Abstract—In this paper, we focus on the design of a Kalman filter-based algorithm to track multiple targets in range and azimuth. Similarly to a Rake receiver, which takes advantage of multipath components for data demodulation, the proposed tracker is aimed at collecting and eventually processing all useful target components in both range and azimuth. To this end, but in the attempt to limit the computational complexity of such an approach, at the \( k \)th scan, a GLRT-based detector is fed by the predicted target position (computed by the tracker using measurements up to the \((k - 1)\)th scan) and returns to the tracker a measurement obtained by taking into account “target strength” over multiple range-azimuth cells. In case of prediction of multiple targets in spatial proximity a multitarget detection approach is pursued to combat mutual interference. A JPDA is employed to handle the data association problem. We model target echoes as coherent pulse trains; moreover, we suppose that targets are embedded in white Gaussian noise with unknown power. Performance analysis assumes a maritime scenario and shows that the proposed algorithm can provide accurate estimates of target trajectories.

Index Terms—Maximum likelihood (ML) estimation, generalized likelihood ratio test (GLRT), multiple targets tracking, Kalman filter, joint probabilistic data association (JPDA).

I. INTRODUCTION

Tracking algorithms are designed assuming that the sensor provides a set of point measurements at each scan. In a radar system such measurements could be obtained by thresholding the output of a matched filter fed by a baseband version of collected data. Then, the tracking algorithm links measurements across time and estimates the parameters of interest [1]. As to the detection stage, most detectors considered so far assume that the target is located exactly “where the matched filter is sampled” and, hence, that there is no spillover of target energy to adjacent matched filter samples. However, the spillover is taken into account in [2] where the case of several closely spaced targets, ruled by a Swerling 2 model, is considered; therein, a maximum likelihood (ML) extractor is developed. The idea is further investigated in order to estimate the angles and ranges of multiple unresolved extended targets in [3]. In [4] an adaptive space-time detector is derived and assessed: it takes advantage of the possible spillover of target energy for the case of coherent target echoes. In [5] this idea has been extended to address localization of multiple point-like targets within the same range cell or range cells in spatial proximity; therein, it is shown that the multitarget approach outperforms the corresponding single target one in the localization of a weak target for the case of targets with different strength. In [6], a Kalman filter-based algorithm, to track a single target in range, has been designed and assessed. The proposed approach exploits spillover of target energy between consecutive matched filter samples.

In this work, we propose and assess an algorithm to track multiple targets in range and azimuth. It is the natural, though not trivial, extension of the Kalman filter-based algorithm proposed in [6] to multitarget scenarios and to range-azimuth tracking. The main novelty of the proposed algorithm is the potential to deal with multiple targets, even in spatial proximity, possibly resorting to the multitarget approach proposed in [5]. To this end, the multiple target detector, as well as the single target one, has been modified to implement the Rake idea in both range and azimuth (and to estimate the unknown Doppler frequency shift of each target).

The paper is organized as follows. Section II is devoted to the problem formulation while Section III focuses on the derivation of the newly-proposed tracker. Section IV deals with the performance assessment.

II. PROBLEM FORMULATION

A mechanically rotating reflector is assumed as radar antenna. It is mounted at a certain height \( d \) above the sea level and illuminates the surveillance area transmitting a bandpass modulated signal whose lowpass equivalent is a train of rectangular pulses of duration \( T_p \). We neglect the dependence of the antenna radiation pattern on the elevation angle within the surveillance area. The rate at which observations are made in azimuth depends on the time (in seconds) for the antenna to make one rotation, namely \( \frac{2\pi}{\omega_a} \) where \( \omega_a \) denotes the angular velocity of the antenna, in radians per seconds (rad/s), which we assume to be positive in case of counterclockwise rotation. Notice also that the number of consecutive pulses on a stationary target, \( N \) say, is (approximately) equal to

\[
N = \frac{\pi}{180} \frac{\Delta \theta_3}{|\omega_a| T_p}
\]

where \( \Delta \theta_3 \) is the (-3 dB) beamwidth (in degrees), namely the beamwidth measured between the half-power points (in
azimuth) of the antenna radiation pattern, and \( T \) is the pulse repetition time\(^{1}\) (PRT). In practice, the pulses of each train are subject to an amplitude modulation process (that at first we neglect in order not to burden too much the introduction of the main ideas involved in the derivation of the system). For instance, assuming the radar parameters of Table 1, it turns out that \( N = 6 \).

Now consider a Cartesian coordinate frame with the \( z \) axis coincident with the rotation axis of the antenna and the \( x - y \) plane locally coincident with the sea surface. Moreover, suppose that the (azimuthal) pointing direction of the antenna at time instant \( t \) is

\[
\theta_a(t) = \theta_{0a} + \omega_a t
\]

with \( \theta_{0a} \in (0, 2\pi] \) the angular position of the antenna at \( t = 0 \). In addition, assume a prospective (point-like and slowly-fluctuating) target moving within the (non ambiguous) surveillance area (and in the antenna far field) with approximately constant radial velocity \( v \) (\( v > 0 \) if the target is approaching the radar) and approximately constant azimuth \( \theta(t) \) over the coherent processing interval; finally, denote by \( r(t) \) the range of the target at time instant \( t \).

Typically, a discrete form for the signal corresponding to each range-azimuth cell is obtained by properly sampling the output of a filter matched to the transmitted pulse (and fed by the received signal). To be more quantitative, the signal transmitted by the radar antenna is given by

\[
\Re \left\{ A e^{j \varphi} \sum_{n=-\infty}^{+\infty} p(t - nT) e^{j 2\pi f_c t} \right\}
\]

where
- \( \Re \{ z \} \) indicates the real part of the complex number \( z \);
- \( A > 0 \) is an amplitude factor related to the transmitted power;
- \( \varphi \in [0, 2\pi] \) is the initial phase of the carrier signal;
- \( p(t) \) is a unit-energy rectangular pulse waveform of duration \( T_p \) (and one-sided bandwidth \( W_p \approx 1/T_p \)), namely
  \[
p(t) = \begin{cases} \frac{1}{\sqrt{2} T_p}, & 0 < t < T_p \\ 0, & \text{elsewhere} \end{cases}
\]
- \( f_c = c/\lambda \) is the carrier frequency with \( c \) the velocity of propagation in the medium and \( \lambda \) the carrier wavelength.

\(^{1}\)Remember also that \( T \) is the inverse of the pulse repetition frequency (PRF).

Neglecting compression or stretching of the time scale, the complex envelope of the received signal, containing \( N \) consecutive useful echoes backscattered by the target (approximately corresponding to the antenna main beam), is given by [7]

\[
z_r(t) = \alpha e^{j 3\pi f_d t} \sum_{n=n_0(h)}^{n_0(h)+N-1} p(t - nT - \tau) + w_r(t), \quad t \in T_o
\]

where
- \( T_o = (n_0(h)T, (n_0(h) + N)T) \);
- \( \alpha \in \mathbb{C} \) is a factor that accounts for \( Ae^{j \varphi} \), the effects of the transmitting (and receiving) antenna gain (remember that, for the time being, we neglect the amplitude modulation process), the two-way path loss, the radar cross section of the target, etc.;
- \( f_d \ll W_p \) is the Doppler frequency shift of the signal backscattered by the target (i.e., \( f_d = \frac{\Delta f}{T_p} \);
- \( n_0(h) \) denotes the first pulse of the \( h \)th train of \( N \) pulses that “illuminates” the target; \( n_0(h) \) is approximately given by (we assume that \( N \) is even)

\[
n_0(h) = \text{round} \left( \frac{\theta(t) + 2\pi \text{sgn}(\omega_a) - \theta_{0a}}{\omega_a T} \right) - \frac{N}{2} \quad (2)
\]

with \( h \) any integer number such that

\[
\theta(t) = \theta_{0a} + \omega_a t - 2\pi \text{sgn}(\omega_a)
\]

\( \text{sgn}(x) \) the signum of \( x \) and \( \text{round}(x) \) the closest integer to \( x \);
- \( \tau \) is the round-trip delay of the received signal, namely
  \[
  \tau = 2r(t_m)/c, \quad t_m \in T_o;
  \]
- \( w_r(t) \) is the complex envelope of the overall disturbance.

The discrete form for the received signal is obtained by sampling the output of a filter matched to \( p(t) \) and fed by \( z_r(t) \). As a matter of fact, the matched filter output is given by [7]

\[
z(t) = \alpha \sum_{n=n_0(h)}^{n_0(h)+N-1} \chi_p(t - nT - \tau, f_d) e^{j 2\pi n\nu} + w(t), \quad t \in T_o
\]

where \( \alpha \) is a modified version of the previous constant, \( \chi_p(\cdot, \cdot) \) is the (complex) ambiguity function of the pulse waveform \( p(t) \) [7], \( w(t) \) is the noise component, and \( \nu = f_d T \) is the normalized target Doppler frequency.

In order to generate the vector of the noisy returns, it is customary to assume that the target is located exactly “where the matched filter is sampled”; in other words, the matched filter output \( z(t) \) is sampled to obtain returns corresponding to a certain number of range-azimuth cells. For the sake of clarity, assume that \( 2\pi/|\omega_a| \) is a multiple integer of \( T \) and define the \( s \)th azimuth cell as the one corresponding to the values of \( \theta \) that belong to the set \( \Theta_s = \{ \theta : \theta_{0a} + s\Delta \theta_3 + s\Delta \theta_3/2 < \theta \leq \theta_{0a} + s\Delta \theta_3 + s\Delta \theta_3/2 \} \). It follows that samples corresponding to the \( (l+1) \)th range gate, \( l = 1, \ldots, L \), and to the \( s \)th azimuth cell \( s = 1, \ldots, H \), are those at time instants

\[
t_{l,n} = t_{\text{min}} + (l - 1)T_p + nT, \quad n = n_{\theta_a}(h), \ldots, n_{\theta_a}(h)+N-1
\]
with \( \theta_s = \theta_{0s} + s\Delta\theta_s \) and \( n_{0s}(h) \) obtained by equation (2) with \( \theta_p \) in place of \( \theta(t) \); in addition, the time samples can be grouped to form an \( N \)-dimensional vector as follows

\[
z_l(n_0) = \left[ z(t_{l,n_0}(h)) \cdots z(t_{l,n_0}+N-1) \right]' = s_l + n_l
\]

where \( s_l \) is the signal component, \( n_l \) the noise component, and \( ' \) denotes transpose. As to \( t_{\text{min}} \), it is the beginning of the sampling process. Remember also that \( T_p \) is the pulse duration and \( T \) is the PRT.

However, assuming that the round-trip delay \( \tau \) of the received signal is different from \((l-1)T_p\) and such that

\[
\tau \in [t_{\text{min}}, t_{\text{min}} + (L-1)T_p]
\]

it follows that

\[
\exists l_0 \in \{1, \ldots, L-1\} : \tau = t_{\text{min}} + (l_0-1)T_p + \epsilon, \quad 0 \leq \epsilon \leq T_p
\]

where \( \epsilon \) (or \( T_p - \epsilon \), if \( \epsilon > T_p/2 \)) represents a residual delay that gives rise to the spillover of target energy. In other words, the target contributes to the \( l_0 \)th and to the \((l_0 + 1)\)th range cells. In addition, the actual samples that contain the signal backscattered by a target with azimuthal position \( \theta \), when it is illuminated through the (-3 dB) antenna beamwidth, are those corresponding to the vectors

\[
z_l(n_0) = \left[ z(t_{l,n_0}) \cdots z(t_{l,n_0}+N-1) \right]' = s_l + n_l + 1
\]

with \( n_0 \) obtained by equation (2) with \( \theta \) in place of \( \theta_l(t) \) (we neglect here the dependence of \( n_0 \) on \( h \)). Thus, it follows that

\[
s_l = \begin{cases} 
\alpha \chi_p(-\epsilon, \nu/T) v(\nu), & l = l_0 \\
\alpha \chi_p(T_p - \epsilon, \nu/T) v(\nu), & l = l_0 + 1 \\
0, & l \neq l_0, l_0 + 1 \end{cases} 
\]

where

\[
v(\nu) = \frac{1}{\sqrt{N}}[e^{j2\pi\nu} \cdots e^{j2\pi(N-1)\nu}]'
\]

is the temporal steering vector. Equation (3) highlights the relationship between the amplitude of the target in the range cell where it is located and that in the adjacent range cell affected by spillover.

As a final remark, notice that generalization to the case that several targets are present, within a given set of range-azimuth cells, is possible by using the superposition principle.

### III. DESIGN OF THE DETECTION AND TRACKING STAGES

In this section, we illustrate how measurements, to be fed to the tracker, are obtained and processed. We assume a multitarget scenario. However, even in a multitarget scenario, it seems reasonable to attack detection and localization (but also the tracking stage) resorting to either single target or multitarget strategies depending on whether targets are well separated (in range and/or in azimuth) or there are two or more closely spaced targets, respectively. The actual scenario being in force impacts the design approach:

- for “non-interfering targets” both detector and tracker process data resorting to the single target approach, namely using algorithms designed under the assumption that a single target is present;
- in case of closely spaced targets, instead, it seems reasonable to take into account, at both the detection and the tracking stage, the presence of mutual interference; for this reason, we resort to a decision scheme designed to address detection and localization of multiple targets (the number of targets is assumed known at the detection stage) and to a tracker that has to handle the data association problem, a point better specified in Section III-B.

The way targets are clustered (and eventually the type of detector used to process each cluster) depends on the available information on each target location (and possibly on how much such information is reliable). To this end, we can use the predicted positions of the targets returned by the tracker. The actual clustering criterion is described in Section III-B.

#### A. Detection stage

1) Single Target Detector: first assume that \( \theta \) and \( \nu \) (and hence \( \nu \)) are known parameters and that we are looking for the presence of an “isolated target” in the \( l \)th range cell. Also remember that \( z_l(n_0) \in \mathbb{C}^{N \times 1}, l = 1, \ldots, L, \) is the vector of returns from the \( l \)th range cell, obtained by stacking the \( N \) samples of the output of the matched filter at the time instants

\[
t_{l,n} = t_{\text{min}} + (l-1)T_p + nT, \quad n = n_0, \ldots, n_0 + N - 1
\]

We have to choose between the \( H_0 \) hypothesis that \( z_l \) contains disturbance only and the \( H_1 \) hypothesis that it also contains a useful target signal, proportional to the (temporal) steering vector \( v(\nu) \).

Now, under the \( H_1 \) hypothesis, there is spillover of target energy to either \( z_{l-1} \) or \( z_{l+1} \) depending on the value of the target delay

\[
t_{l,\epsilon} = t_{\text{min}} + (l-1)T_p + \epsilon
\]

with \(-T_p/2 \leq \epsilon \leq T_p/2\). It seems reasonable to use the above additional vectors, in addition to \( z_l \), to discriminate between \( H_0 \) and \( H_1 \) and to localize the possible target [4]. Finally, we assume that the corresponding noise vectors \( n_i, i = l-1, l, l+1 \), are \( (N\text{-dimensional}) \) independent and identically distributed (iid) complex normal random vectors [4]; in particular, \( n_i \sim \mathcal{CN}(0, \sigma^2 I_N) \), with \( I_N \) the \( N \times N \) identity matrix, and \( \sigma^2 > 0 \) an unknown deterministic quantity.

Following the lead of [4], the hypothesis test is

\[
H_0: z_i(n_0) = n_i, \quad i = l-1, l, l+1
\]

\[
\begin{cases} 
z_{l-1}(n_0) = \alpha \chi_p (-T_p - \epsilon, \frac{\nu}{T}) v_{l-1}(\nu, \theta) + n_{l-1} \\
z_l(n_0) = \alpha \chi_p (-\epsilon, \frac{\nu}{T}) v_l(\nu, \theta) + n_l \\
z_{l+1}(n_0) = n_{l+1} \\
-T_p/2 \leq \epsilon \leq 0
\end{cases}
\]

\[
H_1: \begin{cases} 
z_{l-1}(n_0) = n_{l-1} \\
z_l(n_0) = \alpha \chi_p (-\epsilon, \frac{\nu}{T}) v_l(\nu, \theta) + n_l \\
z_{l+1}(n_0) = \alpha \chi_p (T_p - \epsilon, \frac{\nu}{T}) v_{l+1}(\nu, \theta) + n_{l+1} \\
0 < \epsilon \leq +T_p/2
\end{cases}
\]
where, taking into account the amplitude modulation process due to the antenna gain, we have that
\[
v_l(\nu, \theta) = \frac{1}{\sqrt{N}} [c_{l,\theta}(n_\theta) c_{l,\theta}(n_\theta + 1)e^{2\pi \nu v} \cdots c_{l,\theta}(n_\theta + N - 1)e^{2\pi (N-1)\nu v}]'
\]
with
\[
c_{l,\theta}(n) = \sqrt{G(\theta_a(nT) - \theta)G(\theta_a(t_l,n) - \theta)}
\]
and \(G(\cdot)\) the antenna gain; more explicitly, \(\theta_a(nT)\) is the antenna pointing direction at the time instant \(n\) th pulse is transmitted and \(\theta_a(t_l,n)\) is the antenna pointing direction at the time instant the signal backscattered from the target is sampled (given by equation (4)); as to \(\theta_a(t)\), it is given by equation (1). Obviously, the previous expression for the \(c_{l,\theta}(n)\)'s can be approximated as
\[
c_{l,\theta}(n) \approx G(\theta_a(nT) - \theta) \approx G(\theta_a(nT) - \theta') \equiv c_0(n)
\]
by neglecting dependence on the rotation of the antenna over the time interval \(t_l,n-nT = t_{\min} + (l-1)T_p\), and replacing \(\theta\) with \(\theta'\) given by
\[
\theta' = \left(n_\theta + \frac{N}{2}\right) \omega_a T + \theta_{0a} - 2h\pi \text{ sgn}(\omega_a)
\]
(6)
It turns out that
\[
v_l(\nu, \theta) \approx v(\nu, \theta) = \frac{1}{\sqrt{N}} [c_0(n_\theta) c_0(n_\theta + 1)e^{2\pi \nu v} \cdots c_0(n_\theta + N - 1)e^{2\pi (N-1)\nu v}]'
\]
(7)
Solving the above problem will lead to an adaptive decision scheme that takes advantage of the possible spillover of target energy and that could return more accurate measurements (to be fed to the tracking stage) than a conventional detector. In order to formulate and eventually solve the above problem we define an augmented received vector
\[
\hat{z}(n_\theta) = [z_{l-1}(n_\theta) \ z_l(n_\theta) \ z_{l+1}(n_\theta)]' \in \mathbb{C}^{3N \times 1}
\]
Accordingly, we also define an augmented steering vector, \(\hat{\upsilon}(\nu, \nu, \nu)\) say, and the hypothesis test becomes
\[
\begin{align*}
H_0 &: \hat{z}(n_\theta) = \hat{n} \\
H_1 &: \hat{z}(n_\theta) = \alpha \hat{\upsilon}(\nu, \nu, \nu) + \hat{n}
\end{align*}
\]
Obviously, \(\hat{n} \sim \mathcal{CN}_{3N}(0, \sigma^2 I_{3N})\).

It is not difficult to derive the corresponding GLRT; to this end, it is important to observe that the unknown parameters (leaving aside, for the moment, \(\nu, \nu\) and \(\nu\)) are \(\sigma^2, \alpha, \nu\) and \(\epsilon\) under \(H_1\) and \(\sigma^2\) under \(H_0\). Moreover, maximization of the probability density function (pdf) of the observables under the \(H_1\) hypothesis is performed in closed form with respect to \(\sigma^2\) and \(\alpha\), given \(\epsilon\). It is not difficult to show that the corresponding decision rule is given by [6]
\[
\max_{\epsilon} \frac{||\hat{z}(n_\theta)\hat{\upsilon}(\nu, \nu, \nu)||^2}{\hat{z}(n_\theta)\hat{\upsilon}(\nu, \nu, \nu)\hat{\upsilon}(\nu, \nu, \nu)} \stackrel{H_1}{\geq} \frac{H_1}{H_0} \gamma
\]
(8)
where \(\epsilon \in (-T_p/2, T_p/2)\), and \(||\cdot||\) denote conjugate transpose and the modulus of a complex number, respectively, while \(\gamma\) is the detection threshold to be set according to the desired probability of false alarm (\(P_a\)). As to maximization over \(\epsilon\), it cannot be conducted in closed form and, hence, we resort to a grid search. Maximization over \(\epsilon\) is performed assuming that it takes on values in the finite set
\[
\mathcal{E} = \left\{ \frac{n - N_\epsilon - 1/2}{2N_\epsilon} T_p \right\}_{n=1}^{2N_\epsilon}
\]
(9)
with \(N_\epsilon\) a preassigned positive integer.

It is now a good moment to recall that \(\nu\) and \(\nu\) are unknown parameters, and, hence, that we also have to face with such a priori uncertainty. As to \(\nu\), we assume an uncertainty sector approximately given by \(\Theta = \{\theta : \theta - \Delta\theta_3 < \theta < \theta + \Delta\theta_3\}\) where \(\theta\) is the predicted azimuthal position of the target returned by the tracker. More precisely, we assume that \(n_\theta \in \mathcal{N} = \{n_{\pi - N}, \ldots, n_{\pi + N}\}\). A straightforward approach would be to maximize the above statistic, given by equation (8), also with respect to \(n_\theta \in \mathcal{N}\) and \(\nu \in \{0, 1/M, \ldots, (M - 1)/M\}\) with \(M\) an integer number greater than or equal to \(N\). Notice also that, strictly speaking, the resulting detector is no longer a GLRT because it uses different data for each value of \(n_\theta\). More important, we have now to implement a three-dimensional search. In order to avoid such a computationally-intensive optimization, we propose the following heuristic procedure:

- first, estimate \(\nu \in \{0, 1/M, \ldots, (M - 1)/M\}\), given \(n_\theta \in \mathcal{N}\), resorting to the ML estimator of \(\nu\) (under the \(H_1\) hypothesis) for the case of a target located at the center of the \(l\) th range cell and known \(\theta\), \(\hat{\upsilon}(n_\theta)\) say; it is given by
\[
\hat{\upsilon}(n_\theta) = \arg \max_{\nu} ||z_l'(n_\theta)\upsilon(\nu, \nu, \nu)||^2
\]

- then, resort to the following decision rule
\[
\max_{n_\theta \in \mathcal{N}, \nu \in \mathcal{E}} \frac{||\hat{z}(n_\theta)\hat{\upsilon}(\nu, \nu, \nu)||^2}{||\hat{z}(n_\theta)\hat{\upsilon}(\nu, \nu, \nu)||^2} \stackrel{H_1}{\geq} \frac{H_1}{H_0} \gamma
\]
(10)
with \(||| \cdot |||\) the Euclidean norm of a vector. Notice that this decision statistic is obtained by that for known \(\nu\) and \(\nu\) by maximizing with respect to \(n_\theta\) after replacing \(\nu\) with \(\hat{\upsilon}(n_\theta)\).

The output of the detection stage is a decision on the hypothesis actually in force or, equivalently, a value of the binary variable \(D\) that takes on values zero and one (\(D = 1\) is tantamount to choosing \(H_1, i = 0, 1\)); in addition, if \(D = 1\), the detector also returns the estimated value of \(\epsilon\) and \(\theta\), \(\hat{\epsilon}\) and \(\hat{\theta}\) say, and of \(\nu\), \(\hat{\upsilon}\) (that is not forwarded to the tracking stage). The tracker will use \(\hat{\epsilon}\) and \(\hat{\theta}\) to compute the measurement.

It is worth stressing that the above detection strategy can be extended to attack detection of a single target within a larger area, namely considering a certain number of adjacent range-azimuth cells. Such a generalized approach might be effective when the information on the range cell \(l\) the target belongs to and/or the uncertainty sector we assume (i.e., the
value of $\overline{\theta}$ is the result of a not too reliable prediction on the actual position of a (previously detected) target; in fact, such a prediction might be affected by an estimation error, due, for instance, to a target that has initiated a maneuver or to an adverse scenario in terms of signal-to-noise ratio value. In particular, we will also consider detector (10) with $E'$

$$E' = \left\{ \frac{n - N_e - 1/2}{2N_e} T_p \right\}^{3N_e}_{n = -N_e + 1}$$

As a final comment, observe that the GLRT and its ad hoc implementations guarantee the constant false alarm rate (CFAR) property with respect to $\sigma^2$.

2) Multiple Target Detector: suppose that two targets are present at most. The generalization to an arbitrary, albeit known, number of targets is straightforward. Moreover, assume that the two closely spaced (point-like and slowly-fluctuating) targets are moving with constant radial velocity (and in the antenna far field) within three consecutive range cells, indexed by $l - 1, l, l + 1$; $v_l$ is the radial velocity of the $i$th target ($v_l > 0$ if the $i$th target is approaching the radar). We denote by $r_i(t)$ and $\theta_i(t)$ range and azimuth of the $i$th target, respectively, and suppose that the azimuth is approximately constant, $\theta_i$, say, over the coherent processing interval. We leave aside, for the moment, the fact that the $v_l$s (and, hence, the normalized Doppler frequency shifts $v_l = \frac{2\nu_l f_c}{t}$) are unknown parameters. As to the $\theta_i$s, they are unknown parameters too; however, we suppose that

$$\theta_m = \frac{\theta_{\min} + \theta_{\max}}{2}$$

with $\theta_{\min} = \min\theta_i$ and $\theta_{\max} = \max\theta_i$, is approximately known. Due to the presence of spillover there might be target energy also in the matched filter outputs corresponding to the $(l - 2)$th and the $(l + 2)$th range cells. Then, it is reasonable to process samples at the output of the matched filter at the instants $t_{n,l} = t_{\min} + (j - 1) T_p + n T, j = l - 2, \ldots, l + 2, \quad n = n_{\theta_m} - 3N/2, \ldots, n_{\theta_m} + 5N/2 - 1$, (again, with $n_{\theta_m}$ obtained by equation (2) with $\theta_m$ in place of $\theta(t)$) in order to collect most of the energy backscattered by the targets (the actual pulses to be used depend on the actual targets’ azimuthal location).

Thus, for the case at hand, time samples are grouped to form $4N$-dimensional vectors as follows, $j = l - 2, \ldots, l + 2$

$$z_j = [z(t, n, o_m - 3N/2) \ldots z(t, n, o_m + 5N/2 - 1)]' = s_j + n_j$$

with $s_j$ the signal component and $n_j$ the noise component. It is not difficult to compute, following the lead of [4], [5], the contribution, $s_{i,j} \in \mathbb{C}^{4N \times 1}$ say, of the $i$th target to the $j$th range cell, $j = l - 2, \ldots, l + 2$, up to a complex factor $\alpha_i$ taking into account target and channel effects. Obviously, $s_{i,j}$ depends on the range cell the target belongs to and on $\epsilon_i$ that denotes the “residual delay” from the center of the resolution cell (the $i$th target belongs to). Moreover, it is proportional (up to constants depending on the target position) to a spatio-temporal steering vector, $\nu_i = \nu_j$ say; assuming that

$$\theta_i \in \Theta_d = \left\{ \theta_m' + \frac{\Delta \theta_i}{N} \right\}^{2N}_{r = -\frac{\Delta \theta_i}{N}}$$

where $\theta_m'$ is given by equation (6), with $n_{\theta_m}$ in place of $n_\theta$, the spatio-temporal steering vector can be expressed as

$$\nu_i = \frac{1}{\sqrt{N}} [0 \ldots 0 \ c_{\theta_i}(n_{\theta_m}) c_{\theta_i}(n_{\theta_m} + 1) e^{2\pi i v_i} \cdots c_{\theta_i}(n_{\theta_m} + N - 1) e^{2\pi i (N - 1) v_i} \ 0 \ldots 0]'$$

It follows that the signal component $s_j = \nu_i, j = l - 2, \ldots, l + 2$, is given by

$$s_j(j_1, j_2, \epsilon_1, \epsilon_2, \nu_1, \nu_2, \theta_1, \theta_2) = \sum_{i=1}^{2} \alpha_i s_{i,j}(j_1, \epsilon_1, \nu_1, \theta_i)$$

where $j_i \in \{l - 1, l, l + 1\}$ denotes the range cell the $i$th target belongs to. We also assume that signal returns are buried in zero mean, complex normal noise with covariance matrix $\sigma^2 I_{20N}$ independent from range cell to range cell.

Summarizing, the problem at hand is that of detecting the possible presence of two targets and of estimating the parameters of interest, namely $j_1, j_2, \epsilon_1, \epsilon_2, \theta_1$ and $\theta_2$, together with the nuisance parameters $\alpha_1, \alpha_2, \sigma^2$. In order to formulate and eventually solve the estimation problem we can now define an augmented received vector

$$\tilde{z} = [z_{l-2}' \ z_{l-1}' \ z_{l}' \ z_{l+1}' \ z_{l+2}']' \in \mathbb{C}^{20N \times 1}$$

Accordingly, we can define an augmented steering vector, corresponding to the $i$th target, $h_i(j_1, \epsilon_1, \nu_1, \theta_i) \in \mathbb{C}^{20N \times 1}$ say. In addition, the hypothesis test to be solved is

$$\begin{cases}
H_0 : \tilde{z} = \tilde{n} \\
H_1 : \tilde{z} = \sum_{i=1}^{2} \alpha_i h_i(j_1, \epsilon_1, \nu_1, \theta_i) + \tilde{n} = H(x_1, x_2) x_3 + \tilde{n}
\end{cases}$$

where

$$H(x_1, x_2) = [h_1(j_1, \epsilon_1, \nu_1, \theta_1) \ h_2(j_2, \epsilon_2, \nu_2, \theta_2)]$$

is assumed to be a full-column-rank matrix, with $x_1 = [j_1 \ j_2 \ \epsilon_1 \ \epsilon_2 \ \theta_1 \ \theta_2]'$, $x_2 = [\nu_1 \ \nu_2]'$, and $x_3 = [\alpha_1 \ \alpha_2]'$. Obviously, $\tilde{n} \sim \mathcal{CN}_{20N}(0, \sigma^2 I_{20N})$ with $\sigma^2$ an unknown parameter.

Thus, leaving aside for the moment the unknown parameters $\nu_1, \nu_2$, the above hypothesis test can be solved resorting to the GLRT, which is given by

$$\begin{align}
& \frac{\max_{x_1, x_2, \sigma^2} f\left(\tilde{z}; x_1, x_2, x_3, \sigma^2\right)}{\max_{\sigma^2} f_0\left(\tilde{z}; \sigma^2\right)} \\
& > \gamma \\
& \text{H}_0
\end{align}$$

where $f_i$ denotes the pdf of the observables under the $H_i$ hypothesis, $i = 1, 2$, and $\gamma$ is the threshold to be set according
to the desired $P_{fa}$. It is not difficult to come up with the following form for the GLRT

$$\max_{\hat{x}_1} \frac{\hat{z}^T P_H(x_1, \pi_2) \hat{z}}{\hat{z}^T \hat{z}} \geq \frac{H_1}{H_0} \gamma$$

(13)

where $P_H(x_1, \pi_2)$ is the projection matrix onto the space spanned by the columns of the matrix $H(x_1, \pi_2)$ and $\gamma$ denotes a modification of the original threshold.

It still remains to overcome the uncertainty on $\nu_1$ and $\nu_2$. To this end, we resort to a suboptimum procedure. In fact, neglecting the actual expression of the spillover term to adjacent range cells, we can model the received signal, corresponding to the $j$th range cell, as

$$z_j = [v(\nu_1, \theta_1) v(\nu_2, \theta_2)] u_j + n_j \in \mathbb{C}^{4N \times 1}$$

with $\theta_i \in \Theta_d$ and $v$ given by equations (11) and (12), respectively, $u_j = [\alpha_{ij} \alpha_{ij}]'$, and $n_j \sim \mathcal{CN}(0, \sigma^2 I_{4N})$. It turns out that the ML estimate of $x_2, \hat{x}_2 = [\hat{\nu}_1, \hat{\nu}_2]^T$ say, is given by

$$\hat{x}_2 = \arg \max_{\nu_1, \nu_2, \theta_1, \theta_2} \frac{\sum_{j=-2}^{+2} z_j^T P(\nu_1, \nu_2, \theta_1, \theta_2) z_j}{\hat{z}^T \hat{z}}$$

(14)

where $P(\nu_1, \nu_2, \theta_1, \theta_2)$ is the projection matrix onto the space spanned by the columns of the matrix $[v(\nu_1, \theta_1) v(\nu_2, \theta_2)]$, $\nu_1 \neq \nu_2, \nu_1 \in \{0, 1/M, \ldots, (M-1)/M\}, \theta_1 \neq \theta_2, \theta_i \in \Theta_d$. Then, we plug the estimated pair of normalized Doppler frequencies $\hat{x}_2$ into the above detector (13), in place of $x_2$, thus obtaining

$$\max_{\hat{x}_1} \frac{\hat{z}^T P_H(x_1, \hat{x}_2) \hat{z}}{\hat{z}^T \hat{z}} \geq \frac{H_1}{H_0} \gamma$$

(15)

A few final remarks are in order. First, observe that the maximization over $x_1 = [j_1, j_2, \epsilon_1, \epsilon_2, \theta_1, \theta_2]$ cannot be conducted in closed form. We resort to a grid search to maximize with respect to $\epsilon_1$ and $\epsilon_2$. To this end, we will assume that $\epsilon_1$ and $\epsilon_2$ take on values in the set given by equation (9). As to the possible values assumed by the $\theta_i$s, again we refer to the set obtainable using equation (11) and use the corresponding steering vector given by equation (12). Second, the GLRT and its ad hoc implementation are CFAR with respect to $\sigma^2$. Finally, notice that the generalization of the procedure to an arbitrary (but known) number of targets is in principle straightforward. However, the computational complexity of the grid search would increase exponentially with the number of targets.

B. Tracking stage

Assume that radar data are available over $[0, +\infty)$. Moreover, assume that the motion of each target can be modeled according to a nearly constant velocity model along $x$ and $y$. More precisely, we suppose that the kinematics of the $i$th target is ruled by the following state space model

$$\begin{cases}
\nu_i(k+1) = F \nu_i(k) + w_{1i}(k) \\
y_i(k) = H \nu_i(k) + w_{2i}(k)
\end{cases}$$

where $k$ denotes the $k$th antenna scan and $\nu_i(k) = [x_i(k) v_{x_i}(k) y_i(k) v_{y_i}(k)]'$ with $x_i(k)$ and $y_i(k)$ the cartesian coordinates of the target over the time interval $[t_{k,i}, t_{k+1,i} + NT]$, $v_{x_i}(k)$ and $v_{y_i}(k)$ the target velocity components along $x$ and $y$ over the same time interval; similarly, $x_i(k+1) = [x_i(k+1) v_{x_i}(k+1) y_i(k+1) v_{y_i}(k+1)]'$ is the vector of cartesian coordinates and velocity components over $[t_{k+1,i}, t_{k+1,i} + NT]$, with $t_{k+1,i} - t_{k,i}$ approximately equal to $T_a = 2\pi/|\omega_a|$, $y_i(k) = [x_i(k) y_i(k)]'$, $F = \text{diag}(F_1, F_1)$ with

$$F_1 = \begin{bmatrix} T_a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

As to $w_{1i}(k)$ and $w_{2i}(k)$, they are assumed to be (approximately) independent, zero-mean normal sequences (at least under perfect matching between the nominal and the actual target model). Moreover, we have that

$$E[w_{1i}(k) w_{1j}^T(h)] = Q \delta_{ij} \delta_{kh} \quad E[w_{2i}(k) w_{2j}^T(h)] = R \delta_{ij} \delta_{kh}$$

with $Q = \text{diag}(Q_1, Q_1)$ where [1]

$$Q_1 = q \begin{bmatrix} T_a^2 & T_a \\ T_a & T_a \end{bmatrix}$$

and $\delta_{ij}$ the Kronecker symbol. The parameter $q$ can be set based on the maximum tolerable difference between the nominal and the actual target velocity component (along $x$ or $y$), $\Delta v_{\text{max}}$ say, over the “sampling interval” $T_a$. Similarly, $R = \text{diag}(R_1, R_1)$. As to $R_1$, it can be chosen based on an educated guess on the second-order moments of the measurement errors.

In order to describe the main ideas underlying the proposed multitarget tracker, denote by $T_k, k \geq 2$, the set of (positive) integers indexing the tracks that are active over the $k$th antenna scan (i.e., confirmed and to-be-confirmed ones):

1) At the $k$th scan, for the $i$th track, $i \in T_k$, we compute the one step prediction of the state, $\tilde{x}_i(k+1|k)$ say, and the associated error covariance matrix by implementing the Kalman filter equations.

2) Based on the predicted positions of targets over the $(k+1)$th scan (corresponding to active tracks over the $k$th scan) we construct a partition of $T_k, T_k', \ldots, T_k^{m_k}$ say, in order to determine targets that might be in close spatial proximity and, hence, groups of range-azimuth cells to be jointly processed by the detection stage; the partition is constructed according to a criterion described below.

3) If the cardinality of $T_k^{m_k}, j = 1, \ldots, m_k$, is greater than one, the multitarget decision scheme is used to detect

2We assume that during the time interval $[t_{k,i}, t_{k,i} + NT]$ the target is within the (-3 dB) antenna beamwidth. Obviously, the $t_{k,i}$s depend on the azimuthal position of the $i$th target over time. Remember also that $N$ is the number of processed pulses and $T$ the PRT.
and localize the targets whose indices are defined by \( T^j_k \) (remember that, although we have described the detector to be used in case of two targets in close spatial proximity, its generalization to an arbitrary, but known, number of targets is in principle straightforward); otherwise, we resort to the single target detector. At this point we have a set of measurements for the \((k+1)\)th scan, \( M_{k+1} \), say. 

4) \( M_{k+1} \) is fed to a joint probabilistic data association (JPDA) algorithm [1]. 

5) The state of each track is updated; more precisely, a to-be-confirmed track can be confirmed (otherwise, it is deleted) using an \( m_c/n_c \) rule and, similarly, a confirmed track can be deleted using an \( n_d/m_d \) rule. At the end of track maintenance, \( T_{k+1} \) contains the set of (positive) integers indexing the active tracks over the \((k+1)\)th scan. 

6) After existing tracks have been updated, it is possible to run a more conventional detector over remaining range-azimuth cells belonging to the surveillance region. Range-azimuth cells “occupied by targets” associated with active tracks, but also cells surrounding the above estimated target positions, are not scanned by the conventional detector; notice that it is necessary to set guard cells to avoid that the spillover due to (already detected) strong targets might originate additional false targets. Thus, if \( \hat{l} \) and \( \hat{\theta} \) are the estimated range cell and azimuthal position of the target corresponding to an active track, respectively, the conventional detector does not scan range cells \( \hat{l}, \hat{l}−1, \text{ and } \hat{l}+1 \) together with pulses \( n_\hat{\theta}−N/2, \ldots, n_\hat{\theta}+3N/2−1 \). The more conventional detector is the cascade of two detection stages: the observables corresponding to each range-azimuth cell within the area of interest are fed to a detector that assumes the presence of a target “matched to” the \( l \)th range cell and the \( h \)th azimuth cell, namely at range \( r_l = \frac{1}{2}(t_{min}+(l−1)T_p) \) with azimuth \( \theta_h = \theta_{0a} + h\Delta\theta_h \); if a target is detected, setting the threshold to guarantee \( P_{fa} = 10^{-2} \), the observables are fed to detector (10). 

Again, the threshold is set to guarantee \( P_{fa} = 10^{-2} \). Actually, the area of interest is scanned (in the first stage) by shifting \( N/3 \) pulses in azimuth the data window. Such an approach guarantees a better match to the design assumptions of the first stage detector than a non-overlapping search mode. 

7) The tracker initiates potential tracks over the \((k+1)\)th scan in case of detections over the \( k \)th and the \((k+1)\)th scans corresponding to range-azimuth cells in spatial proximity; more precisely, the presence of a detection at the \( k \)th scan and a detection at the \((k+1)\)th scan, whose distance is less than \( v_{max}T_a \), with \( v_{max} \) denoting the maximum admissible target velocity, can be interpreted as originated by the same target and gives rise to a new (to be confirmed) track. More generally, in presence of multiple detections over the \( k \)th scan and/or the \((k+1)\)th scan, we compute all distances between detections over consecutive scans and, among those compatible with the maximum admissible target velocity, choose the smallest pairs (with distinct first and second detection) to initiate new tracks. The same rationale is used to initiate tracks over the second scan based on possible detections over the first two scans that come from the detector described in the previous item. 

It still remains to specify how to partition \( T_k \). To this end, denote by \( d(x, y) \) the Euclidean distance between \( x \) and \( y \) and by \( d_0 > 0 \) a given threshold. In particular, we investigate the following partition rule that, leading to sets consisting of one or two elements, keeps the computational complexity of the detection stage at a low level. Assume that the cardinality of \( T_k \) is greater than one. 

1) Set \( T' = T_k \) and \( j = 1 \). 

2) Construct the set of distances, \( D \), say, between any two elements \( \hat{x}_{r_1}(k+1|k) \) and \( \hat{x}_{r_2}(k+1|k) \), \( r_1, s \in T' \), that satisfy the following additional conditions \( |r_i − l_s| ≤ 2 \), \( |n_{\theta_i} − n_{\theta_s}| ≤ 3N \) where \( l_i \) and \( \theta_i \) denote the range cell and the azimuthal position of the target corresponding to the \( i \)th track in \( T' \). 

3) Compute the minimum among such distances, i.e., \( d_j = \min D \) (remember that the minimum is equal to \( +\infty \) for an empty set); if such a minimum is less than \( d_0 \), let \( T^j_k = \{ \hat{x}_{r_1}(k+1|k), \hat{x}_{r_2}(k+1|k) \} \), where \( r_1 \) and \( r_2 \) are such that \( d(\hat{x}_{r_1}(k+1|k), \hat{x}_{r_2}(k+1|k)) = d_j \); otherwise construct a singleton for any \( \hat{x}_{r_i}(k+1|k), r \in T' \), and quit the loop. 

4) Compute the new \( T' \) as 

\[ T' = T' \setminus \{ \hat{x}_{r_1}(k+1|k), \hat{x}_{r_2}(k+1|k) \} \] 

if \( T' \) is an empty set quit the loop; if, instead, the cardinality of \( T' \) is one, let 

\[ T^{j+1}_k = \{ \hat{x}_{r_i}(k+1|k), r \in T' \} \] 

and quit the loop; otherwise, let \( j = j + 1 \) and return to step 2. 

IV. PERFORMANCE ASSESSMENT AND CONCLUSIONS

In this section, we assess the performance of the tracker by simulation techniques. In order to evaluate the thresholds necessary to ensure \( P_{fa} = 10^{-2} \), we resort to \( 10/P_{fa} \) independent trials. All the illustrative examples assume the parameters of Table 1, \( N_c = 5, q = 1 \text{ m}^2/\text{s}^3, R_i = 5 \text{ m}^2, m_c = 5, n_c = 8, m_d = 1, n_d = 8, \) and \( d_0 = 15 \text{ m} \). The trajectories of the targets are depicted in Figure 1 and assume a constant acceleration model; markers indicate the position of each target at scans multiple of 5. Notice the crossing of three targets around scan 50 and concerning the two targets moving along parallel trajectories the overtake of the left one on the right one in between scans 45 and 55. Moreover, we define the signal-to-noise ratio (SNR) for the \( i \)th target as \( SNR_i = |\alpha_i|^2/\sigma^2 \) with \( \sigma^2 = 1 \). Figures 2 and 3 refer to the case of targets with the same strength while in
Fig. 1. Simulated trajectories.

Fig. 2. Tracking results assuming targets with equal strength. SNR=20 dB. Actual trajectories in blue, estimated ones in green for the proposed tracker and in red for the competitor.

Fig. 3. Tracking results assuming targets with equal strength. SNR=20 dB. Actual trajectories in blue, estimated ones in green for the proposed tracker and in red for the competitor.

Fig. 4. Tracking results assuming two strong targets (SNR=30 dB) and a weaker one (SNR=25 dB). Actual trajectories in blue, estimated ones in green for the proposed tracker and in red for the competitor.

Fig. 5. Tracking results assuming a strong target (SNR=30 dB) and a weaker one (SNR=25 dB). Actual trajectories in blue, estimated ones in green for the proposed tracker and in red for the competitor.

Figures 4 and 5 one of the three targets whose trajectories cross each other is weaker than the others (the one moving from top-right to bottom-left); similarly, one of the two targets moving along parallel trajectories is weaker than the other (the one on the right). For comparison purposes the performance of a tracker that employs the single target detector even for targets in spatial proximity is considered too. Figures 2-5 show the superiority of the proposed tracker with respect to the competitor: the former is able to estimate the tracks even when targets are in close spatial proximity while the latter may lose one of them.

Ongoing research activity is aimed to extend the proposed approach to bistatic scenarios.

REFERENCES


