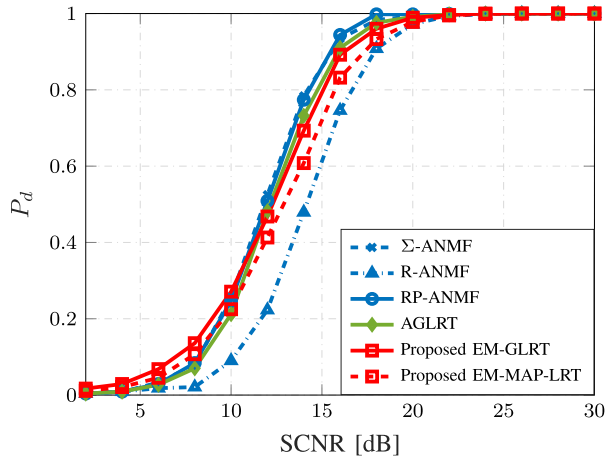

Fig. 6. P_d as a function of the SCNR for the proposed detectors, in comparison with natural competitors, assuming $N = 8$, $K = 96$, and $L = 3$, using a compound-Gaussian model for the clutter.

TABLE I
Average Execution Time (In Seconds) of the Considered Detectors for $N = 8, K = 48$

Σ-ANMF	R-ANMF	RP-ANMF	AGLRT	EM-GLRT	EM-MAP-LRT
$5.29 \cdot 10^{-4}$	0.0016	0.0038	0.21	0.098	0.113

the remaining information related to the other classes is not exploited.

We also assess the performance of the proposed detectors in the presence of a mismatch on the assumed design model, as done for Fig. 4. The obtained results, reported in Fig. 6, confirm that both the proposed EM-MAP-LRT and EM-GLRT are quite robust against the introduced mismatch with a limited loss with respect to the competitors, which however operate under conditions perfectly tailored to their design assumptions.

We conclude the analysis by quantifying the computational load of the proposed algorithms, in comparison with the considered state-of-the-art detectors. To this aim, we have run all the algorithms on the same hardware platform and computed the average execution time on 100 distinct trials. The findings are presented in Table I. Notably, the Σ-ANMF detector is the fastest among all the considered detectors. This can be attributed to its lightweight processing that does not entail any iterative or recursive steps. The R-ANMF and RP-ANMF methods exhibit only slightly higher execution times, attributable to their recursive covariance matrix estimation. Comparatively, the proposed EM-GLRT showcases an average execution time in between one and two orders of magnitude greater than R-ANMF and RP-ANMF. This is noteworthy considering that the EM procedure adaptively estimates a larger number of parameters (within the $\hat{\mathcal{P}}_0$ and $\hat{\mathcal{P}}_1$ sets) to account for the more complex heterogeneous environment at hand. The EM-MAP-LRT is only slightly more complex than the EM-GLRT owing to the two additional 1-D maximizations performed in (10) and (11) to infer the most probable clutter class in the CUT under the two hypotheses. In contrast, the AGLRT algorithm shows the highest average execution time, around two times greater than the proposed EM-GLRT. Overall, it is crucial to emphasize that all recorded average execution times are measured in absolute temporal units, meaning that all detectors necessitate mere fractions of a second for their execution.

C. Results on Real Radar Data

In this section, we assess the effectiveness of the proposed detection schemes when applied on two real radar datasets, each characterized by a different type of heterogeneous clutter. Clearly, such data do not match the underlying design assumptions used to derive the proposed detectors and described in Section II. Therefore, the analyses conducted on these experimental data allow us to unveil to what extent the novel approaches are able to effectively adapt to heterogeneous environments found in real operating scenarios.

1) *Analysis on PhaseOne Data:* We start the analysis by considering the L-band land clutter data, recorded in 1985 using the MIT Lincoln Laboratory Phase One radar at the Katahdin Hill site, MIT Lincoln Laboratory. We process the dataset contained in the file *H067037.2*, which consists of 30 720 temporal returns from 76 range cells with VV polarization. From the 3-D-clutter intensity field of the dataset H067037.2, reported in [29, Fig. 8], it is evident that the monitored area consists of two major regions, the first extending from range bin 1 to 48, and the second one from 49 to 76. As discussed in [41], these two macroregions correspond to range cells containing agricultural fields and windblown vegetation. At a finer scale analysis, it is possible to identify five main categories of terrains spread across the two macroregions. In what follows, we set $r = 2$ [42] and consider different values of L; we refer the reader to [41], [43], [44], and references therein for more details about the PhaseOne datasets.

a) *P_d performance analysis:* We first compute the thresholds for all detectors in order to guarantee exactly the same P_{fa} on the range bin $R = 37$, selected as CUT, with the secondary data picked from range bins adjacent to the CUT, with indices in $[R - K/2, R - 1] \cup [R + 1, R + K/2]$. The number of primary data is set to $N = 8$ pulses, whereas the number of secondary data is set to $K = 72$ so as to cover the whole extent (in range) of the dataset. Given the limited availability of samples, thresholds are set to guarantee a $P_{\text{fa}} = 10^{-2}$, and from one trial to the next one, the data window is slid by three pulses until a sufficient number of $100/P_{\text{fa}} = 10^4$ trials is reached. The P_d is estimated in a similar manner using 10^3 trials, with data generated by adding a synthetic target $\alpha \boldsymbol{v}$ to the CUT at different values of the SCNR, with $\boldsymbol{v} = [1 \ \exp(j2\pi f_d) \cdots \exp(j2\pi(N-1)f_d)]^T$, being f_d the normalized Doppler frequency. For the specific dataset at hand, the value of f_d is set to 0 in accordance with the normalized power spectral density (PSD) curves reported in [29], which show that the clutter exhibits a peak in correspondence of 0 Hz. This is tantamount to considering the average worst case of a target embedded

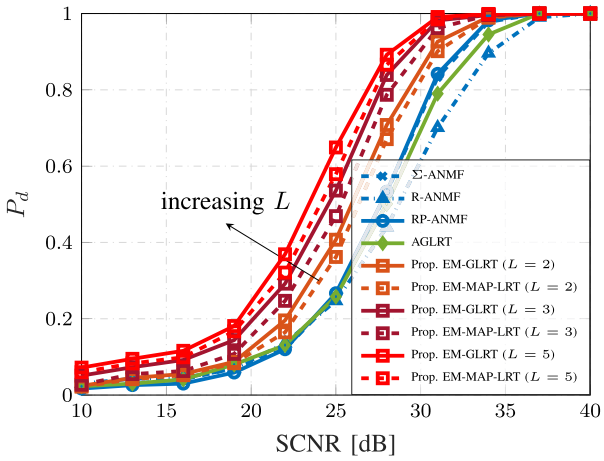

Fig. 7. P_d as a function of the SCNR for the proposed detectors, in comparison with natural competitors, assuming $N = 8$, $K = 72$, for varying settings of L, on the PhaseOne real dataset, using as CUT the range bin 37.

in clutter. Finally, we redefine the SCNR as $|\alpha|^2/\hat{\sigma}^2$, where $\hat{\sigma}^2$ is the average power of the clutter plus noise estimated from the I/Q samples of each range bin.

In Fig. 7, we report the P_d curves of all the detectors as a function of the SCNR, considering three different choices of L, namely $L = 2$, $L = 3$, and $L = 5$. Remarkably, the EM-GLRT and the EM-MAP-LRT confirm their superior performance compared to state-of-the-art detectors, providing an improvement in the detection power that is already about 2 dB for $L = 2$. Under this setup, the proposed algorithms are able to fit the heterogeneous environment only on a larger spatial scale, following the dimension of the two macroregions discussed previously. Accordingly, in the estimation process they try to assign the CUT (both algorithms) and the secondary data (only the EM-GLRT) to one of the two possible clutter classes. When a value of $L = 5$ is used to reparameterize the proposed algorithms, it becomes possible to distinguish between all the five terrains found in the dataset, allowing in turn a better adaptation of the algorithms to the different clutter classes on a much finer scale. This immediately translates into an evident increase of the performance in terms of P_d, leading to an improvement of up to about 5 dB compared to the ANMF schemes and AGLRT.

b) *P_{fa} sensitivity analysis:* For completeness, we also investigate the CFAR behavior of the considered detectors, namely their sensitivity in terms of variations of the actual P_{fa} from its nominal value. The analysis is conducted by assuming that the thresholds for all the algorithms are simply set on synthetic data to guarantee a nominal $P_{\text{fa}} = 10^{-2}$, considering the sole presence of additive white noise in the initial model given in (1) and (2), i.e., $\boldsymbol{R}_l = \boldsymbol{0}$, $l = 1, \ldots, L$, and $\sigma_w^2 = 1$. Since the size of the secondary data ($K = 72$) occupies almost the entire extent (in range) of the PhaseOne dataset to capture the heterogeneity of the observed scenario, we opted to compute different estimates of the P_{fa} using the temporal (slow time) dimension of the CUT (range bin 37). More specifically, we consider three different time windows (denoted as Window A, B, and C and labeled as W-A, W-B, and W-C for short) that divide the first 30 000 samples of range bin 37 in equal parts (10 000 samples per each). Within each window, the N-dimensional vector of primary data is slid by 1 pulse until reaching the end of the considered window. An estimate of the actual P_{fa} of each detector, denoted by $\hat{P}_{\text{fa}}$, is then obtained by counting the total number of false alarms across each individual temporal window. The results reported in Table II reveal that the proposed EM-MAP-LRT is the least sensitive to deviations from the nominal P_{fa}, whereas the Σ-ANMF turns out to be the most sensitive. Notably, the proposed EM-GLRT shares the same weak sensitivity of the R-ANMF and RP-ANMF detectors, while the AGLRT detector appears slightly more sensitive. Overall, the analysis shows that the proposed detectors guarantee P_{fa} values that do not deviate too much from the desired nominal one when operating on the PhaseOne dataset H067037.2, despite the fact that their thresholds were set to a simplified synthetic model of white noise only, which is clearly mismatched with respect to the actual disturbance found in the dataset.

TABLE II
$\hat{P}_{\text{fa}}/P_{\text{fa}}$ for PhaseOne Dataset H067037.2, With $P_{\text{fa}} = 10^{-2}$

	Σ-ANMF	R-ANMF	RP-ANMF	AGLRT	EM-GLRT	EM-MAP-LRT
W-A	38.36	4.81	4.19	7.62	4.74	1.98
W-B	14.19	2.53	2.64	6.08	3.21	1.51
W-C	10.82	2.79	2.48	4.95	2.55	1.87

2) *Analysis on IPIX Data:* We now consider another well-known real dataset collected by the McMaster IPIX radar overlooking Lake Ontario from the shore in Grimsby, in the winter of 1998; the database is freely available in [45]. The acquisition system is a fully coherent X-band radar, with advanced features such as dual transmit/receive polarization. The radar was originally developed for iceberg detection, but after major upgrades between 1993 and 1998, it became a benchmark for testing advanced detection algorithms. The dynamic range is 10 bits, the transmitted power is 8 kW, the carrier frequency is 9.39 GHz (fixed), or ranges from 8.9 to 9.4 GHz (agile), with a bandwidth of 25 MHz [45].

For the analysis purpose, we process the dataset file *19980223_165836_antstep.cdf* (D84 for short), which contains samples corresponding to ranges from 3000 to 3990 m, with a range resolution of 30 m and for a total of 34 different range bins with VV polarization. For each range bin, the signal is recorded for 60 s, corresponding to a total of 60 000 pulses. The dataset has been preprocessed to remove mean and normalize standard deviation from the I and Q channels separately and to compensate the phase imbalance due to hardware imperfections. In Fig. 8, we report the 3-D clutter intensity field from the IPIX D84 dataset. Compared to the PhaseOne dataset, which captures returns from land, the sea clutter returns in the D84 are visibly contaminated by power variations over the different range bins, clutter discretes, and other outliers that introduce variations also over the temporal (slow-time) domain. Overall, two main

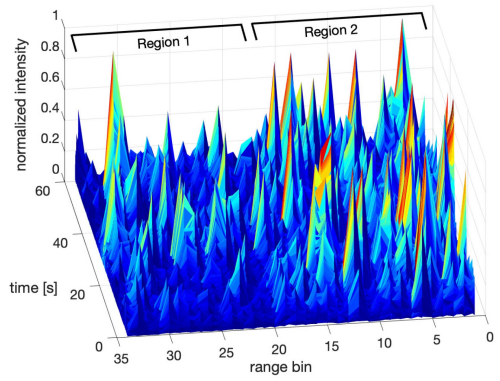

Fig. 8. 3-D normalized intensity field of clutter returns for IPIX D84 dataset.

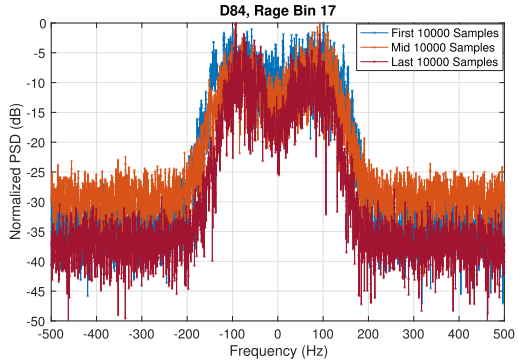

Fig. 9. Normalized PSDs of clutter for IPIX D84 (VV polarization), range bin 17, obtained using only the first, mid, and last 10 s of acquisition.

regions extending from range bin 1 to 17 and from 18 to 34 can be identified. For a more detailed discussion on the IPIX datasets, we refer the interested reader to [9], [46], and [47].

To better inspect the strong heterogeneity of the D84 dataset, in Fig. 9, we report the normalized PSD for the range bin 17. Different curves are obtained by applying the Welch method [48] fed with data over the integer set $[N_{\text{offset}}+1, N_{\text{offset}}+N_d]$, where N_d is the number of processed data and a 50% overlap between segments of length $N_w = 4\,096$ is considered; segments are multiplied by a Blackman window (of length N_w) using the built-in MATLAB function pwelch(x, $window$) (version 2020b). By fixing $N_d = 10\,000$, the parameter N_{offset} is then used to select N_d samples corresponding to the first ($N_{\text{offset}} = 0$), mid ($N_{\text{offset}} = 20000$), and last ($N_{\text{offset}} = 50000$) 10 s of acquisition, respectively, so as to assess the variations of the PSD over the slow time (i.e., intracell variations). Some interesting conclusions can be deduced from the analysis of Fig. 9. First, it is evident that the clutter has a peak around about 80 Hz, and that the gap between the peak and the floor levels at higher frequencies varies between about 25 and 40 dB. More generally, although the general shape of the PSDs is shared among the three intervals of samples, curves visibly differ among each other, revealing that the nature of the clutter is nonstationary over time. Overall, Fig. 9 coupled with other findings in [46] confirms a marked heterogeneity

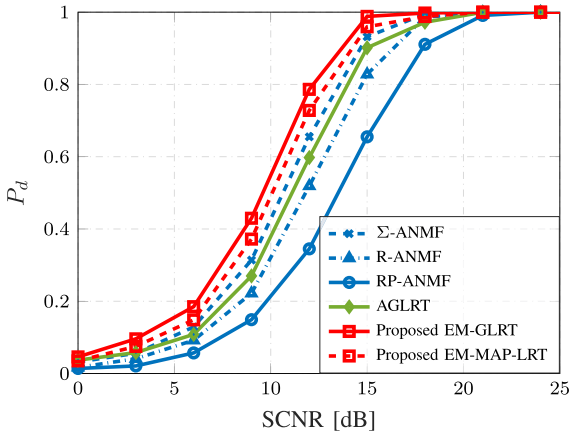

Fig. 10. P_d as a function of the SCNR for the proposed detectors, in comparison with natural competitors, assuming $N = 8$, $K = 32$, for $L = 2$, on the IPIX D84 dataset, using as CUT the range bin 17.

TABLE III
$\hat{P}_{\text{fa}}/P_{\text{fa}}$ for IPIX Dataset D84, With $P_{\text{fa}} = 10^{-2}$

	Σ-ANMF	R-ANMF	RP-ANMF	AGLRT	EM-GLRT	EM-MAP-LRT
W-A	17.94	5.56	2.52	6.61	2.86	3.82
W-B	12.24	3.63	1.49	4.68	3.12	3.42
W-C	7.58	1.87	0.54	3.45	2.74	1.58

of D84 in both fast time (spatial dimension) and slow time, also with different floor levels.

a) *P_d performance analysis:* The thresholds for all the detectors are computed by using as CUT the range bin $R = 17$ previously analyzed in Fig. 9, following the same procedure as for the PhaseOne dataset and for $P_{\text{fa}} = 10^{-2}$. In this case, however, we set the number of secondary data to $K = 32$ since the number of range bins available in the dataset is lower than that of PhaseOne data. Furthermore, we set $f_d = 0.08$ to generate the synthetic target in order to be consistent with the clutter peak found in Fig. 9. Finally, we set $r = 3$.

In Fig. 10, we report the P_d curves of all the detectors as a function of the SCNR, assuming $L = 2$ as observed in Fig. 8. Further analysis not reported here for brevity confirms the aforementioned results using $L = 3$. Interestingly, we observe that the novel detection schemes continue to outperform all the competitors, with a less marked but still visible P_d improvement. Among the competitors, the Σ-ANMF and the AGLRT offer the best performance, whereas the RP-ANMF exhibits an evident loss compared to its previous behavior observed in Fig. 7. The outcome of this analysis confirms the effectiveness of the proposed algorithms, highlighting their ability to correctly deal also with the presence of stronger heterogeneous and nonstationary clutter conditions as those found in D84.

b) *P_{fa} sensitivity analysis:* We conclude the analysis by testing the sensitivity of the algorithms to variations of their actual P_{fa} compared to its nominal value. The analysis is carried out following the same procedure adopted for Table II. The results reported in Table III confirm the goodness of the proposed approaches, which experience small deviations from the nominal P_{fa} value despite the

more challenging operational scenario under consideration. In this case, the RP-ANMF algorithm turns out to be the least sensitive, whereas the Σ-ANMF exhibits significant deviations toward much larger values of P_{fa}, confirming its too high sensitivity.

V. CONCLUSION

We have addressed the detection of a coherent target in heterogeneous environments where both power levels and covariance structures of the clutter may vary in range. To this end, the clutter covariances for the CUT and secondary data are modeled as low-rank matrices belonging to a set of L classes, each representing a different type of clutter return.

To handle the unknown association of each range cell to one of the L possible clutter classes, we introduced a latent variables mixture model and formalized a binary hypothesis test with observations modeled by a multivariate contaminated Gaussian distribution. We tackled the resulting detection problem via a GLRT approach and proposed a novel strategy based on the EM algorithm to adaptively estimate the unknown parameters related to the statistical properties of the disturbance. As a byproduct of the EM procedure, we have also derived an alternative detection scheme that tries to infer the most probable disturbance distribution in the CUT using a MAP classification rule followed by an LRT. When tested on synthetic data, the two detection schemes can achieve significant improvements in terms of detection power compared to state-of-the-art algorithms.

Furthermore, they can also operate under mismatches on the assumed design model. Remarkably, the proposed strategies confirmed their effectiveness also when operating on two rather different experimental datasets, each comprising a different type of clutter (land and sea). From these analyses, we conclude that the novel algorithms are able to fully adapt to heterogeneous environments and offer superior performance compared to existing approaches.

Future research tracks might encompass the detection of range-spread targets whose position and extension are unknown under the same heterogeneous assumptions and/or the exploitation of special structures for the covariance matrices. Moreover, other approaches might be investigated, as for instance those relying on the theory of machine learning. A very simple attempt in this respect has been done in [49] where the potential of a K-nearest neighbor classifier is investigated. We expect that more general approaches based on neural networks [50], [51] might lead to even more powerful decision schemes at the price of a more challenging training stage.

Finally, the presence of outliers in the secondary data is a problem of relevant interest in radar detection, as they can lead to increased false alarms or missed detections. To mitigate the influence of outliers, a possible generalization of the proposed decision scheme could include an initial stage focused on detecting and eliminating cells near the CUT (secondary data) that are affected by outliers. This approach aligns with established methodologies in radar detection contexts such as [40] and [52]. Building on a similar idea, the algorithm introduced in [30], when applied to the secondary data, is able to adaptively exclude range cells that might contain components related to targets (hence representing outliers for secondary data, which are usually assumed free of target components). Developing an effective strategy that incorporates an outlier detection method capable of handling the heterogeneous environments in which the newly proposed detection schemes are deployed thus represents another interesting direction for future work.

APPENDIX A
PROOF OF (22)

As to the estimation of the p_ls, we preliminarily observe that constraints $p_l \geq 0, l = 1, \ldots, L$, will be automatically satisfied by the solution of the following problem:

$$\begin{cases} \max\limits_{p_l, l=1,\ldots,L} \sum_{l=1}^{L} \left[\left(\sum_{k=1}^{K} q_k^{(h-1)}(l) + \tilde{q}^{(h-1)}(l) \right) \log p_l \right] \\ \text{s.t.} \ \sum_{l=1}^{L} p_l = 1 \end{cases}.$$

Using the method of Lagrange multipliers, we construct the function

$$g(p_1, \ldots, p_L, \lambda) = \sum_{l=1}^{L} a_l \log p_l + \lambda \left(\sum_{l=1}^{L} p_l - 1 \right)$$

with

$$a_l = \sum_{k=1}^{K} q_k^{(h-1)}(l) + \tilde{q}^{(h-1)}(l)$$

while λ is the Lagrange multiplier. Computing the derivative of g with respect to p_l and setting such derivative equal to zero, yields

$$\frac{\partial g(p_1, \ldots, p_L, \lambda)}{\partial p_l} = \frac{a_l}{p_l} + \lambda = 0$$

and hence, we have that

$$p_l = -\frac{a_l}{\lambda}.$$

The constraint implies that

$$\begin{aligned} \sum_{l=1}^{L} p_l &= -\frac{1}{\lambda} \sum_{l=1}^{L} a_l \\ &= -\frac{1}{\lambda} \sum_{l=1}^{L} \left(\sum_{k=1}^{K} q_k^{(h-1)}(l) + \tilde{q}^{(h-1)}(l) \right) = 1 \end{aligned}$$

and hence, $\lambda = -(K+1)$ exploiting the conditions

$$\sum_{l=1}^{L} q_k^{(h-1)}(l) = 1, \quad k = 1, \ldots, K$$

and

$$\sum_{l=1}^{L} \tilde{q}^{(h-1)}(l) = 1.$$

Equation (22) follows in a straightforward manner. ■

ACKNOWLEDGMENT

The authors would like to thank Prof. S. Haykin for the IPIX data and Prof. J. B. Billingsley for the Phase One L-band data.

REFERENCES

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

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

[3] S. Kalson, "An adaptive array detector with mismatched signal rejection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 195–207, Jan. 1992.

[4] A. Coluccia, A. Fascista, and G. Ricci, "CFAR feature plane: A novel framework for the analysis and design of radar detectors," *IEEE Trans. Signal Process.*, vol. 68, pp. 3903–3916, 2020.

[5] N. Pulsone and C. Rader, "Adaptive beamformer orthogonal rejection test," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 521–529, Mar. 2001.

[6] A. Coluccia, A. Fascista, and G. Ricci, "Design of customized adaptive radar detectors in the CFAR feature plane," *IEEE Trans. Signal Process.*, vol. 70, pp. 5133–5147, Oct. 2022, doi: 10.1109/TSP.2022.3216372.

[7] K. D. Ward, "Compound representation of high resolution sea clutter," *Electron. Lett.*, vol. 17, no. 16, pp. 561–563, Aug. 1981.

[8] 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.

[9] 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 Navig.*, vol. 144, no. 3, pp. 121–130, 1997.

[10] 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.

[11] K. D. Ward, R. J. A. Tough, and S. Watts, *Sea Clutter, Scattering, The K Distrib. and Radar Perform.*, 2nd ed. Stevenage, UK: Inst. Eng. Technol., 2013.

[12] E. Conte and M. Longo, "Characterisation of radar clutter as a spherically invariant random process," *IEE Proc. F.—Commun., Radar Signal Process.*, vol. 134, no. 2, pp. 191–197, 1987.

[13] E. Conte, M. Longo, and M. Lops, "Modelling and simulation of non-Rayleigh radar clutter," *IEE Proc. F—Radar Signal Process.*, vol. 138, no. 2, pp. 121–130, 1991.

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

[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. 28, no. 1, pp. 208–216, Jan. 1992.

[16] D. Orlando, G. Ricci, and L. L. Scharf, "A unified theory of adaptive subspace detection Part I: Detector designs," *IEEE Trans. Signal Process.*, vol. 70, pp. 4925–4938, Sep. 2022, doi: 10.1109/TSP.2022.3205761..

[17] E. Conte, M. Lops, and G. Ricci, "Adaptive radar detection in compound-Gaussian clutter," in *Proc. 7th Eur. Signal Process. Conf.*, Sep. 1994, pp. 526–529.

[18] E. Conte, M. Lops, and G. Ricci, "Adaptive detection schemes in compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 4, pp. 1058–1069, Oct. 1998.

[19] 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.

[20] M. S. Greco and F. Gini, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Process.*, vol. 82, no. 12, pp. 1847–1859, 2002.

[21] E. Conte and A. De Maio, "Mitigation techniques for non-Gaussian sea clutter," *IEEE J. Ocean. Eng.*, vol. 29, no. 2, pp. 284–302, Apr. 2004.

[22] F. Pascal, Y. Chitour, J. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 34–48, Jan. 2008.

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

[24] 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, Jul. 2020, doi: 10.1109/TSP.2020.3009836.

[25] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, 1977.

[26] K. Murphy, *Machine Learning: A Probabilistic Perspective.* Cambridge, MA, USA: MIT Press, 2012.

[27] J. R. Guerci, "Cognitive radar: A knowledge-aided fully adaptive approach," in *Proc. IEEE Radar Conf.*, 2010, pp. 1365–1370.

[28] Y. Mhiri, M. N. El Korso, L. Bacharach, A. Breloy, and P. Larzabal, "Expectation-maximization based direction of arrival estimation under a mixture of noise," in *Proc. 29th Eur. Signal Process. Conf.*, 2021, pp. 1890–1893.

[29] P. Addabbo, S. Han, D. Orlando, and G. Ricci, "Learning strategies for radar clutter classification," *IEEE Trans. Signal Process.*, vol. 69, pp. 1070–1082, Jan. 2021, doi: 10.1109/TSP.2021.3050985.

[30] L. Yan, S. Han, C. Hao, D. Orlando, and G. Ricci, "Innovative cognitive approaches for joint radar clutter classification and multiple target detection in heterogeneous environments," *IEEE Trans. Signal Process.*, vol. 71, pp. 1010–1022, Mar. 2023, doi: 10.1109/TSP.2023.3250084.

[31] B. Kang, V. Monga, and M. Rangaswamy, "Rank-constrained maximum likelihood estimation of structured covariance matrices," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 501–515, Jan. 2014.

[32] Z. Chen, H. Li, and M. Rangaswamy, "Conjugate gradient adaptive matched filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 1, pp. 178–191, Jan. 2015.

[33] R. Klemm, "Principles of space-time adaptive processing," *IEE Radar, Sonar, Navig. Avionics*, 2006. [Online]. Available: https://digital-library.theiet.org/content/books/ra/pbra021e

[34] C. Peckham, A. Haimovich, T. Ayoub, J. Goldstein, and I. Reid, "Reduced-rank STAP performance analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 664–676, Apr. 2000.

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

[36] L. E. Brennan and F. M. Staudaher, "Subclutter visibility demonstration," Adaptive Sensors Incorporated, Tech. Rep. RL-TR-92-21, 1992.

[37] G. Ginolhac, P. Forster, F. Pascal, and J.-P. Ovarlez, "Performance of two low-rank STAP filters in a heterogeneous noise," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 57–61, Jan. 2013.

[38] Y. Sun, A. Breloy, P. Babu, D. P. Palomar, F. Pascal, and G. Ginolhac, "Low-complexity algorithms for low rank clutter parameters estimation in radar systems," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1986–1998, Apr. 2016.

[39] W. L. Melvin, Ed., *Principles of Modern Radar: Advanced Techniques.* Stevenage, UK: Institution of Engineering and Technology, 2012.

[40] S. Blunt and K. Gerlach, "Doppler-sensitive adaptive coherence estimate detector methods," US Patent 2006/0238408 A1.

[41] E. Conte, A. De Maio, and A. Farina, "Statistical tests for higher order analysis of radar clutter: Their application to L-band measured data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 205–218, Jan. 2005.

[42] D. Xu, P. Addabbo, C. Hao, J. Liu, D. Orlando, and A. Farina, "Adaptive strategies for clutter edge detection in radar," *Signal Process.*, vol. 186, 2021, Art. no. 108127.

[43] J. B. Billingsley, A. Farina, F. Gini, M. V. Greco, and L. Verrazzani, "Statistical analyses of measured radar ground clutter data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 2, pp. 579–593, Apr. 1999.

[44] M. Greco, F. Gini, A. Farina, and J. B. Billingsley, "Validation of windblown radar ground clutter spectral shape," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 2, pp. 538–548, Apr. 2001.

[45] "McMaster University - IPIX radar database," Apr. 1, 2023. [Online]. Available: http://soma.mcmaster.ca/ipix.php
[46] A. Coluccia, A. Fascista, D. Orlando, and G. Ricci, "Radar detectors for heterogeneous environments: A comparison on IPIX data," in *Proc. Int. Conf. Radar Syst.*, 2022, pp. 324–329.
[47] E. Conte, A. De Maio, and C. Galdi, "Statistical analysis of real clutter at different range resolutions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 3, pp. 903–918, Jul. 2004.
[48] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing*, 4th ed. Hoboken, NJ, USA: Prentice-Hall, 2006.
[49] A. Coluccia, A. Fascista, and G. Ricci, "A KNN-Based radar detector for coherent targets in non-Gaussian noise," *IEEE Signal Process. Lett.*, vol. 28, pp. 778–782, Apr. 2021, doi: 10.1109/LSP.2021.3071972.
[50] T. Diskin, U. Okun, and A. Wiesel, "Learning to detect with constant false alarm rate," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun.*, 2022, pp. 1–5.
[51] T. Diskin, Y. Beer, U. Okun, and A. Wiesel, "CFARNet: Deep learning for target detection with constant false alarm rate," 2023, *arXiv:2208.02474*.
[52] J. R. Guerci, *Cognitive Radar: The Knowledge-aided Fully Adaptive Approach.* Norwood, MA, USA: Artech House, 2010.

Angelo Coluccia (Senior Member, IEEE) received the Ph.D. degree in information engineering from University of Salento, Lecce, Italy, in 2011.

He has been a Research Fellow with Forschungszentrum Telekommunikation Wien, Vienna, Austria, and has held a visiting position with the Department of Electronics, Optronics, and Signals, Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, France. He is currently an Associate Professor of telecommunications with the Department of Engineering, University of Salento, Lecce, Italy. His research interests include the area of multichannel, multisensor, and multiagent statistical signal processing for detection, estimation, localization, and learning problems, with relevant application fields in radar, wireless networks (including 5 G and beyond), and emerging network contexts (including intelligent cyber-physical systems, smart devices, and social networks).

Dr. Coluccia is a Member of the Sensor Array and Multichannel Technical Committee for IEEE Signal Processing Society, a Member of the Data Science Initiative for IEEE Signal Processing Society, and a Member of the Technical Area Committee in Signal Processing for Multisensor Systems of European Association for Signal Processing (EURASIP).

Alessio Fascista (Member, IEEE) received the Ph.D. degree in engineering of complex systems from the University of Salento, Lecce, Italy, in 2019.

He has held a visiting position with the Department of Telecommunications and Systems Engineering, Universitat Autonoma de Barcelona, Spain, in 2018, and with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden, in 2022. He is currently an Assistant Professor in telecommunications with the Department of Innovation Engineering, University of Salento. His main research interests include the field of telecommunications with focus on statistical signal processing for detection, estimation, and localization in terrestrial wireless systems.

Dr. Fascista is a Member of the Technical Area Committee in Signal Processing for Multisensor Systems of European Association for Signal Processing (EURASIP). He serves as an Associate Editor for IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY (OJ-COMS).

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 2007 to 2010, he was with the University of Cassino, Cassino, Italy, involved in a research project on algorithms for track-before-detect of multiple targets in uncertain scenarios. In 2009, he was a Visiting Scientist with the NATO Undersea Research Centre, La Spezia, Italy. From 2011 to 2015, he was with Elettronica S.p.A. involved as a System Analyst in the field of electronic warfare. In 2015, he joined Università degli Studi "Niccolo' Cusano," Rome, Italy, where he is currently an Associate Professor. In 2007, he has held visiting positions with the Department of Avionics and Systems, ENSICA (now Institut Supérieur de l'Aéronautique et de l'Espace, ISAE), Toulouse, France, and from 2017 to 2019, with the Chinese Academy of Science, Beijing, China. He is the author or co-author of more than 180 scientific publications in international journals, conferences, and books. His main research interests include statistical signal processing with more emphasis on adaptive detection and tracking of multiple targets in multisensor scenarios.

Dr. Orlando was the Senior Area Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING and currently an Associate Editor for IEEE OPEN JOURNAL OF SIGNAL PROCESSING, *EURASIP Journal on Advances in Signal Processing*, and *MDPI Remote Sensing*.

Giuseppe Ricci (Senior Member, IEEE) was born in Naples, Italy, in 1964. He received the Dr. and the Ph.D. degrees, both in electronic engineering, from the University of Naples "Federico II" in 1990 and 1994, respectively.

He has held visiting positions with the University of Colorado at Boulder, CO, USA, in 1997–1998 and 2001, with the Colorado State University, Fort Collins, CO, in 2003, 2005, 2009, and 2011, with Ensica (Toulouse, France) in 2006, with the University of Connecticut, Storrs CT, USA, in 2008, and with Arizona State University, Tempe, AZ, USA, in 2019. Since 1995, he has been with the University of Salento (formerly University of Lecce), Lecce, Italy, first as an Assistant Professor of telecommunications, and since 2002, as a Professor. His research interests include the field of statistical signal processing with emphasis on radar processing, localization algorithms, and CDMA systems.

Dr. Ricci has been a Member of the Multisensor Target Tracking Working Group (MSTWG), operating under the auspices of the International Society of Information Fusion (ISIF), from 2012 until 2016, and of the Special Area Team (SAT) Eurasip Signal Processing for Multisensor Systems (SPMuS) from 2016 until 2021. He has been tutorial co-chair for the IEEE Radar Conference, Florence, Italy, in 2020. He has also been Rector Delegate for teaching activities and continuing education in 2016, until 2017.

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