Novel Parameter Estimation and Radar Detection Approaches for Multiple Point-Like Targets: Designs and Comparisons

Pia Addabbo, Senior Member, IEEE, Jun Liu, Senior Member, IEEE, Danilo Orlando, and Giuseppe Ricci, Senior Member, IEEE

Abstract—In this work, we develop and compare two innovative strategies for parameter estimation and radar detection of multiple point-like targets. The first strategy, which appears here for the first time, jointly exploits the maximum likelihood approach and Bayesian learning to estimate targets’ parameters including their positions in terms of range bins. The second strategy relies on the intuition that for high signal-to-interference-plus-noise ratio values, the energy of data containing target components projected onto the nominal steering direction should be higher than the energy of data affected by interference only. The adaptivity with respect to the interference covariance matrix is also considered exploiting a training data set collected in the proximity of the window under test. Finally, another important innovation aspect concerns the adaptive estimation of the unknown number of targets by means of the model order selection rules.

Index Terms—Adaptive detection, bayesian learning, model order selection, multiple targets, radar, target localization, unsupervised learning.

I. INTRODUCTION

In the recent years, advances in technology have paved the way for the design of radar processing units where high-performance sophisticated algorithms are executed to comply with the radar tasks and the system requirements. Consider, for example, the space-time adaptive detection algorithms that exploit large volumes of data from sensor arrays and/or pulse trains to take advantage of temporal and spatial integration/diversity [1], [2] at the price of an increased computational load [3]–[9]. Another important example concerns high resolution radars, which can resolve a target into a number of different scattering centers depending on the radar bandwidth and the range extent of the target [5], [10], [11]. In this case, conventional radar detection algorithms process one range bin at a time despite the fact that contiguous cells can contain target energy. As a consequence, they do not collect as much energy as possible to increase the signal-to-interference-plus-noise ratio (SINR). To overcome this drawback, architectures capable of detecting distributed targets by exploiting a certain number of contiguous range bins have been developed [5], [12], [13].

The energy issues described for range-spread targets also hold for multiple point-like targets and detection algorithms, which can take advantage of the total energy associated with all the point-like targets, are highly desirable. However, the problem of jointly detecting multiple point-like targets is very difficult since target positions along with target number (which deserves special attention) are unknown parameters that must be estimated. Existing examples in the open literature share the assumption that the number of targets (or at least an upper bound on it) is known and are based upon the maximum likelihood (ML) approach [14], [15].

In this work, we devise two architectures to detect multiple point-like targets without assuming that the number of targets is known. The first approach jointly exploits ML and Bayesian learning [16] to estimate targets’ parameters including their positions in terms of range bins. The second approach relies on the intuition that for high SINR values, the energy of data containing target components projected onto the nominal steering direction should be higher than the energy of data affected by interference only (it can be shown that it is a special case of [14]). Thus, it is possible to discard range bins with low energy according to a given criterion. In both cases, ad hoc modifications of the generalized likelihood ratio test (GLRT) are used to design adaptive architectures based upon previous estimates. Remarkably, the number of targets is adaptively estimated by means of model order selection (MOS) rules [17], as the Bayesian information criterion (BIC), the Akaike information criterion (AIC), and the generalized information criterion (GIC).

Finally, the performance analysis, conducted on simulated data and in comparison with classical detectors for extended...
and/or multiple targets, is aimed at showing the effectiveness of the proposed architectures in terms of both estimation and detection performance.

The remainder of this letter is organized as follows. In Section II, we formulate the detection problem at hand and define preliminary quantities which are used in Section III to derive two adaptive procedures. Finally, in Section IV we assess the performance of the proposed methods in comparison with their classical counterparts, whereas in Section V we draw some concluding remarks.

II. PROBLEM FORMULATION

Let us consider a radar system that transmits a burst of \( N_p \) pulses by means of a linear array of \( N_a \) antennas to sense the surrounding environment. The received echoes from the environment are suitably conditioned and organized to form \( N \)-dimensional vectors with \( N = N_aN_p \) representing the range bins [2], [18]. Thus, for each range bin belonging to the window under test (WUT), the corresponding \( N \)-dimensional vector is the result of the superposition of an interference component (representative of thermal noise, clutter, etc.) and possible useful signal components. Besides, under the hypothesis that the environment is stationary over range and time (the so-called homogeneous environment), a further set of training samples (secondary data) is collected by the system in proximity of the WUT and used to achieve adaptivity with respect to the interference covariance matrix (ICM) [2]–[4], [18] (data sets to be processed are depicted in Fig. 1).

Thus, denoting by \( Z = [z_1 \cdots z_{K_S}] \in \mathbb{C}^{N \times K_S} \) a matrix whose columns are the vectors belonging to the WUT and by \( R = [r_1 \cdots r_{K_S}] \in \mathbb{C}^{N \times K_S} \) the secondary data matrix, we are interested in solving the following hypothesis test

\[
\begin{align*}
H_1 : & \left\{ \begin{array}{l}
z_h & \sim \mathcal{CN}_N(\alpha_h \nu(\theta_T, \nu), M), \ h \in \Omega_T, \\
r_k & \sim \mathcal{CN}_N(0, M), \ k \in \Omega_S,
\end{array} \right. \\
H_0 : & \left\{ \begin{array}{l}
z_h & \sim \mathcal{CN}_N(0, M), \ h \in \Omega_T, \\
r_k & \sim \mathcal{CN}_N(0, M), \ k \in \Omega_S,
\end{array} \right.
\end{align*}
\]

(1)

where \( \alpha_h \in \mathbb{C} \) is a complex factor accounting for the received energy backscattered by a coherent target; \( \nu(\theta_T, \nu) \) is the space-time steering vector with \( \nu \) the normalized target Doppler frequency and \( \theta_T \) the target angle of arrival measured with respect to the array normal\(^2\); \( z_1, \ldots, z_{K_S}, r_1, \ldots, r_{K_S} \) are statistically independent random vectors; \( \Omega_T = \{1, \ldots, K_T\} \), \( \Omega_S = \{1, \ldots, K_S\} \), and \( \Omega_T \subseteq \Omega_P \) with \( |\Omega_T| = K_T \).

It is important to underline here that the above problem describes an operating situation where the radar system jointly processes a set of contiguous range bins which might contain \( K_T \leq K_P \) point-like targets. The main related problem is that \( \Omega_T \) is unknown and must be somehow estimated from data.

Finally, for future developments, let us write the PDF of \( z_h \) under \( H_1 \), \( i = 0, 1 \), as

\[
f(z_h; i_{\Omega_T}(h) i \alpha_h, M) = \frac{e^{-\frac{1}{2} || z_h - i_{\Omega_T}(h) i \alpha_h ||^2}}{\pi^N \det(M)},
\]

where \( i_{\Omega_T}(h) \) is the indicator function of \( \Omega_T \). It follows that the joint PDF of \( Z \) can be written as

\[
f_1(Z; M, K_T, A, \Omega_T) = \prod_{h \in \Omega_T} f(z_h; \alpha_h, M) \prod_{h \in \Omega_P \setminus \Omega_T} f(z_h; 0, M)
\]

under \( H_1 \), where \( A = \{ \alpha_h : h \in \Omega_T \} \), and as \( f_0(Z; M) = \prod_{h=1}^{K_P} f(z_h; 0, M) \) under \( H_0 \).

III. ADAPTIVE DETECTOR DESIGN

In this section, we drive the reader towards the design of fully-adaptive detection architectures by gradually adding an adaptivity layer at each step. As a matter of fact, in the first two subsections, both estimation procedures along with the related decision schemes are derived assuming that \( K_T \) and \( M \) (namely the ICM) are known (the remaining parameters are clearly estimated from data). In the last subsection, we make the previously developed architectures adaptive with respect to the ICM and \( K_T \). To this end, the ICM is replaced by the sample covariance matrix, whereas the adaptivity with respect to \( K_T \) is achieved by resorting to the MOS rules, which allow us to estimate \( K_T \). Finally, the adopted detector design criteria rely on ad hoc modifications of the GLRT where the unknown parameters are replaced by the estimates returned by the proposed procedures.

Precisely, the GLRT for known \( M \) and \( K_T \) (and based upon the range bins of the WUT) has the following expression

\[
\max_{\Omega_T} \max_{A} \frac{H_1}{\Omega_T} A_1(Z; M, K_T, A, \Omega_T) \geq \eta, \quad \frac{H_0}{H_0}
\]

(2)

\(^1\)Notation: In the sequel, vectors and matrices are denoted by boldface lowercase and upper-case letters, respectively. For a generic vector \( x \), the symbol \( ||x|| \) indicates its Euclidean norm. The symbols \( \det(\cdot) \), \( (\cdot)^T \), and \( (\cdot)^\dagger \) denote the determinant, 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 (or vectors if \( M = 1 \)). Let \( x \in \mathbb{C} \), then \( ||x|| \) denotes the modulus of \( x \), whereas for each set \( A, |A| \) stands for the cardinality of \( A \). The set difference is denoted by \( \setminus \), while \( \binom{x}{\mathbb{C}} \) indicates the binomial coefficient. Symbols \( I \) and \( 0 \) represent the identity matrix and the null vector or matrix, respectively, of proper dimensions. Let \( x \) a random vector, we denote by \( f(x; \theta) \) the probability density function (PDF) of \( x \) with parameter vector \( \theta \). If \( C \) is an event and \( x \) a random vector, \( P(C|x) \) and \( P(C|\theta) \) are the probability of \( C \) and the probability of \( C \) given \( x \), respectively. Finally, we write \( x \sim \mathcal{CN}_N(m, M) \) if \( x \) is a \( N \)-dimensional complex normal random vector with mean \( m \in \mathbb{C}^{N \times 1} \) and positive definite covariance matrix \( M \in \mathbb{C}^{N \times N} \).

\(^2\)In the remainder of the paper for simplicity, we omit the dependency of \( \nu \) on \( \nu \) and \( \theta_T \).

\(^3\)Note that \( \lambda \) depends on \( \Omega_T \), which is function of the actual number of targets. However, we omit these dependences in order to simplify the notation.
where $\Lambda_1(Z; M, K_T, A, \Omega_T) = f_1(Z; M, A, K_T, \Omega_T) / f_0(Z; M)$ and $\eta$ is the threshold$^4$ to be set according to the desired probability of false alarm ($P_{fa}$). Before moving to the heart of the derivations, we briefly outline the reasoning followed in this subsection. Specifically, let us notice that solving problem (2) represents a difficult task (at least to the best of authors’ knowledge) due to the maximization with respect to $\Omega_T$. Indeed, an exhaustive search would deal with $(K_F)^K_T$ subsets of $\Omega_T$, each with cardinality $K_T$. Now, when $K_F / K_T$ becomes large, this approach would be prohibitive from a computational point of view. For this reason, we conceive two alternative and “smart” strategies.

A. Joint Bayesian and ML (BML) Estimation

In the first proposed solution, we follow an alternative route looking at problem (2) from a different perspective. The strategy consists in defining hidden random variables (RVs) which represent the classes to which the range bins belong with a given probability. Specifically, we treat the estimation of $\Omega_T$ like a clustering problem, where the class labeled as “0” corresponds to the absence of target while the bins containing prospective clusters are associated with the classes defined by indices greater than 0. As a consequence, it is possible to isolate the range bins containing interference only from the others. In this context, we assume that each range bin belongs to one of two classes, including labels “0” and “1,” which refer to the absence and the presence of the target, respectively. Specifically, let $b_h, h \in \Omega_T$, be independent RVs, with alphabet $B = \{0, 1\}$ and the related probability mass function $P(b_h = k) = \pi^{BML}_h, k \in \Omega_T, h \in B$. Note that $b_h = 1$, $h \in \Omega_T$, if and only if the $h$-th range bin contains a target echo with the consequence that, if $K_T$ targets are distributed over $K_F$ range bins, the resulting class priors are $\pi^{BML}_h = \frac{K_T}{K_F}$ and $\pi^{BML}_0 = \frac{K_F - K_T}{K_F}$. The PDF of $z_h, h \in \Omega_T$, under $H_1$, can be written as a linear combination of two “sub-PDFs” namely $f(z_h; \pi^{BML}_h, \alpha_T, M) = \pi^{BML}_h f(z_h; 0, M) + \pi^{BML}_0 f(z_h; \alpha_T, M)$, where $\pi^{BML} = [\pi^{BML}_0 \pi^{BML}_1]^T$. Now, we jointly exploit the Bayesian framework and the ML approach to come up with an estimate of $p^{BML}_{rk} = P(b_h = k | z_\hat{h})$ for each $h \in \Omega_T$. Thus, resorting to the Bayes rule, we can write

$$p^{BML}_{rk} = \frac{\pi^{BML}_k f(z_h; k \alpha_T, M)}{\sum_{i=0}^{1} \pi^{BML}_i f(z_h; i \alpha_T, M)} \tag{3}$$

As for the generic $\alpha_T$, we estimate it by solving a classical ML problem [3] to obtain $\hat{\alpha}_T = \hat{v} M^{-1} z_h / \hat{v} M^{-1} v$. Replacing this estimate in (3), we obtain

$$p^{BML}_{rk} = \frac{\pi^{BML}_k f(z_h; k \hat{\alpha}_T, M)}{\sum_{i=0}^{1} \pi^{BML}_i f(z_h; i \hat{\alpha}_T, M)} \tag{4}$$

Finally, $p^{BML}_{rk}$ can be suitably exploited to obtain an estimate of $\Omega_T$, say, by discarding the range bins indexed by $h \in \Omega_T$ such that $\arg \max_k p^{BML}_{rk} = 0$ and, if the remaining set contains more than $K_T$ elements, by selecting the remaining $K_T$ range bins with the highest $p^{BML}_{rk}$. When the number of the remaining bins is less than $K_T$, we select all of them. Now, given the estimates of $\alpha_T, h \in \Omega_T$, and $\Omega_T$, we can build up the following decision rules

$$\Lambda_1^{BML}(Z; M, K_T) = \frac{\prod_{h \in \Omega_T} f(z_h; \hat{\alpha}_T, M) H_1}{\prod_{h \in \Omega_T} f(z_h; 0, M) H_0} \tag{5}$$

that is statistically equivalent to the generalized adaptive matched filter (GAMF) computed over $\Omega_T^{BML}$ [5].

B. Estimation Based Upon Energy

From an intuitive point of view, when $z_h$ contains high-SINR target echoes, the energy of the component along $v$ or $M^{-1}v$ should be higher than the component in the case where the target is not present. Otherwise stated, let $z_h = \alpha_T v + n_h$, and $z_h = \eta h$, with $n_h \sim C_N(0, M)$, $i = 1, 2$, then we expect that $|v^T M^{-1} z_h|^2 > |v^T M^{-1} z_h|^2$. The above insight suggests a strategy for the selection of the elements belonging to $\Omega_T$. Specifically, let us order the range bins according to the energy amount of the component along $M^{-1/2} v$, namely $|v^T M^{-1/2} z_{h_1}|^2 > |v^T M^{-1/2} z_{h_2}|^2 > \ldots > |v^T M^{-1/2} z_{h_{K_F}}|^2$ and, then, select the first $K_T$ range bins $z_{h_1}, \ldots, z_{h_{K_T}}$ to form $\Omega_T^{WEN}$ (where WEN stands for Whitened ENergy). Finally, such estimates can be exploited to come up with the following decision rule$^6$

$$\Lambda_2^{WEN}(Z; M, K_T) = \sum_{k \in \Omega_T^{WEN}} \frac{|v^T M^{-1} z_{k}|^2 H_1}{|v^T M^{-1} v|^2 H_0} \tag{6}$$

C. Adaptivity With Respect $m$ and $k_t$

Assuming that the ICM is known does not have a practical value. As a matter of fact, the a priori information about the ICM is often limited to its specific structure. Such information might come from possible symmetries induced by system geometry or clutter properties [19]-[22]. For this reason, in order to make architectures (5) and (6) adaptive with respect to $M$, we replace the latter with the ubiquitous sample covariance matrix based upon secondary data set [4], [18], [23], namely $\hat{M} = \frac{1}{N} \hat{R} R^\dagger$.

The final step towards adaptivity consists in estimating $K_T$ from the observables. In fact, there exist situations where the a priori information about the actual number of targets is not available or is not exact. For this reason, we resort to the MOS rules whose general expression is

$$K_T = \arg \min_{k=1, \ldots, K_F} -2 \log f_1(Z; \hat{M}, k, \hat{A}(k), \hat{\Omega}_T(k)) + p(k). \tag{7}$$

In the above equation, $\hat{A}(k)$ and $\hat{\Omega}_T(k)$ are the estimates$^8$ of $A$ and $\Omega_T$, respectively, assuming that $K_T = k; p(k) = 3 k \nu$ is the penalty term [17], where the factor 3 represents the number of unknowns for each target, i.e., the complex-valued amplitude and its position, while $\nu = 2$ for AIC, $\nu = (1 + \rho), \rho \geq 1$, for GIC, and $\nu = \log K_F$ for BIC [17].

$^4$Hereafter, we denote by $\eta$ the generic detection threshold.

$^6$Actually MOS rules exploit the compressed likelihood function, where the parameters are replaced by the respective ML estimates. In our development, we replace the parameters with alternative estimates when the former are not available.
Finally, we obtain the adaptive detectors replacing the unknown matrix $M$ and $K_T$ in (5) and (6) with the sample covariance matrix and (7), respectively.

**IV. SIMULATION RESULTS**

In this section, we analyze the performance of the newly proposed unsupervised learning approaches resorting to standard Monte Carlo (MC) counting techniques. More precisely, the detection performance and the thresholds are estimated over $10^2$ and $100/P_f$ independent trials, respectively. In all the illustrative examples, we set $N = N_a = 16$, $K_T = 5$, $K_P = 30$, $K_S = 48$, and $P_f = 10^{-3}$. The ICM is given by $M = I + CNR M_c$, with a clutter to noise ratio (CNR) of 40 dB. The $(i, j)$th entry of the clutter component $M_c$ is given by $M_c(i, j) = \rho_c^{i-j}$ with $\rho_c = 0.9$. The SINR is defined as $\text{SINR} = P_o^{\alpha_p}(0, 0)^T M^{-1} v(0, 0)$, where $P_o^{\alpha_p}$ is the average power of the point-like targets.

In Figure 2(a), we plot the probability of detection ($P_d$) versus SINR for (5) and (6) with known $K_T$. For comparison purposes, we also report the $P_d$ curves of the so-called GAMF and GASD introduced in [5]. It is possible to observe that both the proposed procedures are capable of achieving a gain close to 1.6 dB at $P_d = 0.9$ with respect to the competitors. The remaining subfigures refer to the case where $K_T$ is unknown. In this case, due to the estimation of the latter, the gain over the GAMF and GASD becomes lower than or equal to about 1 dB for all the considered architectures. Specifically, the maximum gain is achieved by the GIC-based architectures, whereas the gain associated with the BIC-based detectors decrease to about 0.5 dB; the AIC-based decision schemes share the same performance as the GASD. Finally, notice that the similar detection performance of (5) and (6) can be explained by the fact that, as shown below, $\Omega_T$ and $\hat{\alpha}^{\text{BML}}$ contain almost the same indices.

The estimation quality of both MOS-based procedures is assessed in Figures 3 and 4. In Figure 3, we consider the following figures of merit as functions of the SINR: the root mean square (RMS) number of missed targets evaluated by verifying that the estimated vector $\hat{\alpha} = [\hat{\alpha}_1 \cdots \hat{\alpha}_K]^T$ has zero elements in the actual target positions; the RMS number of ghosts (false positive targets) defined as the non-zero components of $\hat{\alpha}$ in positions different from that of the targets; the Hausdorff metric [24] between $\alpha = [\alpha_1 \cdots \alpha_K]^T$ and $\hat{\alpha}$. In Figure 4, we show the histograms that represent the probability of selecting a specific number of targets over 1000 MC trials (assuming two different SINR values). Inspection of these figures highlights that both proposed estimation procedures share the same performance since the values of the considered metrics are overlapped. Moreover, AIC- and BIC-based procedures experience a floor for the number ghosts and Hausdorff metric leading to poor estimation performance as corroborated by the histograms. On the contrary, GIC-based procedures are capable of returning the correct number of targets with high probability ($>0.75$). For the above reasons, the recommended architectures for the detection problem at hand are those relying on GIC.

**V. CONCLUSION**

The problem of parameter estimation and adaptive detection for multiple point-like targets has been addressed using two different approaches. In the first case, we jointly resorted to ML estimation and Bayesian learning. The second approach is based on energy considerations. Adaptivity with respect to the ICM and the number of targets is achieved through the exploitation of a training data set and of the MOS rules, respectively. The illustrative examples have shown the superiority of the GIC-based approaches with respect to the other MOS-based solutions and to the competitors.
REFERENCES


