A Formulation of Multitarget Tracking as an Incomplete Data Problem

H. GAUVRIT

J. P LE CADRE
IRISA/CNRS
France

C. JAUFFRET
DCN/Ingénierie/Sud
France

Traditional multihypothesis tracking methods rely upon an enumeration of all the assignments of measurements to tracks. Pruning and gating are used to retain only the most likely hypotheses in order to drastically limit the set of feasible associations. The main risk is to eliminate correct measurement sequences.

The Probabilistic Multiple Hypothesis Tracking (PMHT) method has been developed by Streit and Luginbuhl in order to reduce the drawbacks of "strong" assignments. The PMHT method is presented here in a general mixture densities perspective. The Expectation-Maximization (EM) algorithm is the basic ingredient for estimating mixture parameters. This approach is then extended and applied to multitarget tracking for nonlinear measurement models in the passive sonar perspective.

I. INTRODUCTION

Multitarget tracking is concerned with states estimation of an unknown number of targets, in a surveillance region. Available measurements may have originated from the clutter or the targets of interest if detected. Due to this uncertainty, multitarget tracking does not correspond to a traditional estimation problem. Two different problems have to be solved jointly: data-association and estimation. Extension of standard estimation algorithms as Nearest Neighbor methods generally provide poor results in dense clutter environment. In these algorithms, the prediction is based upon the measurements which lie in a suitable neighborhood of the predicted mesurements. So, these methods are essentially focused on estimation. On the contrary, the probabilistic data association filter (PDAF) and its extension to multiple targets, the joint PDAF (JPDAF), rely on an assignment step through the probability that a measurement originates from a target. So, it is worth stressing that the fundamental difficulty in multitarget tracking lies in data-association.

Since the mid-sixties, this subject has attracted much attention especially because performance of the algorithms used to solve the problem conditions greatly the overall performance of the SONAR or RADAR system. A natural framework consists to face the combinatorial complexity of the data-association problem. Morefield [10] first, and more recently with the arrival of increasingly powerful computers, Pattipati, et al. [13], Castañon [3], Poore, et al. [14], formulate this problem as a multidimensional assignment one. This problem is known to be NP-hard, that is to say that the complexity grows exponentially with the number of scans and measurements. That is why these methods were not intensively exploited at the beginning and some alternatives were proposed. In 1979, Reid [16] presented a real-time algorithm, the Multiple Hypothesis Tracker (MHT) in which measurements received at a scan are assigned to initialized targets, new targets or false alarms. Pruning and gating techniques are used to retain the most likely hypotheses and so limit their number. The main risk is to eliminate correct measurement sequences and it is all the more likely since the more interesting sources may be weak and fluctuating.

Another way to treat this problem is to consider data-association from a probabilistic viewpoint in order to compute the likelihood function. But the uncertainty in the measurements and the number of estimated parameters make it difficult to compute directly. Multitarget tracking can be viewed as an estimation problem *but* with partial observations. If the assignment vector of measurements to targets is available, it is an easy task to calculate the likelihood function for each target. Then, multitarget tracking becomes a traditional estimation problem. In this way, we can apply the Expectation-Maximization (EM) algorithm to find the maximum likelihood (ML) estimator. Avitzour [1] pioneered the use of this algorithm for multitarget tracking. In their seminal papers, Streit and Luginbuhl[1] [18, 19] gave a more rigorous application of the EM algorithm and proposed a batch Maximum A Posteriori

---

[1]For the sequel, we refer to Streit and Luginbuhl as Streit, et al.

(MAP) algorithm, called the Probabilistic Multiple Hypothesis Tracking (PMHT), which coupled EM iterations and Kalman filtering in the case of linear measurements. Based upon the duality of multitarget tracking, the measurement assignments are modeled as random variables which must be estimated jointly with target states. This probabilistic approach to the problem is original and does not require any enumeration of measurement to target associations. In fact, by avoiding the enumeration problem, the probabilistic approach to the problem is not NP-hard.

The approach used is quite similar in its spirit to the one of Streit, et al. [19]. In this work, a general formulation of multitarget tracking in the framework of an incomplete data problem is proposed. First, this incomplete data problem is applied to finite mixture densities. The measurement-to-target assignment vector is naturally interpreted as the missing information. Thus, we define the couple of measurements and assignments as the complete data (Section IV). Inclusion of data association in estimation is done through the extended state vector, composed of target states and assignment probabilities. We develop two classes of algorithms (Section V): one in the case of the ML and the other in the case of the MAP. For the latter, the general mixture framework allow us to retrieve the results obtained by Streit, et al. [19]. Then, we apply in Section VI, the ML approach to the passive SONAR for which, by definition, we must face nonlinear measurement equations. The maximization procedures are solved by means of a modified Newton algorithm using the Levenberg-Marquardt method [21]. In the simulations, measurements were generated through the classical beamforming in order to take into account track coalescence in data association. The results demonstrate the good ability of the algorithm to separate tracks and they bring out that multitarget tracking performance is closely related to signal processing since estimation of powerful target states may be severely deteriorated due to biased measurements provided by the classical beamforming in crossing area. This work is organized as follows. In Section II a general presentation of multitarget tracking is provided, followed in Section III by a review of the assumptions made in classical multitarget tracking approaches in comparison with the assumptions retained in the probabilistic approach. Section IV deals with a general presentation of the EM algorithm and its application to finite mixture densities. Results of this section are then extended to multitarget tracking in Section V. The ML approach is applied in Section VI to state estimation of nonmaneuvering targets for a narrowband passive SONAR. We show finally in Section VII that these results can be extended to the multisensor data association problem.

The following standard notations are used throughout the paper. Calligraphic letters indicate sets of batch length; upper case letters are used to denote sets for a given time; lower case letters indicate vectors; Z denotes the measurements; X denotes target states; K denotes the assignment vector; Π denotes the assignment probabilities; O denotes the parameter vector; $\mathcal{T}$ is a track; $\mathcal{P}$ is a partition of the measurements into tracks; $\mathcal{N}(\cdot)$ denotes the multivariate Gaussian probability density function.

II. GENERAL PROBLEM OF MULTITARGET TRACKING, NOTATIONS

The purpose of this section is to present a general framework for multitarget tracking in order to compare the different alternatives with the probabilistic one. Based on the ideas of Streit, et al., a general formulation is provided and extended by adding track and partition definitions. This definition of the problem is quite similar to that of Mori, et al. [11] where a general and mathematical formulation of the problem is presented. In their approach, a very general multisensor multitarget tracking problem was considered. Here, the formulation of our problem is restricted to a single sensor and a batch formulation of the multitarget tracking problem is considered.

Let $\mathcal{Z}$ be the set of cumulative measurements and $\mathcal{X}$ the set of cumulative state vector, both of length T:

$$\mathcal{X} = (X(1),\ldots,X(T))$$
$$\mathcal{Z} = (Z(1),\ldots,Z(T)).$$

For this batch length, we consider a set of M models of targets moving in the surveillance region. This number is not restrictive and one model can incorporate false alarms. Target motions are modeled by Markov process which can be expressed in a general way for discrete time as follows:

$$x_s(t) = f_s[x_s(t-1),v_s(t)], \qquad t = 1,\ldots,T$$

where $v_s(t)$ is a white Gaussian noise with covariance matrix $Q_s(t)$ and the subscript s indicates the model index.

The associated measurement equation is given below

$$z(t) = h_s[x_s(t),n_s(t)], \qquad t = 1,\ldots,T \qquad (1)$$

where $n_s(t)$ is the measurement Gaussian white noise with covariance matrix $N_s(t)$ and $z(t)$ is the measurement which originates from model s.

In the linear Gaussian case, these equations result in the classical ones

$$\begin{cases} x_s(t) = \Phi_s(t,t-1)x_s(t-1) + v_s(t), & t = 1,\ldots,T \\ z(t) = H_s(t)x_s(t) + w_s(t), & t = 1,\ldots,T. \end{cases}$$
$$(2)$$

At each time t, a new vector of measurements is received

$$Z(t) = (z_1(t),\ldots,z_{m_t}(t)).$$

Its size m_t varies in time due to false alarms, nondetected targets or track initalizations. The corresponding state vector is constituted of the states of the M models of targets at time t. It is denoted

$$X(t) = (x_1(t),\ldots,x_M(t)).$$

In order to take into account the uncertainty in measurement origin, we introduce an assignment vector. The concatenated assignment vector is noted $\mathcal{K}$ and defined by

$$\mathcal{K} = (K(1),\ldots,K(T)).$$

Each element $K(t)$ expresses an assignment hypothesis at time t:

$$K(t) = (k_1(t), \ldots, k_{m_t}(t))$$

where $k_j(t) = s$ indicates that target s produces measurement j at time t.

A track $\mathcal{T}_i$ is thus defined as a set of measurements assigned to the same model i:

$$\mathcal{T}_i = \{z_j(t) \mid k_j(t) = i, \ 1 \leq t \leq T\}, \qquad 1 < i \leq M.$$

In this definition, a track may contain missed detections (detection probability less than unity). As a model is allocated to false alarms, this definition includes false-alarms "track."

A partition of the measurements into tracks is defined as the set of nonempty tracks:

$$\mathcal{P} = \{\mathcal{T}_i \mid \mathcal{T}_i \neq \emptyset\}.$$

To each partition corresponds a data-association hypothesis of the meaurements $\mathcal{Z}$. This notion of partition [10] implies an exhaustive enumeration of all the measurement to target assignments. Thus, data-association consists in finding the most probable partition. It is obvious that if we consider a recursive hypothesis generation [16], their number will grow exponentially with time. So, pruning and gating are used to reduce the computational burden with a risk of eliminating some correct hypotheses. The probabilistic approach proposed by Streit, et al. avoids this enumeration by assigning all the measurements to each target in a probabilistic sense.

Although many different approaches were proposed, they use a common formalism. Consider for example two different approaches: the JPDAF and the MHT. We do not detail these algorithms but just show how they can be expressed in term of our definitions. In the JPDAF, hypotheses are built for the current measurement vector. An hypothesis is defined as the event associated with the corresponding assignment vector. The major aim of gating is to limit their number to the most likely ones. On the other hand, the probabilities of these hypotheses [6] are calculated independently of gating.

Suppose a new set of measurements be received at time n, $Z(n)$, the probability of the event associated with the assignment vector, $K(n)$ stands as follows:[2]

$$p(K(n) \mid Z^n) = \frac{1}{c} p(Z(n) \mid K(n), Z^{n-1}) p(K(n) \mid Z^{n-1})$$

where c represents the normalization constant.

In this writing, we confused the assignment vector with the associated event. The first term in the right-hand side identifies with the probability density function (pdf) of the current scan conditionned on the assignment vector and the past measurements. The last term of this side is the prior probability of such an event. It depends on the target detection probability and the false-alarm density. By summing all the events for which measurement j originates from target s, the event probabillity $p(k_j(n) = s \mid Z^n)$ is calculated. These probabilities are used to update Kalman filters.

In MHT, when a new set of measurements is received at time n, each of the previously established hypotheses contained in the partition $\mathcal{P}(n-1)$, are extended by adding the current hypothesis associated with the assignment vector $K(n)$ to form a new partition $\mathcal{P}(n) = (\mathcal{P}(n-1), K(n))$ [16]. The probability of such a partition is computed recursively:

$$p(\mathcal{P}(n) \mid Z^n) = \frac{1}{c} p(Z(n) \mid \mathcal{P}(n-1), K(n), Z^{n-1})$$

$$\times \, p(K(n) \mid \mathcal{P}(n-1), Z^{n-1})$$

$$\times \, p(\mathcal{P}(n-1) \mid Z^{n-1})$$

where we identify the first two terms with those of the JPDAF and the last one with the probability of the past measurements.

Conversely, in the approach of Streit, et al., partitionning is avoided because measurements are not assigned to particular targets but to all the targets. The assignment vector is considered as a random variable. Streit, et al. defines a parameter vector called the observer state $\mathcal{O}$ with continous and discrete components corresponding to the targets states and the assignment vector, respectively. This idea is original and powerful in the sense that data-association is basically included in the estimation problem. However in multitarget tracking, the missing data are the assignment vectors. So, they cannot be included directly in the estimation vector. Instead, the assignment probability of a measurement to a model is added to the target state vector.

Let Π denote this new vector:

$$\Pi = (\Pi(1), \ldots, \Pi(T)) \qquad \text{and}$$

$$\Pi(t) = (\pi_1(t), \ldots, \pi_M(t))$$

where the notation $\pi_m(t)$ indicates the probability that a measurement originates from the model m. In other words, this probability is independent of the measurement, i.e.,

$$\pi_m(t) = p(k_j(t) = m), \qquad \text{for all} \quad j = 1, \ldots, m_t.$$

Finally, let us define the parameter vector as

$$\mathcal{O} \stackrel{\Delta}{=} (\mathcal{X}, \Pi).$$

III. HYPOTHESES IN MULTITARGET TRACKING

In the previous section the general concepts of multitarget tracking were introduced. We now focus our attention on the constraints associated with the assignment vector and the related statistical assumptions. In traditional approaches [6, 10, 16] the following holds true.

1) One measurement originates from one target or the clutter (no merged measurement [11]) which means that the association is exclusive and exhaustive, i.e.,

$$\bigcup_{i=1}^{p} \mathcal{T}_i = \mathcal{Z} \qquad \text{and} \qquad \mathcal{T}_i \cap \mathcal{T}_j = \emptyset,$$

$$i \neq j, \quad i = 1, \ldots, p \quad \text{and} \quad j = 1, \ldots, p$$

with $p \leq M$ the number of tracks in the partition $\mathcal{P}$.

---

[2]We adopt traditional notations for the cumulative data sets up to time n, denoted Z^n. We note the tracks built up to time n, $T_i(n)$, and the corresponding partition $P(n)$.

2) A track produces at most one measurement (no split measurement [11]) which underscores the dependence in assigning measurements inside a scan; suppose model M is allocated to false alarms, this assumption implies:

$$t = 1 \ldots T, \qquad j \neq j', \qquad j = 1, \ldots, m_t,$$

$$j' = 1, \ldots, m_t \qquad m \in \{1, \ldots, M\}$$

$$k_j(t) = m \Rightarrow k_{j'}(t) \neq m.$$

The first assumption results in the following constraint on the assignment probabilities:

$$\sum_{m=1}^{M} \pi_m(t) = 1. \tag{3}$$

In PMHT, only this assumption is specified, the second one is ignored so that measurements may have the same origin. Furthermore, it is assumed that the components of the assignment vector $K(t)$ are independent. Consequently, the probability of the associated event is

$$p(K(t)) = \prod_{j=1}^{m_t} p(k_j(t)).$$

Actually, this assumption is crucial for the derivation of the EM algorithm. It corresponds to practical cases as in multiple frequency line tracking. On the contrary, in classical approaches, the constraints on the assignment vector impose on the calculation of the joint probability of such events. But, inside a scan, they consider measurements independent. This latter assumption is also retained in our approach. Furthermore, we assume that the assignment vectors are independent of target states at each time t, and the target tracks are independent of each other.

IV. EM ALGORITHM FOR MIXTURE DENSITIES

This section is devoted to a brief presentation of the EM algorithm and its application to mixture densities. The reference paper is [4]. Some interesting additional information can be found in [9, 23] for the convergence criteria, in [17] for mixture densities, and in [22] for application to exponential families. EM algorithms are used in many applications to solve incomplete data problems. Application of the EM algorithm to the estimation of mixture parameters is considered in this section. Using a similar formalism, we show in the next section that the PMHT equations can be obtained in this way.

A. Brief Review of EM Algorithm

Consider two spaces $\mathcal{X}$ and $\mathcal{Y}$ and a mapping from $\mathcal{X}$ to $\mathcal{Y}$ and the set

$$\mathcal{X}(y) = \{x \mid x \in \mathcal{X} \text{ and } y(x) = y\}.$$

Let us assume that the pdfs of x and y are parameterized by ϕ. They are denoted $f(x \mid \phi)$ and $g(y \mid \phi)$ respectively, where

$$\phi = (\varphi_1, \ldots, \varphi_d) \in \Omega \subset R^d.$$

The problem consists in estimating ϕ thanks to the ML based on the observation of y. We stress that x is not directly observable. This problem is well known as incomplete data problem. Thereafter, we refer to x as the complete data and to y as the observed incomplete data. Finally, we introduce the conditional density of x given y and ϕ, namely:

$$h(x \mid y, \phi) = \frac{f(x \mid \phi)}{g(y \mid \phi)}. \tag{4}$$

This density represents the pdf of the missing data conditioned on the observed data and the parameter vector. It plays an important part in the EM algorithm. Since the density of the complete data is unknown, it is estimated through the expectation of the pdf of the missing data (5).

EM algorithm is a method to find this maximum given an observation y. Each iteration comprises an expectation step (E-step) and a maximization step (M-step). Denote ϕ^i the parameter estimated during the (i)th iteration, the updated parameter ϕ^{i+1} at the $(i + 1)$th iteration is obtained via the following (EM) recursion.

1) E-step: Compute the expectation defined like this

$$Q(\phi \mid \phi^i) = E\{\log[f(x \mid \phi)] \mid y, \phi^i\}$$
$$= L(\phi) + H(\phi \mid \phi^i) \tag{5}$$

where

$$L(\phi) = \log[g(y \mid \phi)]$$
$$H(\phi \mid \phi^i) = E\{\log[h(x \mid y, \phi)] \mid y, \phi^i\}.$$

2) M-step: Find ϕ which maximizes $Q(\phi \mid \phi^i)$

$$\phi^{i+1} = \arg\max_{\phi} Q(\phi \mid \phi^i).$$

Unknowing the likelihood of the complete data, its expectation (E-step) is calculated and maximized at the current step based upon the previous value of ϕ and the observation y. Using Jensen's inequalities, it can be shown that the log-likelihood $L(\phi)$ increases at each iteration [4].

B. Mixture Densities

The problem of estimating parameters from mixture densities has been the subject of an important literature. A brief review of different approaches is presented in the article of Redner, et al. [17] (other references can be found there). Parameter estimation of mixture densities may be achieved by means of the EM algorithm. Following the guidelines of Redner and Walker, the EM calculations is now detailed. Their utility for deriving the PMHT equations is demonstrated in the next section.

As previously, X denotes the complete data vector and Y the incomplete data vector. Suppose the set of N independent data be observed, $Y = \{y_j\}_{j=1 \ldots N}$. Each measurement y_j belongs to a family of parametric density functions with probabilities $\{\pi_i\}_{i=1 \ldots m}$:

$$p(y_j \mid \phi) = \sum_{i=1}^{m} \pi_i p_i(y_j \mid \varphi_i)$$

where

$$\{\pi_i \geq 0\}_{i=1}^m \qquad \text{and} \qquad \sum_{i=1}^m \pi_i = 1.$$

Each p_i is a pdf parametrized by φ_i. Denote ϕ the parameter vector, $\phi = (\pi_1, \ldots, \pi_m, \varphi_1, \ldots, \varphi_m)$. Using the independence assumption for the measurements, the likelihood function and the log-likelihood express as follows:

$$g(Y \mid \phi) = \prod_{j=1}^N \sum_{i=1}^m \pi_i p_i(y_j \mid \varphi_i)$$

$$L(\phi) = \sum_{j=1}^N \log \left[\sum_{i=1}^m \pi_i p_i(y_j \mid \varphi_i) \right].$$

To derive EM iterations, we need to introduce the density function of the complete data and the density function of the missing data given the observed data. Let us define the complete data vector as $X = \{x_j\}_{j=1,\ldots,N}$ where each complete data x_j is defined by $x_j = (y_j, k_j)$ and where the missing data k_j takes value in $\{1, \ldots, m\}$, indicating which density the observed data y_j comes from. Notice that X is a mixed continuous and discrete random variable. The probability mass function of the missing data vector, $K = \{k_j\}_{j=1,\ldots,N}$, is

$$p(K \mid \phi) = \prod_{j=1}^N \pi_{k_j}.$$

Elementary calculations yield

$$f(X \mid \phi) = p(Y \mid K, \phi) p(K \mid \phi)$$

$$= \prod_{j=1}^N \pi_{k_j} p_{k_j}(y_j \mid \varphi_{k_j}). \qquad (6)$$

The probability mass function of the missing data conditioned on the observed data and the parameter vector is easily deduced based on (4) and (6):

$$h(X \mid Y, \phi) = p(K \mid Y, \phi) = \prod_{j=1}^N \frac{\pi_{k_j} p_{k_j}(y_j \mid \varphi_{k_j})}{p(y_j \mid \phi)}. \qquad (7)$$

1) *E-step*:   Suppose that the parameters ϕ^i have been estimated at iteration (i). To update the estimation, we must calculate the following expectation:

$$Q(\phi \mid \phi^i) = E\{\log[f(X \mid \phi)] \mid Y, \phi^i\}.$$

The expectation $Q(\phi \mid \phi^i)$ may thus be deduced using (6) and (7) and stands as follows:

$$Q(\phi \mid \phi^i) = \sum_k \log[f(X \mid \phi)] p(K \mid Y, \phi^i)$$

$$= \sum_{k_1=1}^m \cdots \sum_{k_N=1}^m \left\{ \sum_{j=1}^N \log[\pi_{k_j} p_{k_j}(y_j \mid \varphi_{k_j})] \right\}$$

$$\times \prod_{j'=1}^N \frac{\pi_{k_{j'}}^i p_{k_{j'}}(y_{j'} \mid \varphi_{k_{j'}}^i)}{p(y_{j'} \mid \phi^i)}. \qquad (8)$$

Using straightforward calculations the following result (see Appendix A) is obtained

$$Q(\phi \mid \phi^i) = \sum_{k=1}^m \left[\sum_{j=1}^N \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)} \right] \log[\pi_k]$$

$$+ \sum_{k=1}^m \sum_{j=1}^N \log[p_k(y_j \mid \varphi_k)] \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}. \qquad (9)$$

REMARKS.
1) Notice that (9) is decomposed in two terms each one containing only one kind of parameters. This attractive expression allows considerable simplification of the maximization step.
2) Moreover, suppose parameters $\{\varphi_k\}_{k=1\cdots m}$ mutually independent, maximization with respect to these parameters is again decomposed in m maximizations.
3) Finally, as $\log[\pi_1], \ldots, \log[\pi_m]$ appears linearly in the first term in (9), an explicit solution to the first maximization can be determined.

2) *M-step*:   As said above, M-step is solved by maximizing (9) with respect to $\{\pi_k\}_{k=1,\ldots,m}$ on the one hand, and $\{\varphi_k\}_{k=1,\cdots,m}$ on the other.
Thus, the first maximization stands as

$$\text{Maximize} \quad g(\pi) = \sum_{k=1}^m \left[\sum_{j=1}^N \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)} \right] \log[\pi_k]$$

$$(10)$$

$$\text{subject to} \quad \sum_{k=1}^m \pi_k = 1.$$

Forming the Lagrangian

$$L(\pi, \lambda) = g(\pi) + \lambda \left(1 - \sum_{k=1}^m \pi_k \right)$$

a necessary condition is

$$\nabla_\pi L(\pi, \lambda) = 0.$$

The probability π_k^{i+1} is straightforwardly deduced

$$\pi_k^{i+1} = \frac{1}{\lambda} \sum_{j=1}^N \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}$$

where the Lagrange multiplier λ is determined by using the constraint (3). Since

$$p(y_j \mid \phi^i) = \sum_{s=1}^m \pi_s^i p_s(y_j \mid \varphi_s^i)$$

elementary calculation yields

$$\lambda = N.$$

Then, the new probability is updated at iteration $(i + 1)$ according to

$$\pi_k^{i+1} = \frac{1}{N} \sum_{j=1}^N w_{j,k}^{i+1} \qquad (11)$$

where

$$w_{j,k}^{i+1} = \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}. \tag{12}$$

In a second time, the second term of (9) is maximized with respect to $\{\varphi_k\}_{k=1\cdots m}$. Let us recall the expression to maximize:

$$\sum_{k=1}^{m}\sum_{j=1}^{N}\log[p_k(y_j \mid \varphi_k)]\frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}.$$

Since parameters φ_k are mutually independent in the previous expression, the parameter vector φ_k is updated by

$$\varphi_k^{i+1} \in \underset{\varphi_k}{\operatorname{Arg\,max}} \sum_{j=1}^{N}\log[p_k(y_j \mid \varphi_k)]w_{j,k}^{i+1},$$

$$\forall \quad k = 1\cdots m. \tag{13}$$

The algorithm then takes the following form:

1. update of π_k :

$$\pi_k^{i+1} = \frac{1}{N}\sum_{j=1}^{N}w_{j,k}^{i+1}$$

where $w_{j,k}^{i+1} = \dfrac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}$;

2. update of φ_k according to:

$$\varphi_k^{i+1} \in \underset{\varphi_k}{\operatorname{Arg\,max}} \sum_{j=1}^{N}\log[p_k(y_j \mid \varphi_k)]w_{j,k}^{i+1},$$

$$\forall \quad k = 1\cdots m.$$

Each weight $w_{j,k}^{i+1}$ (12) corresponds to the posterior probability that measurement y_j originates from the kth density given the current estimation ϕ^i. Equation (9) takes advantage of all the information available at current iteration. In some applications, maximization with respect to φ_k admits analytic updating solutions. Consider for example the case where $p_k(y \mid \varphi_k)$ is a Gaussian density:

$$p_k(y \mid \varphi_k) = \frac{1}{(2\pi)^{n/2}(\det \Sigma_k)^{1/2}} e^{-1/2(y-\mu_k)^T \Sigma_k^{-1}(y-\mu_k)},$$

$$\varphi_k = (\mu_k, \Sigma_k)$$

where $y \in R^n$, $\mu_k \in R^n$ and Σ_k is an $n \times n$ positive defined matrix.

Then, the updated expression of φ_k^{i+1} at iteration $(i+1)$ is given below

$$\mu_k^{i+1} = \frac{\sum_{k=1}^{N} y_k \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}}{\sum_{k=1}^{N} \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}},$$

$$\Sigma_k^{i+1} = \frac{\sum_{k=1}^{N}(y_k - \mu_k^{i+1})(y_k - \mu_k^{i+1})^T \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}}{\sum_{k=1}^{N} \frac{\pi_k^i p_k(y_j \mid \varphi_k^i)}{p(y_j \mid \phi^i)}}.$$

We have considered the application of the EM method to the estimation of mixture parameters. This approach is now extended to multitarget tracking.

V. EM ALGORITHM APPLIED TO MULTITARGET TRACKING

A batch algorithm as in [19] is considered. A set of measurements $\mathcal{Z}$ is available. The assignment vector is not observable. So, we consider quite naturally that $\mathcal{Z}$ is the incomplete data set and that $(\mathcal{Z}, \mathcal{K})$ is the complete data set. Let us now define the different densities needed to derive EM algorithm. Firstly, The probability mass function of the missing data $\mathcal{K}$ is now defined as:[3]

$$p(\mathcal{K} \mid \mathcal{O}) = \prod_{t=1}^{T}\prod_{j=1}^{m_t} \pi_{k_j}(t). \tag{14}$$

Then, the likelihood of the complete data set is given below by means of Bayes' formula and independence assumptions:[4]

$$p(\mathcal{Z}, \mathcal{K} \mid \mathcal{O}) = \prod_{t=1}^{T} p(Z(t), K(t) \mid O(t))$$

$$= \prod_{t=1}^{T}\prod_{j=1}^{m_t} p(z_j(t) \mid x_{k_j}(t))\pi_{k_j}(t). \tag{15}$$

Next, the likelihood functional of the incomplete data stands as follows:

$$p(\mathcal{Z} \mid \mathcal{O}) = \prod_{t=1}^{T} p(Z(t) \mid X(t), \Pi(t))$$

$$= \prod_{t=1}^{T}\prod_{j=1}^{m_t} p(z_j(t) \mid O(t))$$

$$= \prod_{t=1}^{T}\prod_{j=1}^{m_t}\sum_{m=1}^{M} p(z_j(t) \mid x_m(t))\pi_m(t). \tag{16}$$

Expression (16) is deduced from the independence assumption about the assignment vector components at the current scan.

Finally, a generalization of (7) to multitarget tracking is

$$p(\mathcal{K} \mid \mathcal{Z}, \mathcal{O}) = \frac{p(\mathcal{Z}, \mathcal{K} \mid \mathcal{O})}{p(\mathcal{Z} \mid \mathcal{O})} = \prod_{t=1}^{T}\prod_{j=1}^{m_t} p(k_j(t) \mid z_j(t), O(t)). \tag{17}$$

The algorithm consists not only in estimating the targets states but also the assignment probabilities.

A. E-Step

Denote $\mathcal{O}^*$ the estimated vector at the previous iteration. This step is devoted to calculate the expectation

[3] $\mathcal{O} \overset{\Delta}{=} (\mathcal{X}, \Pi)$, see p. 7.

[4] Independence from scan to scan and from assignment to assignment.

of the log-likelihood of complete data based upon the knowledge of the parameters at previous step and measurements. By analogy with (8), we consider the function $Q(\mathcal{O} \mid \mathcal{O}^*)$:

$$Q(\mathcal{O} \mid \mathcal{O}^*) = E\{\log[p(\mathcal{Z}, \mathcal{K} \mid \mathcal{O})] \mid \mathcal{Z}, \mathcal{O}^*\}$$

$$= \sum_{\mathcal{K}} \log[p(\mathcal{Z}, \mathcal{K} \mid \mathcal{O})] p(\mathcal{K} \mid \mathcal{Z}, \mathcal{O}^*). \tag{18}$$

This calculation is an extension of (8). Denote $nb = \sum_{t=1}^{T} m_t$. The sum comprises M^{nb} terms. Substituting (15) and (17) in (18), the independence assumption of $K(t)$ yields

$$Q(\mathcal{O} \mid \mathcal{O}^*) = \sum_{\mathcal{K}} \left\{ \sum_{t=1}^{T} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_{k_j}(t)) \pi_{k_j}(t)] \right\}$$

$$\times \left\{ \prod_{t=1}^{T} \prod_{j=1}^{m_t} p(k_j(t) \mid z_j(t), \mathcal{O}^*(t)) \right\}.$$

Simplifying the sums, we obtain

$$Q(\mathcal{O} \mid \mathcal{O}^*) = \sum_{t=1}^{T} \sum_{j=1}^{m_t} \sum_{k_j(t)=1}^{M} \log[\pi_{k_j}(t)] w_{j,k_j}(t)$$

$$+ \sum_{t=1}^{T} \sum_{j=1}^{m_t} \sum_{k_j(t)=1}^{M} \log[p(z_j(t) \mid x_{k_j}(t))] w_{j,k_j}(t). \tag{19}$$

As $k_j(t)$ takes value in $\{1, \ldots, M\}$ whatever $t = 1, \ldots, T$ and $j = 1, \ldots, m_t$, we can invert sums in (19), yielding

$$Q(\mathcal{O} \mid \mathcal{O}^*) = \sum_{k=1}^{M} \sum_{t=1}^{T} \left[\sum_{j=1}^{m_t} w_{j,k}(t) \right] \log[\pi_k(t)]$$

$$+ \sum_{k=1}^{M} \sum_{t=1}^{T} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_k(t))] w_{j,k}(t). \tag{20}$$

The remarks from the previous section still remain valid. The independence assumption of targets allow us to solve M maximizations, each coupled thanks to the weighting

$$w_{j,k}(t) = \frac{\pi_k^* p(z_j(t) \mid x_k^*(t))}{p(z_j(t) \mid O^*(t))}.$$

B. M-Step

The first maximization problem reverts to T maximizations:

Maximize for all t, $\quad g(\Pi(t)) = \sum_{k=1}^{M} \sum_{j=1}^{m_t} \log[\pi_k(t)] w_{j,k}(t)$

$$\tag{21}$$

subject to $\quad \sum_{k=1}^{M} \pi_k(t) = 1.$

Solution to each maximization is given by (11):

$$\pi_k(t) = \frac{1}{m_t} \sum_{j=1}^{m_t} w_{j,k}(t). \tag{22}$$

The second maximization problem with respect to $\mathcal{X}$ is equivalent to (13) generalized for all $t \in \{1, \ldots, T\}$. Hence, for all $t = 1, \ldots, T$ and $k = 1, \ldots, M$,

$$x_k(t) \in \underset{x_k}{\operatorname{Arg\,max}} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_k(t))] w_{j,k}(t). \tag{23}$$

This last expression does not include tracking in states estimation. If we suppose that targets obey determinist trajectories (see Section VI), state parameters reduce to their initial states, denoted $X_0 = (x_{1,0}, x_{2,0}, \ldots, x_{M,0})$. So the maximization step becomes: for all $k = 1, \ldots, M$,

$$x_{k,0} \in \underset{x_{k,0}}{\operatorname{Arg\,max}} \sum_{t=1}^{T} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_{k,0})] w_{j,k}(t). \tag{24}$$

C. EM in the Case of Maximum A Posteriori

Throughout the discussion, we only deal with the ML. An easy extension can be introduced to take advantage of knowledge of prior probability relative to the parameters. We show that a MAP derivation of the algorithm corresponds to the PMHT of Streit, et al. which results in Kalman filtering in the linear Gaussian case [19]. Let us denote $P(\mathcal{O}) \overset{\Delta}{=} \log[p(\mathcal{X}, \Pi)]$. Suppose assignment probabilities don't have prior probabilities as in [19], and the different models follow a Markov process[5]

$$P(\mathcal{O}) \overset{\Delta}{=} \log[p(\mathcal{X}, \Pi)]$$

$$= \log\left[p(X(0)) \prod_{t=1}^{T} p(X(t) \mid X(t-1)) \right]$$

$$= \log\left[\prod_{k=1}^{M} p(x_k(0)) \prod_{k=1}^{M} \prod_{t=1}^{T} p(x_k(t) \mid x_k(t-1)) \right]$$

$$= \sum_{k=1}^{M} \log[p(x_k(0))]$$

$$+ \sum_{k=1}^{M} \sum_{t=1}^{T} \log[p(x_k(t) \mid x_k(t-1))]. \tag{25}$$

The MAP can be straightforwardly embedded into the EM framework. The E-step of the algorithm is changed

---

[5]For the sequel, only target models are considered in prior probability. The model devoted to false alarms has been dropped since no kinematic parameter has to be estimated. Of course, such a model is taken into account in assignment probabilities (see Section VI).

in [4]:

$$M(\mathcal{O} \mid \mathcal{O}^*) \stackrel{\Delta}{=} Q(\mathcal{O} \mid \mathcal{O}^*) + P(\mathcal{O}). \qquad (26)$$

Substituting (20) and (25) in (26) leads to

$$M(\mathcal{O} \mid \mathcal{O}^*) = \sum_{k=1}^{M} \sum_{t=1}^{T} \left[\sum_{j=1}^{m_t} w_{j,k}(t) \right] \log[\pi_k(t)]$$

$$+ \sum_{k=1}^{M} \sum_{t=1}^{T} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_k(t))] w_{j,k}(t)$$

$$+ \sum_{k=1}^{M} \log[p(x_k(0))]$$

$$+ \sum_{k=1}^{M} \sum_{t=1}^{T} \log[p(x_k(t) \mid x_k(t-1))].$$

The maximization with respect to Π and to $\mathcal{X}$ is always decomposed in independent maximizations coupled with weightings $w_{j,k}(t)$. Assignment probabilities are updated according to (22). The state vectors for next iteration are given by

$$(x_k^0 \cdots x_k^T) \in$$

$$\operatorname*{arg\,max}_{X_k} \left\{ \begin{array}{l} \displaystyle\sum_{t=1}^{T} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_k(t))] w_{j,k}(t) \\[2ex] \displaystyle + \log[p(x_k(0))] + \sum_{t=1}^{T} \log[p(x_k(t) \mid x_k(t-1))] \end{array} \right\},$$

$$\text{for all} \quad k = 1, \ldots, M. \qquad (27)$$

An interesting case is the linear Gaussian Markov process described by (2). Instead of maximizing directly (27), consider the exponential of this expression [19]:

$$p(x_k^0) \prod_{t=1}^{T} \left\{ p(x_k(t) \mid x_k(t-1)) \prod_{j=1}^{m_t} p(z_j(t) \mid x_k(t))^{w_{j,k}(t)} \right\},$$

$$k = 1, \ldots, M$$

and notice that

$$\prod_{j=1}^{m_t} p(z_j(t) \mid x_k(t))^{w_{j,k}^{i+1}(t)}$$

$$\propto \prod_{j=1}^{m_t} \mathcal{N}(z_j(t) \mid H_k x_k(t), (w_{j,k}^{i+1}(t))^{-1} R_k)$$

$$\propto \mathcal{N}(\tilde{z}_k(t) \mid H_k x_k(t), (m_t \pi_k^{i+1}(t))^{-1} R_k)$$

where

$$\tilde{z}_k(t) = \frac{1}{m_t \pi_k^{i+1}(t)} \sum_{j=1}^{m_t} w_{j,k}^{i+1}(t) z_j(t).$$

This maximization can be solved by using the Kalman filtering where the measurement is replaced with the measurement centroid $\tilde{z}_k(t)$ with covariance matrix $\tilde{R}_k$ defined as

$$\tilde{R}_k = \frac{R_k}{m_t \pi_k^{i+1}(t)}.$$

D. Conclusion

In conclusion, based upon an initialization of the parameters denoted $\mathcal{O}^{(0)} = (\Pi^{(0)}, \mathcal{X}^{(0)})$, the new parameters at step $(i+1)$ are updated iteratively according to the following.

1) Update of Π:

$$\pi_k^{i+1}(t) = \frac{1}{m_t} \sum_{j=1}^{m_t} w_{j,k}^{i+1}(t) \qquad \text{where}$$

$$w_{j,k}^{i+1}(t) = \frac{\pi_k^i p(z_j(t) \mid x_k^i(t))}{p(z_j(t) \mid \mathcal{O}^i(t))}$$

$$\text{for all} \quad t = 1, \ldots, T \quad \text{and} \quad k = 1, \ldots, M. \qquad (28)$$

2) Update of $\mathcal{X}$:
For the MAP:

$$(x_k^{i+1}(0) \cdots x_k^{i+1}(T))$$

$$\in \operatorname*{arg\,max}_{X_k} \left\{ p(x_k(0)) \prod_{t=1}^{T} p(x_k(t) \mid x_k(t-1)) \right.$$

$$\left. \times \prod_{j=1}^{m_t} p(z_j(t) \mid x_k(t))^{w_{j,k}^{i+1}(t)} \right\}$$

$$\text{for all} \quad k = 1, \ldots, M.$$

For the ML:

$$x_{k,0} \in \operatorname*{Arg\,max}_{x_{k,0}} \sum_{t=1}^{T} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_{k,0})] w_{j,k}(t)$$

$$(29)$$

$$\text{for all} \quad k = 1, \ldots, M. \qquad (30)$$

Maximizations with respect to $\mathcal{X}$ depend on the application and require generally the use of iterative optimization algorithms. In the next section, multitarget tracking is applied to passive SONAR. The corresponding maximization problems are detailed.

VI. MULTITARGET TRACKING IN PASSIVE SONAR

By considering measurement history composed of Cartesian positions for targets moving in horizontal plane, multitarget tracking is considerably simplified since observability is ensured and measurement equation is a linear function of target states. On the contrary, receiver maneuvers are required for ensuring observability in the bearings-only tracking context. Furthermore, target state estimation is much more sensitive because optimizations in the M-step are no longer linear. In this section, we assume that the observability conditions are fulfilled.

A. Parameter Estimation

In this section, targets move with constant velocity vectors. The parameter vector is thus denoted $\mathcal{O} = (X_0, \Pi)$ where X_0 indicates the M initial target states as in the
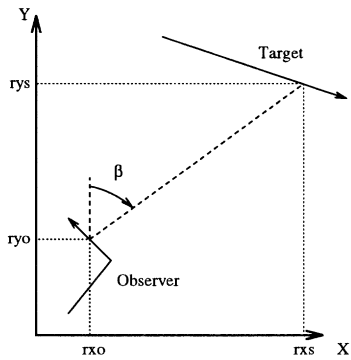
Fig. 1. Scenario in TMA.

E-step of the previous section:

$$\mathcal{O} = (x_{1,0}, \ldots, x_{M,0}, \Pi(1), \ldots, \Pi(T)).$$

We focus on the update of X_0^{i+1} based upon $\mathcal{O}^i$ and drop the subscript on the target model since the maximization with respect to X_0 is decomposed into M maximizations.

Denote the receiver state (Fig. 1) as

$$x_{\text{obs}}(t) = (r_{xo}, r_{yo}, v_{xo}, v_{yo})^T$$

and the target state

$$x_s(t) = (r_{xs}, r_{ys}, v_{xs}, v_{ys})^T.$$

The relative state vector is thus defined as

$$x(t) = x_s(t) - x_{\text{obs}}(t) = (r_x, r_y, v_x, v_y)^T.$$

The measurement equation (2) is replaced with

$$z = \beta(t) + n(t) \tag{31}$$

where

$$\beta(t) = \arctan\left(\frac{r_x(t)}{r_y(t)}\right).$$

In (31), $n(t)$ is a white Gaussian noise with variance $\sigma^2(t)$. Since targets are supposed to move at constant velocity, a ML approach is used. The M-step of the algorithm consists in computing (24). Using (31), the functional $f(Z \mid x_0)$ which has to be maximized for each model is

$$f(Z \mid x_0) = \sum_{t=1}^{T} \sum_{j=1}^{m_t} \log[p(z_j(t) \mid x_0)] w_j^{i+1}(t) \tag{32}$$

$$= \sum_{t=1}^{T} \sum_{j=1}^{m_t} -\frac{1}{2\sigma^2(t)} (z_j(t) - \beta(x_0,t))^2 w_j^{i+1}(t) + \cdots. \tag{33}$$

In (33), $\beta(x_0,t)$ is the target azimuth at time t for initial target vector x_0. Terms independent of x_0 are dropped from (33). Taking the partial derivative with respect to x_0 yields the gradient vector:

$$\nabla_{x_0} f = \sum_{t=1}^{T} \frac{1}{\sigma^2(t)} \frac{\partial \beta(x_0,t)}{\partial x_0} \sum_{j=1}^{m_t} (z_j(t) - \beta(x_0,t)) w_j^{i+1}(t)$$

which can be expressed in a vectorial form as

$$\nabla_{x_0} f = (R_{x_0}^{-1} A_{x_0})^T N^{-1} B_{x_0} \tag{34}$$

with

$$A_{x_0} =$$

$$\begin{bmatrix} \cos\beta(x_0,1) & -\sin\beta(x_0,1) & \delta t \cos\beta(x_0,1) & -\delta t \sin\beta(x_0,1) \\ \vdots & \vdots & \vdots & \vdots \\ \cos\beta(x_0,T) & -\sin\beta(x_0,T) & T\delta t \cos\beta(x_0,T) & -T\delta t \sin\beta(x_0,T) \end{bmatrix}$$

$$B_{x_0} = \begin{pmatrix} \sum_{j=1}^{m_1} (z_j(1) - \beta(x_0,1)) w_j^{i+1}(1) \\ \vdots \\ \sum_{j=1}^{m_T} (z_j(T) - \beta(x_0,T)) w_j^{i+1}(T) \end{pmatrix}$$

$$N = \text{diag}(\sigma^2(t)), \qquad t = 1, \ldots, T$$

$$R_{x_0} = \text{diag}(r(t)), \qquad t = 1, \ldots, T.$$

In (34), R_{x_0} is the matrix of the range sources at sampling time (δt is the sampling period) for inital target state x_0 while N is the measurement variance matrix if we suppose that all measurements within a scan have the same variance.

Notice that the gradient of the expression we need to maximize in M-step differs from the gradient of the classical MLE in bearings-only TMA [12] only by the measurement term B_{x_0}. In cluttered environment, it is replaced by a mixing proportion of all the measurements with their assignment probabilities that they originate from the target of interest given the previous estimation.

A solution to this maximization problem is obtained numerically with a Gauss–Newton algorithm improved by the Levenberg–Marquadt method [5]. This method uses a trust region strategy. Since each M-step may face "weak" estimability, this method is required to avoid unrealistic estimates. Each iteration of this algorithm consists in computing

$$x_0^{l+1} = x_0^l - [A_{x_0^l}^T R_{x_0^l}^{-1} N^{-1} R_{x_0^l}^{-1} A_{x_0^l} + \mu_{x_0^l} \text{Id}]^{-1}$$

$$\times A_{x_0^l}^T R_{x_0^l}^{-1} N^{-1} B_{x_0^l}$$

where Id is the matrix identity and $\mu_{x_0^l}$ is the current value used to perturb the Hessian $\nabla_{x_0}^2 f$. Note that in evaluating $\nabla_{x_0}^2 f$, terms $\partial^2 \beta(x_0,t)/\partial^2 x_0$ have been dropped.

Summarizing the previous results, the PMHT algorithm combines EM and Gauss–Newton iterations. Inside each EM iteration, the algorithm have to compute M maximizations which can be easily implemented on parallel computer since calculations of the weightings $w_j^{i+1}(t)$ can be computed separately.

B. Numerical Results

Whatever the detection system (SONAR or RADAR), the data association step is a high level one whose inputs are the outputs of low level processing including signal
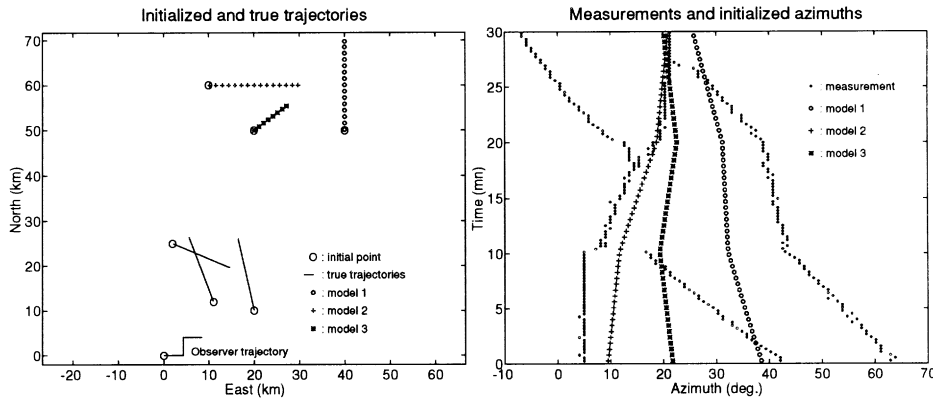
Fig. 2. Initialized trajectories and corresponding azimuths (case 1).

processing (filtering, discrete Fourier transform (DFT)) and array processing (beamforming, predetection, etc.). In a narrowband passive SONAR, roughly speaking, the signal processing step is composed as follows: signals received from a sensor array are filtered, then sampled and analyzed (e.g., spectral analysis). A classical beamforming is performed to estimate instantaneous target parameters (bearings, etc.). Array processing also includes a predetection step.

Our aim is now to define a realistic sonar simulation. So, the basic level of our simulations is the vector formed from a single DFT bin from each sensor; this vector is often called the snapshot vector. It is denoted X.

An assumption, common in the passive array processing context, asserts that this vector is zero-mean, Gaussian circular with covariance Γ given by

$$\Gamma = \sum_{i=1}^{M} \sigma_i D_{\beta_i} D_{\beta_i}^* + b\,\text{Id}$$

where * indicates transposition and conjugation.

In the above formula, D_{β_i} denotes the steering vector of target i (M targets present in the surveillance region) and, in the case of a linear array,[6] is given by the classical formula:

$$D_{\beta_i} \triangleq [1, \exp(-2i\pi f_s), \ldots, \exp(-2i\pi(n_c - 1)f_s)]^T$$

where f_s stands for the spatial frequency ($f_s = \cos\beta_i/2$, intersensor distance $d = \lambda/2$). The parameter σ_i represents the power of target i while b is the noise power.

In classical beamforming the spatial density of the field impinging the array in the direction β is estimated by the value of the following quadratic form:

$$P(\beta) = D_\beta \hat{\Gamma} D_\beta^*.$$

In this expression, $\hat{\Gamma}$ is the estimated cross-spectral matrix of the sensor outputs. If the averaged periodogram is considered, $\hat{\Gamma}$ takes the following form:

$$\hat{\Gamma} = \frac{1}{N} \sum_{t=0}^{N-1} X_t X_t^*$$

---

[6]Linear array constituted of n_c equispaced sensors.

where X_t is the tth vector of sensor DFT (N is the number of integrations).

1) *Case* 1 : *Pd* = 1 *and Pfa* = 0: For all the simulations, the array processing parameters take the following values: $n_c = 32$, $N = 125$. In order to evaluate the ability of the data-association algorithm for multiple sources, false alarms are not included at first. Extension of measurement generation to cluttered environment is straightforward.

Finally, to initialize the parameters of the algorithm, we consider that each model is equally probable and that no strategy is applied for target state X_0^0. However, poor guess of the initial state may have dramatic effect on the estimated source states and assignment probabilities. To avoid this drawback, we replace the direct maximization of the likelihood functional $L(\mathcal{O}) \triangleq \log[p(\mathcal{Z} \mid \mathcal{O})]$ by a sequence of intermediate maximixations of functionals $\{L_i(\mathcal{O})\}_{i=1\cdots N}$. In these functionals the measurement covariance is increased to take into account more measurements in the evaluation of the weightings $w_{j,k}^{i+1}(t)$. This idea of minimizing the influence of target states initiation have been previously derived in the probabilistic data association (PDA) context [7]. More precisely, this sequence is defined as follows: $L_i(\mathcal{O}) = L(\mathcal{O}, a_i\sigma)$ where the $\{a_i\}_{i=1\cdots N}$ is a decreasing sequence such that $a_N = 1$. The state estimate obtained from $L_i(\mathcal{O})$ serves as the target state initiation for $L_{i+1}(\mathcal{O})$.

Fig. 2 presents a scenario for three targets with different signal to noise ratios. The first target is travelling from position $(2000\text{ m}, 25000\text{ m})^T$ with speed $(7\text{ m/s}, -3\text{ m/s})^T$, the second one from $(11000\text{ m}, 12000\text{ m})^T$ with speed $(-3\text{ m/s}, 8\text{ m/s})^T$ and the third one from $(20000\text{ m}, 10000\text{ m})^T$ with speed $(-2\text{ m/s}, 9\text{ m/s})^T$. The signal-to-noise ratios are assumed constant during the simulation, equal to $(-10\text{ dB}, -15\text{ dB}, -15\text{ dB})^T$, respectively. Corresponding measurements are also shown in Fig. 2. Even if the target powers are quite similar, it is worth stressing the absorption of weak targets by powerful ones, which results in a track coalescence in crossing areas. Initialized target trajectories and corresponding azimuths (values see Table I) are displayed in Fig. 2.

The receiver trajectory is composed of three legs to ensure system observability. Its speed vector is $(7\text{ m/s}, 0\text{ m/s})^T$ for odd leg and $(0\text{ m/s}, 7\text{ m/s})^T$ for even leg. The duration of the simulation is 1800s and the sampling period is $\delta t = 16$ s. The results are presented

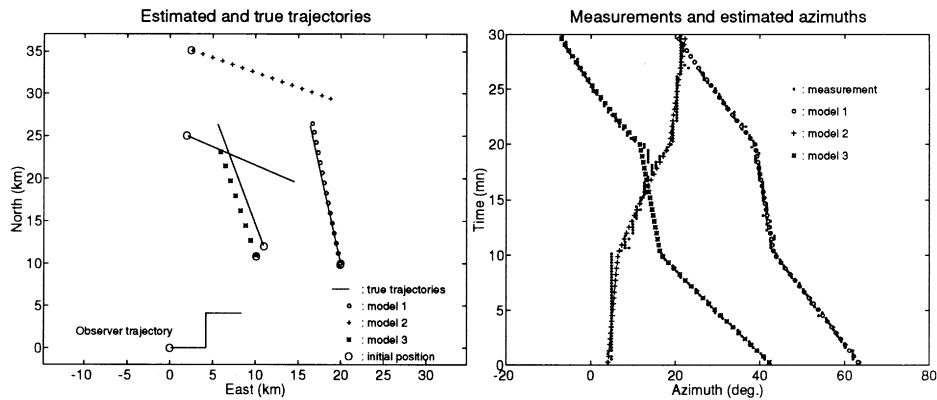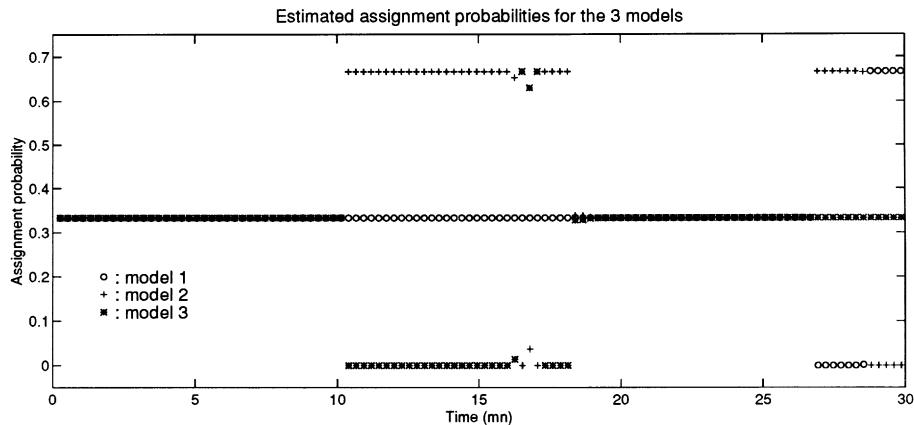Fig. 3. Estimated trajectories and corresponding azimuths (case 1).

Fig. 4. Estimated probabilities $\{\pi_m(t)\}_{t=1\cdots T, m=1\cdots M}$ (case 1).

TABLE I
Initiation of Algorithm X_0^0

	model 1	model 2	model 3
r_x (m)	40000	10000	20000
v_x (m/s)	0	11	4
r_y (m)	50000	60000	50000
v_y (m/s)	11	0	3

in Figs. 3 and 4. Despite a poor initiation, the algorithm converges to the global solution. Although local solutions or stationary points are possible, the use of the sequence of functions defined above reduces their numbers. Despite its simplicity, this simulation demonstrates the ability of the algorithm to resolve data-association in difficult crossing areas. Some interesting remarks can be made by considering simultaneously Figs. 3 and 4.

REMARKS.

1) For tracks well separated, the probability that a measurement originates from each model is equally probable as it was expected and is verified up to $t \simeq$ 11 mn, and from $t \simeq 18$ mn to $t \simeq 27$ mn.

2) From $t \simeq 10$ mn to $t \simeq 18$ mn, model 2 and model 3 share the same measurements which are all assigned to the first one, thus assignment probability of model 2

is equal to 2/3, for model 3 to zero, and consequently for model 1 to 1/3. The same remark can be made for model 2 and model 1 in the other crossing area defined for $t \simeq 25$ mn up to the end of the scenario. Note in these areas, the instant when models separate and the change in their corresponding probabilities.

3) Finally, as a conclusion of these remarks, we can analyze the track coalescence effects with respect to target powers. In the simulation, the most powerful track (model 2) is poorly estimated in comparison with weaker ones (model 1 and model 3). In other words, state estimate of the most powerful target is biased due to the bias of bearing estimates in crossing area. The probabilities of model 1 and model 3 are very weak in these areas, so their initial states are estimated based upon unbiased measurements. The less the difference between target powers is, the more significant the bias is. Moreover, this phenomena is increased with the status of "weak" estimability.

2) *Case 2: Multitarget Tracking in the Presence of False Detections*: In the previous section, the target's probability of detection, Pd, was assumed equal to unity and no false alarms were generated. In this section, the measurements that do not originate from targets are independently and uniformly distributed in the measurement space, denoted U_β. The number of false measurements in each scan followed a Poisson distribution with parameter λ, where λ was equal to the expected number of false alarms per unit volume. The
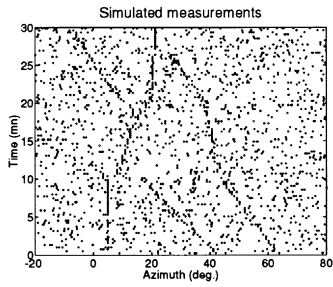
Fig. 5. Simulated measurements (case 2).

same measurement vector was used in this section with a target detection probability fixed to $Pd = 0.8$ for all the targets during the batch. Although this simulation does not integrate the detection step, it gives a good approximation of reality. Fig. 5 depicts the simulated measurements for the same scenario as in the previous section.

Let us now describe the modifications of the algorithm to include false alarms. Suppose the $(M-1)$ first models describe target models and the Mth one is devoted to false alarms. Since no kinematic parameter has to be estimated for such a model, the M-step of the algorithm always consists in the same number of optimizations. Moreover, for each target the same functional has to be maximized. It is straightforward to show that the modification of the algorithm only lies in the evaluation of the assignment probabilities. The conditional probability density of a measurement, given the previous parameter estimate and the model from which it is originated, is given below

$$
p(z_j(t) \mid x_k(t))
$$
$$
= \begin{cases} \dfrac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{ -\dfrac{1}{2\sigma^2}(z_j(t) - \beta(x_{k0},t))^2 \right\}, & 1 \le k < M \\ \dfrac{1}{U_\beta}, & k = M. \end{cases}
$$
(35)

Thus, the a posteriori probabilities that a measurement originates from a target conditioned on the previous parameter estimates is as follows:

by

$$
P_{Us} = \int_{U_\beta} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{ -\frac{1}{2\sigma^2}(z - \beta(x_{k0},t))^2 \right\} dz.
$$

For the given scenario, initiation of the algorithm is presented in Fig. 6 and the estimated trajectories in Fig. 7. This simulation demonstrates that the algorithm is very good at solving the data association even in densely cluttered environment. Although the global solution is reached in this example, the algorithm may converge to local solutions or stationary points of the likelihood function. It is worth stressing that poor initializations can lead to poor data-association estimates, and thus, poor target state estimates.

VII. EXTENSION TO MULTISENSOR DATA ASSOCIATION

From previous sections, the data association problem has focussed on temporal association of measurements. It is an easy task to extend the approach to spatial data association of measurements from multiple sensors. The purpose is to study if we favour the solution to the data association problem by providing informations from S sensors. We show that the incomplete data framework extend easily to multisensor tracking. For that purpose we introduce new notations:

- Suppose we have S sensors.
- $\mathcal{Z} = \{\mathcal{Z}_s\}_{s=1\ldots S}$ where $\mathcal{Z}_s$ denote the set of measurements received from sensor s.
- $\mathcal{X}$ denote always the state vector of the M models.
- $m_{s,t}$ is the number of measurements received by sensor s at time t.
- K_s is the assignment vector of measurements received by sensor s.
- Π denote assignment probabilities of the measurements to the M models.

We suppose that the set of measurements received from different sensors are statistically independent. As a consequence, we consider the assignment vector of measurements $\mathcal{Z}_s$ independent from the assignment vector of $\mathcal{Z}_{s'}$ $(s \ne s')$. Thus, the assignment problem becomes a 2-dimensional data association problem

$$
w_{j,k}^{i+1}(t) = \frac{\pi_k^i p(z_j(t) \mid x_k^i(t))}{p(z_j(t) \mid O^i(t))}
$$

$$
w_{j,k}^{i+1}(t) = \begin{cases} \dfrac{\dfrac{\pi_k^i(t)}{P_{Uk}} \exp\left\{ -\dfrac{1}{2\sigma^2}(z_j(t) - \beta(x_{k0},t))^2 \right\}}{\pi_M^i(t)\dfrac{(2\pi\sigma^2)^{1/2}}{U_\beta} + \sum_{s=1}^{M-1} \dfrac{\pi_s^i(t)}{P_{Us}} \exp\left\{ -\dfrac{1}{2\sigma^2}(z_j(t) - \beta(x_{s0},t))^2 \right\}}, & 1 \le k < M \\ \\ \dfrac{\pi_M^i(t)\dfrac{(2\pi\sigma^2)^{1/2}}{U_\beta}}{\pi_M^i(t)\dfrac{(2\pi\sigma^2)^{1/2}}{U_\beta} + \sum_{s=1}^{M-1} \dfrac{\pi_s^i(t)}{P_{Us}} \exp\left\{ -\dfrac{1}{2\sigma^2}(z_j(t) - \beta(x_{s0},t))^2 \right\}}, & k = M. \end{cases}
$$
(36)

In (36), P_{Us} is the normalization probability (i.e., the probability that the measurement originates from the given model and lies in the surveillance region) defined

(temporal and spatial). Moreover, we consider a centralized fusion where the assignment probabilities are estimated based on all the measurements received by the
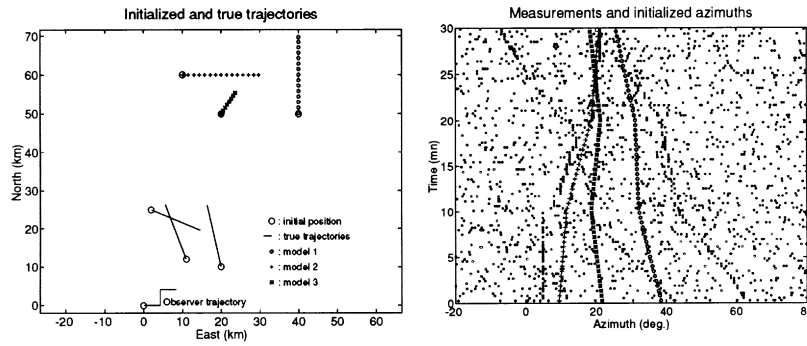
Fig. 6. Initialized trajectories and corresponding azimuths (case 2).
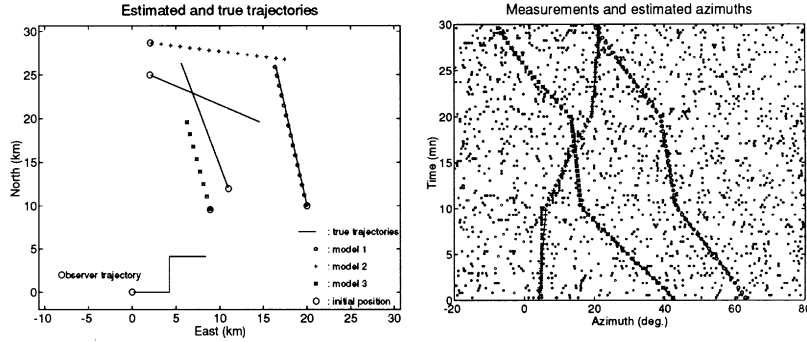

Fig. 7. Estimated trajectories and corresponding azimuths (case 2).

centralized system. From this viewpoint, the necessary likelihoods are redefined so as to derive EM algorithm. The likelihood of the complete data expresses as:

$$p(\mathcal{Z},\mathcal{K} \mid \mathcal{O}) = p(\mathcal{Z}_1 \ldots \mathcal{Z}_S, \mathcal{K}_1 \ldots \mathcal{K}_S \mid \mathcal{O})$$

$$= \prod_{s=1}^{S} p(\mathcal{Z}_s \mid \mathcal{K}_s, \mathcal{O}) p(\mathcal{K}_s \mid \mathcal{O}) \quad (37)$$

$$= \prod_{s=1}^{S} \prod_{t=1}^{T} \prod_{i_s=1}^{m_{s,t}} p(z_{i_s}(t) \mid x_{k_{i_s}}(t)) \pi_{k_{i_s}}(t). \quad (38)$$

Then, the likelihood of the incomplete data is obtained by extending (16) to S sensors:

$$p(\mathcal{Z} \mid \mathcal{O}) = \prod_{s=1}^{S} p(\mathcal{Z}_s \mid \mathcal{O})$$

$$= \prod_{s=1}^{S} \prod_{t=1}^{T} \prod_{i_s=1}^{m_{s,t}} p(z_{i_s}(t) \mid O(t))$$

$$= \prod_{s=1}^{S} \prod_{t=1}^{T} \prod_{i_s=1}^{m_{s,t}} \sum_{m=1}^{M} p(z_{i_s}(t) \mid x_m(t)) \pi_m(t). \quad (39)$$

Finally, the density of the missing data is deduced from (38) and (39):

$$p(\mathcal{K} \mid \mathcal{Z}, \mathcal{O}) = \frac{p(\mathcal{Z}, \mathcal{K} \mid \mathcal{O})}{p(\mathcal{Z} \mid \mathcal{O})}$$

$$= \prod_{s=1}^{S} \prod_{t=1}^{T} \prod_{i_s=1}^{m_{s,t}} p(k_{i_s}(t) \mid z_{i_s}(t), O(t)) \quad (40)$$

where

$$p(k_{i_s}(t) \mid z_{i_s}(t), O(t)) = \frac{p(z_{i_s}(t) \mid x_{k_{i_s}}(t)) \pi_{k_{i_s}}(t)}{\sum_{m=1}^{M} p(z_{i_s}(t) \mid x_m(t)) \pi_m(t)}. \quad (41)$$

Thus, the calulation of $Q(\mathcal{O} \mid \mathcal{O}^*)$ during E-step becomes:

$$Q(\mathcal{O} \mid \mathcal{O}^*) = \sum_{\mathcal{K}} \log[p(\mathcal{Z}, \mathcal{K} \mid \mathcal{O})] p(\mathcal{K} \mid \mathcal{Z}, \mathcal{O}^*)$$

$$= \sum_{\mathcal{K}} \left\{ \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{i_s=1}^{m_{s,t}} \log[p(z_{i_s}(t) \mid x_{k_{i_s}}(t)) \pi_{k_{i_s}}(t)] \right\}$$

$$\times \left\{ \prod_{s=1}^{S} \prod_{t=1}^{T} \prod_{i_s=1}^{m_{s,t}} p(k_{i_s}(t) \mid z_{i_s}(t), O(t)) \right\}.$$

The calculation derived in appendix extends straightforward from the previous expressions since the decomposition required for this calculation remains valid. The simplified expression is therefore:

$$Q(\mathcal{O} \mid \mathcal{O}^*) = \sum_{k=1}^{M} \sum_{t=1}^{T} \left[\sum_{s=1}^{S} \sum_{i_s=1}^{m_{s,t}} w_{i_s,k_{i_s}}(t) \right] \log[\pi_k(t)]$$

$$+ \sum_{k=1}^{M} \sum_{t=1}^{T} \sum_{s=1}^{S} \sum_{i_s=1}^{m_{s,t}} \log[p(z_{i_s}(t) \mid x_{k_{i_s}}(t))] w_{i_s,k_{i_s}}(t). \quad (42)$$

In (42), $w_{i_s,k_{i_s}}(t)$ is the weight of the measurement i_s from sensor s and for the model k_{i_s}, or more precisely the a posteriori probability to assign this measurement to
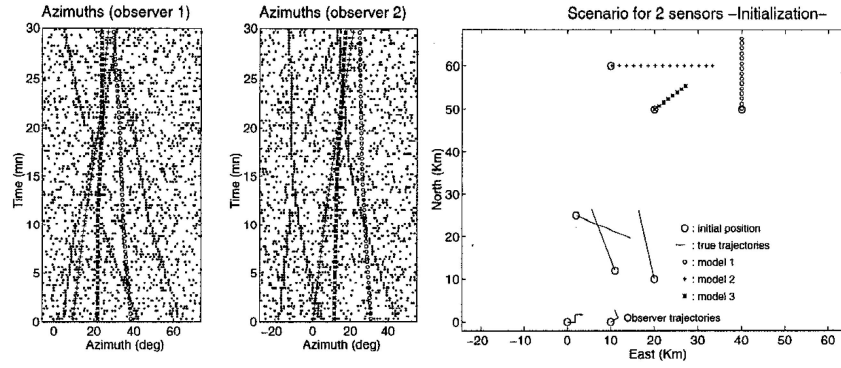
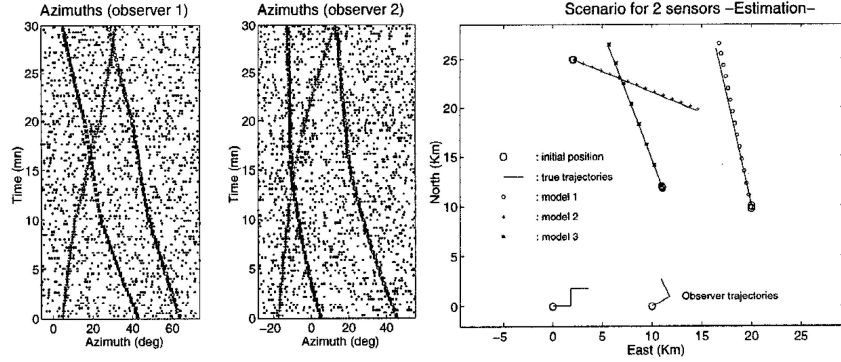Fig. 8. Measurements and initialized trajectories for two sensors.

Fig. 9. Estimated trajectories and their corresponding azimuths for two sensors.

model k_{i_s}:

$$w_{i_s,k_{i_s}}(t) = \frac{\pi_{k_{i_s}}^* \, p(z_{i_s}(t) \mid x_{k_{i_s}}^*(t))}{p(z_{i_s}(t) \mid O^*(t))}.$$

In the passive sonar context, if we suppose model M is devoted to false-alarms, we show that the updated probability of model k, $\pi_k(t)$, is obtained by summing the a posteriori probabilities to assign all the measurements within a scan and from all the sensors to model k

$$\pi_k^{i+1}(t) = \frac{1}{\left(\sum_{s=1}^{S} m_{s,t}\right)} \sum_{s=1}^{S} \sum_{i_s=1}^{m_{s,t}} w_{i_s,k_{i_s}}^{i+1}(t). \tag{43}$$

In (43), $w_{i_s,k_{i_s}}^{i+1}$ is calculated based on (36) where the measurement equation is replaced by the measurement equation of sensor s. Moreover, the maximization with respect to all the target state vectors is modified by adding the measurement equation for each sensor. We notice that the maximization of (42) can again be decomposed in a maximization for each model. Since sources are moving at constant velocity, the function to maximize for a model $k \in \{1 \ldots M-1\}$ defined by (33) is modified for S sensors in:

$$f(Z \mid x_0) = \sum_{t=1}^{T} \sum_{s=1}^{S} \sum_{i_s=1}^{m_{s,t}} \log[p(z_{s,i_s}(t) \mid x_0)] w_{i_s}^{i+1}(t) \tag{44}$$

$$= \sum_{t=1}^{T} \sum_{s=1}^{S} \sum_{i_s=1}^{m_{s,t}} -\frac{1}{2\sigma_s^2(t)} (z_{s,i_s}(t) - \beta_s(x_0,t))^2 w_{i_s}^{i+1}(t)$$

$$+ \cdots. \tag{45}$$

The gradient of (45) expressed in a vectorial form is

then:

$$\nabla_{x_0} f = \sum_{s=1}^{S} (R_{s,x_0}^{-1} A_{s,x_0})^T N_s^{-1} B_{s,x_0} \tag{46}$$

where the terms in (46) for a sensor s have the same meaning as in (34).

We apply again a Gauss–Newton algorithm modified by the Levenberg–Marquadt method to solve the maximization problem. Each iteration consists in computing:

$$x_0^{l+1} = x_0^l - \left[\sum_{s=1}^{S} A_{s,x_0^l}^T R_{s,x_0^l}^{-1} N_s^{-1} R_{s,x_0^l}^{-1} A_{s,x_0^l} + \mu_{x_0^l} \mathrm{Id} \right]^{-1}$$

$$\times \sum_{s=1}^{S} A_{s,x_0^l}^T R_{s,x_0^l}^{-1} N_s^{-1} B_{s,x_0^l}.$$

In summary, we have proved that the method extends easily to multisensor problems. More precisely, we have proposed a general framework to solve the difficult multitarget multisensor tracking problem. However, the numerous simulations to which the algorithm was submited underscore that adding new sensors do not contribute to the improvement of the solution of the data association problem. The algorithm converges regularly to some stationary points or local solutions preventing an accurate estimation of target parameters. We present a simulation on Figs. 8 and 9 where the algorithm converges to the global solution. This scenario was built based on the scenario of the previous section in the case of one sensor. Here, the detection probability was equal to unity and 15 false alarms were generated in each scan. It brings to us the advantage of adding another sensor in terms of the quality of target parameter estimation as soon as data association is correctly estimated.

VIII. CONCLUSION

The multitarget tracking has been considered in the natural framework of incomplete data problems. Based on the approach of Streit, et al., we have developed a multitarget multisensor tracking method for the passive sonar context. More precisely, the formalism of Streit, et al. has been extended by using the general framework of mixture parameter estimation. The assignment vector is not incorporated into the parameter vector but is interpreted as the missing data information. Thus, target state vector is estimated jointly with the probabilities that a measurement comes from a target. The use of the EM algorithm to compute the likelihood function or the MAP, allowed us to derive two different algorithms which appear very appealing because the global maximization with respect to all the target states is divided into M maximizations with respect to each target state, each coupled thanks to the weightings $w_j(t)$. An extension of the temporal data association to the spatial data association has been investigated. This framework allows us to formulate the multisensor multitarget tracking problem. Simulation results for realistic scenarios have been considered. They demonstrate the good ability of the PMHT algorithm to perform multitarget tracking in crossing areas. Such simulations permit us to study signal processing effects on multitarget tracking performance in an integrated way. Some more work need to be pursued in the evaluation of the tracking performance of the algorithm and in the improvement of its robustness.

APPENDIX A

This Appendix is devoted to the calculation of the expectation of the E-step in mixture densities. Extension to multitarget tracking is straightforward. At the E-step, we need to evaluate:

$$Q(\phi \mid \phi^i) = \sum_k \log[f(x \mid \phi)] p(k \mid y, \phi^i)$$

$$= \sum_{k_1=1}^{m} \cdots \sum_{k_N=1}^{m} \left\{ \sum_{j=1}^{N} \log[\pi_{k_j} p_{k_j}(y_j \mid \varphi_{k_j})] \right\}$$

$$\times \prod_{j'=1}^{N} \frac{\pi_{k_{j'}}^i p_{k_{j'}}(y_{j'} \mid \varphi_{k_{j'}}^i)}{p(y_{j'} \mid \phi^i)}.$$

This expression can be simplified by considering the N sums on one of the element of the expression inter braces:

$$Q(\phi \mid \phi^i)$$

$$= \sum_{j=1}^{N} \sum_{k_1=1}^{m} \cdots \sum_{k_N=1}^{m} \log[\pi_{k_j} p_{k_j}(y_j \mid \varphi_{k_j})]$$

$$\times \prod_{j'=1}^{N} \frac{\pi_{k_{j'}}^i p_{k_{j'}}(y_{j'} \mid \varphi_{k_{j'}}^i)}{p(y_{j'} \mid \phi^i)}$$

$$= \sum_{j=1}^{N} \left\{ \sum_{k_j=1}^{m} \log[\pi_{k_j} p_{k_j}(y_j \mid \varphi_{k_j})] \frac{\pi_{k_j}^i p_{k_j}(y_j \mid \varphi_{k_j}^i)}{p(y_j \mid \phi^i)} \right.$$

$$\left. \times \sum_{k_1=1}^{m} \cdots \sum_{k_{j-1}=1}^{m} \sum_{k_{j+1}=1}^{m} \cdots \sum_{k_N=1}^{m} \prod_{\substack{j'=1 \\ j' \neq j}}^{N} \frac{\pi_{k_{j'}}^i p_{k_{j'}}(y_{j'} \mid \varphi_{k_{j'}}^i)}{p(y_{j'} \mid \phi^i)} \right\}. \tag{47}$$

The previous expression comes from a factorization of the terms in k_j.

Using (7), the calculation of B_j defined as

$$B_j = \sum_{k_1=1}^{m} \cdots \sum_{k_{j-1}=1}^{m} \sum_{k_{j+1}=1}^{m} \cdots \sum_{k_N=1}^{m} \prod_{\substack{j'=1 \\ j' \neq j}}^{N} \frac{\pi_{k_{j'}}^i p_{k_{j'}}(y_{j'}(y_{j'} \mid \varphi_{k_{j'}}^i)}{p(y_{j'} \mid \phi^i)}$$

is an easy task since

$$B_j = \sum_{k_1=1}^{m} \cdots \sum_{k_{j-1}=1}^{m} \sum_{k_{j+1}=1}^{m} \cdots \sum_{k_N=1}^{m}$$

$$\times p(k_1, \ldots, k_{j-1}, k_{j+1}, \ldots, k_N \mid y, \phi^i) = 1 \tag{48}$$

and then by substituting (48) in (47) and inverting sums yields (9).

ACKNOWLEDGMENTS

REFERENCES

[1]   Avitzour, D. (1992)
       A maximum likelihood approach to data association.
       IEEE Transactions on Aerospace and Electronics Systems, **28**, 2 (Apr. 1992), 560–565.
[2]   Bar-Shalom, Y. (Ed.) (1990)
       Multitarget-Multisensor Tracking: Advanced Applications.
       Dedham, MA: Artech House, 1990.
[3]   Castañon, D. A. (1992)
       New assignment algorithms for data association.
       SPIE Proceedings, **1698-313**, Orlando, FL, Apr. 1992.
[4]   Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977)
       Maximum likelihood from incomplete data via the EM algorithm.
       Journal of Royal Statistical Society, Series B, **39** (1977), 1–38.
[5]   Dennis, J. E., Jr., and Schnabel, R. B. (1983)
       Numerical Methods for Unconstrained Optimization and Nonlinear Equations.
       Englewood Cliffs, NJ: Prentice Hall, 1983.
[6]   Fortmann, T. E., Bar-Shalom, Y., and Scheffe, M. (1983)
       Sonar tracking of multiple targets using joint probabilistic data association.
       IEEE Journal of Oceanic Research, **OE-8** (July 1983), 173–184.
[7]   Jauffret, C., and Bar-Shalom, Y. (1990)
       Track formation with bearing and frequency measurements in the presence of false detection.
       IEEE Transactions on Aerospace and Electronics Systems, **26**, 6 (1990), 999–1010.
[8]   Le Cadre, J. P., and Zugmeyer, O. (1993)
       Temporal integration for array processing.
       Journal of the Acoustical Society of America, **3** (Mar. 1993), 1471–1481.
[9]   Louis, T. A. (1982)
       Finding the observed information matrix when using the EM algorithm.
       Journal of the Royal Statistical Society, Ser. B, **44**, 2 (1982), 226–233.

[10] Morefield, C. L. (1977)
Application of 0–1 integer programming to multitarget tracking problems.
IEEE Transactions on Automatic Control, **AC-22**, 3 (June 1977), 302–311.

[11] Mori, S., Chong, C., Tse, E., and Wishner, R. P. (1986)
Tracking and classifying multiple targets without a priori identification.
IEEE Transactions on Automatic Control, **AC-31**, 5 (May 1986), 401–409.

[12] Nardone, S. C., Lindgren, A. G., and Gong, K. F. (1984)
Fundamental properties and performance of conventional bearings-only target motion analysis.
IEEE Transactions on Automatic Control, **AC-29**, 9 (Sept. 1984), 775–787.

[13] Pattipati, K. R., Deb, S., Bar-Shalom, Y., and Washburn, R. B. (1992)
A new relaxation algorithm and passive sensor data association.
IEEE Transactions on Automatic Control, **AC-37**, 2 (Feb. 1992), 198–213.

[14] Poore, A. B., and Rijavec, N. (1991)
Multi-target tracking and multi-dimensional assignment problems.
In *Proceedings of SPIE International Symposium, Signal and Data Processing of Small Targets 1991*, Vol. 1481, Orlando, FL, Apr. 1991.

[15] Rago, C., Willett, P., and Streit, R. (1995)
A comparison of the JPDAF and PMHT tracking algorithms.
In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, **5** (May 9–12, 1995), 3571–3574.

[16] Reid, D. B. (1979)
An algorithm for tracking multiple targets.
IEEE Transactions on Automatic Control, **AC-24**, 6 (Dec. 1979), 843–854.

[17] Redner, R. A., and Walker, H. F. (1984)
Mixture densities, maximum likelihood and the EM algorithm.
Society for Industrial and Applied Mathematics, **26**, 2 (Apr. 1984).

[18] Streit, R. L., and Luginbuhl, T. E. (1993)
A probabilistic multi-hypothesis tracking algorithm without enumeration and pruning.
In *Proceedings of the Sixth Joint Service Data Fusion Symposium*, Laurel, MD, June 14–18, 1993, 1015–1024.

[19] Streit, R. L., and Luginbuhl, T. E. (1994)
Maximum likelihood method for probabilistic multi-hypothesis tracking.
In *Proceedings of SPIE International Symposium, Signal and Data Processing of Small Targets 1994*, SPIE Proceedings Vol. 2335-24, Orlando, FL, Apr. 5–7, 1994.

[20] Streit, R. L., and Luginbuhl, T. E. (1994)
Maximum likelihood training of probabilistic neural networks.
IEEE Transactions on Neural Networks, **5**, 5 (Sept. 1994), 764–783.

[21] Streit, R. L., and Luginbuhl, T. E. (1995)
Probabilistic multi-hypothesis tracking.
Technical report 10428, Naval Undersea Warfare Center, Newport, RI, Feb. 15, 1995.

[22] Sundberg, R. (1976)
An iterative method for solution of the likelihood equations for incomplete data from exponential families.
Comm. Statist. Simulation Comput., **B5** (1976), 55–64.

[23] Wu, C. F. (1983)
On the convergence of the EM algorithm.
Annals of Statistics, **11**, 1 (1983), 95–103.
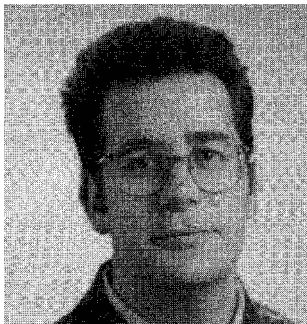
Hervé Gauvrit graduated from the *Ecole Supérieure d'Electronique de l'Ouest*, Angers, in 1992, and received his *Diplôme d'Etude Approfondie* in signal processing from the University of Rennes 1 in 1994.

He joined IRISA in 1994 and is currently in the Ph.D. program at the University of Rennes 1. His current research interests are in data association, estimation theory and its applications in multitarget multisensor tracking.
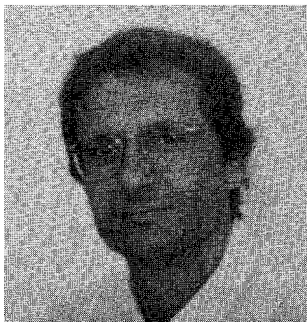
Claude Jauffret was born on March 29, 1957, in Ollioules, France. He received the Diplôme d'Etudes Approfondies in applied mathematics from the Saint Charles University, Marseille, France in 1981 and the Diplôme D'ingénieur from the Ecole Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble in 1983, the title of "docteur de l'université de Toulon" in 1993 and the "habilitation à diriger des recherches" in 1996 from the University of Toulon.

From Nov. 1983 to Sept. 1984, he worked on image processing used in passive sonar systems at the GERDSM, Six-Fours Les Plages, France. From 1984 to 1988, he worked on target motion analysis (TMA) problems at the GERDSM. A sabbatical year at the University of Connecticut, Storrs, allowed him to work on tracking problems in a cluttered environment in 1989. His research interests are mainly estimation and detection theory.

He has published papers about observability and estimation in nonlinear systems and TMA. Since September 1996, he is with the University of Toulon (France) where he is a Professor.

J. P. Le Cadre (M'93) received the M.S. degree in mathematics in 1977, the Doctorat de 3^{-eme} cycle in 1982 and the Doctorat d'Etat in 1987, both from INPG, Grenoble, France.

From 1980 to 1989, he worked at the GERDSM (Groupe d'Etudes et de Recherche en Detection Sous-Marines), a laboratory of the DCN (Direction des Constructions Navales), mainly on array processing. Since 1989, he has been with IRISA/CNRS, where he is a director of research. His interests have now moved toward topics like system analysis, detection, data association and operations research.

He received (with O. Zugmeyer) the Eurasip *Signal Processing* best paper award in 1993. He is a member of various societies of IEEE and of SIAM.