Sequential along-track integration for early detection of moving targets

Emanuele Grossi and Marco Lops, Senior Member, IEEE

Abstract

This paper concerns the joint multi-frame sequential target detection and track estimation in early-warning radar surveillance systems. The rationale for applying sequential procedures in such a scenario is that they promise a sensitivity increase of the sensor or, alternatively, a reduction in the time needed to take a decision. Unlike previous works on sequential radar detection, the attention is not restricted to stationary targets, namely position changes during the illumination period are allowed. Starting from previous sequential rules, different truncated sequential strategies are proposed and assessed: they are aimed at orienting the sensor resources towards either the detection or the track estimation or the position estimation. Bounds on the performances of the proposed procedures in terms of the system parameters are derived and computational complexity is examined. Also, numerical experiments are provided to elicit the interplay between sensor-target parameters and system performances and to quantify the gain with respect to other fixed-sample-size procedures.

Index Terms

Sequential detection and estimation, hidden Markov models, early warning, surveillance radar.

I. INTRODUCTION

Target detection and tracking is a statistical decision problem where estimation of the target trajectory has to be performed under uncertainty as to signal presence. Application of sequential decision rules to this scenario arouses much interest since they promise a considerable gain in sensitivity, measured by the reduction in the average sample number (ASN), with respect to fixed-sample-size (FSS) procedures. These advantages are particularly attractive in remote radar surveillance, where the signal amplitude is

The authors are with the DAEIMI, Università degli Studi di Cassino, via G. Di Biasio 43, Cassino (FR), 03043 Italy. Phone: +39 0776 2993742, Fax: +39 0776 2993987, E-mail: {e.grossi, lops}@unicas.it.
weak compared to the background noise and, in general, in the detection of dim fluctuating targets: in such a situation, stringent detection specifications can be met only by processing multiple frames so as to integrate the backscattered energy of the target along its (unknown) trajectory, with no intermediate thresholding taking place [1]–[3]. In this case, FSS techniques are usually inefficient, while sequential procedures are known to increase the sensitivity of power-limited systems (such as, for example, airborne radars) or, alternatively, to reduce the ASN [4], [5].

The extension of sequential procedures to the radar problem has been addressed in [5], where, with reference to a single target scenario, a sequential probability ratio test (SPRT) is applied to discriminate between a simple hypothesis and a composite alternative. A generalized SPRT is instead proposed in [6] to solve a joint target detection and position classification problem, while in [7] and [8] the authors considered as many SPRT as the resolution cells, each SPRT being allowed to independently complete the test. All of the previous approaches, however, assume a stationary target persisting in the same resolution cell while being illuminated by the radar: this condition may be too restrictive, especially in airborne applications where the relative radial velocity between target and radar may exceed Mach-2.

The paper provides a generalization in this sense. A joint multi-frame sequential detection and trajectory estimation is performed in long-range, early-warning surveillance radar systems. Previous limitations on target mobility are removed in order to account for the case of fast moving and/or manoeuvring targets. In particular the target has been assumed to follow a Markov evolution so that, if the signal is present, the pair signal–observation forms a hidden Markov process. As concerns the joint detection and tracking aspects, rather than adopting a ‘true’ track-before-detect technique of the kind of [1]–[3], the basic approach taken here is to consider a number of potential target paths in the measurement space and to integrate signal intensity along these paths during the time-on-target (TOT): this is an “along-track integration” rather than a conventional tracking. The goal of this procedure, thus, is to obtain reliable detections and/or reliable track segments from the raw data: the track initiation and maintenance processes are handled by a higher-level, conventional tracking algorithm which has the task of combining the track segments or the positions into long continuous tracks [9], [10].

Different sequential frame acquisition plans are proposed. In particular, the sequential rules of [11]–[15], stated for a general multi-hypotheses testing problem with finite number of hypotheses, are adapted to the present radar framework, where the target mobility allows for an unbounded number of possible tracks (i.e. hypotheses). Starting from these sequential procedures, three different strategies are proposed, in order to improve the system performance either in the detection or in the track segment estimation or in the final position estimation. The possibility of occasionally long tests is avoided by imposing a cut-off
stage at which the procedure is truncated. In this way, a control over the maximum dwell time on an angular sector is maintained thus avoiding the possibility that targets may traverse undetected through the surveillance region. Furthermore, a complexity analysis is given, showing that the statistics needed in the different proposed strategies are all amenable to a dynamic programming (DP) implementation, which makes the complexity linear in the number of integrated frames. Also, bounds on the system performances in terms of the thresholds of the proposed sequential rules are given, along with useful indications on threshold settings. Finally, a thorough performance analysis is given, aimed at both comparing the different strategies and investigating the effects of system parameters, such as target mobility and/or signal-to-noise ratio (SNR). The superiority of sequential detection and estimation rules with respect to FSS techniques is also shown.

The rest of the paper is organized as follows. Next section is devoted to the derivation of the sensor measurement and target models. In Section III, the joint target detection and tracking problem is formalized in terms of statistical decision theory and the simple FSS case is first taken under investigation. Section IV presents the proposed sequential rules for three different strategies. Bounds on the probabilities of error and useful hints on threshold setting are given in Section V while complexity issues are addressed in Section VI. Finally numerical results are presented in Section VII and concluding remarks are given in Section VIII.

II. SENSOR MEASUREMENT MODEL AND TARGET MODEL

The physical situation considered here is shown in Figure 1. A low pulse repetition frequency (PRF) radar is used for monitoring a given portion of the sky with an electronically scanned, pencil beam antenna. The surveillance area is divided into smaller angular regions, visited by the antenna beam in cyclic manner. An angular region is composed of $N_a$ azimuthal beam-pointing positions and a single elevation, each pointing position being revisited on a recurrent basis. The pulse repetition time is $T_p$ and the pulses are processed in groups of $N_d$: this implies that the time spent to scan through an angular region is $T_R = N_a N_d T_p$. The complex envelope of the return received from the $m$-th azimuthal sector, $m \in \{1, \ldots, N_a\}$, at the $\ell$-th scan, $\ell \in \mathbb{N}$, is written as

$$r_{m,\ell}(t) = A_{m,\ell} e^{j \phi_{m,\ell}} \sqrt{2 N_d \bar{E}_p} \psi_{m,\ell}(t, \tau_\ell, f_\ell) + w_{m,\ell}(t),$$

if a target is present and $\phi_\ell \in [(m - 1)\Phi, m\Phi)$,

$$r_{m,\ell}(t) = w_{m,\ell}(t),$$
Fig. 1. Search zone: example with 10 angular regions, each composed of 4 azimuthal beam-pointing positions and 1 elevation.

otherwise, where \( \psi_{m,\ell}(t, \tau_\ell, f_\ell) = \frac{1}{\sqrt{2N_d E_p}} \sum_{n=0}^{N_a-1} p(t - nT_p - \tau_\ell - (m - 1)N_d T_p - (\ell - 1)T_R) e^{2\pi i f_\ell t} \)

and

- \( \{A_{m,\ell}e^{j\theta_{m,\ell}}\}_{\ell \in \mathbb{N}, m \in \{1, \ldots, N_a\}} \) is the target response; it is supposed that both \( A_{m,\ell} \) and \( \theta_{m,\ell} \) are i.i.d. stochastic process, independent of each other, the former Rayleigh distributed with unitary mean square value, the latter uniformly distributed on \( [0, 2\pi) \);\(^1\)
- \( \phi_\ell \in [0, N_a \Phi) \) is the target azimuth at the \( \ell \)-th scan, \( \Phi \) being the antenna beamwidth;
- \( \tau_\ell \) and \( f_\ell \) are the target delay and Doppler frequency at the \( \ell \)-th scan, respectively;
- \( p(t) \) is the baseband non-sophisticated pulse waveform with energy \( 2E_p \), \( \tau_c \) being the (one-sided) duration of its autocorrelation function;
- \( w_{m,\ell}(t) \) is the complex white Gaussian thermal noise; it is assumed to be a zero mean proper process, with power spectral density \( 2N_0 \), i.i.d. for all \( m,\ell \).\(^2\)

In this way, the possibility of having more targets in the surveillance area can be easily managed, similarly to [5], [7], [16].

\(^1\)This is the case where the target has a Swerling-I fluctuation model and frequency agility is used to achieve scan-to-scan target amplitude decorrelation.

\(^2\)Notice that, since it was assumed that the sensor is pointing towards the sky, no other interference source is accounted for.
The process of signal discretization is a standard procedure in radar problems and is related to the fact that the target parameters \( \{ \phi_\ell, \tau_\ell, f_\ell \}_{\ell \in \mathbb{N}} \), which are inherently continuous, can be estimated up to an uncertainty dictated by the beamwidth of the antenna and by the ambiguity function of the transmitted signal [17]. That is, the region that must be considered is divided into a grid, whose resolution is given by the azimuthal sector size \( \Phi \), the duration of the pulse autocorrelation function \( \tau_c \) and the pulse train repetition frequency \( 1/(N_d T_p) \). The continuous-time received signal is thus discretized through projection onto the orthonormal basis

\[
\left\{ \psi_{m,\ell}(t, (n+I)\tau_c, \frac{\nu-1}{N_d T_p}) : m \in \{1, \ldots, N_a\}, n \in \{1, \ldots, N_r\}, \nu \in \{1, \ldots, N_d\} \right\},
\]

(1)

where \( \{I + 1, \ldots, I + N_r\} \) are the range bins under inspection.\(^3\) At the design stage it is supposed that, for every \( \ell \in \mathbb{N} \), the target parameters \( \tau_\ell \) and \( f_\ell \) are integer multiples of \( \tau_c \) and \( \nu/(NT_p) \), respectively: if the grid is sufficiently fine, losses due to possible mismatches (scalloping losses) may be neglected. As concerns \( \phi_\ell \), while it is desirable to integrate as much signal energy as possible, the choice to neglect the signal return from adjacent azimuthal sectors has been made in order not to increase the noise contribution. Under these hypotheses, the projection of \( r_{m,\ell}(t) \) onto the element \( (m, n, \nu) \) of the basis in (1) is

\[
\int_{\mathbb{R}} r_{m,\ell}(t)\psi_{m,\ell}^*(t, (n+I)\tau_c, \frac{\nu-1}{N_d T_p}) dt = \begin{cases} A_{m,\ell}e^{i \theta_{m,\ell}} \sqrt{2N_d E_p} + w'_{m,\ell}, & \text{if a target is present in } (m, n, \nu), \\ w'_{m,\ell}, & \text{otherwise,} \end{cases}
\]

(2)

where \( w'_{m,\ell} = \int_{\mathbb{R}} w_{m,\ell}(t)\psi_{m,\ell}^*(t, (n+I)\tau_c, \frac{\nu-1}{N_d T_p}) dt \), \((\cdot)^*\) denoting conjugate.

Finally the measurement at stage \( \ell \in \mathbb{N} \), a frame from now on, is the set of returns received from all of the radar resolution elements, i.e.

\[
Z_\ell = \left\{ Z_\ell(m, n, \nu) = \left| \frac{1}{2N_0} \int_{\mathbb{R}} r_{m,\ell}(t)\psi_{m,\ell}^*(t, (n+I)\tau_c, \frac{\nu-1}{N_d T_p}) dt \right|^2 : m \in \{1, \ldots, N_a\}, \\ n \in \{1, \ldots, N_r\}, \nu \in \{1, \ldots, N_d\} \right\}
\]

(3)

and the set of measurement available up to epoch \( k \) is \( Z_k = \{Z_1, \ldots, Z_k\}, k \in \mathbb{N} \).

Observation 2.1: As usual, the discrete-time measurement process can be obtained sampling at rate \( 1/(N_d T_p) \), for each hypothesis of the target delay, the output of \( N_d \) filters matched to the transmitted pulse train \( \frac{1}{\sqrt{2N_0 E_p}} \sum_{n=0}^{N_d-1} p(t - nT_p) \) for each possible Doppler shift \( \nu \).

\(^3\) Obviously, \( I \) and \( N_r \) are integers such that \( I + N_r \leq T_p/\tau_c \) in order to account for non-ambiguous ranges only.
are given by trajectory. A first-order Gaussian-Markov random walk is used to model the transition probabilities, which

\[ x \in S \text{ space consists of the set of all the resolution cells, i.e. } S = \{1, \ldots, N_a\} \times \{1, \ldots, N_r\} \times \{1, \ldots, N_d\} = S_a \times S_r \times S_d, \text{ with } M = \text{card}(S) = N_aN_rN_d. \]

The target state vector at epoch \( \ell \) is denoted \( X_\ell \), \( x_\ell = (m_\ell, n_\ell, \nu_\ell) \) being a realization of \( X_\ell \). A sequence of states \( \{X_\ell\}_{\ell=1}^k \) is denoted \( X_k \) and called trajectory. A first-order Gaussian-Markov random walk is used to model the transition probabilities, which are given by \( a(x_\ell, x_{\ell+1}) = a_a(m_\ell, m_{\ell+1}, n_\ell) a_r(n_\ell, n_{\ell+1}) a_d(\nu_\ell, \nu_{\ell+1}), \) where

\[
\begin{align*}
a_a(m_\ell, m_{\ell+1}, n_\ell) &= Q\left(\frac{m_\ell-m_{\ell+1}-1/2}{\sigma_a(n_\ell)}\right) - Q\left(\frac{m_\ell-m_{\ell+1}+1/2}{\sigma_a(n_\ell)}\right), \\
a_r(n_\ell, n_{\ell+1}) &= Q\left(\frac{n_\ell-n_{\ell+1}-1/2}{\sigma_r}\right) - Q\left(\frac{n_\ell-n_{\ell+1}+1/2}{\sigma_r}\right), \\
a_d(\nu_\ell, \nu_{\ell+1}) &= 1/N_d.
\end{align*}
\]

In the above equation, \( a_a, a_r \) and \( a_d \) represent the azimuth, range and Doppler transition probabilities, respectively; \( Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-t^2/2} dt \); \( \sigma_a(\cdot) \) and \( \sigma_r \) are parameters related to the target mobility along the azimuth and range dimensions: large values of these parameters allow large target manoeuvres but decrease, at the same time, target detection and estimation capabilities. Notice that the azimuthal transition probabilities depend on the starting range position in that farther targets necessarily exhibit lower azimuthal mobility than closer targets. Notice also that Doppler transitions have been assumed equally likely: indeed a low PRF (LPRF) radar is being considered and, thus, Doppler measurements are highly ambiguous.\(^4\) As concerns the initial probability, \( \pi \) say, if no other prior information is available (for example previous detections), it is reasonable to force \( \pi(x) = 1/M \), for all \( x \in S \).

Remark 2.3: A more complex state space can be considered (for example one involving also velocities).

In fact, this is not a limitation since all of the following derivations can be easily extended to the case of a larger state space. However, it should be pointed out that a state space enlargement would imply an increase of the computational complexity which would be deleterious since simple algorithms are

\(^4\)If a staggered LPRF radar is used, then a transition probability similar to those for azimuth and range could be adopted for the Doppler dimension as well.
required in multi-frame radars. Indeed, the task of such a kind of sensors is not to perform a long time
tracking but rather to obtain reliable detections and/or reliable track segments which are to be forwarded
to a higher level, conventional tracking algorithm in order to combine them into long continuous tracks
[9], [10].

Remark 2.4: Even if a Brownian motion model is considered for target kinematics and, thus, all
transitions are theoretically admissible, real targets necessarily need to satisfy physical constraints, such
as limitations on the maximum velocity and acceleration. Constraints actually reduce the computational
burden and increase the detection and tracking capabilities by decreasing the admissible transitions (i.e.
some of the elements of the transition probabilities are set to zero and a truncated Gaussian distribution
is adopted). Typical constraints on the target motion are those related to maximum velocity in both
range and azimuth. Indeed, given the maximum azimuthal velocity $v_a$ and the maximum radial velocity
for approaching and moving away targets $v_r^+ \quad \text{and} \quad v_r^-$, respectively,$^5$ the maximum admissible azimuth
transition corresponding to the $n$-th range bin, $\Delta_a(n)$ say, and the maximum admissible range transition
in a scan, $\Delta_r^\pm$ say, are given by $\Delta_a(n) = \left[ \frac{v_a T_R}{c \sigma_a(n)} \right]$ and $\Delta_r^\pm = \left[ \frac{v_r^\pm T_R}{c \sigma_r(n)} \right]$, respectively, where $c$ denotes
the speed of light. In this case, for example, the azimuthal transition probabilities would be

$$a_a(m_\ell,m_{\ell+1},n_\ell) = \frac{Q \left( \frac{m_\ell-m_{\ell+1}-1/2}{\sigma_a(n_\ell)} \right) - Q \left( \frac{m_\ell-m_{\ell+1}+1/2}{\sigma_a(n_\ell)} \right)}{Q \left( - \frac{\Delta_a(n_\ell)}{2\sigma_a(n_\ell)} \right) - Q \left( \frac{\Delta_a(n_\ell)}{2\sigma_a(n_\ell)} \right)} 1 \{ m_\ell-m_{\ell+1}<\Delta_a(n_\ell) \},$$

where $1\{\cdot\}$ denotes the indicator function. The range transitions are defined similarly.

Remark 2.5: The sensor measurement model introduced above is pretty general and it is also suited for
subtasks of a radar with agile beam (such as an electronically scanned antenna). This kind of sensors is
required to handle a variety of tasks, such as tracking a set of targets and searching a sector for new targets.
In this case, sequential procedures can be used for both searching and tracking initiation/maintenance
purposes. In the latter case, the sensor is always provided with some prior information on the pointing
angles of the target (they can come from the raw data if a detection has occurred and a track confirmation
process needs to be performed or they can come from the overall tracking algorithm if a track update
is required). If the prior information on angles is sufficient to allow the radar to dwell at a single beam
position, then a single azimuth can be considered in the multi-frame procedure (i.e. $N_a = 1$). On the
other hand, if such angular information is only good enough to limit the amount of angular space to be
searched, then $N_a > 1$.

$^5$These two velocities are, in general, different since the radar may be placed on a moving platform: this is the case of airborne
radars, for example.
III. JOINT DETECTION AND ESTIMATION

Given these elements, one is to sample the process \(\{Z_\ell\}_{\ell \in \mathbb{N}}\) sequentially and decide, after each observation, whether to stop sampling and take an action or to continue and take an action sometimes later. The action to take is to decide, as soon as possible, if measurements are generated by noise alone or if they come from a target and, in the latter case, it is also required to estimate the target trajectory which has generated such measurements and/or its final position. The sequential nature of the decision process allows now a trade-off between quickness of decision and decision accuracy.

It is not difficult to see that this is a problem of decision theory. Indeed, the parameter space is \(\Theta = \Theta_0 \cup \Theta_1\), \(\Theta_0 = \{\theta_0\}\) characterizing the case that observations are generated by noise alone and \(\Theta_1 = \times_{i \in \mathbb{N}} S\) being the set of possible trajectories and corresponding to the case that a target is present. A prior on \(\Theta_1\) remains defined since \(\{X_\ell\}_{\ell \in \mathbb{N}}\) is a discrete-time, homogeneous Markov chain with state space \(S\) and distribution given by

\[
p_k(x_k) = \pi(x_1) \prod_{\ell=2}^k a(x_{\ell-1}, x_\ell), \quad \forall k \in \mathbb{N}. \tag{4}
\]

The discrete-time process available for inference is \(\{Z_\ell\}_{\ell \in \mathbb{N}}\). Based on observation 2.2, each component \(Z_\ell(x)\) of the measurement \(Z_\ell\), for \(x \in S\), is an exponentially distributed random variable with density

\[
h_1(\alpha) = \frac{e^{-\frac{\alpha}{1+\rho}}}{1+\rho} u(\alpha), \quad \text{if the target is present in location } x, \tag{5a}
\]

\[
h_0(\alpha) = e^{-\alpha} u(\alpha), \quad \text{otherwise,} \tag{5b}
\]

where \(\rho = N_d E_p / N_0\) denotes the SNR per frame and \(u\) is the unit step function. In this case, for each \(\ell \in \mathbb{N}\), \(Z_\ell\) has density

\[
f(z|x_\ell) = h_1(z(x_\ell)) \prod_{x \in S; x \neq x_\ell} h_0(z(x)), \quad \text{if the target is present in location } x_\ell,
\]

\[
f(z_\ell|\theta_0) = \prod_{x \in S} h_0(z_\ell(x)), \quad \text{if no target is present.}
\]

Thus, the joint distribution of \(Z_k\), \(k \in \mathbb{N}\), has density

\[
f_k(z_k|x_k) = \prod_{\ell=1}^k f(z_\ell|x_\ell), \quad \text{if the target is present with trajectory } x_k,
\]

\[
f_k(z_k|\theta_0) = \prod_{\ell=1}^k f(z_\ell|\theta_0), \quad \text{if no target is present,}
\]
and the likelihood ratio of \( f_k(z_k|x_k) \) to \( f_k(z_k|\theta_0) \) is

\[
\Lambda_k(z_k|x_k) = \prod_{\ell=1}^{k} \frac{\ell_1(z_\ell(x_\ell))}{\ell_0(z_\ell(x_\ell))} = \prod_{\ell=1}^{k} \frac{e^{z_\ell(x_\ell)\rho/(1+\rho)}}{1 + \rho}, \quad \forall \ k \in \mathbb{N}.
\] (6)

This is a statistical decision problem where estimation of \( \theta \in \Theta_1 \) has to be performed under uncertainty as to the signal presence (\( \theta \in \Theta_0 \) or \( \theta \in \Theta_1 \)). As in [18], [19], whose focus, however, was on non-sequential decision rules, there is a mutual coupling of detection and estimation and different strategies may be adopted. The problem of devising such a decision rule, however, is considerably more difficult than that of finding a FSS procedure. For this reason, it is worth first analyzing two different situations in the case that a fixed sample of \( k \) observations is available for inference.

Assume for the moment that a prior on the target presence is given, \( \alpha/(1 + \alpha) \) say, with \( \alpha > 0 \), and consider a Bayesian framework with the canonical 0-1 loss. Then, it is easy to see that the Bayes decision rule for testing the simple hypothesis \( H_0: \text{‘noise alone’} \) against the alternative \( H_1: \text{‘target present, no matter of its trajectory’} \) is

\[
\sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \begin{cases} H_1 \geq c \quad \text{if } H_1 \\ H_0 \end{cases}
\]

It is not diffucult to show [20] that, for a suitably chosen \( \alpha \), the above test solves the problem

\[
\text{maximize} \quad P_d = \sum_{x_k \in S^k} p_k(x_k) P_d(x_k),
\]

subject to \( P_{fa} \leq \alpha \),

where \( P_d \) is the average probability of detection, \( P_d(x_k) \) is the probability of detection given that the target has the trajectory \( x_k \) and \( P_{fa} \) denotes the probability of false alarm. I this case, \( \alpha \) is chosen so that \( P_{fa} = \alpha \) and, thus, it is only a constant which controls the false alarm.

Using the same setting, it can be easily shown also that the Bayes decision rule for the \( M^k + 1 \) classification problem among ‘noise alone’, ‘target present with trajectory \( x_k \)’ for all \( x_k \in S^k \) is

\[
\hat{x}_k(z_k) = \arg \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \begin{cases} H_1 \geq c \quad \text{if } H_1 \\ \theta_0 \end{cases}
\]

As before, the above procedure solves the problem of maximizing the average probability of correct trajectory classification, \( P_{\text{track}} \) say, subject to a false alarm constraint [20] and, again, \( \alpha \) is a constant that regulates the false alarm.
IV. SEQUENTIAL RULES

The procedures described in the previous section are now to be extended to a sequential scenario. As for the detection oriented strategy, Wald's SPRT can be used [11]. However, since the log-likelihood ratio has no longer the structure of a random walk, all of the optimality property as to the ASN is lost. As concerns the classification oriented strategy, the problem of devising a sequential decision rule for a classification problem is considerably more difficult than that of finding a non-sequential decision rule. A recursive solution has been obtained in [21] but it is very complex except in a few special cases. Furthermore, any sequential procedure, if considered unbounded, turns out to be impractical, since occasionally long observations may be needed so that the radar beam can become ‘hung-up’ along a particular direction and the surveillance along other directions is neglected. Thus a cut-off stage must be specified and its effect often cannot be neglected while looking for an optimal sequential rule. For the same reason, discussions as to the asymptotical optimality have no great meaning.

The approach here taken, then, is to extend and generalize the SPRT to the present setting designing practical, possibly suboptimal, sequential decision rules. A sequential decision rule is the pair \((\varphi, \delta)\), where \(\varphi = \{\varphi_k\}_{k \in \mathbb{N}}\) is a stopping rule and \(\delta = \{\delta_k\}_{k \in \mathbb{N}}\) a terminal decision rule [22]. Since a joint detection and estimation problem is being analyzed, the terminal decision rule is itself composed of a detection rule \(d = \{d_k\}_{k \in \mathbb{N}}\), a trajectory estimator \(\hat{x} = \{\hat{x}_k\}_{k \in \mathbb{N}}\) and a final position estimator \(\hat{w} = \{\hat{w}_k\}_{k \in \mathbb{N}}\), i.e. \(\delta = (d, \hat{x}, \hat{w})\). The stopping rule \(\varphi\) determines the stopping time \(\tau\), whose conditional distribution is \(\psi = \{\psi_k\}_{k \in \mathbb{N}}\). With these definitions the probabilities of false alarm and average miss are expressed by

\[
P_{fa} = \sum_{k \in \mathbb{N}} E_{\theta_0} \left[ \psi_k(Z_k) d_k(Z_k) \right], \tag{7a}\]

\[
P_{miss} = \sum_{k \in \mathbb{N}} \sum_{x_k \in S_k} p_k(x_k) E_{x_k} \left[ \psi_k(Z_k) (1 - d_k(Z_k)) \right], \tag{7b}\]

while the probabilities of average detection, average correct track classification and average correct final position classification by

\[
P_d = \sum_{k \in \mathbb{N}} \sum_{x_k \in S_k} p_k(x_k) E_{x_k} \left[ \psi_k(Z_k) d_k(Z_k) \right], \tag{8a}\]

\(\psi_k(z_k)\) is the probability that \(\tau = k\) given a realization \(z_k\) of \(Z_k\), for any \(k \in \mathbb{N}\); the relationship between \(\psi\) and \(\varphi\) is:

\[
\psi_1(z_1) = \varphi_1(z_1) \text{ and } \psi_k(z_k) = \varphi(z_k) \prod_{i=1}^{k-1} \left(1 - \varphi(z_i)\right), \text{ for } k > 1.
\]

Notice that, if the stopping rule is non-randomized, \(\psi_k(z_k) \in \{0, 1\}, \forall z_k \in \mathbb{R}^{kM} \text{ and } k \in \mathbb{N}\).
\[
P_{\text{track}} = \sum_{k \in \mathbb{N}} \sum_{x_k \in S^k} p_k(x_k) \mathbb{E}_{x_k} \left[ \psi_k(Z_k) d_k(Z_k) 1\{x_k = x_k\} \right], \quad (8b)
\]
\[
P_{\text{pos}} = \sum_{k \in \mathbb{N}} \sum_{x_k \in S^k} p_k(x_k) \mathbb{E}_{x_k} \left[ \psi_k(Z_k) d_k(Z_k) 1\{x_k = x_k\} \right], \quad (8c)
\]
respectively. Here and in the following, \( \mathbb{E}_{\theta_0} \) and \( \mathbb{E}_{x_k} \) denote the operators of expectation when the observables \( Z_k \) have density \( f_k(\cdot|\theta_0) \) and \( f_k(\cdot|x_k) \), respectively. On the other hand, the ASN under the noise alone hypothesis, \( \text{ASN}_{H_0} \), and under the signal plus noise hypothesis, \( \text{ASN}_{H_1} \), are expressed by
\[
\text{ASN}_{H_0} = \sum_{k \in \mathbb{N}} k \mathbb{E}_{\theta_0} \left[ \psi_k(Z_k) \right],
\]
\[
\text{ASN}_{H_1} = \sum_{k \in \mathbb{N}} \sum_{x_k \in S^k} p_k(x_k) \mathbb{E}_{x_k} \left[ \psi_k(Z_k) \right].
\]

As in the FSS case, the figures of merit to consider when designing a sequential procedure should be those defined in equations (7) and (8). In the following sections, then, the stopping rule \( \varphi \) (or, equivalently, the stopping time \( \tau \)) will be designed with the intent of maximizing \( P_0 \), \( P_{\text{track}} \) or \( P_{\text{pos}} \): the corresponding strategies will be named detection oriented (DO), track classification oriented (TCO) and position classification oriented (PCO), respectively.

A. DO strategy

The detection part of the procedure realizes an SPRT while, in order to estimate target trajectory and/or its final position, a gated estimator is used, in the sense that estimation is enabled by the result of the detection operation. The adoption of an SPRT for target detection has been previously used in [5] but the target was supposed to be stationary during the TOT. Since \( H_1 \) corresponds to the case that the observations \( Z_k \) have density \( \sum_{x_k \in S^k} p_k(x_k) f_k(z_k|x_k) \), the sequential decision rule is
\[
\varphi_k(z_k) = \begin{cases} 
1, & \text{if} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \notin (\gamma_0, \gamma_1), \\
0, & \text{otherwise}, 
\end{cases} \quad (9a)
\]
\[
d_k(z_k) = \begin{cases} 
1, & \text{if} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \gamma_1, \\
0, & \text{if} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \leq \gamma_0, 
\end{cases} \quad (9b)
\]
\[
\hat{x}_k(z_k) = \arg \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k), \quad \text{if} \ d_k(z_k) = 1, \quad (9c)
\]
\[
\hat{w}_k(z_k) = \arg \max_{w \in S} \sum_{x_k \in S^k \mid x_k = w} p_k(x_k) \Lambda_k(z_k|x_k), \quad \text{if} \ d_k(z_k) = 1, \quad (9d)
\]
where $\Lambda_k(z_k|x_k)$ is the likelihood ratio of $f_k(z_k|x_k)$ to $f_k(z_k|\theta_0)$. The boundaries of the test, $\gamma_0$ and $\gamma_1$ are chosen to satisfy $0 < \gamma_0 < 1 < \gamma_1 < +\infty$.

B. TCO strategy

The sequential procedures introduced in [12]–[15] have to be restated to cover the case of non stationary targets. As in Section III, let $c \in (0, +\infty)$, and consider the prior $c/(1 + c)$ for the target presence.\footnote{Again it will be seen that, rather then a prior, $c$ is just a threshold to set in order to have the required specifications.} In this case the equivalent procedure of [12] would be:

“stop sampling whenever

$$\frac{c}{1 + c} f_k(z_k|\theta_0) \geq \frac{1}{1 + A_0},$$

in which case declare noise alone, or

$$\frac{1}{1 + c} p_k(y^k) f_k(z_k|y^k) \geq \frac{1}{1 + A_1},$$

for some $y^k \in S^k$, in which case declare target present with trajectory $y^k$."

$A_0$, $A_1$ should be constants taken from the interval $(0, 1)$: with this choice, indeed, the condition $\geq 1/(1 + A_i)$, $i = 0, 1$, can be satisfied by at most one trajectory or by the noise alone hypothesis, since the posterior probabilities must sum to one. However, in the present dynamic environment, two different trajectories may share one or more positions. In such a situation, the sensor in general rejects the noise-alone hypothesis, (i.e. the threshold $1/(1 + A_0)$ is not crossed) but, at the same time, does not reach the required accuracy on the trajectory estimation (i.e. also the threshold $1/(1 + A_1)$ is not crossed): this implies that the antenna may remain blocked monitoring some region without taking any decision. To cope with such a situation, $A_1$ is extended to take on values in $(0, +\infty)$ and the terminal decision rule is modified to take as estimate the trajectory with the largest posterior, even if none of them exceeds $1/2$: this modification corresponds to requiring a lower accuracy in discriminating between different tracks than that specified in the detection process. After some manipulations, the sequential rule,
in the sequel referred to as TCO1, can be recast as follows

\[
\varphi_k(z_k) = \begin{cases} 
1, & \text{if } (1 + A_1) \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) + c \\
\text{or } \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \leq cA_0, \\
0, & \text{otherwise},
\end{cases}
\]

(10a)

\[
d_k(z_k) = \begin{cases} 
1, & \text{if } (1 + A_1) \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) + c, \\
0, & \text{if } \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \leq cA_0,
\end{cases}
\]

(10b)

\[
\hat{x}_k(z_k) = \arg \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k), \text{ if } d_k(z_k) = 1,
\]

(10c)

\[
\hat{w}_k(z_k) = \arg \max_{x_k \in S^k} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k), \text{ if } d_k(z_k) = 1.
\]

(10d)

As concerns the sequential rule of [13] (a modification of the combination of one-sided SPRTs earlier introduced in [23]–[25]), it can be adapted to the present framework as follows

“stop sampling whenever

\[
\frac{c}{1 + c f_k(z_k | \theta_0)} \max_{x_k \in S^k} p_k(x^k) f_k(z_k | x_k) \geq \frac{1}{B_0},
\]

in which case declare noise alone, or

\[
\frac{1 + c p_k(y^k) f_k(z_k | y^k)}{\max \left\{ \frac{c}{1 + c f_k(z_k | \theta_0)}, \frac{1}{1 + c} \max_{x_k \in S^k} p_k(x_k) f_k(z_k | x_k) \right\}} \geq \frac{1}{B_1},
\]

for some \(y^k \in S^k\), in which case declare target present with trajectory \(y^k\).”

Notice that each of the above terms represents the generalized likelihood ratio between the hypothesis “noise alone” (or “target present with trajectory \(y^k\)”) and the remaining hypotheses. If \(B_0\) and \(B_1\) are taken from \((0, 1)\), the threshold can be crossed by at most one ratio. However, similarly to TCO1, \(B_0\) is extended to take values on \((0, +\infty)\). After some manipulations, the rule, which will be referred to as TCO2, takes the form

\[
\varphi_k(z_k) = \begin{cases} 
1, & \text{if } B_1 \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \max \left\{ \max_{x_k \in S^k} 2 p_k(x_k) \Lambda_k(z_k|x_k), c \right\} \\
\text{or } \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \leq cB_0, \\
0, & \text{otherwise},
\end{cases}
\]

(11a)
\[ d_k(z_k) = \begin{cases} 
1, & \text{if } B_1 \max_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \max \left\{ \max_{x_k \in S_k} 2 p_k(x_k) \Lambda_k(z_k|x_k), c \right\}, \\
0, & \text{if } \max_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \leq cB_0, 
\end{cases} \]  

(11b)

\[ \hat{x}_k(z_k) = \arg \max_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k), \text{ if } d_k(z_k) = 1, \]  

(11c)

\[ \hat{w}_k(z_k) = \arg \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k), \text{ if } d_k(z_k) = 1, \]  

(11d)

where \( \max 2 \) denotes the operator which extracts second largest element of a set. The alternative final position estimator, \( \hat{w}_k(z_k) = \hat{x}_k(z_k) \), can be possibly preferred to (11d) as it allows, for this strategy only, to have a lower computational complexity. Further details will be given in Section VI.

C. PCO strategy

This solution represents a sort of compromise between the two strategies introduced above. Adopting the same setting as in Section IV-B, the quantities involved in this case are the likelihoods \( f_k(z_k|\theta_0) \) and \( \sum_{x_k \in S_k} p_k(y^k) f_k(z_k|y^k), w \in S \), where the latter is the density of the observables given that a target is present with final position \( w \). The equivalent strategies of TCO1, then, is

\[ \varphi_k(z_k) = \begin{cases} 
1, & \text{if } (1 + A_1) \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) + c \\
or \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \leq cA_0, \\
0, & \text{otherwise,} 
\end{cases} \]  

(12a)

\[ d_k(z_k) = \begin{cases} 
1, & \text{if } (1 + A_1) \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) + c, \\
0, & \text{if } \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \leq cA_0, 
\end{cases} \]  

(12b)

\[ \hat{x}_k(z_k) = \arg \max_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k), \text{ if } d_k(z_k) = 1, \]  

(12c)

\[ \hat{w}_k(z_k) = \arg \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k), \text{ if } d_k(z_k) = 1, \]  

(12d)

and will be referred to as PCO1, while the equivalent for the TCO2 is

\[ \varphi_k(z_k) = \begin{cases} 
1, & \text{if } B_1 \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \max \left\{ \max_{w \in S} 2 \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k), c \right\}, \\
or \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k|x_k) \leq cB_0, \\
0, & \text{otherwise,} 
\end{cases} \]  

(13a)
\[ d_k(z_k) = \begin{cases} 
1, & \text{if } B_1 \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k | x_k) \geq \max \left\{ \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k | x_k) \right\}, \\
0, & \text{if } \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k | x_k) \leq cB_0, 
\end{cases} \] (13b)

\[ \hat{x}_k(z_k) = \arg \max_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k | x_k), \text{ if } d_k(z_k) = 1, \] (13c)

\[ \hat{w}_k(z_k) = \arg \max_{w \in S} \sum_{x_k \in S_k} p_k(x_k) \Lambda_k(z_k | x_k), \text{ if } d_k(z_k) = 1, \] (13d)

and will be referred to as PCO2.

D. Truncation strategies

A truncated sequential decision rule is requested, in the present scenario, for a number of reasons. From previous discussions, no kind of optimality (such as that of the minimization of the ASN under all hypotheses among sequential and non-sequential tests with a fixed strength [26]) can be granted in this framework. Moreover, due to possible mismatches between the design and actual values of some parameters (for example, the SNR) the resulting TOT can be very large, especially if small error probabilities are requested. In surveillance radars, for example, this can cause the beam antenna remaining ‘hung-up’ along a particular direction so that a target could ”snake through” some angular region in the surveillance zone without its echo being examined.

Truncation of sequential procedures can be used to prevent such a problem: this is a compromise between an entirely sequential procedure and a classical FSS procedure. Several truncation strategies have been adopted in the past for the canonical SPRT. Basically, they can be of two kinds: abrupt, i.e. no change in the procedure is introduced until the truncation stage itself, and gradual, i.e. the procedure is modified at every stage so that the boundaries monotonically converge [27]–[29]. In this study the former approach is chosen, i.e. a regular sequential procedure is carried out until either a decision is made or a fixed stage \( K \) is reached, in which case detection and estimation take place comparing the statistic with a truncation threshold. In particular, the detection rule (9b) for the DO strategy is modified at stage \( K \) as follows

\[ d_K(z_K) = \begin{cases} 
1, & \text{if } \sum_{x_K \in S_K} p_k(x_K) \Lambda_K(z_K | x_K) > \gamma_{DO}, \\
0, & \text{otherwise}, 
\end{cases} \]
while (10b) and (11b) lead, for the TCO strategy, to
\[
d_K(z_K) = \begin{cases} 
1, & \text{if } \max_{x_K \in S_K} p_K(x_K)A_K(z_K|x_K) > \gamma_{TCO}, \\
0, & \text{otherwise},
\end{cases}
\]
Finally, as the PCO strategy, (12b) and (13b) result in the detection rule
\[
d_K(z_K) = \begin{cases} 
1, & \text{if } \max_{w \in S_k} \sum_{x_K \in S_K} p_K(x_K)A_K(z_K|x_K) > \gamma_{PCO}, \\
0, & \text{otherwise}.
\end{cases}
\]

Final threshold setting is a very difficult task since it affects the error probabilities, unless the probability of truncation is negligible. Quantitative analysis and/or feedbacks from simulation results are in general needed. In this framework, on the other hand, practical, possibly sub-optimal, solutions are chosen. As for the DO strategy, \( \gamma_{DO} \) is simply taken to be the geometric mean between \( \gamma_0 \) and \( \gamma_1 \), i.e. \( \gamma_{DO} = \sqrt{\gamma_0 \gamma_1} \). As for TCO1, the detection rule can be equivalently rewritten as
\[
d_k(z_k) = \begin{cases} 
(1 + A_1) \max_{x_k \in S_k} p_k(x_k)A_k(z_k|x_k) + c \geq 1, \\
c(1 + A_0) \sum_{x_k \in S_k} p_k(x_k)A_k(z_k|x_k) + c \geq 1,
\end{cases}
\]
so that, if \( \tau = K \), target presence can be claimed upon comparing \((1 + A_1) \max_{x_k \in S_k} p_k(x_k)A_k(z_k|x_k)\) to \(c(1 + A_0)\); this immediately gives \( \gamma_{TCO} = c(1 + A_0)/(1 + A_1) \). Similarly for TCO2, considering the inequality
\[
\frac{cB_0}{\max_{x_k \in S_k} p_k(x_k)} \geq \frac{cB_0}{\max \left\{ \max_{x_k \in S_k} 2p_k(x_k)A_k(z_k|x_k), c \right\}},
\]
one gets \( \gamma_{TCO} = cB_0/B_1 \). The expressions of the final thresholds for the PCO strategies are identical.

V. LOWER BOUNDS ON THE PROBABILITIES OF ERROR AND THRESHOLDS SETTING

Consider non-truncated procedures first. Even if this situation is of limited interest in practical applications, the results that can be derived may be relevant: on one hand, the bounds that can be established may still be considered to hold as long as the probability of truncation is negligible, on the other, they offer hints on threshold setting. In this context we give

**Lemma 5.1:** The following bounds hold (recall that \( \text{card}(S) = M \)).

**DO:** \( P_{fa} \leq P_d/\gamma_1 \leq 1/\gamma_1 \), \( P_{miss} \leq \gamma_0(1 - P_{fa}) \leq \gamma_0 \);

**TCO1:** \( P_{fa} \leq P_dA_1/c \leq A_1/c \), \( P_{miss} \leq cA_0(1 - P_{fa}) \leq cA_0 \), \( P_{track} \geq P_d/(1 + A_1) \);
TCO2: $P_{fa} \leq P_d B_1/c \leq B_1/c$;

PCO1: $P_{fa} \leq P_d A_1/c \leq A_1/c, P_{miss} \leq cA_0(1 - P_{fa}) \leq cA_0, P_{pos} \geq P_d/(1 + A_1)$;

PCO2: $P_{fa} \leq P_d B_1/c \leq B_1/c, P_{miss} \leq cMB_0(1 - P_{fa}) \leq cB_0 M$.

Proof: For the DO case, these are the bounds of Wald [11], which are quite general and require neither the i.i.d. structure of the observations nor that the procedure terminates with probability one [30].

As concerns the TCO1 strategy, the bound on $P_{fa}$ can be demonstrated as follows.

$$P_{fa} = \sum_{k \in \mathbb{N}} E_{\theta_0} \left[ \psi_k(Z_k) d_k(Z_k) \right]$$

$$\leq \sum_{k \in \mathbb{N}} E_{\theta_0} \left[ \left( \frac{A_1}{c} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(Z_k|x_k) \right) \psi_k(Z_k) d_k(Z_k) \right]$$

$$= \frac{A_1}{c} \sum_{k \in \mathbb{N}} \sum_{x_k \in S^k} p_k(x_k) E_{x_k} \left[ \psi_k(Z_k) d_k(Z_k) \right]$$

$$= P_d A_1/c \leq A_1/c,$$

where in the first inequality it has been exploited the fact that, on the set $\{z_k : d_k(z_k) = 1\}$,

$$(1 + A_1) \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \sum_{y^k \in S^k} p_k(y^k) \Lambda_k(z_k|y^k) + c,$$

i.e.,

$$A_1 \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \sum_{y^k \in S^k, y^k \neq x_k} p_k(y^k) \Lambda_k(z_k|y^k) + c \geq c,$$

and $\frac{A_1}{c} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq 1$. As for $P_{miss}$, the following holds.

$$P_{miss} = \sum_{k \in \mathbb{N}} \sum_{x_k \in S^k} p_k(x_k) E_{x_k} \left[ \psi_k(Z_k) (1 - d_k(Z_k)) \right]$$

$$= \sum_{k \in \mathbb{N}} E_{\theta_0} \left[ \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(Z_k|x_k) \psi_k(Z_k) (1 - d_k(Z_k)) \right]$$

$$\leq \sum_{k \in \mathbb{N}} E_{\theta_0} \left[ cA_0 \psi_k(Z_k) (1 - d_k(Z_k)) \right]$$

$$= cA_0 \left( \sum_{k \in \mathbb{N}} E_{\theta_0} \left[ \psi_k(Z_k) \right] - \sum_{k \in \mathbb{N}} E_{\theta_0} \left[ \psi_k(Z_k) d_k(Z_k) \right] \right)$$

$$= cA_0 \left( P(\{\tau < +\infty\}|H_0) - P_{fa} \right)$$

$$\leq cA_0 (1 - P_{fa}).$$
where in the first inequality it has been exploited the fact that, on the set \( \{ z_k : d_k(z_k) = 1 \} \), \( \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \leq c A_0 \). As for \( P_{\text{track}} \), it results

\[
P_{\text{track}} = \sum_{k \in \mathbb{N}} \sum_{x_k \in S^k} p_k(x_k) E_{x_k} \left[ \psi_k(Z_k) d_k(Z_k) 1_{\{ \hat{X}_k(Z_k) = x_k \}} \right]
\]

\[
= \sum_{k \in \mathbb{N}} \sum_{x_k \in S^k} \sum_{y_k \in S^k} p_k(y^k) E_{y^k} \left[ \psi_k(Z_k) \left( \frac{p_k(x_k|Z_k|x_k)}{\sum_{w^k \in S^k} p_k(w^k|Z_k|w^k)} \psi_k(Z_k) d_k(Z_k) 1_{\{ \hat{X}_k(Z_k) = x_k \}} \right) \right]
\]

\[
\geq \frac{1}{1 + A_1} \sum_{k \in \mathbb{N}} \sum_{y_k \in S^k} p_k(y^k) E_{y^k} \left[ \psi_k(Z_k) d_k(Z_k) \right] = P_{fa}(1 + A_1).
\]

In the TCO2 case, on the other hand, the single bound on \( P_{fa} \) can be derived exactly as in TCO1, exploiting the fact that, on the set \( \{ z_k : d_k(z_k) = 1 \} \),

\[
B_1 \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq \max_{y^k \in S^k} \{ \max_{x^k \in S^k} p_k(x^k) \Lambda_k(z_k|y^k), c \} \geq c,
\]

i.e.,

\[
\frac{B_1}{c} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) \geq 1.
\]

The PCO1 strategy can be managed similarly to the TCO1 case.

Finally, for the PCO2 strategy, the bound on \( P_{fa} \) can be derived similarly to the TCO2 case while that on \( P_{\text{miss}} \) can be obtained starting from the TCO1 case (see also [6]), i.e.

\[
P_{\text{miss}} = \cdots
\]

\[
= \sum_{k \in \mathbb{N}} E_{\theta_k} \left[ \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(Z_k|x_k) \psi_k(Z_k)(1 - d_k(Z_k)) \right]
\]

\[
\leq \sum_{k \in \mathbb{N}} E_{\theta_k} \left[ M \max_{w \in S} \sum_{x_k \in S^k} p_k(x_k) \Lambda_k(Z_K|x_K) \psi_k(Z_k)(1 - d_k(Z_k)) \right]
\]

\[
\leq \sum_{k \in \mathbb{N}} E_{\theta_k} \left[ c B_0 M \psi_k(Z_k)(1 - d_k(Z_k)) \right]
\]

\[
\leq c B_0 M (1 - P_{fa}).
\]

It is emphasized that the above lemma holds under general assumptions, nor does it require independence or homogeneity of the observed data. Notice that these bounds allow selection of the thresholds in such a way that the tests belong to a determined class \( (\alpha, \beta) \), i.e. the error of first kind \( P_{fa} \) and that of the second kind \( P_{\text{miss}} \) do not exceed levels \( \alpha \) and \( \beta \), respectively.
Lemma 5.1 also offers insightful hints on the interplay between the system parameters and the performances. In particular it is apparent that the DO strategy require choosing \( \gamma_0 \) and \( \gamma_1 \) so as to match design values of \( P_d \) and \( P_{fa} \), which in turn determine the ASN, \( P_{\text{track}} \) and \( P_{\text{pos}} \). For examples, small \( \gamma_0 \) and large \( \gamma_1 \) result in better test performances at the price of an increase of the ASN. For TCO1, conversely, the system parameters are the constants \( A_0 \), \( A_1 \) and \( c \), to be set so as to comply with pre-assigned values of \( P_{\text{track}} \) and \( P_{fa} \), which in turn determine the ASN, \( P_d \) and \( P_{pos} \). For example, lowering the ratio \( A_1/c \) and the product \( cA_0 \) ensures smaller false alarm and miss probabilities, respectively; on the other hand, the “amplifying factor” \( A_1 \) is directly tied to the system capability of discriminating between distinct tracks. Similar considerations apply to PCO1, PCO2 and TCO2: for the latter no bound on \( P_{\text{miss}} \) and \( P_{\text{track}} \) is available, but its analogy with TCO1 and PCO suggests (and the results confirm) that the trends outlined for TCO1 holds for TCO2 as well.

VI. COMPLEXITY ISSUES

The sequential rules introduced in Section IV require to evaluate the statistics

\[
\sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k), \quad \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k), \quad \sum_{x_k \in S^k: x_k \neq w} p_k(x_k) \Lambda_k(z_k|x_k),
\]

(14)

for \( k = 1, \ldots, r \), where, by equations (4) and (6),

\[
p_k(x_k) \Lambda_k(z_k|x_k) = \pi(x_1) f(z_1|x_1) \prod_{\ell=2}^{k} a(x_{\ell-1}, x_\ell) f(z_\ell|x_\ell) / f(z_\ell|\theta_0),
\]

\( \forall z_k \in \mathbb{R}^M \). These statistics, then, have the form of a a stage-separated function on the algebraic system \((\mathbb{R}, +, \cdot)\) and \((\mathbb{R}, \max, \cdot)\), respectively and, thus, they can be computed through the following dynamic programming algorithms [31], which are known to lower the computational complexity from exponential to linear in the number of scans \( k \). The first one, similar to the forward-backward procedure [32], is

**Algorithm 6.1:** Let \( \{F^k\}_{k \in \mathbb{N}} \) be a sequence of real-valued functions on \( S \). Then the algorithm proceeds as follows.

1) Initialization: \( F^1(x) = \pi(x) f(z_1|x) / f(z_1|\theta_0), \quad \forall x \in S \).
2) Recursion: for every \( k \geq 2 \), \( F^k(x) = f(z_k|x) / f(z_k|\theta_0) \sum_{y \in S} a(y, x) F^{k-1}(y), \quad \forall x \in S \).

The second one, similar to the Viterbi algorithm [33], is

**Algorithm 6.2:** Let \( \{G^k\}_{k \in \mathbb{N}} \) be a sequence of real-valued functions on \( S \) and \( \{b_k\}_{k \geq 2} \) be a sequence of endomorphisms on \( S \). Then the algorithm proceeds as follows.

1) Initialization: \( G^1(x) = \pi(x) f(z_1|x) / f(z_1|\theta_0), \quad \forall x \in S \).

Copyright (c) 2008 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.
2) Recursion: for \( k \geq 2 \),

\[
G^k(x) = \frac{f(z_k|x)}{f(z_k|\theta_0)} \max_{y \in S} \{ a(y, x)G^{k-1}(y) \},
\]

\[
b^k(x) = \arg \max_{y \in S} \{ a(y, x)G^{k-1}(y) \}, \quad \forall \ x \in S.
\]

At this point, the variables \( \{F^k(x), G^k(x)\}_{x \in S} \) can be used to compute at each epoch the needed statistics in (14) as follows

\[
\sum_{x_k \in S^k, \ x_k = x} p_k(x_k)\Lambda_k(z_k|x_k) = F^k(x), \quad \forall \ x \in S,
\]

\[
\sum_{x_k \in S^k} p_k(x_k)\Lambda_k(z_k|x_k) = \sum_{x \in S} F^k(x),
\]

\[
\max_{x_k \in S^k} p_k(x_k)\Lambda_k(z_k|x_k) = \max_{x \in S} G^k(x).
\]

As concerns the estimation process, after the decision rule has stopped sampling, the track estimation proceeds backwards using functions \( \{G^k\}_{k \in \mathbb{N}} \) and \( \{b_k\}_{k \geq 2} \) of algorithm 6.2 while the final position estimation is based on functions \( \{F^k\}_{k \in \mathbb{N}} \) of algorithm 6.1. Suppose the procedure has stopped at stage \( k \), then

\[
\hat{x}_\ell(z_k) = \begin{cases} 
\arg \max_{x \in S} G^\ell(x), & \text{if } \ell = k \\
b_{\ell+1}(\hat{x}_{\ell+1}(z_k)), & \text{if } \ell = k - 1, k - 2, \ldots, 1,
\end{cases}
\]

\[
\hat{w}_k(z_k) = \arg \max_{x \in S} F^k(x).
\]

Notice that the running time of algorithms 6.1 and 6.2, and, thus, of all of the sequential procedures introduced, is \( \mathcal{O}(kM^2) \), i.e. it is linear in the number of scans.

**Observation 6.3:** The alternative final position estimator, \( \hat{w}_k = \hat{x}_k \), to be possibly preferred to (11d), is motivated by the inherent complexity reduction. On one hand, indeed, not computing the statistics \( \sum_{x_k \in S^k} p_k(x_k)\Lambda_k(z_k|x_k) \) halves the computational complexity since algorithm 6.1 has not to be implemented. On the other, complexity is further lowered by the fact that, since the observations come from an exponential family, the computation of the statistics \( \max_{x_k \in S^k} p_k(x_k)\Lambda_k(z_k|x_k) \) only can be equivalently carried on without requiring exponential transformation of the observations \( \{z_\ell\}_{\ell=1}^k \). This makes the adoption of TCO2 convenient with respect to TCO1 for practical realization and simulation. However, thresholds setting for TCO1 is easier, based on the bounds on \( P_{\text{track}} \) and \( P_{\text{pos}} \), unavailable for TCO2 (see lemma 5.1). As to the PCO strategies, no complexity reduction is granted by PCO2 with
respect to PCO1, but the latter offers the possibility of controlling \( P_{\text{pos}} \) through an additional bound. Thus, the adoption of PCO1 is in general preferred.

**Observation 6.4:** Notice that, in all of the introduced sequential rules, the estimation procedure is triggered by the detection one: for this reason, the estimator can be called a ‘gated’ estimator. In any case, the estimation procedure may be carried on along the detection one with the final estimate being discarded if the detection test has accepted the null hypothesis. Notice also that, since maximization in (9c), (10c), (11c), (12c), (13c) is equivalent to \( \arg \max \) the estimation procedure may be carried on along the detection one with the final estimate being discarded triggered by the detection one: for this reason, the estimator can be called a ‘gated’ estimator. In any case,

Thus, the adoption of PCO1 is in general preferred.

**Observation 6.5:** In the target model of Section II it has been assumed that all of the frequency transitions are allowed to take place with equal probability. This permits a reduction in the computational complexity of the DP algorithms for detection and tracking since Doppler detections may be actually carried out frame-by-frame. This may be seen by rewriting summation and maximization in (14) as

\[
\sum_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) = \sum_{x_k \in S^k} \sum_{\nu_k \in S_d} \tilde{p}_k(m_k, n_k) \frac{1}{N_d} \prod_{i=1}^k \frac{h_1(z\ell(m_\ell, n_\ell, \nu_\ell))}{h_0(z\ell(m_\ell, n_\ell, \nu_\ell))} \\
= \sum_{(m_k, n_k)} \tilde{p}_k(m_k, n_k) \prod_{\ell=1}^k \left[ \frac{1}{N_d} \sum_{\nu \in S_d} h_1(z\ell(m_\ell, n_\ell, \nu)) \right] \\
= \max_{x_k \in S^k} p_k(x_k) \Lambda_k(z_k|x_k) = \cdots = \max_{(m_k, n_k) \in S^k \times S_d} \tilde{p}_k(m_k, n_k) \prod_{\ell=1}^k \left[ \frac{1}{N_d} \max_{\nu \in S_d} h_1(z\ell(m_\ell, n_\ell, \nu)) \right],
\]

respectively, where \( \tilde{p}_k \) denotes the density of the target trajectory along the first two dimensions (i.e. azimuth and range). Otherwise stated, the 3-dimensional (azimuth-range-Doppler) detection and tracking algorithms operating on statistics \( h_1(z\ell(x_\ell))/h_0(z\ell(x_\ell)) \) can be replaced by 2-dimensional (azimuth-range) algorithms operating on the statistics

\[
\frac{1}{N_d} \sum_{\nu \in S_d} h_1(z\ell(m_\ell, n_\ell, \nu))/h_0(z\ell(m_\ell, n_\ell, \nu)), \quad \max_{\nu \in S_d} \frac{h_1(z\ell(m_\ell, n_\ell, \nu))}{h_0(z\ell(m_\ell, n_\ell, \nu))}.
\]

**Observation 6.6:** In some LPRF conventional radars the Doppler parameter is eliminated before scan-to-scan integration by performing maximization over the Doppler shifts for each received pulse train. This choice can be useful in order to limit the system complexity; moreover, the lost target velocity information.
can be effectively recovered from the estimated trajectory. For this scenario, the measurement set is

\[ Z_\ell = \left\{ Z_\ell(m,n) = \max_{\nu \in \{1,...,N_d\}} \left[ \frac{1}{2N_d} \int_\mathbb{R} r_{m,\ell}(t) \psi_{m,\ell}(t, n+I_\ell \tau_c, \nu - 1 \frac{N_d}{N_d}) dt \right] \right\} : m \in \{1,\ldots,N_a\}, n \in \{1,\ldots,N_r\}, \]

and the state space \( S = \{1,\ldots,N_a\} \times \{1,\ldots,N_r\} \), i.e. no Doppler tracking is performed. In this case, the densities of \( Z_\ell(x) \) in (5) will be

\[
\begin{align*}
    h_1(\alpha) &= (N_d - 1) e^{-\alpha} (1 - e^{-\alpha})^{N_d - 2} (1 - e^{-\frac{\alpha}{1 + \rho}}) u(\alpha) + \frac{e^{-\frac{\alpha}{1 + \rho}}}{\rho} (1 - e^{-\alpha})^{N_d - 1} u(\alpha), \\
    h_0(\alpha) &= N_d e^{-\alpha} (1 - e^{-\alpha})^{N_d - 1} u(\alpha),
\end{align*}
\]

and the likelihood ratio of in (6)

\[
\Lambda_k(z_k|x_k) = \prod_{\ell=1}^{k} \frac{h_1(z_\ell(x_\ell))}{h_0(z_\ell(x_\ell))} = \prod_{\ell=1}^{k} \left( \frac{N_d - 1 - e^{-\frac{z_\ell(x_\ell)}{1 + \rho}}}{N_d - 1 - e^{-z_\ell(x_\ell)}} + \frac{1}{N_d} \frac{e^{-\frac{z_\ell(x_\ell)}{1 + \rho}}}{1 + \rho} \right), \quad \forall k \in \mathbb{N}.
\]

At this point all of the proposed sequential procedures can be adopted starting with this new likelihood ratios.

**Observation 6.7:** It could be interesting to analyze the case where the surveillance area is not partitioned into smaller regions so that these sequential procedures have to cope with a multi-target scenario. In this case two different situations have to be necessarily considered. Precisely, if the number of targets present in the region is known and equal to \( L \), everything remains unchanged once an enlarged state vector \( X_\ell \), taking on values in is \( S^L \), is defined. If, instead, the number of targets is not know, one can extend the proposed rules as follows. At each time \( \ell \), all of the candidate trajectories having the same root pixel are associated to a single target. The proposed sequential rules, then, can be used for each set-of-trajectories/target and the procedures stop sampling when all of the tests have terminated. This can be an interesting approach where the number of set-of-trajectories/target dynamically varies with time.

**VII. numerical results**

The detection and estimation performances of the proposed sequential rules have been studied considering the sensor parameters shown in table I. The search zone is that of Figure 1 and all of the analysis is restricted to a single region, in turn composed of 4 azimuth cells. The SNR per frame \( \rho \) has been set at the design value of \( N_d \), which corresponds to an SNR per pulse \( \rho_p = 0 \) dB. As to the target kinematic, since the detection and tracking algorithm has been restricted to the furthest \( N_r \) range bins, given the sensor parameters of table I, azimuth transitions can be neglected, i.e. \( \sigma_a = 0 \). On the other hand, a
TABLE I
SENSOR PARAMETERS.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>pulse shape</td>
<td>rectangular</td>
</tr>
<tr>
<td>carrier frequency</td>
<td>30 GHz</td>
</tr>
<tr>
<td>PRF</td>
<td>820 Hz</td>
</tr>
<tr>
<td>$\tau_c$</td>
<td>0.5 $\mu$s</td>
</tr>
<tr>
<td>$N_r$ ($I$)</td>
<td>100 (2339)</td>
</tr>
<tr>
<td>$N_a$ ($\Phi$)</td>
<td>4 (4$^\circ$)</td>
</tr>
<tr>
<td>$N_d$</td>
<td>16</td>
</tr>
<tr>
<td>$K$</td>
<td>20</td>
</tr>
<tr>
<td>$\rho$</td>
<td>16</td>
</tr>
<tr>
<td>$\nu_t^{\pm}$</td>
<td>Mach-3, Mach-2</td>
</tr>
</tbody>
</table>

TABLE II
THRESHOLDS VALUES.

<table>
<thead>
<tr>
<th>DO</th>
<th>Sequential</th>
<th>FSS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma_0 = 0.18$</td>
<td>$\gamma_{DO} = 3.15$</td>
</tr>
<tr>
<td>$\gamma_1 = 55$</td>
<td>$\gamma_{DO} = 40$</td>
<td>$\gamma_{TCO} = 4.87$</td>
</tr>
<tr>
<td>$\gamma_{DO} = 3$</td>
<td>$\gamma_{TCO} = 10$</td>
<td>$\gamma_{PCO} = 9.14$</td>
</tr>
</tbody>
</table>

mobility parameter $\sigma_r$ equal to 0.8 has been chosen for range transitions, with $\Delta_r^{\pm} = 3$ range bins. The target delay and Doppler pair $\{\tau_\ell, f_\ell\}_{\ell \in \mathbb{N}}$ has been generated as continuous parameter so as to account for possible mismatches in the discretization process. The behavior of the proposed strategies have been tested through MonteCarlo simulations ($10^6$ and $5 \cdot 10^6$ runs under $H_1$ and $H_0$, respectively) in terms of $P_{fa}, P_{pos}, P_{track}, \text{ASN}$ and coefficient of variation (cv) of the sample size.

The first set of plots is aimed at comparing the different proposed strategy when there is a mismatch between the actual signal-to-noise ratio per pulse $\text{SNR}_p$ and the design value $\rho_p$; all of the curves are given as a function of $\text{SNR}_p$ and the performances of the FSS competitors are also included for comparison purposes. The thresholds, listed in table II, have been set so as to ensure a common $P_{fa}$ of $10^{-3}$ and, for the sequential case, the same ASN under both hypotheses, i.e. 3 under $H_0$ and 2.5 under $H_1$ at $\text{SNR}_p = \rho_p$; the FSS procedures always integrate 4 samples. Figure 2 shows the ASN under $H_1$: it can be noticed the characteristic peak at intermediate values of the SNR, where there is no pronounced trend to cross either boundaries (and indeed the deleterious effect of the antenna remaining blocked on a particular direction has been avoided by truncation). Notice also that all of the strategies exhibit

---

8The coefficient of variation of a random variable is the ratio $\sigma/\mu$ of its standard deviation $\sigma$ and its mean $\mu \neq 0$. 

Copyright (c) 2008 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.
approximately the same trend: DO and TCO1 are nearly coincident, TCO2 has the lowest ASN over the entire SNR range and TCO1 and PCO2 exhibit the highest peak at -4 dB. Figure 3 shows the coefficient of variation under $H_1$ (cv details under $H_0$ are also included). TCO2 exhibits the lowest dispersion while DO and PCO1 are almost equivalent: in both cases, cv remains limited in the range $[0.3, 0.8]$. The same trend is exhibited in the probability of truncation, plotted versus the SNR in Figure 4. The probability of detection is instead reported in Figure 5. It can be seen that for both the sequential and the FSS case all of the strategies exhibit almost the same behavior with the only exception of PCO2. Notice also that all the procedures achieve the same $P_d$ at the design SNR, sequential ones showing some performances improvements at lower SNRs and FSS being slightly better at larger SNRs, consistent with the ASN trend shown in Figure 2. However, it can be shown that, giving up a small amount of frame size saving, the DO sequential procedure achieves larger $P_d$ over all the inspected range of SNRs. The behavior of PCO2 can be itself explained in the light of its ASN, which is slightly lower than that of the other strategies at higher SNRs. Figures 6 and 7 show $P_{\text{track}}$ and $P_{\text{pos}}$ for both sequential and FSS rules. Again all of the strategies in the same class (sequential or FSS) exhibit almost the same performances. Notice, in particular, the massive gain granted by sequential procedures. This is mainly due to the low ASN required by sequential procedures, especially for high SNRs: in this case, indeed, the larger sample size of FSS techniques causes a larger number of candidate false tracks or final positions to be generated and to be possibly confused with the true ones. However, since it is assumed that a more complex and refined tracking algorithm is run at higher level, the goal is in fact to provide reliable short tracks or positions, leaving to the second stage the task of forming long continuous trajectories. Moreover, while the comparrison in terms of $P_{\text{track}}$ must be appropriately interpreted considering the average track size, that in terms of $P_{\text{pos}}$ reveals a substantial gain in the estimation accuracy. It is important to emphasize that Figures 2 – 7 show that sequential procedures always result in a full samples size saving with respect to their FSS counterparts. Indeed, with the conservative choice of 4 samples (which, from Figure 2, is uniformly larger than the ASN of the truncated sequential rules), FSS techniques exhibit equal or poorer performances with respect to sequential rules under both $H_0$ and $H_1$ (at the design SNR). In particular, the latter integrate on the average 3 and 2.5 samples which result in a 25% frame size saving under hypothesis $H_0$ and 37.5% under $H_1$.

Finally, in Figure 8 the performances of the sequential and FSS rules are tested in terms of the target mobility. $P_d$, $P_{\text{pos}}$, $P_{\text{track}}$ and ASN are represented for the DO strategy only, while the target mobility parameter $\sigma_r$ ranges from 0.1 to 100, i.e. from the case of a steady target to that of a maximally agile target. Notice that even if the latter case corresponds to an almost uniform transition probability which,
given the system parameters of table I, fits no real target, this situation can be of interest for two reasons: it can be considered as a “worst case” on the system performances and it can model the measurements of sensors with higher range accuracy (i.e. with a lower PRF). It can be seen, that, while the probabilities of detection and position estimation of the FSS procedure impair as the target mobility increases, those of the sequential rules remain almost unchanged, in that large values of $\sigma_r$ are counterbalanced by higher ASNs. As to $P_{\text{pos}}$, it obviously decreases in both cases, but sequential techniques retain their superiority over all the range of $\sigma_r$, even for very high values, where the almost uniform transition probabilities make past observation useless for position estimation purposes.\(^9\)

**VIII. Conclusions**

The general problem of joint sequential target detection and estimation has been considered for multi-frame surveillance radar systems. Limitations on target mobility imposed by previous works have been removed to account for fast moving targets. Different truncated sequential strategies have been proposed, in order to orient the sensor efforts either to the detection or to the track estimation or to the position estimation tasks. Bounds to control the test strength have been provided, along with a complexity analysis which has shown that all of the involved statistics can be computed through DP algorithms, whereby their complexity is always linear in the number of integrated frames. Finally, numerical results have shown a uniform and visible gain of sequential procedures with respect to the corresponding FSS: this gain becomes particularly important in the target position and/or track estimation step. As to the relative behavior of the different sequential strategies, it has been shown that all exhibit similar performances, DO and PCO1 featuring slightly better performances, TCO2 lower computational complexity.

**REFERENCES**


\(^9\)In sequential procedures upper threshold crossing is in general caused by a favorable realization of the target fluctuation, which facilitates final position estimation. FSS techniques, instead, may incur more often in unfavorable realization so that the signal amplitude at the (fixed) time of decision can be not strong enough for reliable final position estimation.


Fig. 2. Average sample number versus the SNR per pulse under \( H_1 \) and for different strategies. The number of integrated frames for fixed-sample-size procedures has been also reported.
Fig. 3. Coefficient of variation versus the SNR per pulse under $H_1$ and for different strategies. cv details under $H_0$ are also included.
Fig. 4. Probability of truncation under $H_1$ versus the SNR per pulse for different strategies.

Fig. 5. Probability of detection versus the SNR per pulse for different strategies, both sequential and fixed-sample-size.
Fig. 6. Probability of correct track estimation versus the SNR per pulse for different strategies, both sequential and fixed-sample-size.
Fig. 7. Probability of correct final position estimation versus the SNR per pulse for different strategies, both sequential and fixed-sample-size.
Fig. 8. Probabilities of detection, final position estimation and track estimation versus the target range mobility parameter for both sequential and FSS DO procedures (top); average sample number in the lower figure.