

# The Probabilistic Data Association Filter

## ESTIMATION IN THE PRESENCE OF MEASUREMENT ORIGIN UNCERTAINTY

YAAKOV BAR-SHALOM, FRED DAUM, and JIM HUANG



**D**ata association uncertainty occurs when remote sensing devices, such as radar, sonar, or electro-optical devices, yield measurements whose origin is uncertain, that is, not necessarily the target of interest. This uncertainty occurs when the signal from the target is weak and, to detect it, the detection threshold has to be lowered, in which case background signals and sensor noise may also be detected, yielding clutter or spurious measurements. This situation can also occur when several targets are present in the same neighborhood. Using spurious measurements in a tracking filter leads to divergence of the estimation error and, thus, track loss. Consequently, the first problem is to select measurements to be used to update the state of the target of interest in the tracking filter, which can be a Kalman filter (KF) or an

extended KF (EKF). The second problem is to determine whether the filter has to be modified to account for this data association uncertainty. The goal is to obtain the minimum mean square error (MMSE) estimate of the target state and the associated uncertainty.

The literature on tracking targets in the presence of data association uncertainty is extensive [1]–[4]. This tutorial starts with an illustration of the data association uncertainty stemming from the ambiguity as to which measurements are appropriate for use in the tracking filter. Then we present a brief discussion on the optimal state-estimation algorithm in the MMSE sense in the presence of data association uncertainty and how other approaches stand in relationship to the optimum. The optimal estimator consists of the recursive computation of the conditional probability density function



PHOTOS  
COURTESY OF  
NASA KENNEDY  
SPACE CENTER  
(NASA-KSC)

Digital Object Identifier 10.1109/MCS.2009.934469

(pdf) of the state. The conditions under which this pdf is a sufficient statistic in the presence of data association uncertainty are detailed. We then discuss several practical algorithms that carry out estimation and data association, namely, the probabilistic data association filter (PDAF), joint PDAF (JPDAF), mixture reduction PDAF (MXPDAF), particle filter (PF), and multiple hypothesis tracker (MHT), together with the approximations made by these algorithms. PDAF and JPDAF are then described in detail along with two illustrative examples that show the reduced track loss probability of PDAF compared to KF in the presence of false measurements. Finally, several real-world applications of PDAF and JPDAF are discussed, together with some lessons learned from the selection and design of tracking and data association algorithms.

Table 1 lists the acronyms used in this article.

## THE DATA ASSOCIATION PROBLEM IN TARGET TRACKING

In target-tracking applications, the signal-detection process yields measurements from which the measurements to be incorporated into the target state estimator are selected. In a radar, the reflected signal from the target of interest is sought within a time interval determined by the anticipated range of the target when it reflects the energy transmitted by the radar. A *range gate* is set up, and the detections within this gate can be associated with the target of interest. These measurements can consist of range, azimuth, elevation, or direction cosines, possibly also range rate for radar or active

sonar; bearing, possibly also frequency, time difference of arrival, and frequency difference for passive sonar; line-of-sight angles or direction cosines for optical sensors. A multidimensional gate is then set up for detecting the signal from the target. This procedure avoids searching for the signal from the target of interest in the entire measurement space. However, a measurement in the gate, while not guaranteed to have originated from the target associated with the gate, is a valid association candidate, and such a gate is called a *validation region*.

The validation region is set up to guarantee that the target measurement falls in it with high probability, called the *gate probability*, based on the statistical characterization of the predicted measurement. If more than one detection appears in the gate, then we face an association uncertainty, namely, we must determine which measurement is target originated and thus to be used to update the track, which consists of the state estimate and covariance or, more generally, the sufficient statistic for this target. Measurements outside the validation region can be ignored because they are too far from the predicted measurement and thus unlikely to have originated from the target of interest. This situation occurs when the gate probability (the probability that the target-originated measurement falls in the gate) is close to unity and the statistical model used to obtain the gate is correct.

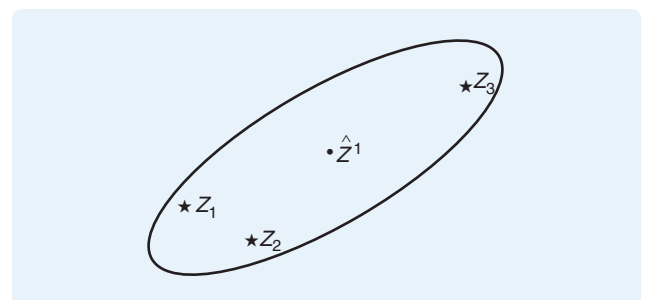
### A Single Target in Clutter

The problem of tracking a single target in clutter arises when several measurements occur in the validation region. The set of validated measurements consists of the correct measurement, if detected in this region, and the spurious measurements, which are clutter or false-alarm originated. In air traffic control systems, with cooperative targets, each measurement also includes a target identifier, called the squawk number. If the identifier is perfectly reliable, there is no data association uncertainty. However, a potentially hostile target is not expected to be cooperative, and then data association uncertainty is a problem.

A situation with several validated measurements is depicted in Figure 1. The two-dimensional validation

**TABLE 1** List of acronyms.

C/FA	Clutter/false alarms
C-K	Chapman-Kolmogorov
EKF	Extended Kalman filter
HOMHT	Hypothesis-oriented MHT
KF	Kalman filter
JPDAF	Joint probabilistic data association coupled filter
JPDAF	Joint probabilistic data association filter
JVC	Jonker-Volgenant-Castanon
LR	Likelihood ratio
MAP	Maximum a posteriori
MHT	Multiple hypothesis tracker
ML	Maximum likelihood
MMSE	Minimum mean square error
MXPDAF	Mixture probabilistic data association filter
NN	Nearest neighbor
NNSF	Nearest neighbor standard filter
OTHR	Over-the-horizon radar
PDAF	Probabilistic data association filter
pdf	Probability density function
PF	Particle filter
pmf	Probability mass function
TOMHT	Track-oriented MHT



**FIGURE 1** Several measurements  $z_i$  in the validation region of a single target. The validation region is an ellipse centered at the predicted measurement  $\hat{z}^1$ . Any of the shown measurements could have originated from the target or none if the target is not detected.

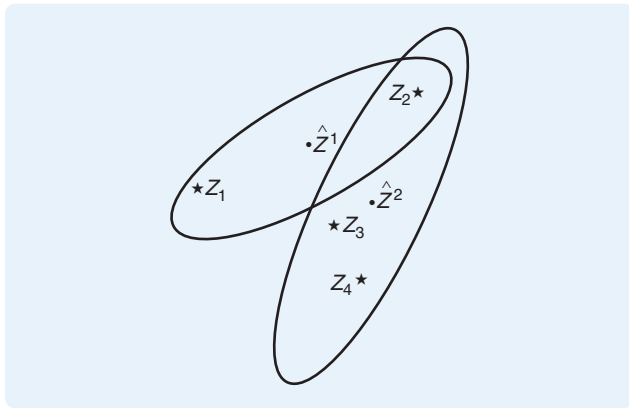
region in Figure 1 is an ellipse centered at the predicted measurement  $\hat{z}$ . The elliptical shape of the validation region is a consequence of the assumption that the error in the target's predicted measurement, that is, the innovation, is Gaussian. The parameters of the ellipse are determined by the covariance matrix  $S$  of the innovation.

All the measurements in the validation region may have originated from the target of interest, even though at most one is the true one. Consequently, the set of possible *association events* are the following:  $z_1$  originated for the target, and  $z_2$  and  $z_3$  are spurious;  $z_2$  originated for the target, and  $z_1$  and  $z_3$  are spurious;  $z_3$  originated for the target, and  $z_2$  and  $z_1$  are spurious; all are spurious. The fact that these association events are mutually exclusive and exhaustive allows the use of the total probability theorem for obtaining the state estimate in the presence of data association uncertainty.

Under the assumption that there is a single target, the spurious measurements constitute a *random interference*. A typical model for such false measurements is that they are uniformly spatially distributed and independent across time. This model corresponds to residual clutter. The constant clutter, if any, is assumed to have been removed.

### Multiple Targets in Clutter

When several targets as well as clutter or false alarms are present in the same neighborhood, the data association becomes more complicated. Figure 2 illustrates this case, where the predicted measurements for the two targets are denoted as  $\hat{z}^1$  and  $\hat{z}^2$ . In Figure 2 three measurement origins, or association events, are possible,  $z_1$  from target 1 or clutter;  $z_2$  from either target 1 or target 2 or clutter; and  $z_4$  and  $z_5$  from target 2 or clutter. However, if  $z_2$  originated from target 2 then it is likely that  $z_1$  originated from target 1. This situation illustrates the interdependence of the



**FIGURE 2** Two targets with a measurement  $z_2$  in the intersection of their validation regions. The validation regions are the ellipses centered at the predicted measurements  $\hat{z}^1$  and  $\hat{z}^2$ . Each of the measurements in the validation region of one of targets could have originated from the corresponding target or from clutter. The measurement  $z_2$  in the intersection of the validation regions could have originated from either target or both.

associations in a situation where a *persistent interference* from a neighboring target is present in addition to random interference or clutter. In this case, *joint association events* must be considered.

A more complicated situation can arise as follows. Up to this point each measurement is assumed to originate from either one of the targets or from clutter. However, in view of the fact that any signal processing system has an inherent finite resolution capability, an additional possibility has to be considered, that  $z_2$  could be the result of the merging of the detections from the two targets, namely, that this measurement is an unresolved one. The unresolved measurement constitutes a fourth origin hypothesis for a measurement that lies in the intersection of two validation regions.

The above discussion illustrates the difficulty of associating measurements to tracks. The full problem, as discussed below, consists of associating measurements at each time instant, updating the track sufficient statistic, and propagating it to the next time instant.

## STATE ESTIMATION IN THE PRESENCE OF DATA ASSOCIATION UNCERTAINTY

When estimating the states of several targets using measurements with uncertain origin, it is not known which measurement originated from which target or from clutter. The goal is to obtain the MMSE estimate of the vector  $x$ , of known dimension, which might be the state of a single target or a stacked vector consisting of the states of several targets. The approaches to target tracking in the presence of data association uncertainty can be classified into several categories, discussed below.

### Pure MMSE approach

The pure MMSE approach to tracking and data association is obtained using the smoothing property of expectations (see [5, Sec. 1.4.12]). In other words, the conditional mean of the state is obtained by averaging over all the association events as

$$\begin{aligned}\hat{x}^{\text{MMSE}} &= E[x | Z] = E\{E[x | A, Z] | Z\} \\ &= \sum_{A_i \in \mathcal{A}} E[x | A_i, Z] P\{A_i | Z\} \\ &\triangleq \sum_{i=1}^M \hat{x}_i \beta_i,\end{aligned}\quad (1)$$

where  $Z$  is the available measurement data,  $A_i$  is an association event with  $\hat{x}_i$  the corresponding state estimate, and the summation is over all events  $A_i$  in the set  $\mathcal{A}$  of  $M$  mutually exclusive and exhaustive association events. A Bayesian model is assumed, with prior probabilities, from which posterior probabilities can be calculated. The expression given in (1), which requires the evaluation of all the conditional, or posterior, probabilities

$$\beta_i \triangleq P\{A_i | Z\}, \quad (2)$$

is a direct consequence of the total probability theorem (see [5, Sec. 1.4.10]).

The covariance associated with the mixed estimate (1) is, according to the moment equations for mixture pdfs (see [5, Sec. 1.4.16]), given by

$$P^{\text{MMSE}} = \text{cov}[x|Z] = \sum_{A_i \in \mathcal{A}} [P_i + (\hat{x}_i - \hat{x}^{\text{MMSE}})(\hat{x}_i - \hat{x}^{\text{MMSE}})'] \beta_i, \quad (3)$$

where  $(\cdot)'$  denotes transpose and

$$P_i \triangleq \text{cov}[x|A_i, Z]. \quad (4)$$

The conditional mean (1) and the conditional covariance (3) require the conditional pdf

$$p(x|Z) = \sum_{A_i \in \mathcal{A}} p(x|A_i, Z) \beta_i, \quad (5)$$

which is a mixture of association-conditioned pdfs.

### MMSE-MAP Approach

The MMSE-MAP approach, instead of enumerating and summing over all the association events, selects the one with highest posterior, or maximum a posteriori (MAP), probability, namely,

$$A^{\text{MAP}} = A_{i_{\text{MAP}}}, \quad (6)$$

where

$$i_{\text{MAP}} = \arg \max_i P\{A_i|Z\}, \quad (7)$$

and then

$$\hat{x}^{\text{MMSE-MAP}} = E[x|A^{\text{MAP}}, Z]. \quad (8)$$

Obtaining the MAP event in (7) requires the evaluation of the posterior probabilities of all the association events. This evaluation is done using Bayes's formula

$$P\{A_i|Z\} = \frac{p(Z|A_i)P\{A_i\}}{p(Z)}, \quad (9)$$

starting with the prior probabilities  $P\{A_i\}$ . Note that  $p(Z)$  is independent of  $i$ , and thus (7) is equivalent to

$$i_{\text{MAP}} = \arg \max_i [p(Z|A_i)P\{A_i\}]. \quad (10)$$

The covariance associated with (8) is

$$P^{\text{MMSE-MAP}} \triangleq \text{cov}[x|A^{\text{MAP}}, Z] = P_{i_{\text{MAP}}}, \quad (11)$$

which is conditioned on event  $A_{i_{\text{MAP}}}$  being the correct one.

### MMSE-ML Approach

The MMSE-ML approach does not assume priors for the association events and relies on the maximum likelihood (ML) approach to select the event, that is,

$$A^{\text{ML}} = A_{i_{\text{ML}}}, \quad (12)$$

where

$$i_{\text{ML}} = \arg \max_i p(Z|A_i), \quad (13)$$

and  $p(Z|A_i)$  is the likelihood of event  $A_i$ , given by the pdf of the observations conditioned on  $A_i$ . It then follows that

$$\hat{x}^{\text{MMSE-ML}} = E[x|A^{\text{ML}}, Z]. \quad (14)$$

Note that the MAP approach (10) coincides with the ML approach (13) for a uniform prior for the association events  $A_i \in \mathcal{A}$ .

Similarly to (11), the covariance associated with (14) is

$$P^{\text{MMSE-ML}} \triangleq \text{cov}[x|A^{\text{ML}}, Z] = P_{i_{\text{ML}}}. \quad (15)$$

This covariance is conditioned on the selected association (12) being the correct one.

The MMSE-MAP estimate (8) and the MMSE-ML estimate (14) are obtained assuming that the selected association is correct, that is, both rely on a *hard decision*. This hard decision is sometimes correct and sometimes wrong. On the other hand, the pure MMSE estimate (1) relies on a *soft decision* since it averages over all the possibilities. This soft decision is never totally correct but never totally wrong. The uncertainty, quantified as covariances associated with the estimates (8) and (14), might be optimistic in view of the above observation. The covariance associated with the estimate (1) is increased in view of the fact that it includes the data association uncertainty. This increase consists of the spread of the means term given by the dyads on the right-hand side of (3), which are positive semidefinite.

### Heuristic Approaches

In addition to the above approaches, some simpler heuristic techniques are available. The simplest technique relies on the Mahalanobis distance metric, which is the square of the norm of the error with respect to its covariance, see [5, Sec. 5.4], for sequential measurement-to-track association and [3, Sec. 8.4], for track-to-track association. This approach associates each measurement with a track based on the nearest neighbor (NN) rule, which is a local or greedy selection [1], [4]. The same criterion can be used in a global cost function, which is minimized by an assignment algorithm using binary optimization, which can be the Jonker-Volgenant-Castanon algorithm (JVC) [6] or auction [7]. These approaches result in the local and global nearest neighbor standard filter (NNSF) approaches, respectively. This filter is designated as standard because it assumes the assignment as perfect, without accounting for the possibility that it might be erroneous.



## ESTIMATION AND DATA ASSOCIATION IN NONLINEAR DYNAMIC STOCHASTIC SYSTEMS

### The Model

We consider the discrete-time stochastic system

$$x(k+1) = f[k, x(k), u(k), v(k)], \quad (16)$$

where  $x \in \mathcal{R}^n$  is the state vector of the targets under consideration,  $u$  is a known input (included here for the sake of generality), and  $v$  is the process noise with a known pdf. The measurements at time  $k+1$  are described by the vector

$$z(k+1) = h[k+1, x(k+1), A(k+1), w(k+1)], \quad (17)$$

where  $A(k+1)$  is the data association event at  $k+1$  that specifies i) which measurement component originated from which components of  $x(k+1)$ , namely, from which target, and ii) which measurements are false, that is, originated from the clutter process. The vector  $w(k)$  is the observation noise, consisting of the error in the true measurement and the false measurements. The pdf of the false measurements and the probability mass function (pmf) of their number are also assumed to be known.

The noise sequences and false measurements are assumed to be white with known pdf and mutually independent. The initial state is assumed to have a known pdf and to be independent of the noises. Additional assumptions are given below for the optimal estimator, which evaluates the pdf of the state conditioned on the observations.

The optimal state estimator in the presence of data association uncertainty consists of the computation of the conditional pdf of the state  $x(k)$  given all the information available at time  $k$ , namely, the prior information about the initial state, the intervening inputs, and the sets of measurements through time  $k$ . The conditions under which the optimal state estimator consists of the computation of this pdf are presented in detail.

### The Optimal Estimator for the Pure MMSE Approach

The information set available at  $k$  is

$$I^k = \{Z^k, U^{k-1}\}, \quad (18)$$

where

$$Z^k \triangleq \{z(j)\}_{j=1}^k \quad (19)$$

is the cumulative set of observations through time  $k$ , which subsumes the initial information  $Z^0$ , and  $U^{k-1}$  is the set of known inputs prior to time  $k$ .

For a stochastic system, an information state [8] is a function of the available information set that summarizes the past of the system in a probabilistic sense. In other

words, the information state is an exact substitute for the conditioning on the past data in the pdf of every present and future quantity related to the state of the system. We show below that the conditional pdf of the state

$$p_k \triangleq p[x(k) | I^k] \quad (20)$$

is an information state if i) the two noise sequences (process and measurement) are white and mutually independent, and ii) the target detection and clutter/false measurement processes are white. Once the conditional pdf (20) is available, the pure MMSE estimate, that is, the conditional mean, as well as the conditional variance, or covariance matrix, can be obtained.

The optimal estimator, which consists of the recursive functional relationship between the information states  $p_{k+1}$  and  $p_k$ , is given by

$$p_{k+1} = \psi[k+1, p_k, z(k+1), u(k)], \quad (21)$$

where

$$\begin{aligned} \psi[k+1, p_k, z(k+1), u(k)] &= \frac{1}{c} \sum_{i=1}^{M(k+1)} p[z(k+1) | x(k+1), A_i(k+1)] \\ &\times \int p[x(k+1) | x(k), u(k)] p_k dx(k) \\ &\times P\{A_i(k+1)\} \end{aligned} \quad (22)$$

is the transformation that maps  $p_k$  into  $p_{k+1}$ ; the integration in (22) is over the range of  $x(k)$  and  $c$  is the normalization constant. The derivation of this functional recursion is given in "The Recursion of the Conditional Density of the State in the Presence of Data Association Uncertainty."

Recursion (22) shows that the optimal MMSE estimator in the presence of data association uncertainty has the following properties:

- » P1: The pdf  $p_{k+1}$  is a weighted sum of pdfs, conditioned on the current time association events  $A_i(k+1)$ ,  $i = 1, \dots, M(k+1)$ , where  $M(k+1)$  is the number of mutually exclusive and exhaustive association events at time  $k+1$ .
- » P2: If the exact previous pdf, which is the sufficient statistic, is available, then only the most recent association event probabilities are needed at each time.

However, the number of terms of the mixture in the right hand side of (22) is, by time  $k+1$ , given by the product

$$M^{k+1} = \prod_{i=1}^{k+1} M(i), \quad (23)$$

which amounts to an exponential increase in time. This increase is similar to the increase in the number of the branches of the MHT hypothesis tree. Mixture (22) is equivalent to the MHT hypothesis tree.

## The Conditional Density of the State as the Information State

We now show that  $p_k$  is an information state, that is, the pdf of every future state at  $j > k$  depends only on  $p_k$  and the intervening known inputs. These inputs are denoted as

$$U_k^{j-1} \triangleq \{u(i)\}_{i=k}^{j-1}. \quad (24)$$

This result is shown below to follow from the total probability theorem (see [5, Eq. (1.B.10-10)]). For  $j > k$  the state prediction pdf is

$$p[x(j) | I^k, U_k^{j-1}] = \mu[j, p_k, U_k^{j-1}], \quad (25)$$

where the transformation that maps  $p_k$  and the known input into the pdf of a future state  $x(j)$  is given by

$$\begin{aligned} \mu[j, p_k, U_k^{j-1}] &= \int p[x(j) | x(k), I^k, U_k^{j-1}] p[x(k) | I^k] dx(k) \\ &= \int p[x(j) | x(k), U_k^{j-1}] p_k dx(k). \end{aligned} \quad (26)$$

In (26) the whiteness of the process noise sequence and its independence from the measurement noises are used.

Thus, from (26) it follows that  $p_k$  is sufficient to predict the pdf of every future state. Since (22) shows that  $p_k$  can be obtained recursively, and thus is computable, it follows that  $I^k$  is summarized by  $p_k$ . Therefore, the whiteness and mutual independence of i) the two noise sequences and ii) the detection and the clutter processes are a sufficient condition for  $p_k$  to be an information state. We emphasize that whiteness is the crucial assumption.

The above conditions are equivalent to the requirement that  $x(k)$  be an incompletely observed Markov process, that is, a Markov process observed partially or in the presence of noise. If, for example, the process noise sequence is not white, then  $p_k$  does not summarize the past data. In this case the vector  $x$  is not a state, and thus it has to be augmented. This discussion explains why the formulation of stochastic estimation and control problems is done with mutually independent white noise sequences.

## PRACTICAL ESTIMATORS

Two of the practical estimators that handle data association are PDAF for a single target in clutter and JPDAF for multiple targets. These estimators are based on the pure MMSE approach, but they are suboptimal. The suboptimality of PDAF follows from the fact that it approximates the conditional pdf of the target's state at every stage as a Gaussian with moments matched to the mixture as in (5). The exact pdf under the linear-Gaussian assumption for the target and its measurement model, and with uniformly and independently distributed clutter (false measurements), is a Gaussian mixture with an exponentially

growing number of terms. This mixture follows directly from (22).

Similarly, JPDAF [3, Sec. 6.2] approximates each target state as an independent random variable with Gaussian pdf. The coupled JPDAF [3, Sec. 6.2.7] is closer to the optimum by approximating the joint pdf of the targets' state vectors as a Gaussian. The effectiveness of this single-Gaussian approximation of the Gaussian mixture is illustrated in [3] and by the fielded systems that utilize it, as described in the sequel.

The approach of [9] is to improve on the PDAF by keeping a limited number of terms in the Gaussian mixture, resulting in the mixture PDAF (MXPDAF). The particle filter approach is used in [10] to approximate the conditional-state pdf using a weighted sum of delta functions (particles). The resulting JPDA is shown to handle multiple maneuvering targets in clutter. The key is a suitable control that keeps the number of particles limited while approximating the state's mixture pdf well enough.

The MHT, which makes hard association decisions over multiple frames of measurements, that is, on sequences of association hypotheses, has two versions: i) the hypothesis oriented MHT (HOMHT) [11] (see also [3]), which uses the MMSE-MAP approach, and ii) the track oriented MHT (TOMHT) [12], [1], [4], which uses the MMSE-ML approach. For details, see [13].

A common feature of both MHT algorithms is the exponentially increasing number of hypotheses, which necessitates pruning. This pruning is accomplished by a sliding window implementation as well as discarding low probability/likelihood hypotheses.

The MHT uses a Gaussian posterior of the target states at the rear end of its sliding window, conditioned on the chosen hypothesis up to that point in time (MAP for the HOMHT, ML for the TOMHT), while it ignores all other association hypotheses. This use of a Gaussian posterior implicitly assumes a dominant hypothesis, namely, that the data association uncertainties from the past have been resolved. However, within its window, the MHT can carry a large number of association hypotheses, whose probabilities or likelihoods are recomputed, and the hypotheses are propagated in time as the window slides forward. Another advantage of the MHT is the fact that, by making hard decision associations, this method can use the measurements unassociated with existing tracks to start new tracks.

## THE PDAF

### Overview of PDA

The PDA algorithm calculates the association probabilities to the target being tracked for each validated measurement at the current time. This probabilistic or Bayesian information is used in the PDAF tracking algorithm, which accounts for the measurement origin uncertainty. Since the state and measurement equations are assumed to be linear, the

## The Recursion of the Conditional Density of the State in the Presence of Data Association Uncertainty

To obtain a recursion for the conditional density of  $x(k+1)$ , we rewrite (22) using the total probability theorem (see [5, Eq. (1.4.10-6)]) and Bayes's formula as

$$\begin{aligned}
 p_{k+1} &\triangleq p[x(k+1)|I^{k+1}] \\
 &= \sum_{i=1}^{M(k+1)} p[x(k+1)|A_i(k+1), I^{k+1}] P\{A_i(k+1)|I^{k+1}\} \\
 &= \sum_{i=1}^{M(k+1)} p[x(k+1)|A_i(k+1), I^k, z(k+1), u(k)] \beta_i(k+1) \\
 &= \sum_{i=1}^{M(k+1)} \frac{1}{c} p[z(k+1)|x(k+1), A_i(k+1), I^k, u(k)] \\
 &\quad \cdot p[x(k+1)|A_i(k+1), I^k, u(k)] \beta_i(k+1) \\
 &= \frac{1}{c} \sum_{i=1}^{M(k+1)} p[z(k+1)|x(k+1), A_i(k+1), I^k, u(k)] \\
 &\quad \times p[x(k+1)|I^k, u(k)] \beta_i(k+1), \tag{S1}
 \end{aligned}$$

where  $I^{k+1}$  is defined in (18),  $A_i$  is given by (1),  $c$  is the normalization constant (to avoid additional notational complexity, the normalization constants in different expressions are denoted by the same symbol), and

$$\beta_i(k+1) \triangleq P\{A_i(k+1)|I^{k+1}\} \tag{S2}$$

is the posterior probability of the  $i$ th joint association event at  $k+1$ . The summation in (S1) is over the  $M(k+1)$  mutually exclusive and exhaustive association events at  $k+1$ . In the last line of (S1) we use the fact that the state prediction pdf from  $k$  to  $k+1$  is independent of the association events at  $k+1$ .

The first term in the last line of (S1) is the pdf of the set of measurements  $z(k+1)$ , conditioned on the state at  $k+1$ , the association event  $A_i(k+1)$ , and the past information. The randomness of  $z(k+1)$  is, according to (17), due to i) the association event, which manifests itself in the detection process, namely, which measurement originates from which component of the state  $x$ , that is, from which target, ii) the clutter/

false alarm measurement (C/FA) process (the number and locations of C/FA measurements, both random), and iii) the measurement noise, which determines the location of the measurements away from the perfect measurements.

It is assumed that the target detection process, the clutter/false measurement process, and the measurement noise sequence, conditioned on the most recent true state, are independent of the past data (past detections, number, and locations of the false measurements, past measurements, and process noises). Then

$$\begin{aligned}
 p[z(k+1)|x(k+1), A_i(k+1), I^k, u(k)] \\
 = p[z(k+1)|x(k+1), A_i(k+1)]. \tag{S3}
 \end{aligned}$$

Note that the known input is irrelevant in the conditioning in (S3). Expression (S3) is the likelihood function of  $x(k+1)$  and  $A_i(k+1)$  given  $z(k+1)$ . In practice, this pdf is approximated as a product of Gaussian pdfs for the target-originated measurements centered at the predicted measurements and uniform pdfs for the clutter/false measurements. The choice of pdf for each measurement is determined by the association event in the conditioning.

For an arbitrary value of the known input  $u(k)$ , the state prediction pdf, which is a function rather than a point prediction, is given by

$$\begin{aligned}
 p[x(k+1)|I^k, u(k)] \\
 = \int p[x(k+1)|x(k), I^k, u(k)] p[x(k)|I^k, u(k)] dx(k), \tag{S4}
 \end{aligned}$$

where the integration is over the entire domain of  $x(k)$ . Equation (S4), which is an immediate consequence of the total probability theorem, is the Chapman-Kolmogorov (C-K) equation.

It is further assumed that the process noise sequence is white and independent of the measurement noises in the sense that  $v(k)$  conditioned on  $x(k)$  has to be independent of  $v(j-1)$  and  $w(j)$ ,  $\forall j \leq k$ . State-dependent process noise is allowed. The whiteness of  $v(k)$  is equivalent to requiring the state vector  $x(k)$  to be a Markov process.

resulting PDAF algorithm is based on KF. If the state or measurement equations are nonlinear, then PDAF is based on EKF.

### Assumptions

- 1) Only one target of interest is present, whose state  $x \in \mathcal{R}^n$  is assumed to evolve in time according to the equation

$$x(k) = F(k-1)x(k-1) + v(k-1), \tag{27}$$

with the true (target-originated) measurement  $z(k) \in \mathcal{R}^n$  given by

$$z(k) = H(k)x(k) + w(k), \tag{28}$$

where  $v(k-1)$  and  $w(k)$  are zero-mean mutually independent, white Gaussian noise sequences with known covariances matrices  $Q(k-1)$  and  $R(k)$ , respectively.

- 2) The track has been initialized. For details on track initialization, see [14]–[16].
- 3) The past information through time  $k-1$  about the target is summarized approximately by a sufficient statistic in the form of the Gaussian posterior

$$\begin{aligned}
 p[x(k-1)|Z^{k-1}] \\
 = \mathcal{N}[x(k-1); \hat{x}(k-1|k-1), P(k-1|k-1)]. \tag{29}
 \end{aligned}$$

Equation (29), which is the *basic assumption* of PDAF, is the key simplification of the mixture (22). This

Note that  $I^k$  is irrelevant in the conditioning of the first term on the right-hand side of (S4), that is,

$$p[x(k+1)|x(k), I^k, u(k)] = p[x(k+1)|x(k), u(k)]. \quad (S5)$$

The expression in (S5) is the state transition pdf. Furthermore, since the input  $u(k)$  enters the system after the realization of  $x(k)$  occurs, we can make the simplification

$$p[x(k)|I^k, u(k)] = p[x(k)|I^k] \triangleq p_k. \quad (S6)$$

Inserting (S5) and (S6) into (S4), the state prediction pdf can be written as

$$p[x(k+1)|I^k, u(k)] = \int p[x(k+1)|x(k), u(k)] p_k dx(k) \triangleq \phi_1[k+1, p_k, u(k)], \quad (S7)$$

where  $\phi_1$  maps the conditional density  $p_k$  of the state at time  $k$  into another function, namely, the state-prediction pdf.

Using (S3) and (S7), (S1) can be rewritten as

$$p_{k+1} = \frac{1}{c} \sum_{i=1}^{M(k+1)} p[\mathbf{z}(k+1)|x(k+1), A_i(k+1)] \phi_1[k+1, p_k, u(k)] \beta_i(k+1). \quad (S8)$$

The posterior probabilities of the association events at  $k+1$  are obtained with Bayes's formula as

$$\begin{aligned} \beta_i(k+1) &= P\{A_i(k+1)|Z^{k+1}, U^k\} \\ &= P\{A_i(k+1)|\mathbf{z}(k+1), Z^k, U^{k-1}, u(k)\} \\ &= \frac{1}{c} p[\mathbf{z}(k+1)|A_i(k+1), I^k, u(k)] \\ &\quad \times P\{A_i(k+1)|I^k, u(k)\} \end{aligned}$$

$$\begin{aligned} &= \frac{1}{c} \int p[\mathbf{z}(k+1)|x(k+1), A_i(k+1), I^k, u(k)] \\ &\quad \times p[x(k+1)|A_i(k+1), I^k, u(k)] \\ &\quad \times dx(k+1) P\{A_i(k+1)\}, \end{aligned} \quad (S9)$$

where the C-K equation and the independence of the association event  $A_i(k+1)$  from  $I^k$  and  $u(k)$  are used. Note that the unconditional probability of  $A_i(k+1)$  can be expressed in terms of the target detection event probabilities and the number of C/FA measurements for this event, both of which are assumed to be independent of the past.

Since the measurements  $\mathbf{z}(k+1)$ , conditioned on  $x(k+1)$  and  $A_i(k+1)$ , are independent of the past and the state prediction pdf given by (S5) does not depend on  $A_i(k+1)$ , it follows that

$$\begin{aligned} \beta_i(k+1) &= \frac{1}{c} \int p[\mathbf{z}(k+1)|x(k+1), A_i(k+1)] \\ &\quad \times p[x(k+1)|I^k, u(k)] dx(k+1) P\{A_i(k+1)\}. \end{aligned} \quad (S10)$$

Using (S7), (S10) can be rewritten as

$$\begin{aligned} \beta_i(k+1) &= \frac{1}{c} \int p[\mathbf{z}(k+1)|x(k+1), A_i(k+1)] \phi_1 \\ &\quad \times [k+1, p_k, u(k)] dx(k+1) P\{A_i(k+1)\} \\ &\triangleq \phi_{2,i}[k+1, p_k, u(k)]. \end{aligned} \quad (S11)$$

Using (S7) and (S11) in (S1) yields

$$\begin{aligned} p_{k+1} &= \frac{1}{c} \sum_{i=1}^{M(k+1)} p[\mathbf{z}(k+1)|x(k+1), A_i(k+1)] \phi_1 \\ &\quad \times [k+1, p_k, u(k)] \phi_{2,i}[k+1, p_k, u(k)] \\ &\triangleq \psi[k+1, p_k, \mathbf{z}(k+1), u(k)], \end{aligned} \quad (S12)$$

where  $\psi$  maps  $p_k$  into  $p_{k+1}$ . Equation (S12) is the compact version of the functional recursion (22) for the information state.

assumption amounts to an approximation of property P1 of the optimal estimator.

- 4) At each time a measurement validation region is set up around the predicted measurement to select the candidate measurements for association to the target of interest.
- 5) If the target was detected and the corresponding measurement fell into the validation region, then, according to (28), at most one of the validated measurements can be target originated.
- 6) The remaining measurements are assumed to be due to false alarms or clutter and are modeled as independent and identically distributed (i.i.d.) with uniform spatial distribution, and the number of false

alarms or clutter points obeys either a Poisson distribution, that is, a spatial Poisson process, with known spatial density  $\lambda$ , or a diffuse prior.

- 7) The target detections occur independently over time with known probability  $P_D$ .

These assumptions make it possible to obtain a state-estimation scheme that is almost as simple as the KF, but more effective in clutter. Its computational requirements are approximately 50% higher than those of the KF.

### Outline of the Algorithm

Figure 3 summarizes one cycle of a PDAF, which is similar to KF with the following additional features:



- 1) A PDAF has a selection procedure for the validated measurements at the current time.
- 2) For each such measurement, an association probability is computed for use as the weighting of this measurement in the combined innovation. The resulting combined innovation is used in the update of the state estimate; this computation conforms to property P2 of the pure MMSE estimator even though P2 is conditioned on satisfying P1 exactly; nevertheless, P2 is still used for the sake of simplicity when P1 is satisfied approximately.
- 3) The final updated state covariance accounts for the measurement origin uncertainty.

The stages of the algorithm are presented next.

### Prediction

Within the PDAF algorithm, the state vector, the measurement, and the state covariance matrices are predicted ahead in time from  $k-1$  to  $k$  as in the standard KF, namely,

$$\hat{x}(k|k-1) = F(k-1)\hat{x}(k-1|k-1), \quad (30)$$

$$\begin{aligned} \hat{z}(k|k-1) &= H(k)\hat{x}(k|k-1), \\ P(k|k-1) &= F(k-1)P(k-1|k-1)F(k-1)' + Q(k-1), \end{aligned} \quad (31)$$

$$(32)$$

where  $P(k-1|k-1)$  is available from (29). The innovation covariance matrix corresponding to the correct measurement is also computed as in the standard filter, namely,

$$S(k) = H(k)P(k|k-1)H(k)' + R(k). \quad (33)$$

Note that (33) is the covariance of the innovation corresponding to the correct measurement, even though it is not known which measurement is the correct one. This property is a direct consequence of (29), which lumps all past data  $Z^{k-1}$  into a single state estimate for the target.

### Measurement Validation

Following (29), the validation region is the elliptical region

$$\mathcal{V}(k, \gamma) = \{z: [z - \hat{z}(k|k-1)]' S(k)^{-1} [z - \hat{z}(k|k-1)] \leq \gamma\}, \quad (34)$$

where  $\gamma$  is the gate threshold corresponding to the gate probability  $P_G$ , which is the probability that the gate contains the true measurement if detected, and  $S(k)$ , given in (33), is the covariance of the innovation corresponding to the true measurement. The volume of the validation region (34) is

$$V(k) = c_{n_z} |\gamma S(k)|^{1/2} = c_{n_z} \gamma^{n_z/2} |S(k)|^{1/2}, \quad (35)$$

where the volume  $c_{n_z}$  of the  $n_z$ -dimensional unit hypersphere depends on the dimension  $n_z$  of the measurement. For example,  $c_1 = 2$ ,  $c_2 = \pi$ , and  $c_3 = 4\pi/3$ . The set of validated measurements according to (34) is

$$\mathbf{z}(k) = \{z_i(k)\}_{i=1}^{m(k)}. \quad (36)$$

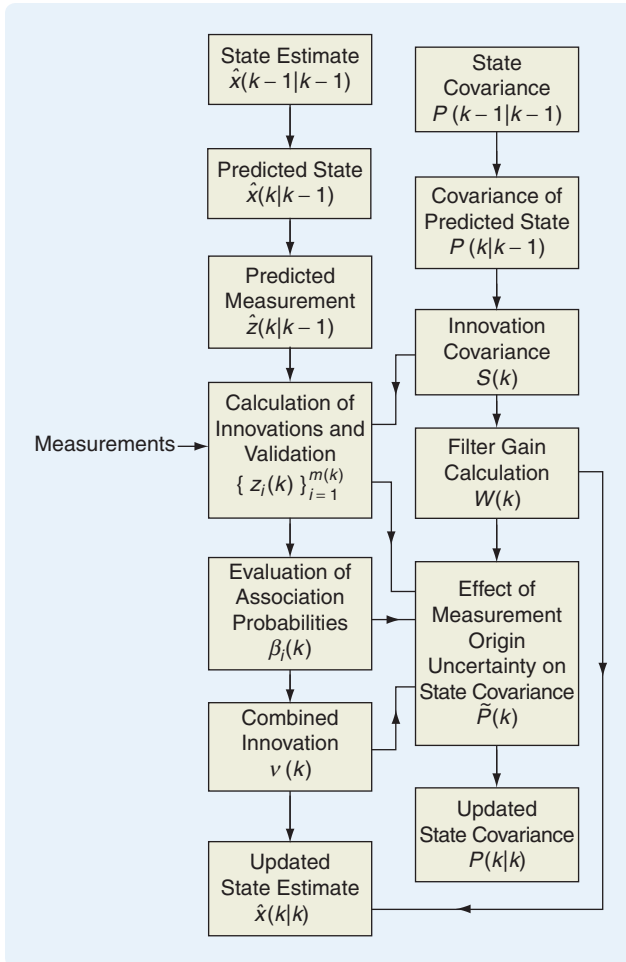
### Data Association

The parametric PDA with the Poisson clutter model with spatial density  $\lambda$  yields the association probability for  $z_i(k)$  being the correct measurement, as

$$\beta_i(k) = \begin{cases} \frac{\mathcal{L}_i(k)}{1 - P_D P_G + \sum_{j=1}^{m(k)} \mathcal{L}_j(k)}, & i = 1, \dots, m(k), \\ \frac{1 - P_D P_G}{1 - P_D P_G + \sum_{j=1}^{m(k)} \mathcal{L}_j(k)}, & i = 0, \end{cases} \quad (37)$$

where  $i=0$  means that none is correct,  $P_D$  is the target detection probability,  $P_G$  is the gate probability [3] corresponding to (34), and

$$\mathcal{L}_i(k) \triangleq \frac{\mathcal{N}[z_i(k); \hat{z}(k|k-1), S(k)] P_D}{\lambda} \quad (38)$$



**FIGURE 3** One cycle of the probabilistic data association filter. The state-estimation and covariance calculations are two-way coupled because of the dependence of the covariances on the innovations.

is the likelihood ratio of the measurement  $z_i(k)$  originating from the target rather than from clutter. Note that the term  $\lambda$  (the density of the spatial Poisson process that models the clutter) in the denominator of (38) plays the role of the uniform pdf of the location of a false measurement [13].

The nonparametric PDA is obtained assuming a diffuse pmf [5] for the number of clutter measurements, that is, all numbers of clutter measurements are equally likely. Note that, with  $m(k)$  validated measurements, the only possibilities for the number of clutter measurements are  $m(k)$  and  $m(k) - 1$ . In this case, the association probabilities are the same as above except for replacing  $\lambda$  in (38) by  $m(k) / V(k)$ , where  $V(k)$  is given in (35).

The discrimination capability of the PDA relies on the difference between the Gaussian and uniform densities. Additional discrimination capability can be obtained by using feature variables; see [3, Sec. 3.4.7]. The PDA yields association probabilities that depend on the location of the corresponding innovation on the likelihood ratio's Gaussian exponential (38) relative to the other innovations.

### State Estimation

The state update equation of PDAF is

$$\hat{x}(k|k) = \hat{x}(k|k-1) + W(k)\nu(k), \quad (39)$$

where the combined innovation is

$$\nu(k) = \sum_{i=1}^{m(k)} \beta_i(k) \nu_i(k), \quad (40)$$

and the gain  $W(k)$  is the same as in the standard (conventional Kalman) filter

$$W(k) = P(k|k-1)H(k)'S(k)^{-1}. \quad (41)$$

The covariance associated with the updated state (39) is

$$P(k|k) = \beta_0(k)P(k|k-1) + [1 - \beta_0(k)]P^c(k|k) + \tilde{P}(k), \quad (42)$$

where the covariance of the state updated with the correct measurement is

$$P^c(k|k) = P(k|k-1) - W(k)S(k)W(k)', \quad (43)$$

and the spread of the innovations term, similar to the spread of the means term in a mixture [5, Eq. (1.4.16-9)], is

$$\tilde{P}(k) \triangleq W(k) \left[ \sum_{i=1}^{m(k)} \beta_i(k) \nu_i(k) \nu_i(k)' - \nu(k) \nu(k)' \right] W(k)'. \quad (44)$$

Covariance (42) is the counterpart of (3) for the PDAF. Since with probability  $\beta_0(k)$  none of the measurements is correct, in which case there is no update of the state estimate, the prediction covariance  $P(k|k-1)$  appears in (42) with weighting  $\beta_0(k)$ . With probability  $1 - \beta_0(k)$  the cor-

rect measurement is available, and the updated covariance  $P^c(k|k)$  appears with weighting  $1 - \beta_0(k)$ . However, since it is not known which of the  $m(k)$  validated measurements is correct, the term  $\tilde{P}$ , which is positive semidefinite, increases the covariance of the updated state. This increase is the effect of the measurement origin uncertainty. Note in (44) the dependence of the estimation accuracy on the data, which is typical of nonlinear estimators. The PDAF is a nonlinear estimator; while the estimate update in (39) appears linear, it is nonlinear because the association probabilities  $\beta_i(k)$  depend on the innovations according to (37).

### THE JOINT PROBABILISTIC DATA ASSOCIATION FILTER

The joint probabilistic data association (JPDA) approach is the extension of the PDA. The following are the assumptions of the JPDA:

- 1) The number of established targets in the clutter is known.
- 2) Measurements from one target can fall in the validation region of a neighboring target. This situation can happen over several sampling times, and acts as a *persistent interference*.
- 3) The past of the system is summarized by an approximate sufficient statistic consisting of state estimates, which are given by approximate conditional means, along with covariances for each target.
- 4) The states are assumed to be Gaussian distributed with means and covariances according to the above approximate sufficient statistic.
- 5) Each target has a dynamic and a measurement model as in (27), (28). The models for the various targets do not have to be identical.

PDAF models all incorrect measurements as *random interference* with uniform spatial distribution. The performance of PDAF degrades significantly when a neighboring target gives rise to persistent interference.

JPDAF consists of the following steps:

- 1) The measurement-to-target association probabilities are computed jointly across the targets.
- 2) In view of the assumption that a sufficient statistic is available, the association probabilities are computed only for the latest set of measurements. This approach conforms to the results from the section "The Optimal Estimator for the Pure MMSE Approach."
- 3) The state estimation is done either separately for each target as in PDAF, that is, in a decoupled manner, resulting in JPDAF, or in a coupled manner using a stacked state vector, resulting in JPDAF.

The key feature of the JPDA is that it evaluates the conditional probabilities of the following joint association events

$$A(k) = \bigcap_{j=1}^m A_{j_t}(k), \quad (45)$$

pertaining to the current time  $k$ , where  $A_{jt}(k)$  is the event that measurement  $j$  at time  $k$  originated from target  $t$ ,  $j = 1, \dots, m$ ,  $t = 0, 1, \dots, N_T$ ;  $t_j$  is the index of the target to which measurement  $j$  is associated in the event under consideration, and  $N_T$  is the known number of targets.

### The Parametric and Nonparametric JPDA

As in the case of the PDA, the JPDA has two versions, according to the model used for the pmf  $\mu_F(\phi)$  of the number of false measurements.

The parametric JPDA uses the Poisson pmf

$$\mu_F(\phi) = e^{-\lambda V} \frac{(\lambda V)^\phi}{\phi!}, \quad (46)$$

for a volume  $V$ , which requires the spatial density  $\lambda$  of the false measurements. Following derivations presented in [3, Sec. 6.2.3], the joint association probabilities of the parametric JPDA are

$$P\{A(k) | Z^k\} = \frac{1}{c_2} \prod_j \{\lambda^{-1} f_{t_j}[z_j(k)]\}^{\tau_j} \prod_t (P_D^t)^{\delta_t} (1 - P_D)^{1 - \delta_t}, \quad (47)$$

where

$$f_{t_j}[z_j(k)] = \mathcal{N}[z_j(k); \hat{z}^t_j(k | k - 1), S^t_j(k)], \quad (48)$$

and  $\hat{z}^t_j(k | k - 1)$  is the predicted measurement for target  $t_j$ , with associated innovation covariance  $S^t_j(k)$ ;  $P_D^t$  is the detection probability of target  $t$ ; and  $\tau_j$  and  $\delta_t$  are the target detection and measurement association indicators, respectively.

The nonparametric JPDA uses the diffuse prior for the number  $\phi$  of false measurements, that is,

$$\mu_F(\phi) = \epsilon \quad \text{for all } \phi, \quad (49)$$

which leads to the joint association probabilities

$$P\{A(k) | Z^k\} = \frac{1}{c_4} \phi! \prod_j \{V f_{t_j}[z_j(k)]\}^{\tau_j} \prod_t (P_D^t)^{\delta_t} (1 - P_D)^{1 - \delta_t}, \quad (50)$$

where  $V$  is the volume of the surveillance region in which the measurements not associated with a target are to be assumed uniformly distributed,  $\phi$  is the number of false measurements in event  $A$ , and  $c_4$  is the normalization constant.

The term  $\phi! / V^\phi$  in the nonparametric JPDA expressions can be called a “pseudosample spatial measurement density” and plays the role of  $\lambda^\phi$  in (47) in the parametric JPDA. This result is similar to the PDA, where the nonparametric version contains  $m(k) / V$  in place of  $\lambda$  from the parametric version.

### The State Estimation

The state-estimation or filtering algorithm can be carried out in two ways. Assuming that the states of the targets

conditioned on the past observations are mutually independent, the problem becomes one of decoupled estimation for the targets under consideration, that is, JPDAF. In this case the marginal association probabilities are needed. These marginal probabilities are obtained from the joint probabilities by summing over all joint events in which the marginal event of interest occurs. This summation can be written as

$$\begin{aligned} \beta_{jt}(k) &\triangleq P\{A_{jt}(k) | Z^k\} \\ &= \sum_{A: A_{jt} \in A} P\{A(k) | Z^k\}. \end{aligned} \quad (51)$$

The state-estimation equations are then decoupled among the targets and exactly the same as in PDAF, presented in (39)–(44).

### JPDAF with Coupling

The more realistic alternative assumption that the states of the targets, given the past, are correlated yields the JPDA coupled filter (JPDACF), which performs coupled estimation for the targets under consideration. The JPDA is developed assuming that, conditioned on the past, the target states and, thus, the target-originated measurements, are independent. Consequently, the joint association is followed by decoupled filtering of the targets' states, which simplifies the resulting algorithm.

For targets that share measurements in the JPDA events for several sampling times, a statistical dependence of their estimation errors ensues, which can be taken into account by calculating the state cross covariances. The resulting algorithm, named JPDACF, does the data association as well as the filtering in a coupled manner for the targets with shared measurements, yielding a covariance matrix with off-diagonal blocks, called cross covariances, which reflect the correlation between the state-estimation errors of the targets.

The conditional probability for a joint association event becomes

$$\begin{aligned} P\{A(k) | Z^k\} &= \frac{1}{c} \frac{\phi!}{m(k)!} \mu_F(\phi) V^{-\phi} f_{t_{j_1}, t_{j_2}, \dots}[z_j(k), j; \tau_j = 1] \\ &\quad \times \prod_t (P_D^t)^{\delta_t} (1 - P_D^t)^{1 - \delta_t}, \end{aligned} \quad (52)$$

where  $f_{t_{j_1}, t_{j_2}, \dots}$  is the joint pdf of the measurements of the targets under consideration,  $t_{j_1}$  is the target to which  $z_{j_1}(k)$  is associated in event  $A$ ,  $\phi$  is the number of false measurements in event  $A$ , and  $c$  is the normalization constant.

The joint probabilities are not reduced to the marginal association probabilities as in (51) for use in decoupled PDA filters. Instead, these joint probabilities are used directly in a coupled filter.

Denote the stacked vector of the predicted states of the targets under consideration, assumed here to be two, and the associated covariance matrix as

$$\hat{x}^T(k|k-1) = \begin{bmatrix} \hat{x}^1(k|k-1) \\ \hat{x}^2(k|k-1) \end{bmatrix}, \quad (53)$$

$$P^T(k|k-1) = \begin{bmatrix} P^{11}(k|k-1) & P^{12}(k|k-1) \\ P^{21}(k|k-1) & P^{22}(k|k-1) \end{bmatrix}, \quad (54)$$

where  $\hat{x}^t$  and  $P^{tt}$  correspond to target  $t$ , and  $P^{t_1 t_2}$  is the cross covariance between targets  $t_1$  and  $t_2$ . This cross covariance is zero before these targets become coupled.

The coupled filtering is given by

$$\begin{aligned} \hat{x}^T(k|k) &= \hat{x}^T(k|k-1) + W^T(k) \sum_A P\{A(k)|Z^k\} \\ &\quad \times [z^T(k, A) - \hat{z}^T(k|k-1)], \end{aligned} \quad (55)$$

where

$$z^T(k, A) = \begin{bmatrix} z_{j_t(A)}(k) \\ z_{j_s(A)}(k) \end{bmatrix}, \quad (56)$$

and  $j_t(A)$  is the index of the measurement associated with target  $t$  in event  $A$  at time  $k$ . The filter gain in (55) is

$$\begin{aligned} W^T(k) &= P^T(k|k-1)H^T(k)' [H^T(k) \\ &\quad \times P^T(k|k-1)H^T(k)' + R^T(k)]^{-1}, \end{aligned} \quad (57)$$

where

$$H^T(k) = \begin{bmatrix} H^1(k) & 0 \\ 0 & H^2(k) \end{bmatrix}, \quad R^T(k) = \begin{bmatrix} R^1(k) & 0 \\ 0 & R^2(k) \end{bmatrix} \quad (58)$$

are the block-diagonal measurement matrix and noise covariance matrix, respectively, for the two targets under consideration. The predicted stacked measurement vector is

$$\hat{z}^T(k|k-1) = H^T(k)\hat{x}^T(k|k-1). \quad (59)$$

The state covariance update is as in (42).

The JPDACF approach is a more realistic approximation of (22) due to the use of the stacked state vector of the targets under consideration, as in (16). Other improvements on JPDAF to avoid its tendency of track coalescence can be found in [17] and [18]. This tendency of tracks of closely spaced targets to become overlapped is due to the shared measurements across tracks, when this sharing lasts for many frames or scans.

## A SIMPLE EXAMPLE OF TRACKING IN CLUTTER

Consider a nearly constant velocity model in two dimensions with state

$$x = [\xi \ \eta \ \dot{\xi} \ \dot{\eta}]', \quad (60)$$

in Cartesian coordinates  $\xi$  and  $\eta$ . This model assumes the acceleration to be a discrete-time zero-mean white noise (see [5, Sec. 6.3]). The state equation is

$$x(k+1) = Fx(k) + \Gamma v(k) \quad (61)$$

with the process noise  $v(k)$  a zero-mean white acceleration sequence with

$$E[v(k)v(k)'] = Q(k). \quad (62)$$

The transition matrix, decoupled between the two coordinates, is given by

$$F = \begin{bmatrix} F_\xi & 0 \\ 0 & F_\eta \end{bmatrix}, \quad (63)$$

where

$$F_\xi = F_\eta = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \quad (64)$$

and  $T$  is the sampling interval.

The vector gain multiplying the acceleration process noise is

$$\Gamma = \begin{bmatrix} \Gamma_\xi \\ \Gamma_\eta \end{bmatrix}, \quad (65)$$

where

$$\Gamma_\xi = \Gamma_\eta = \begin{bmatrix} \frac{1}{2}T^2 \\ T \end{bmatrix}. \quad (66)$$

The process noise covariance is

$$Q(k) = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}. \quad (67)$$

The target observation vector, which consists of target position, is given by

$$z(k) = Hx(k) + w(k), \quad (68)$$

where

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (69)$$

and the measurement noise is assumed zero-mean and white with

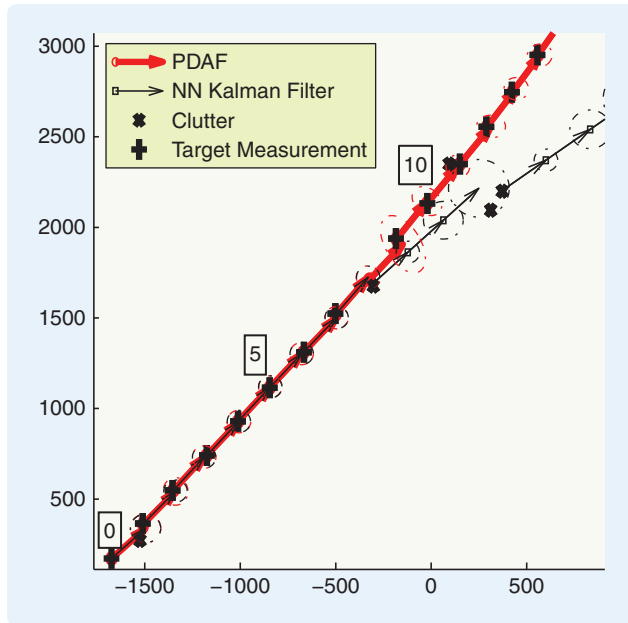
$$E[w(k)w(k)'] = R(k) = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_w^2 \end{bmatrix}. \quad (70)$$



The parameter values used for the simulation are  $P_D = 0.8$ ,  $T = 30$  s, and initial conditions  $\xi(0) = -1.85 \cdot 10^3$  m,  $\dot{\xi}(0) = 5.94$  m/s,  $\eta(0) = 0$ , and  $\dot{\eta}(0) = 5.94$  m/s. The clutter is generated uniformly with the number of clutter points, or false measurements, obtained from a Poisson distribution with spatial density  $10^{-5}/\text{m}^2$ . The noise variances are  $\sigma_v^2 = 0.01^2 \text{ m}^2/\text{s}^4$  and  $\sigma_w^2 = 7^2 \text{ m}^2$ . The gating region used is 99.97%, with  $\gamma = 16$ .

Figure 4 shows one sample run with PDAF and NNKF, where the latter is a KF, for the system specified by (60)–(70). Each estimated position is at the base of an arrow that represents the estimated velocity vector multiplied by the sampling interval. The tips of the arrows therefore indicate the predicted positions, where the corresponding validation region ellipses are centered. Every fifth sampling time is indicated in a rectangular box.

Up to time  $k = 7$ , both filters perform equally well. At time  $k = 8$ , there is no target detection but there is a false measurement. PDAF attaches a low probability that this measurement is target originated and, consequently, its prediction covariance at time  $k = 9$  is significantly increased according to (42). This covariance increase can be seen from its validation region ellipse, which is much larger than the previous ones. The NNKF, however, uses the false measurement at time  $k = 8$  as if it were true one. Because of this overconfidence, the NNKF is sidetracked and has a small updated covariance, to the extent that its validation regions at times  $k = 9$  and  $k = 10$  are empty.



**FIGURE 4** Illustration of the operation of the probabilistic data association filter and the nearest neighbor Kalman filter in the presence of missing target detections and clutter. The measurements are positions in  $\xi$ - $\eta$  Cartesian coordinates. Every fifth sampling time is shown in a box next to the corresponding measurement.

At time  $k = 11$  the NNKF has two measurements in its validation region, which has grown considerably quite a bit because of lack of updates and, using the nearest of them, has again at the next time  $k = 12$  a small validation region, which is empty and far enough from the target that the track is irretrievably lost.

The above example illustrates the ability of the PDAF to keep the target in track despite the less-than-unity detection probability and the presence of clutter. The key to this ability is the limited trust the PDAF places in its validated measurements, as reflected in (40), and the consequent increase in the PDAF state-error covariance according to (42).

## A MISSILE TRACKING EXAMPLE WITH SPACE-BASED SENSORS

We consider a missile moving in a three-dimensional space with axial thrust. The state in Earth-centered inertial Cartesian coordinates  $\xi$ ,  $\eta$ , and  $\zeta$  is

$$x = [\xi \ \eta \ \zeta \ \dot{\xi} \ \dot{\eta} \ \dot{\zeta} \ \alpha]', \quad (71)$$

where  $\alpha$  is the net specific thrust, that is, thrust less drag per unit mass, given by

$$\alpha = \frac{T - D}{M} \quad (72)$$

modeled by the differential equation

$$\dot{\alpha}(t) = \frac{\alpha(t)^2}{c}, \quad (73)$$

with  $c = 5$  km/s and  $\alpha(0) = 20$  m/s<sup>2</sup>. Equations (72)–(73) describe the axial acceleration of the missile during the burn period assuming a constant thrust and a constant mass ejection rate. The solution to (73) is  $\alpha(t) = c / [c\alpha(0)^{-1} - t]$ .

The continuous-time state equation of the target is

$$\dot{x}(t) = f[x(t)], \quad (74)$$

where

$$f_1(x) = x_4 = \dot{\xi}, \quad (75)$$

$$f_2(x) = x_5 = \dot{\eta}, \quad (76)$$

$$f_3(x) = x_6 = \dot{\zeta}, \quad (77)$$

$$f_4(x) = \alpha \frac{\dot{\xi} + \Omega \eta}{v_R} - \mu \frac{\xi}{R^3}, \quad (78)$$

$$f_5(x) = \alpha \frac{\dot{\eta} - \Omega \xi}{v_R} - \mu \frac{\eta}{R^3}, \quad (79)$$

$$f_6(x) = \alpha \frac{\dot{\zeta}}{v_R} - \mu \frac{\zeta}{R^3}, \quad (80)$$

with  $\mu$  the gravitational constant,  $\Omega$  the spin rate of the Earth, and

$$v_R^2 = (\dot{\xi} + \Omega\eta)^2 + (\dot{\eta} - \Omega\xi)^2 + \dot{\zeta}^2, \quad (81)$$

$$R^2 = \xi^2 + \eta^2 + \zeta^2. \quad (82)$$

The target trajectory is generated without process noise. Note that (75)–(77) are standard kinematic equations representing the integration of the velocity into position, while (78)–(80) show the acceleration components as a function of the net thrust and local gravity.

The measurements are obtained from two passive sensors that measure the azimuth and elevation angles of the line of sight to the target with standard deviation 0.1 mrad. The sensors are located on satellites. The sampling period is 1 s, and 100 frames are available from each sensor.

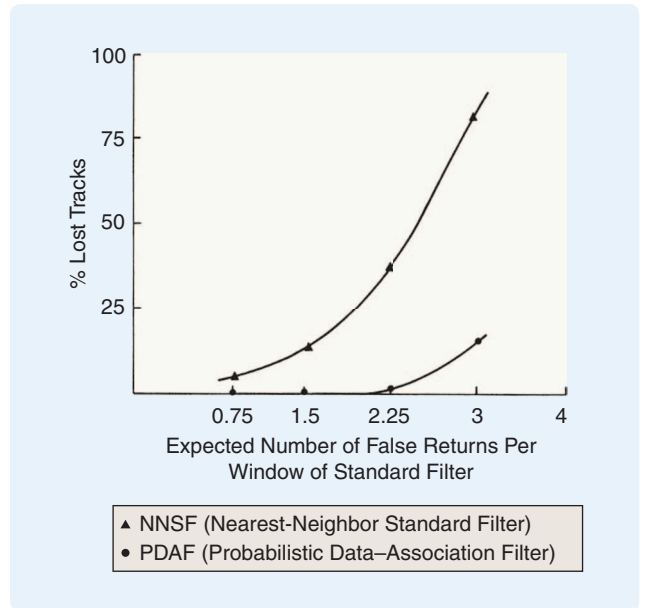
An EKF is used to track this target in a clutterless environment. The state equation is integrated numerically between sampling times to predict the state. First-order linearization of both dynamic and measurement equations is used for covariance calculations. To make the resulting filter consistent, a process noise equal to 1% of the updated state covariance is used in the EKF (see [5, Sec. 10.4]). The measurements from the two sensors are processed sequentially according to the algorithm presented in [3, Sec. 2.2].

The EKF reaches steady state after about ten measurements. The resulting 99% validation region for the measurements, with each measurement from each sensor having dimension two, has gating threshold  $\gamma = 9.2$ , which is designated as the standard gate. This gate is used as the basis for defining the clutter density  $\bar{r}$ . In the simulations, false measurements, that is, the clutter points, are generated, starting at  $k = 10$ , with a uniform pdf for their location with various spatial densities. Various measures of performance are evaluated versus  $\bar{r}$ , the expected number of false measurements in the standard gate defined above.

The nonparametric version of the PDAF is used in this example, so that no knowledge of the spatial density of false measurements is needed. A diffuse prior models the number of false measurements in the validation region. The purpose of the simulations is to examine the tracking capability of the PDAF in comparison with the NNSF as the clutter density increases, which leads to track loss. The percentage of lost tracks is estimated from Monte Carlo simulations.

Figure 5 presents the percentage of lost tracks from 50 Monte Carlo runs for the PDAF and NNSF. A track is considered lost when the correct measurement is not in the 99% validation region of at least one of the sensors for the previous 20 sampling times. This definition is reasonable since in this case the final errors are large.

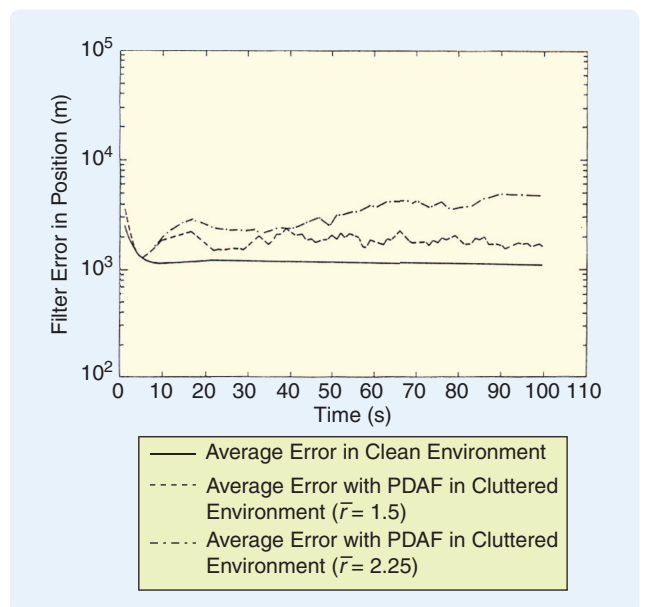
The simulation results show that the PDAF can track reliably up to a clutter density for which the expected number of false measurements in the standard (NNSF) gate is  $\bar{r} = 2$ . The average number of false measurements in the PDAF gate is larger than in the NNSF gate because the PDAF covariance is increased according to (42). The NNSF is “optimistic” (its calculated covariances are too small)



**FIGURE 5** Comparison of tracking capability in terms of percentage of lost tracks between the probabilistic data association filter (PDAF) and the nearest neighbor standard extended Kalman filter (NNSF). The PDAF allows reliable tracking up to a clutter level of two false measurements in the validation gate, at which level NNSF has a track loss probability of about 1/3, which makes NNSF unreliable.

and has a substantially higher percentage of lost tracks than the PDAF due to its unwarranted optimism.

The second measure of performance for this problem is the filter estimation error. Figures 6 and 7 show the average position and velocity estimation errors for PDAF, respectively, from 50 runs, for two values of  $\bar{r}$ , compared to the average error in the clutterless environment. It is notable



**FIGURE 6** Position estimation errors for PDAF. The error increases in the presence of clutter because of the added measurement origin uncertainty.

## Real-World Applications of PDAF and JPDAF

The first real-world application of PDAF is likely the Australian Jindalee over-the-horizon radar (OTHr) [14]. Due to the large number of false alarms in this system, which relies on ionospheric reflection, the KF/EKF is unsuitable as a tracking filter. Initially, PDAF was implemented using the covariance update (42) without the positive-semidefinite spread of the innovations term (44). The result was that the system kept losing the target tracks. The reason for the track loss is that, without the additional terms (44) in the covariance update equation, the

innovation covariance and, consequently, the measurement validation gate are too small. Once the full version of (42) was implemented, the system was able to keep its targets in track.

Experience with the Raytheon radars, described below, is that, for problems with small range errors and very large cross-range errors, the state-covariance matrix is ill-conditioned, and thus special means are necessary to overcome the ill-conditioning. In the case of ill-conditioned covariances, the term (44) can worsen the ill-conditioning. Since the OTHr has large range errors (due to ionospheric reflection) compared to direct line-of-sight radars, and thus additional uncertainty, the state covariance in an OTHr tracking system is not ill conditioned.

Many Raytheon radars use JPDA or MHT, as shown in figures S1–S6. Most of these radars are very-long-range (several thousand kilometer) solid-state, phased-array radars, but many short-range radar applications are also shown. The long-range, solid-state phased array radars are expensive, owing to the high transmit power required to detect small targets at long range. In particular, each solid-state, phased-array radar might cost several hundred million dollars. Reliable data association is thus crucial



**FIGURE S1** Relocatable over the horizon radar (ROTHr) high-frequency phased-array radar. This radar carries out long-range surveillance for drug interdiction against aircraft and ships. ROTHr relies on ionospheric reflection and has a range of several thousand kilometers. The tracking algorithm used by this radar is the classic probabilistic data association with interacting multiple-model Kalman filters, which provides reliable tracking in extremely severe ionospheric clutter. (Courtesy of Raytheon.)



**FIGURE S2** Theater high altitude area defense (THAAD) solid-state, phased-array radar. The THAAD radar uses nearest neighbor joint probabilistic data association with extended Kalman filters. (Courtesy of Raytheon.)



**FIGURE S3** Airport surface detection equipment, model X (ASDE-X) air traffic control radar. The ASDE-X performs surveillance and tracks low-altitude aircraft near the airport, aircraft on the ground, and vehicles in the airport area. This radar operates in a difficult environment with several reflections from a target as well as multipath reflections due to buildings and other man-made metal clutter. The ASDE-X uses a classic joint probabilistic data association filter tracker with interacting multiple-model Kalman filters. (Courtesy of Raytheon.)





**FIGURE S4** Cobra Dane phased-array radar. This radar carries out long-range surveillance against intercontinental ballistic missiles. The algorithm used for tracking is the nearest neighbor joint probabilistic data association with extended Kalman filters. (Courtesy of Raytheon.)

for these radar applications. Moreover, the quantification of uncertainty in data association is also important for these applications.

In all cases the JPDA and MHT algorithms provide reliable data association performance in dense multiple target environments and are very robust with respect to real-world phenomena. MHT is much more sophisticated than JPDA but requires orders of magnitude more real-time computer throughput and lines of source code than JPDA, which is a simple single-scan single-hypothesis algorithm. In contrast, JPDA runs quickly in real time and is easy to code. The ASDE-X radar for air traffic control shown in Figure S3 uses JPDA as the data association algorithm in an extremely challenging environment, namely, airports, whose runways and taxiways are surrounded by man-made clutter, including metal signs, lights, large buildings, trucks, people, and other aircraft as well as natural clutter, including rain, snow, dust, sea clutter, and birds. Multipath propagation between large metal buildings and large metal aircraft is profuse. Also, aircraft can start, stop, turn, and accelerate. Despite these difficulties, the ASDE-X-based air traffic control system with the JPDA association algorithm has been in use at many airports around the world for the last decade (see [20] for more details).

One version of JPDA that is often used in real-world applications is the nearest neighbor JPDA (NNJPDA), in which one measurement is selected to associate with a given track, rather than averaging over several measurements. The a posteriori probability of correct data association computed approximately by JPDA is used to perform a MAP decision as in (8), rather than computing the overall conditional mean as in (1) [21]. This method avoids the problem of coalescence of tracks as well as bias errors (see [17], [18], [25] for more details on coalescence of tracks). A review of NNJPDA is given in [22], and a quantitative evaluation of the performance in terms of



**FIGURE S5** Sea-based X-band (SBX) solid-state phased-array radar. This radar, which is mounted on a movable platform that can be relocated on the ocean, provides long-range surveillance against intercontinental ballistic missiles. The algorithm used for tracking is the nearest neighbor joint probabilistic data association with extended Kalman filters. (Courtesy of Raytheon.)



**FIGURE S6** Upgraded early warning radar (UEWR) solid-state, phased-array radar. This radar serves the National Missile Defense system for long-range ground-based surveillance against intercontinental ballistic missiles. The algorithm used for tracking is the multiple hypothesis tracker with extended Kalman filters, which are carefully designed to mitigate both ill-conditioning and nonlinear effects. (Courtesy of Raytheon.)

probability of correct data association and real-time throughput requirements for NNJPDA compared with global assignment using JVC or auction, and NN (local/greedy) data association algorithms is given in [24]. It turns out that the probability of correct data association is roughly the same over a wide range of target densities and prediction accuracies for all of these single-scan, single-hypothesis algorithms, with JVC slightly better than NNJPDA, and NN the worst, as expected.

Moreover, in real-world applications of the JPDA algorithm, the KF covariance matrix is generally left unmodified to account for uncertainty in data association, owing to problems with ill-conditioning as well as nonlinear problems such as false



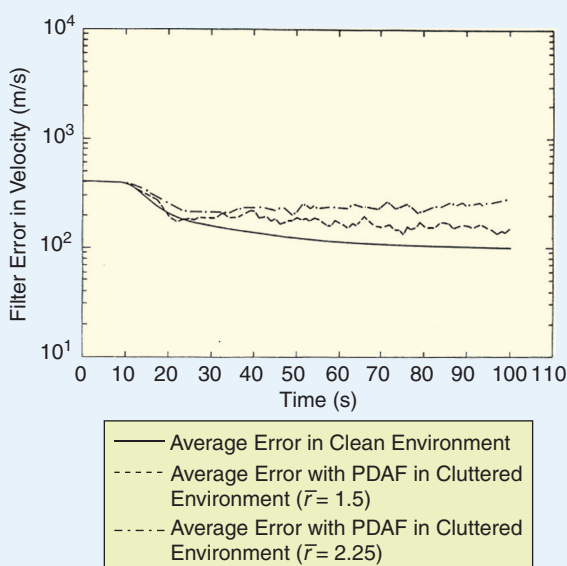
cross-range observability. Adding rank-one updates to a covariance matrix as in (42)–(44) tends to worsen already ill-conditioned covariance matrices. Details on ill-conditioning and false cross range observability in KFs for long-range phased-array radar applications are given in [23] and [30].

A typical radar system using JPDA evolves as follows. A software requirements specification is written with the classic rank-one covariance matrix corrections to account for uncertainty in data association, given in (42)–(44); this algorithm is then coded, debugged, and tested (and debugged again in real time if necessary); next, the software developers or radar system engineers notice that something is wrong with the KF performance; phone calls and e-mail messages asking for help are sent; the suggestion is made to zero the classic rank-one updates of the covariance matrix and then see whether this approach helps; the software team gladly takes this advice because it is so quick and easy to follow, and they learn that the problem disappears. It turns out that whether ill-conditioning is a problem in any given application depends on the radar system parameters as well as the target characteristics, including range-measurement accuracy versus cross-range measurement accuracy; slant range; target-crossing angle, which is the angle between the target velocity vector and the radar line of sight; unmodeled target acceleration (and thus process noise in the KF); temporal variation in SNR; track times; and the desired accuracy of the velocity vector estimate.

For radar applications that are ill-conditioned to begin with, classic JPDA rank-one updates of the covariance matrix do not help. Intuitively, if the KF is tuned with ample process noise, then the problem is not so ill-conditioned, and, moreover, the KF rapidly forgets the rank-one updates, whereas, with small

process noise, the KF remembers these rank-one updates, making the covariance matrix increasingly ill-conditioned as time goes on. It is possible to plot the condition number of the covariance matrix versus time to try to understand this phenomenon, but this condition number is not a reliable diagnostic because the condition number often provides a pessimistic upper bound on ill-conditioning. These issues also arise in MHT or data association algorithms that attempt to correct the covariance matrix to account for uncertainty in data association.

As emphasized in [26], [29], and [31], the most important issue in tracking in a dense multiple target environment is not data association but rather limited sensor resolution. Few of the classic data association algorithms described in the literature, including JPDA and MHT, explicitly model sensor resolution. Notable exceptions include [32], [27], and [29]. Nevertheless, in the real world as well as in detailed Monte Carlo simulations, NNJPDA performs remarkably well in dense multiple target scenarios with unresolved measurements, compared with much more sophisticated algorithms such as classic MHT. This performance difference is due to the assumption in classic MHT that detection events of a given target are statistically independent from one time sample to the next, which causes the calculated probability of track-to-measurement association to be very inaccurate, resulting in the unnecessary loss of track in situations with unresolved data. More specifically, if the a priori probability of detection of a given target at a given time is  $p = 0.9$  and we miss  $N = 10$  measurements in a row on a given track owing to lack of resolution with another target, then the naïve calculation of the probability that this scenario occurs yields



**FIGURE 7** Velocity estimation errors for the probabilistic data association filter. The unavoidable increase in the error due to the clutter is the price paid for the ability to keep the target in track.

that for  $\bar{r} = 0.75$  the performance of the PDAF (not shown) is the same as in the clutterless environment. For  $\bar{r} = 1.5$  there is only a slight degradation and even for  $\bar{r} = 2.25$  it is still not very large. On the other hand, for  $\bar{r} = 2.25$ , NNSF has 40% track loss as indicated in Figure 5.

Therefore, the PDAF can significantly extend the tracking capability into the region of high clutter density where NNSF becomes unreliable because of its high probability of track loss. Of course, no tracker can work in arbitrarily dense clutter. If there is too much clutter it is unavoidable that the target will be lost. Results on the stability limit of PDAF can be found in [19].

## CONCLUSIONS

The measurement selection for updating the state estimate of a target's track, known as data association, is essential for good performance in the presence of spurious measurements or clutter. A classification of tracking and data association approaches has been presented, as a pure MMSE approach, which amounts to a soft decision, and single best-hypothesis approach, which amounts to a hard decision.

$(1 - p)^N = (1 - 0.9)^{10} = 10^{-10}$ , for instance, that it has an extremely low probability. Similar erroneous results are obtained for any reasonable value of  $p$ . Such a track, which has missing detections due to resolution, would be dropped owing to the very low probability of its occurrence as calculated naïvely and erroneously by classic MHT. This extremely inaccurate calculation is caused by the failure to tell the algorithm about the possibility of unresolved measurements, which cause dependence of non-detection events in time. Likewise, multipath fades, radar cross-section fades, target obscuration, clutter obscuration, chaff, and jamming all cause signal loss that is highly correlated over time, contrary to the naïve and erroneous assumption of statistical independence. Such an erroneous assumption cannot be made in JPDA, because JPDA is a single-scan, single-hypothesis algorithm, and thus inherently does not make such an assumption. This phenomenon is an example of a situation in which a simpler algorithm such as JPDA is more robust than a complex algorithm such as MHT.

Furthermore, both classic MHT and JPDA explicitly use the assumption that one measurement corresponds to at most one object, contrary to physical reality; such nonphysical assumptions are not used in [27]–[29], resulting in improved performance. Simple but effective modifications of JPDA or MHT for modeling unresolved measurements result in further improved performance. For example, it is easy to predict all tracks to the time of the next measurement and see whether two tracks are sufficiently close together in range and angle so as to be unresolved; if so, the algorithm can modify its calculations accordingly. This approach was taken in many real radar applications over the last four decades. Moreover, extremely sophisticated

algorithms, such as classic MHT, which do not consider unresolved measurements, can have a worse performance than simple algorithms such as NN, JVC, or JPDA, which have been modified to explicitly model unresolved data.

MHT provides reliable and fast-track initiation in dense multiple target environments, whereas JPDA, JVC, and NN do not provide track initiation. Accordingly, for JPDA or NN we generally use a poor man's version of MHT with carefully designed radar waveforms, such as up-down chirp pulse-pairs, for track initiation [31]. As mentioned above, the real-time computational complexity of MHT is orders of magnitude greater than JPDA; however, we can run MHT in real time for what are considered very stressing multiple target scenarios on a single PC. JPDA was invented about 30 years ago, which corresponds to roughly a five orders-of-magnitude increase in real-time computer throughput and fast random access memory per dollar, owing to Moore's law. The radars, designed in the 1970s, used computers with roughly 5 Mflops of throughput at a cost of about US\$10 million for one computer; whereas a good PC today has roughly 5 Gflops of throughput and costs of about US\$1000. Therefore, although the urgent need for a very fast algorithm such as JPDA, JVC, or NN has disappeared, the need for an algorithm with good performance that can be coded and tested rapidly has not. A carefully designed MHT that accounts for unresolved measurements provides vastly superior data association and track initiation performance in dense multiple target environments compared with single-scan, single-hypothesis algorithms such as JPDA, JVC, or NN. We are thus forced to use MHT for applications that have very challenging performance requirements.

It has been shown that the optimal state estimator in the presence of data association uncertainty consists of the computation of the conditional pdf of the state  $x(k)$  given all information available at time  $k$ , namely, the prior information about the initial state, the intervening known inputs, and the sets of measurements through time  $k$ . It has also been pointed out that if the exact conditional pdf, which is a mixture, is available, then its recursion requires only the probabilities of the most recent association events. The conditions under which this result holds, namely, whiteness of the noise, detection and clutter processes, were presented.

The PDAF and JPDAF algorithms, which carry out data association and state estimation in clutter, have been described. A simple example was given to illustrate how the clutter and occasional missed detections can lead to track loss for a standard tracking filter, and how PDAF can keep the target in track under such circumstances. A simulated space-based surveillance example showed, using Monte Carlo runs, that PDAF can track a target in a level of clutter in which the NNSF loses the target with high probability. The main approximation for both these algorithms is the single Gaussian used as the state's conditional pdf. The fact that PDAF and

JPDAF do not need to recompute past associations is a consequence of the assumption that the information state is available, even though the assumption of a Gaussian information state is only an approximate one. The fact that this approximation is quite good is proven by the numerous successful real-world systems that use it. The MHT, on the other hand, uses an implicit finite mixture through its hypothesis tree, which is a better approximation of the exact information state. The complexity of the MHT in terms of computation time, memory, and code length/debugging is orders of magnitude higher than that of the PDAF/JPDAF.

The numerous applications of the PDAF/JPDAF illustrated in "Real-World Applications of PDAF and JPDAF" show the potential pitfalls of using sophisticated algorithms for tracking in difficult environments as well as how to overcome them. The effect of finite sensor resolution can be a more severe problem than the data association [26] and deserves special attention.

## AUTHOR INFORMATION

**Yaakov Bar-Shalom** (ybs@ee.uconn.edu) received the B.S. and M.S. degrees from the Technion and the Ph.D. degree

from Princeton University. Currently he is Board of Trustees Distinguished Professor in the Department of Electrical and Computer Engineering and Marianne E. Klewin Professor in Engineering at the University of Connecticut. His current research interests are in estimation theory and target tracking. He has published over 380 papers and seven books. He is a Fellow of the IEEE. He served as associate editor of *IEEE Transactions on Automatic Control* (1976–1977) and *Automatica* (1978–1981). He was general chair of the 1985 ACC and chair of the IEEE Control Systems Society Conference Activities Board. He was president of the International Society of Information Fusion (2000 and 2002) and is currently vice president for Publications. He is a Distinguished Lecturer of the IEEE Aerospace and Electronic Systems Society. He is corecipient of the M. Barry Carlton Award for the best paper in *IEEE Transactions on Aerospace and Electronic Systems* in 1995 and 2000. In 2002 he received the Joseph Mignona Data Fusion Award from the DoD JDL Data Fusion Group. He is a member of the Connecticut Academy of Science and Engineering and the recipient of the 2008 IEEE Dennis J. Picard Medal for Radar Technologies and Applications. He can be contacted at the University of Connecticut, ECE Department, Storrs, CT 06269-2157 USA.

**Fred Daum** received the B.S. degree from the New Jersey Institute of Technology and the M.S. degree from Harvard University and is an IEEE Fellow. He is a senior principal fellow at Raytheon Integrated Defense Systems in Woburn, Massachusetts, where he has worked since 1969. He was awarded the Tom Phillips prize for technical excellence. He has published nearly 100 technical papers in nonlinear filtering, target tracking, and discrimination.

**Jim Huang** received the Ph.D. in systems engineering from Boston University in 1995. He is currently a senior principal systems engineer at Raytheon Integrated Defense Systems in Woburn, Massachusetts. His current research interests are in the areas of target tracking, target discrimination, and nonlinear filtering.

## REFERENCES

- [1] S. S. Blackman, *Multiple Target Tracking with Radar Applications*. Norwood, MA: Artech House, 1986.
- [2] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. San Diego, CA: Academic, 1988.
- [3] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.
- [4] S. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.
- [5] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*. New York: Wiley, 2001.
- [6] O. E. Drummond, "Comparison of 2-D assignment algorithms for sparse, rectangular, floating point cost matrices," *J. SDI Panels on Tracking*, vol. 4, pp. 4-81–4-97, Dec. 1990.
- [7] D. P. Bertsekas, *Network Optimization*. Belmont, MA: Athena Scientific, 1998.
- [8] C. Striebel, "Sufficient statistics in the optimum control of stochastic systems," *J. Math. Anal. Appl.*, vol. 12, pp. 576–592, Dec. 1965.
- [9] D. J. Salmond, "Mixture reduction algorithms for target tracking in clutter," in *Proc. SPIE Conf. Signal and Data Processing of Small Targets*, Apr. 1990, vol. 1305, pp. 874–887.
- [10] H. A. P. Blom and E. Bloem, "Joint particle filtering of multiple maneuvering targets from possibly unassociated measurements," *J. Advances Inform. Fusion*, vol. 1, no. 1, pp. 15–34, July 2006.
- [11] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Automat. Contr.*, vol. AC-24, no. 6, pp. 843–854, Dec. 1979.
- [12] T. Kurien, "Issues in the design of practical multitarget tracking algorithms," in *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed. Norwood, MA: Artech House, 1990, ch. 3, pp. 43–84.
- [13] Y. Bar-Shalom, S. S. Blackman, and R. J. Fitzgerald, "The dimensionless score function for measurement to track association," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 392–400, Jan. 2007.
- [14] S. B. Colegrove, A. W. Davis, and J. K. Ayliffe, "Track initiation and nearest neighbors incorporated into probabilistic data association," *J. Electr. Electron. Eng. Australia*, vol. 6, no. 3, pp. 190–198, Sept. 1986.
- [15] D. Musicki, R. Evans, and S. Stankovic, "Integrated probabilistic data association," *IEEE Trans. Automat. Contr.*, vol. 29, no. 6, pp. 1237–1241, June 1994.
- [16] R. Chakravorty and S. Challa, "Augmented state integrated probabilistic data association smoothing for automatic track initiation in clutter," *J. Advances Inform. Fusion*, vol. 1, no. 1, pp. 63–74, July 2006.
- [17] H. A. P. Blom and E. Bloem, "Probabilistic data association avoiding track coalescence," *IEEE Trans. Automat. Contr.*, vol. 45, no. 17, pp. 247–259, 2000.
- [18] H. A. P. Blom and E. Bloem, "Interacting multiple model JPDA avoiding track coalescence," in *Proc. IEEE Conf. Decision and Control*, Dec. 2002, pp. 3408–3415.
- [19] X. R. Li and Y. Bar-Shalom, "Stability evaluation and track life of the PDAF for tracking in clutter," *IEEE Trans. Automat. Contr.*, vol. AC-36, pp. 588–602, May 1991.
- [20] P. Lanzkron and E. Brookner, "Solid state X-band airport surface surveillance radar," in *Proc. IEEE Radar Conf.*, Apr. 2007, pp. 670–676.
- [21] R. J. Fitzgerald, "Development of practical PDA logic for multitarget tracking by microprocessor," in *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed. Norwood, MA: Artech House, 1990, ch. 1, pp. 1–24.
- [22] R. Helmick, "IMM estimator with nearest-neighbor joint probabilistic data association," in *Multitarget-Multisensor Tracking: Applications and Advances*, vol. III, Y. Bar-Shalom and D. Blair, Eds. Norwood, MA: Artech House, 2000, ch. 3, pp. 161–198.
- [23] F. Daum and R. J. Fitzgerald, "Decoupled Kalman filters for phased array radars," *IEEE Trans. Automat. Contr. (Special Issue on Kalman Filters)*, vol. 28, pp. 269–283, Mar. 1983.
- [24] R. J. Fitzgerald, "Performance comparisons of some association algorithms," in *Proc. ONR/GTRI Workshop Target Tracking and Sensor Fusion*, Key West, FL, June 2004.
- [25] R. J. Fitzgerald, "Track biases and coalescence with probabilistic data association," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 21, no. 6, pp. 822–825, Nov. 1985.
- [26] F. Daum and R. J. Fitzgerald, "The importance of resolution in multiple target tracking," in *Proc. SPIE Conf.*, Orlando, FL, Apr. 1994, vol. 2235, pp. 329–338.
- [27] W. Koch and G. van Keuk, "Multiple hypothesis track maintenance with possibly unresolved measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 3, pp. 883–892, July 1997.
- [28] H. A. P. Blom and E. Bloem, "Bayesian tracking of two possibly unresolved maneuvering targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 2, pp. 612–627, Apr. 2007.
- [29] D. Musicki and M. R. Morelande, "An integrated particle filter for a finite resolution sensor," in *Proc. IEEE Conf. Decision and Control*, Dec. 2004, pp. 2100–2105.
- [30] F. Daum, "Nonlinear filters: Beyond the Kalman filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 20, no. 8, pp. 269–283, Aug. 2005.
- [31] F. Daum, "A system approach to multiple target tracking," in *Multitarget-Multisensor Tracking*, vol. II, Y. Bar-Shalom, Ed. Norwood, MA: Artech House, 1992, ch. 6. Reprinted by YBS Publishing, Storrs, CT, 1998.
- [32] K. C. Chang and Y. Bar-Shalom, "Joint probabilistic data association for multitarget tracking with possibly unresolved measurements and maneuvers," *IEEE Trans. Automat. Contr.*, vol. AC-29, no. 7, pp. 585–594, July 1984.

