14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

JOINT SEQUENTIAL DETECTION AND TRAJECTORY ESTIMATION WITH RADAR APPLICATIONS

Stefano Buzzi, Emanuele Grossi and Marco Lops

DAEIMI, Università degli Studi di Cassino
Via G. Di Biasio 43, 03043, Cassino (FR), Italy
phone: +39 07762993742, fax: +39 07762993987, email: {buzzi, e.grossi, lops}@unicas.it

ABSTRACT

The problem of signal detection and trajectory estimation of a dynamic system when a variable number of measurements can be taken is here considered. A sequential probability ratio test (SPRT) when the parameter space has infinite cardinality is proposed for the detection problem while trajectory estimation relies upon a maximum-a-posteriori (MAP) estimate. The computational costs of the proposed algorithm, whose statistics are computed through a dynamic programming (DP) algorithm, are considered and applications to radar surveillance problems are inspected.

1. INTRODUCTION

Many statistical decision problems in engineering applications require to perform state estimation of a system under uncertainty of signal presence. This includes fault detection and diagnosis in a dynamical system control [1], target detection and tracking [2], image and speech segmentation [3], blind deconvolution of communication channels. Application of sequential decision rules to this problem arouses much interest since it promises 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 weak compared to the background noise and stringent detection specifications can be met only by processing multiple image frames (thus integrating the backscattered energy of the target along its trajectory [4]). In this case, FSS techniques usually result to be inefficient while sequential procedures are known to increase the sensitivity of power-limited radar systems or, alternatively, to reduce the average sample number (ASN). The extension of such sequential procedures, however, poses some difficulties: since the instant when the procedure stops sampling is not determined in advance (it is a random stopping time, indeed) the set of trajectories of the dynamic system to be considered (i.e. the parameter space) has an infinite cardinality. On the other hand, sequential testing rules have been already extended to the case of composite hypothesis in [5, 6] while sub-optimal sequential classification procedures have been proposed over the past years [7, 8, 9, 10, 11]. However, all of the proposed solutions are restricted to the case where the parameter space consists of a finite number of elements. This condition, when applied to the problem of detection and tracking, results to be too restrictive: indeed, it corresponds to requiring that the dynamic system may only lie in a determined state, with no transition allowed [6, 8, 11]. Only few works in the past have addressed this topic: in [12] the problems of SPRTs for parametrized hidden Markov models (HMMs) are studied while in [13] the change point detection problem for HMM is analyzed. None of them, however, have considered the situation of joint detection and tracking and possible applications to the radar framework.

This paper provides a generalization in this sense. An SPRT for the detection task when the parameter space has infinite cardinality is proposed while a MAP estimator is adopted for trajectory estimation. A DP-based algorithm for efficient implementation of the signal detector and of the trajectory estimator is adopted. New analytical bounds on the ASN are derived and specific applications to radar surveillance are considered. Moreover, a thorough performance analysis is provided in order to validate the correctness of such bounds and to highlight the effects of system parameters, both for the general case and for radar applications.

The rest of the paper is organized as follows. Next section presents the elements of the problem while section 3 addresses the problem of sequential joint detection and estimation and section 4 covers the radar surveillance problem. Finally, section 5 is devoted to the presentation of numerical results, while concluding remarks are given in section 6.

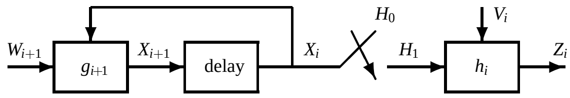

Figure 1: measurement generating process.

2. PROBLEM FORMULATION

The problem is formulated with reference to the dynamic system shown in figure 1, whose elements are listed below [14].

1. The known first-order difference equation describing the dynamic system
$$X_{i+1} = g_i(X_i, W_i) \in \mathcal{S}, \quad \forall\, i \in \mathbb{N}, \tag{1}$$
where $\mathcal{S}$ is a discrete non-empty set referred to as the state space with cardinality M; X_i is the state vector for the dynamic system at the i-th stage (or sampling instant) while W_i is the random forcing function at the i-th stage.
2. The probability density of the initial state p_{X_1}; it is assumed that X_1 and $\{W_i\}_{i\in\mathbb{N}}$ are independent.
3. The statistics of $\{W_i\}_{i\in\mathbb{N}}$, i.e. the probability density p_{W_i} for all $i \in \mathbb{N}$; $\{W_i\}_{i\in\mathbb{N}}$ is assumed to be a zero-mean independent process.
4. The known equation describing the measurement process $\{Z_i\}_{i\in\mathbb{N}}$
$$Z_i = h_i(X_i, V_i) \in \mathbb{R}^d, \quad \forall\, i \in \mathbb{N},$$
if Z_i is originated from the dynamic system described by the difference equation (1) and by the noise or
$$Z_i = h_i(V_i) \in \mathbb{R}^d,$$
if Z_i is generated by noise alone, $\{V_i\}_{i\in\mathbb{N}}$ being the measurement noise process.
5. The statistics of $\{V_i\}_{i\in\mathbb{N}}$, i.e. the probability density f_{V_i} for all $i \in \mathbb{N}$; $\{V_i\}_{i\in\mathbb{N}}$ is assumed to be independent of X_1 and $\{W_i\}_{i\in\mathbb{N}}$.

It is easy to verify that the state sequence $\{X_i\}_{i\in\mathbb{N}}$ is a Markov process and that the measurements given the state $\{Z_i|X_i\}_{i\in\mathbb{N}}$ are independent; furthermore, the joint densities of $X^k = \{X_i\}_{i=1}^k$ and $Z^k|X^k = \{Z_i\}_{i=1}^k|\{X_i\}_{i=1}^k$ are

$$p_{X^k}(x^k) = p_{X_1}(x_1)\prod_{i=2}^{k} P_{x_{i-1}x_i}(i), \tag{2}$$

$$f_{Z^k|X^k}(z^k|x^k) = \prod_{i=1}^{k} f_{Z_i|X_i}(z_i, x_i), \tag{3}$$

respectively, where $P_{x_{i-1}x_i}(i) = p_{X_i|X_{i-1}}(x_i, x_{i-1})$ denotes the transition probability from state x_{i-1} to x_i at stage i.

Given these elements, one is to sample the process $\{Z_i\}_{i\in\mathbb{N}}$ 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 dynamic system and, in the latter case, it is also required to estimate the system trajectory which has generated such measurements. The sequential nature of the decision process allows now a trade-off between quickness of decision and decision accuracy.

As in [15], this is a problem where estimation has to be performed under uncertainty as to the signal presence. Thus, there is a mutual coupling of detection and estimation and two different strategies may be adopted. The first one (detection-oriented) consists in choosing a test function for the detection problem and then, if the latter has decided that the signal is present, design an estimator. The second (classification-oriented) concerns the design of a classification rule where detection is embedded in the rule itself. Since it exhibits certain optimal properties and since detection is the primary interest in surveillance applications later discussed, the former only will be discussed in the following sections.

3. DETECTION AND ESTIMATION PROCEDURES

The problem is to to find an 'optimal' sequential testing function. After the sampling process has stopped, if a signal has been declared to be present, an estimator operating on a fixed-size sample may be used to provide an estimate of the state trajectory.

3.1 Sequential test

It is not difficult to see the detector design as a sequential testing problem of a simple hypothesis versus a composite alternative where the parameter space is $(\Theta, \mathscr{U})$, $\Theta = \Theta_0 \cup \Theta_1$ and $\mathscr{U}$ being a σ-algebra of subsets of Θ. $\Theta_0 = \{\theta_0\}$ characterizes the simple hypothesis H_0:'noise only' while $\Theta_1 = \times_{i\in\mathbb{N}}\mathcal{S}$ characterizes the alternative H_1:'signal present'. The prior μ on Θ_1 is given through the sequence of increasing joint densities $\{p_{X^k}\}_{k\in\mathbb{N}}$. Thus, a sequential rule (φ, ϕ) has to be found, i.e. a stopping rule $\varphi = \{\varphi_k\}_{k\in\mathbb{N}}$ and a terminal decision rule $\phi = \{\phi_k\}_{k\in\mathbb{N}}$, for testing H_0. The strength of such a sequential test is the pair of errors of first and second kind. Often, in detection problems, these errors are referred to as probability of false alarm and probability of miss, respectively. Denoting with τ the stopping time and with $\psi = \{\psi_k\}_{k\in\mathbb{N}}$ its conditional distribution[1], these errors are defined as follows.

(i) The probability of false alarm, P_{fa}, is the error of first kind, i.e. $P_{\text{fa}} = \alpha = \sum_{k\in\mathbb{N}} \mathrm{E}_{\theta_0}\left[\psi_k(Z^k)\phi_k(Z^k)\right]$[2].

(ii) The probability of miss given the system is in state $\theta \in \Theta_1$, $P_{\text{miss}}(\theta)$, is the error of second kind, i.e. $P_{\text{miss}}(\theta) = \beta(\theta) = \sum_{k\in\mathbb{N}} \mathrm{E}_{\theta}\left[\psi_k(Z^k)\big(1-\phi_k(Z^k)\big)\right], \quad \forall\, \theta \in \Theta_1$.

(iii) P_{miss} is the average probability of miss, i.e. $P_{\text{miss}} = \beta = \sum_{k\in\mathbb{N}} \sum_{x^k\in\mathcal{S}^k} p_{X^k}(x^k)\, \mathrm{E}_{x^k}\left[\psi_k(Z^k)\big(1-\phi_k(Z^k)\big)\right]$. On the other hand, $P_{\text{d}} = 1 - P_{\text{miss}}$ denotes the average probability of detection.

Suppose that it is requested to find a sequential test whose probability of error of first kind is equal to $P_{\text{fa}} = \alpha$ and such that the weighted average of the probability of second kind is equal to $P_{\text{miss}} = \beta$. To this end, define first

$$f_{Z^k|H_1}(z^k) = \sum_{x^k\in\mathcal{S}^k} p_{X^k}(x^k)\, f_{Z^k|X^k}(z^k|x^k), \quad \forall\, k \in \mathbb{N}.$$

Then the sequential probability ratio test of strength (α, β) for testing $f_{Z^k|H_0}$ against $f_{Z^k|H_1}$ gives a solution to the problem, i.e.

$$\varphi_k(z^k) = \begin{cases} 0, & \text{if } \Lambda_k(z^k) \in (\gamma_0, \gamma_1), \\ 1, & \text{otherwise}, \end{cases} \tag{4a}$$

$$\phi_k(z^k) = \begin{cases} 0, & \text{if } \Lambda_k(z^k) \leq \gamma_0, \\ 1, & \text{if } \Lambda_k(z^k) \geq \gamma_1, \end{cases} \tag{4b}$$

where Λ_k is the likelihood ratio of $f_{Z^k|H_1}$ and $f_{Z^k|H_0}$. The boundaries of the test, γ_0 and γ_1, with $0 < \gamma_0 \leq 1 \leq \gamma_1 < +\infty$, are chosen in order to have the required strength (α, β).

The testing rule of (4) requires to evaluate the likelihood ratios

$$\Lambda_k(z^k) = \sum_{x^k\in\mathcal{S}^k} p_{X^k}(x^k) \frac{f_{Z^k|X^k}(z^k|x^k)}{f_{Z^k|H_0}(z^k)}, \tag{5}$$

for $k = 1, \ldots, \tau$. From equations (2) and (3) it results that (5) is a stage separated function on the algebraic system $(\mathbb{R}, +, \cdot)$ and, thus, it can be computed through the following dynamic programming algorithm [16], which is known to lower the computational complexity from exponential to linear in k.

Algorithm 3.1. *Let $\{F^i\}_{i\in\mathbb{N}}$ be a sequence of real-valued functions on $\mathcal{S}$. Then the algorithm proceeds as follows.*

1. *Initialization.*

$$F^1(x) = p_{X_1}(x) \frac{f_{Z_1|X_1}(z_1|x)}{f_{Z_1|H_0}(z_1)}, \quad \forall\, x \in \mathcal{S},$$
$$\Lambda_1(z^1) = \sum_{x\in\mathcal{S}} F^1(x).$$

2. *Recursion. For every $i \geq 2$*

$$F^i(x) = \frac{f_{Z_i|X_i}(z_i|x)}{f_{Z_i|H_0}(z_i)} \sum_{x_{i-1}\in\mathcal{S}} P_{x_{i-1}x}(i) F^{i-1}(x_{i-1}), \quad \forall\, x \in \mathcal{S},$$
$$\Lambda_i(z^i) = \sum_{x\in\mathcal{S}} F^i(x).$$

3.2 Gated estimator

Suppose the detection test has stopped sampling at stage $\tau = k$ and it has decided in favor of hypothesis H_1. The problem of finding an 'optimal' classification rule starting from the set of measurements Z^k is immediately solved resorting to the MAP estimator

$$\hat{x}^k = \arg\max_{x^k\in\mathcal{S}^k} p_{X^k}(x^k)\, f_{Z^k|X^k}\big(z^k|x^k\big). \tag{6}$$

The objective of (6) is a stage separated function on the algebraic system $(\mathbb{R}, \max, \cdot)$ and it can be computed through the following DP algorithm.

Algorithm 3.2. *Let $\{F^i\}_{i=1}^k$ be a sequence of real-valued functions on $\mathcal{S}$ and $\{\delta_i\}_{i=2}^k$ be a sequence of endomorphisms on $\mathcal{S}$. Then the algorithm proceeds as follows.*

1. *Initialization. $\forall\, x \in \mathcal{S}$*

$$F^1(x) = p_{X_1}(x)\, f_{Z_1|X_1}(z_1|x).$$

2. *Recursion. For $i = 2, \ldots, k$ and $\forall\, x \in \mathcal{S}$*

$$F^i(x) = f_{Z_i|X_i}(z_i|x) \max_{x_{i-1}\in\mathcal{S}} \left\{P_{x_{i-1}x}(i) F^{i-1}(x_{i-1})\right\},$$
$$\delta^i(x) = \arg\max_{x_{i-1}\in\mathcal{S}} \left\{P_{x_{i-1}x}(i) F^{i-1}(x_{i-1})\right\}.$$

3. *Termination.*

$$\hat{x}_k = \arg\max_{x\in\mathcal{S}} F^k(x).$$

[1] $\psi_k(z^k)$ is the probability that $\tau|\{Z^k = z^k\} = k$; the relationship between ψ_k and φ_k is: $\psi_k(z^k) = \varphi(z^k)\prod_{l=1}^{k-1}\big(1-\varphi(z^l)\big)$, $\forall k \in \mathbb{N}$.

[2] Throughout the paper, with E_θ it is denoted the operator of expectation on the observations given that θ is the true state of nature.

4. *Backtracing. For* $i = k-1, k-2, \dots, 1$

$$\widehat{x_i} = \delta_{i+1}(x_{i+1}).$$

Notice that maximization in (6) is equivalent to $\arg\max_{x^k \in \mathcal{S}^k} p_{X^k}(x^k) f_{Z^k|X^k}(z^k|x^k)/f_{Z^k|H_0}(z^k)$, which means that the estimator may work on the same data as the detector, thus lowering the required computation complexity. Furthermore, the estimation procedure of algorithm 3.2 may be carried on along the detection one in algorithm 3.1 with the final estimate being discarded if the detection test has accepted the null hypothesis. Finally notice that the running time of algorithms 3.1 and 3.2 is $\mathcal{O}(kM^2)$.

3.3 Bounds on the ASN and approximations for boundaries

Approximations of the ASN are difficult to obtain, since, in general, they would depend on the prior μ on Θ_1, i.e. on the sequence of densities $\{p_{X^k}\}_{k\in\mathbb{N}}$, as it can be seen observing their definition:

$$\begin{aligned}\mathrm{ASN}_{H_0} &= \mathrm{E}_{\theta_0}[\tau] = \sum_{k\in\mathbb{N}} \mathrm{E}_{H_0}\left[\psi_k(Z^k)\right]k,\\ \mathrm{ASN}_{H_1} &= \sum_{\theta\in\Theta_1} \mu(\theta)\,\mathrm{E}_{\theta_0}[\tau] = \sum_{k\in\mathbb{N}}\sum_{x^k\in\mathcal{S}^k} p_{X^k}(x^k)\,\mathrm{E}_{x^k}\left[\psi_k(Z^k)\right]k.\end{aligned}$$

Nevertheless, some bounds may be still given. In particular, it can be demonstrated that, under rather mild conditions, the ASN of the SPRT detector, as a function of the prior μ, always lies between two extrema: the deterministic case, for which $\mu(\theta) = 1$ for some $\theta \in \Theta_1$, and the 'maximum uncertainty' case, for which μ is the uniform distribution over Θ_1. The derivation of these bounds is carried out in multiple steps.

First, two auxiliary sequences of functions are defined: $\{h_{10}^k\}_{k\in\mathbb{N}}$ and $\{h_{01}^k\}_{k\in\mathbb{N}}$, with $h_{10}^k, h_{01}^k : H^k \to \mathbb{R}$, for all $k \in \mathbb{N}$, and $H^k = \{q \in [0,1]^{M^k} : \sum_{i=1}^{M^k} q_i = 1\}$. Since every probability density over $\mathcal{S}^k$ can be uniquely represented by a vector in H^k, H^k can be thought of as the set of densities over $\mathcal{S}^k$. On the other hand, h_{10}^k, h_{01}^k are defined, for all $k \in \mathbb{N}$, as

$$\begin{aligned}h_{10}^k(q) &= D\big(\textstyle\sum_{i=1}^{M^k} q_i f_{Z^k|X^k}(\cdot|\theta_i^k) \parallel f_{Z^k|H_0}(\cdot)\big),\\ h_{01}^k(q) &= D\big(f_{Z^k|H_0}(\cdot) \parallel \textstyle\sum_{i=1}^{M^k} q_i f_{Z^k|X^k}(\cdot|\theta_i^k)\big),\end{aligned}$$

where $D(f \parallel g)$ denotes the Kullback-Leibler divergence between f and g and θ_i^k the i-th element of the set $\mathcal{S}^k$ (i.e. the i-th trajectory up to epoch k). With this definition, it results that

$$\begin{aligned}h_{10}^k\big(\{p_{X^k}(x^k)\}_{x^k\in\mathcal{S}^k}\big) &= D(f_{Z^k|H_1} \| f_{Z^k|H_0}) = \mathrm{E}_{H_1}\left[\ln(\Lambda_k)\right],\\ h_{01}^k\big(\{p_{X^k}(x^k)\}_{x^k\in\mathcal{S}^k}\big) &= D(f_{Z^k|H_0} \| f_{Z^k|H_1}) = -\mathrm{E}_{H_0}\left[\ln(\Lambda_k)\right].\end{aligned}$$

It is now given the following.

Condition 3.3. *For every* $k \in \mathbb{N}$ *and for every permutation matrix* P[3]*, functions* h_{10}^k *and* h_{01}^k *satisfy*

$$h_{10}^k(Pq) = h_{10}^k(q), \quad h_{01}^k(Pq) = h_{10}^k(q), \quad \forall q \in H^k.$$

Condition 3.3 can be interpreted as follows. Since the Kullback-Leibler divergence is a measure of the 'distance' or 'divergence' between two statistical populations drawn from different probability distributions, condition 3.3 requires that, after a weighting vector $q \in H^k$ has been chosen, any two mixtures $f_{Z^k|H_1,q} = \sum_{i=1}^{M^k} q_i f_{Z^k|X^k}(\cdot|\theta_i^k)$ and $f_{Z^k|H_1,Pq} = \sum_{i=1}^{M^k}(Pq)_i f_{Z^k|X^k}(\cdot|\theta_i^k)$ are equally distant from density $f_{Z^k|H_0}$ and vice versa (as Kullback-Leibler divergence is not symmetric). This means that, at stage k, if two different priors $\{p_{X^k}(x^k)\}_{x^k\in\mathcal{S}^k}$ are considered, whose only difference is a permutation in their elements, the two dynamical system described would be equally detectable using likelihood ratio techniques.

Now it can be seen that, if condition 3.3 holds, under the hypotheses of i.i.d. measurement noise and possibility of neglecting the excesses of Λ_τ over boundaries γ_0 and γ_1, the ASN of the SPRT detector defined in (4) always satisfies the following bounds:

$$\frac{\beta\ln\gamma_0 + (1-\beta)\ln\gamma_1}{D(f_{Z_i|X_i} \| f_{Z_i|H_0})} \le \mathrm{ASN}_{H_1} \le \frac{\beta\ln\gamma_0 + (1-\beta)\ln\gamma_1}{D\big(\frac{1}{M}\sum_{j=1}^{M} f_{Z_i|X_i}(\cdot,x_j) \parallel f_{Z_i|H_0}\big)}, \tag{8a}$$

$$\frac{(1-\alpha)\ln\gamma_0 + \alpha\ln\gamma_1}{-D(f_{Z_i|H_0} \| f_{Z_i|X_i})} \le \mathrm{ASN}_{H_0} \le \frac{(1-\alpha)\ln\gamma_0 + \alpha\ln\gamma_1}{-D\big(f_{Z_i|H_0} \parallel \frac{1}{M}\sum_{j=1}^{M} f_{Z_i|X_i}(\cdot,x_j)\big)}. \tag{8b}$$

The demonstration, whose details are omitted due to lack of space, is based on the fact that h_{10}^k and h_{01}^k are convex functions on H^k and that they both admit a minimum and a maximum for the deterministic and the uniform prior, respectively.

Since the ASN is finite, the test ends almost surely and approximation for the boundaries comes directly from [5]. Indeed, it results that, ignoring overshoots, the boundaries

$$\gamma_1 = (1-\beta)/\alpha \quad \text{and} \quad \gamma_0 = \beta/(1-\alpha), \tag{9}$$

result in an SPRT detector with strength approximately given by (α,β). As concerns the optimal property of the Wald-Wolfowitz theorem, it generally does not hold for the test in (4). Nevertheless, this test may represent a good choice among available procedures in many problems. Furthermore, it can be demonstrated that, under the additional hypotheses in [17] or [12], this procedure is asymptotically optimal.

3.4 Truncated SPRT

Even if the SPRT exhibits the smallest ASN under both H_0 and H_1 and it terminates with probability one, occasionally long observations can be needed. Moreover, if there are mismatches between design and actual values of some parameters, typically the signal-to-noise ratio (SNR), the resulting ASN can be very large, especially for small error probabilities. Truncation of the SPRT can be used to prevent such a problem: a regular sequential test is carried out until either a decision is made or a fixed stage K is reached, in which case hypothesis H_0 or H_1 is accepted if $\Lambda_K \lessgtr \gamma_K$, respectively. Truncation is, then, a compromise between an entirely sequential test and a classical FSS test. It can be noticed that, as long as the probability of truncation is negligible, the previously derived approximations on error probabilities and ASN are still valid.

4. RADAR APPLICATIONS

The radar problem is characterized by the inherent presence of multiple-resolution elements, which correspond to range 'bins' as well as Doppler, azimuth and elevation cells. This problem has been solved in [6, 8, 11] but all of these approaches concern the case where the target is not allowed to change its position 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 go beyond Mach-2.

The physical situation considered in order to guarantee the presence of at most one target in the search region is that the surveillance area is divided into smaller angular regions, each visited in turn by the antenna beam in cyclic manner. In each region a sequential procedure is used to accept or reject the hypothesis that a single target is present. Because in most sectors of the sky no target is present, sequential procedures can result in a high saving of the average total time spent scanning the surveillance region.

Since in surveillance radar applications the main objective concerns early detections, the detection-oriented sequential strategy presented in section 3 is applied. On the other hand, the possibility of occasionally long tests is ward off imposing a cut-off stage at

[3] A matrix P is a permutation matrix if exactly one entry in each row and column is equal to 1 and all other entries are 0.

which truncate the procedure: in this way, a control over the maximum dwell time on each angular sector is maintained thus avoiding the possibility that targets may traverse undetected through the surveillance region.

4.1 Sensor measurement model

The process of signal discretization is a standard calculation in radar applications and is related to the fact that the target parameters (azimuth, elevation, delay and Doppler shift), which are inherently continuous, can be estimated up to an uncertainty dictated by the beamwidth of the transmit antenna and by the ambiguity function of the transmitted signal. That is, the region that must be considered is divided into a grid and the continuous-time received signal is discretized accordingly (if the grid is sufficiently fine, losses due to possible mismatches may be neglected). The measurement at stage ℓ is, for all $\ell \in \mathbb{N}$,

$$Z_\ell = \big\{ Z_\ell(n) : n \in \{1,\dots,N_a\}\times\{1,\dots,N_e\}\times\{1,\dots,N_r\}\times\{1,\dots,N_d\}\big\},$$

where N_a, N_e, N_r, N_d are the number of resolution elements in azimuth, elevation, range and Doppler, respectively.

4.2 Target model

The target state variable is $X_\ell \in \mathcal{S}$ and the target space consists of the set of all the resolution cells, i.e. $\mathcal{S} = \{1,\dots,N_a\}\times\{1,\dots,N_e\}\times\{1,\dots,N_r\}\times\{1,\dots,N_d\}$, with $M = N_a N_e N_r N_d$. A more complex state space could be considered (for example one involving also velocities): as known, enlarging the state space leads to an improvement of the system performances at the price of an increase of the computational complexity. Since simple algorithms are required in multi-frame radars, this section focus on the case of limited dimensions state space, but all of the discussions can be easily extended to the case of a larger state space.

A first-order Gaussian-Markov random walk model is used to derive the transition probabilities, which are given by $P_{mn}(i) = \prod_{j=1}^4 P_{m_j n_j}$, where $P_{m_j n_j} = Q\Big(\frac{n_j - m_j - 1/2}{\sigma_j\sqrt{\Delta T}}\Big) - Q\Big(\frac{n_j - m_j + 1/2}{\sigma_j\sqrt{\Delta T}}\Big)$[4] for $j = 1,2,3$, where σ_j is a parameter related to the target mobility. On the other hand, for $j = 4$ (i.e. for the transitions Doppler), $P_{m_j n_j} = 1/N_d$ since in surveillance applications low pulse repetition frequency radars are usually adopted and, thus, Doppler measurements are highly ambiguous. As concerns the initial probability, if no other prior information is available (for example previous detections), it is reasonable to put $p_{X_1} = 1/M$.

4.3 Sequential detection and tracking algorithms

It is supposed that each component $Z_\ell(n)$ of the measurement Z_i, for $n \in \mathcal{S}$, be an exponentially distributed random variable with density[5]

$$f_1(y) = \frac{e^{-\frac{y}{1+\rho}}}{1+\rho} u(y), \quad \text{if the target is present in location } n, \qquad (10a)$$

$$f_0(y) = e^{-y} u(y), \qquad \text{otherwise}, \qquad (10b)$$

where ρ denotes the SNR and $u(y)$ is the Heaviside step function. Supposing $\{V_i\}_{i\in\mathbb{N}}$ to be an i.i.d. process (a condition commonly verified), the density of the set of measurements up to epoch k is

$$f_{Z^k|X^k}(z^k|x^k) = \prod_{\ell=1}^{k} f_1\big(z_\ell(x_\ell)\big) \prod_{\substack{x\in\mathcal{S}\\ x\neq x_\ell}} f_0\big(z_\ell(x)\big), \qquad (11)$$

$$f_{Z^k|H_0}(z^k) = \prod_{\ell=1}^{k} \prod_{x\in\mathcal{S}} f_0\big(z_\ell(x)\big), \qquad (12)$$

[4] $Q(x) = \int_x^{+\infty} \frac{1}{2\pi} e^{-x^2/2} dx$. Notice that, even if all transitions are theoretically admissible, real targets necessarily need to satisfy physical constraints, such as limitations on the maximum velocity and acceleration. In this case, a truncated Gaussian density can be used.

[5] This is the case, for example, if measurements comes from a square law envelope detector (commonly adopted in radar applications).

and the likelihood ratio

$$\frac{f_{Z^k|X^k}(z^k|x^k)}{f_{Z^k|H_0}(z^k)} = \prod_{\ell=1}^{k} \frac{f_1\big(z_\ell(x_\ell)\big)}{f_0\big(z_\ell(x_\ell)\big)} = \prod_{\ell=1}^{k} \frac{e^{z_\ell(x_\ell)\rho/(1+\rho)}}{1+\rho}. \qquad (13)$$

At this point, algorithms 3.1 and 3.2 can be used to efficiently compute the statistics in (5) and (6).

In the target model of section 4.2 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 easily seen rewriting summation in (5) maximization in (6) as

$$\sum_{x^k\in\mathcal{S}^k} p_{X^k}(x^k) \frac{f_{Z^k|X^k}(z^k|x^k)}{f_{Z^k|H_0}(z^k)} =$$

$$= \sum_{n_1^k,n_2^k,n_3^k} \sum_{n_4^k} p_{(X^k)_1}(n_1^k) p_{(X^k)_2}(n_2^k) p_{(X^k)_3}(n_3^k) \frac{1}{N_d^k} \prod_{\ell=1}^{k} \frac{f_1\big(z_\ell(x_\ell)\big)}{f_0\big(z_\ell(x_\ell)\big)} =$$

$$= \sum_{n_1^k,n_2^k,n_3^k} p_{(X^k)_1}(n_1^k) p_{(X^k)_2}(n_2^k) p_{(X^k)_3}(n_3^k) \prod_{\ell=1}^{k} \left[\frac{1}{N_d} \sum_{n_4} \frac{f_1\big(z_\ell(x_\ell)\big)}{f_0\big(z_\ell(x_\ell)\big)} \right],$$

$$\max_{x^k\in\mathcal{S}^k} p_{X^k}(x^k) \frac{f_{Z^k|X^k}(z^k|x^k)}{f_{Z^k|H_0}(z^k)} = \cdots =$$

$$= \max_{n_1^k,n_2^k,n_3^k} p_{(X^k)_1}(n_1^k) p_{(X^k)_2}(n_2^k) p_{(X^k)_3}(n_3^k) \prod_{\ell=1}^{k} \left[\frac{1}{N_d} \max_{n_4} \frac{f_1\big(z_\ell(x_\ell)\big)}{f_0\big(z_\ell(x_\ell)\big)} \right],$$

where $(X^k)_i = \big\{(X_1)_i,\dots,(X_k)_i\big\}$, i.e. the target trajectory along the i-th dimension. This corresponds to say that two 4-dimensional (azimuth-elevation-range-Doppler) algorithms operating on statistics $f_1\big(z_\ell(x_\ell)\big)/f_0\big(z_\ell(x_\ell)\big)$ can be replaced by two 3-dimensional (azimuth-elevation-range) algorithms operating on

$$\frac{1}{N}\sum_{n_4} \frac{f_1\big(z_\ell(x_\ell)\big)}{f_0\big(z_\ell(x_\ell)\big)} \quad \text{and} \quad \frac{1}{N}\max_{n_4} \frac{f_1\big(z_\ell(x_\ell)\big)}{f_0\big(z_\ell(x_\ell)\big)}.$$

Finally, the threshold γ_K at the cut-off stage K, needed to the truncated algorithm, is simply chosen to be the geometric mean of γ_0 and γ_1.

5. NUMERICAL RESULTS

First, a general target detection and trajectory estimation problem is considered in order to corroborate the discussion in section 3. The measurement model is that of equations (10)-(13), while the state space is $\mathcal{S} = \{1,\dots,M\}$. It can be easily checked that the functions h_{10}^k and h_{01}^k of equations (7) constructed starting upon these densities satisfy conditions 3.3. The first two curves presented are given as a function of the target prior, which ranges from the deterministic case to the uniform one. The boundaries γ_0 and γ_1 have been set as in equation (9), with design error probability $P'_{\text{fa}} = 10^{-3}$ and $P'_{\text{miss}} = 10^{-3}$. Figure 2 shows the ASN under both H_0 and H_1 as a function of M, with $\rho = 0$ dB. First, it can be seen that both ASN_{H_0} and ASN_{H_1} lie between the values corresponding to the deterministic and the uniform distributions. Furthermore, since the SNR is not large and the error probability are requested to be sufficiently low, the hypothesis that the excesses over boundaries could be neglected is verified and, thus, approximations in (8) are confirmed to be tight. It can be also noticed the deleterious effect of the increase of M confirming the intuitive idea that it becomes more difficult to search a target if it is allowed to wander in a larger state space. Furthermore, it can be seen that more compact priors allows easier detections and, in general, that priors with smaller entropy permit easier detections. Figure 3 shows the effect of the SNR on $P_{\text{c,last}}$ (the probability to correctly classify the last target position with an accuracy of two cells) and $P_{\text{c,track}}$ (the probability to correctly classify the target trajectory with an accuracy of five cells). Notice that, while P_{d} is not

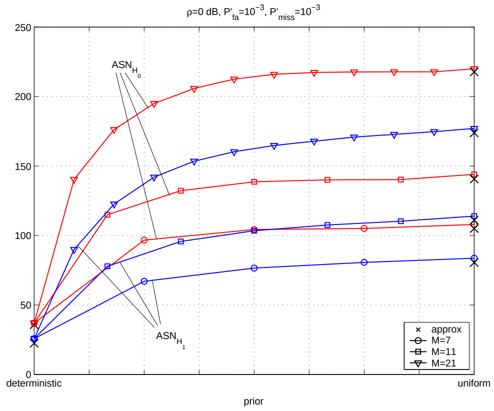

Figure 2: ASN under both hypotheses versus the prior for different values of state space cardinality and SNR of 0 dB.

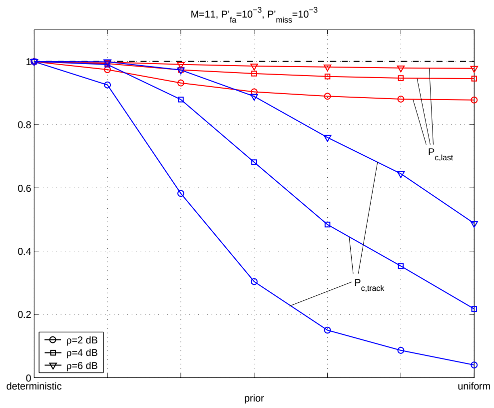

Figure 3: probability of correct classification of the last stage $P_{\mathrm{c,last}}$ and of the trajectory $P_{\mathrm{c,track}}$ versus the prior for different values of the SNR.

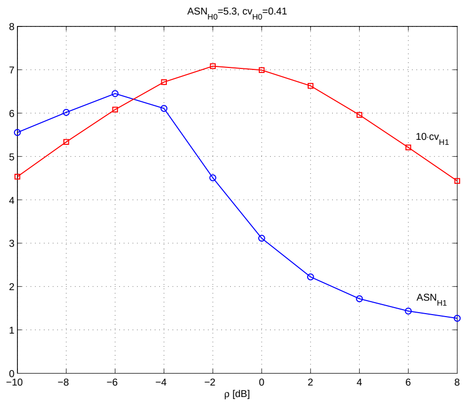

Figure 4: ASN and coefficient of variation versus the SNR for both the hypotheses (radar environment).

influenced by the strength of the received signal (indeed the lower SNRs are traded with a larger ASNs), $P_{\mathrm{c,last}}$ and even more $P_{c,track}$ decreases as the SNR is lowered.

Finally last plot concerns with the radar case of section 4. The parameter of the radar considered are: $N_a = 3$, $N_e = 3$, $N_r = 100$, $N_d = 16$, $K = 20$. The transition probabilities of the target have been set to be uniform among the admissible state transitions (in turn equal to a single resolution cell) while constant false alarm rate statistics have been used to cope with the uncertainty as to the noise power. In figure 4, the ASN and the coefficient of variation[6] is represented versus the SNR under both the hypotheses. Notice the characteristic peak at intermediate values of the SNR, where there is no pronounced tendency to cross either boundaries: yet the deleterious effect of the beam antenna remaining 'hung-up' along a particular direction has been avoided by truncation.

6. CONCLUSIONS

The general problem of sequential detection and trajectory estimation of a dynamic system observed through a set of noisy measurements has been considered and possible applications to radar surveillance have been inspected. Previous limitation on the target mobility imposed by other works present in literature have been removed and a deep analysis on the ASN as a function of the prior distribution has been given. Simulation results have shown correctness of the bounds on the ASN and that the system performance essentially follows the entropy of the prior distribution.

[6]The coeffi cient of variation of a random variable is the ratio σ/μ of its standard deviation σ and its mean $\mu \neq 0$.

REFERENCES

[1] M. A. Massoumnia, G. C. Verghese and A. S. Willsky, "Failure detection and identification," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 316–321, Mar. 1989.

[2] H. Im and T. Kim, "Optimization of multiframe target detection schemes," *IEEE Trans. Aerosp. Elec. Sys.*, vol. 35, pp. 176–187, Jan. 1999.

[3] N. Merhav and Y. Ephraim, "A Bayesian classification approach with application to speech recognition," *IEEE Trans. Sign. Proc.*, vol. 39, pp. 2157–2166, Oct. 1991.

[4] S. Blackman and R. Popoli, *Design and analysis of modern tracking systems*. Address: Artech House, 1999.

[5] A. Wald, "Sequential tests of statistical hypotheses," *Ann. Math. Stat.*, vol. 16, pp. 117–186, Jun. 1945.

[6] M. B. Marcus and P. Swerling, "Sequential detection in radars with multiple resolution elements," *IEEE Trans. Inform. Theory*, vol. 8, pp. 237–245, Apr. 1962.

[7] P. Armitage, "Sequential analysis with more than two alternative hypotheses and its relation to discriminant function analysis," *J. Roy. Statist. Soc.*, vol. 12, pp. 137–144, 1950.

[8] C. W. Helstrom, "A range-sampled sequential detection system," *IEEE Trans. Inform. Theory*, vol. 8, pp. 43–47, Jan. 1962.

[9] G. Lorden, "Likelihood ratio tests for sequential k-decision problems," *Ann. Math. Stat.*, vol. 43, pp. 1412–1427, Oct. 1972.

[10] C. W. Baum and V. Veeravalli, "A sequential procedure for Multihypothesis testing," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1994–2007, Nov. 1994.

[11] A. G. Tartakovsky and X. R. Li and G. Yaralov, "Sequential detection of targets in multichannel systems," *IEEE Trans. Inform. Theory*, vol. 49, pp. 425–445, Feb. 2003.

[12] C. D. Fuh, "SPRT and CUSUM in hidden Markov models," *Ann. Stat.*, vol. 31, pp. 942–977, Apr. 2003.

[13] B. Chen and P. Willett, "Detection of hidden Markov model transient signals," vol. 36, pp.1253–1268, Oct. 2000.

[14] R. E. Larson and J. Peschon, "A dynamic programming approach to trajectory estimation," *IEEE Trans. Automat. Contr.*, vol. 11, pp. 537–540, Jul. 1966.

[15] B. Baygün and A. O. Hero, "Optimal simultaneous detection and estimation under a false alarm constraint," *IEEE Trans. Inform. Theory*, vol. 41, pp. 688–703, May 1995.

[16] S. Verdú and H. V. Poor, "Abstract dynamic programming models under commutativity conditions," *SIAM J. Contr. Opt.*, vol. 25, pp. 990–1006, Jul. 1987.

[17] T. L. Lay, "Asymptotic optimality of invariant sequential probability ratio tests," *Ann. Stat.*, vol. 9, pp. 318–333, Mar. 1981.