Integration of Bayes Detection with Target Tracking

Peter Willett, *Senior Member, IEEE*, Ruixin Niu, and Yaakov Bar-Shalom, *Fellow, IEEE*

***Abstract*—Existing detection systems generally are operated using a fixed threshold and optimized to the Neyman–Pearson criterion. An alternative is Bayes detection, in which the threshold varies according to the ratio of prior probabilities. In a recursive target tracker such as the probabilistic data association filter (PDAF), such priors *are* available in the form of a predicted location and associated covariance; however, the information is not at present made available to the detector. Put another way, in a standard detection/tracking implementation, information flows only one way: from detector to tracker. Here, we explore the idea of two-way information flow, in which the tracker instructs the detector where to look for a target, and the detector returns what it has found. More specifically, we show that the Bayesian detection threshold is lowered in the vicinity of the predicted measurement, and we explain the appropriate modification to the PDAF. The implementation is simple, and the performance is remarkably good.**

I. Introduction

MOST TARGET tracking systems work with the data they are given. By this, we mean that measurements from a detection front-end processor are interrogated for threshold exceedances, and these "hits" are delivered to the tracking algorithm. For the most part, the threshold is set and fixed according to a false-alarm criterion that indicates that there should be, on average, a specified number of false hits per unit volume. There have been studies relating the tracking performance to this threshold and suggesting *global* threshold-settings for optimized performance for a given expected signal-to-noise ratio (SNR) [5]. Further, there has been some research indicating that considerably improved performance is achievable when some amplitude information (AI) is delivered to the tracker along with the measurements and their locations [6], [8].

The above two points have largely been investigated as they pertain to the probabilistic data association filter (PDAF) [2]. The PDAF is a particularly simple and successful target tracking algorithm. It is predicated on the assumptions that the best one-step estimation of the target's location should be sufficient and that once this estimation is accomplished, the target's true location should be afforded a Gaussian distribution about its estimated value. The key to this paper is in this "posterior" distribution on the target's location; in the PDAF case, this is Gaussian and easy to specify.

Communication between the signal processing front-end and the PDAF is presently one way. In this paper, we allow two-way communication or, perhaps more appropriately, "feedback" from the tracker to the detector. The form of this feedback is of the posterior distribution on the target's location. From the detector's point of view, this is prior information for its hypothesis tests (i.e., its matched filters), as represented in Fig. 1. Thus, a detector using this configuration no longer operates in a Neyman–Pearson mode and instead becomes Bayesian, and from a practical point of view, this amounts to a threshold that is depressed near where a target is expected to be and elevated where it is unexpected—this is illustrated in Fig. 2.

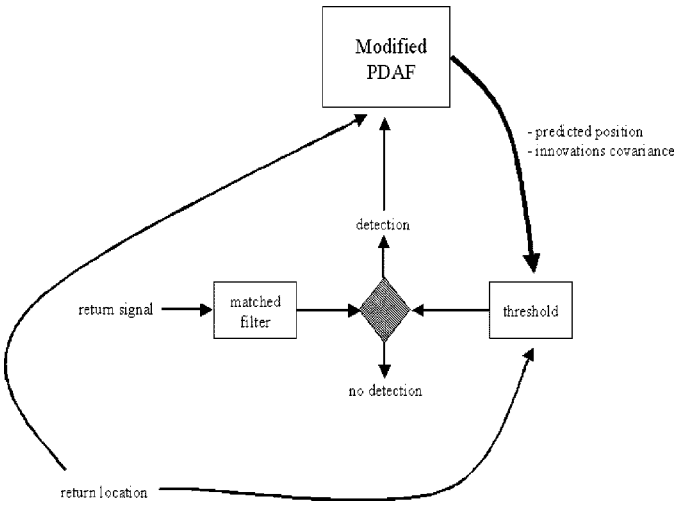

Fig. 1. Representation of flow of data within proposed system. A signal return from a known location is matched filtered and its magnitude compared with a threshold—a threshold exceedance, along with its location, is passed to the tracker, which is a modified PDAF. The threshold itself is determined as a function of the predicted location of the target, the innovation covariance, and the location of the return.

In this new approach, there are fewer false alarms than previous, and these are no longer uniformly distributed in space as they would be for the PDAF. Thus, the PDAF must be modified accordingly, which we do in this paper; the resulting algorithm (the PDAF-BD referring to the Bayesian detector) is arguably simpler than the PDAF, and its performance is considerably better.

As indicated above, there has emerged a new PDAF that uses *amplitude information*, which has not unnaturally been coined the PDAF-AI. In such an implementation, the Neyman–Pearson detector structure of the PDAF is preserved, but in addition to the locations of threshold exceedances, the corresponding amplitudes are reported to the tracker. The original PDAF must assign threshold exceedances as true or false based only on their location relative to that which is expected; amplitude information refines this by functioning as a discriminating feature, and the improvement (relative to the original PDAF) can be dramatic. By contrast, the PDAF-BD communicates only the *lo-*

Manuscript received January 17, 2000; revised September 13, 2000. This work was supported by AFOSR under Contract F49620-95-1-0299 and by ONR under Contract N00014-97-1-0502. The associate editor coordinating the review of this paper and approving it for publication was Prof. Lang Tong.

The authors are with Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269 USA (e-mail: willett@engr.uconn.edu; rxniu@engr.uconn.edu; ybs@engr.uconn.edu).

Publisher Item Identifier S 1053-587X(01)00069-1.

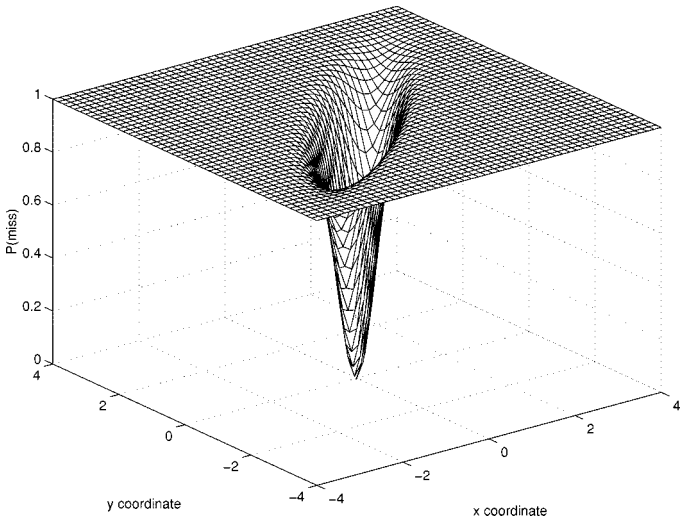

Fig. 2. Illustration of the effect of a position-dependent threshold. The x and y coordinates are those of the innovation, that is, of the one-step predicted measurement subtracted from the return location. The z coordinate shows the probability that a return of a given strength will be missed as a function of its normalized innovation.

cations of threshold exceedances, but owing to the location-dependent threshold, there is a form of amplitude information preserved in that a large innovation must have been accompanied by a high amplitude. It is therefore not surprising that the performance of the PDAF-BD lies between that of the PDAF and the more information-rich PDAF-AI.

Emboldened by this, we also present a modification on the PDAF-BD in which amplitude information is *also* reported, and (with apologies) we term this the PDAF-BDAI. The PDAF-BDAI is considerably better than the PDAF-BD in terms of performance, and indeed, in terms of lost tracks, this version outperforms the PDAF-AI in the cases investigated.

It is worth mentioning that there are trackers other than the PDAF, for example, the multihypothesis tracker (MHT) (e.g., [4]), the EM-based probabilistic multihypothesis tracker (e.g. [1], [7], [9], [15]), and the assignment-based trackers (e.g., [12], [14]). The idea behind the PDAF-BD, of a tracking-dependent Bayesian detection threshold, could probably be applied to any of these—our focus is on the PDAF as an example and due to the nice resulting structure, which will be seen shortly.

It is also reasonable to note that present day detection/tracking systems operate in the PDAF (Neyman–Pearson detection) mode, and modification to a PDAF-BD (or for that matter PDAF-AI) structure may or may not be straightforward. Similarly, the extant tracking systems as above have evolved to deal with a number of complications such as of multiple targets, target maneuver, and multiple sensors. It would thus be necessary to extend the PDAF-BD and its ideas to incorporate these before it could be a serious competitor. The goal of this paper is, consequently, not bravely to suggest a wholesale switch to the PDAF-BD, but rather to propose it as a promising alternative worthy of further development and perhaps to provoke designers of detection systems to consider incorporating a Bayesian-thresholding capability in future generations of their products.

In Section II, we present our assumed target tracking model, the original PDAF, and the PDAF-AI. In Section III, we first explain the Bayesian thresholding and then develop the PDAF-BD and the PDAF-BDAI—most of the theory is in the Appendices. In Section IV, the four trackers are compared, and Section V offers concluding remarks.[1]

II. Background

A. Model of Tracking

Let us agree on the standard tracking terminology that

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{v}_k \quad \mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{w}_k \tag{1}$$

where

- $\mathbf{x}$ target state (to be estimated);
- $\mathbf{y}$ measurement;
- k time index.

The transition and observation matrices $\mathbf{F}$ and $\mathbf{H}$ are assumed known, and the respective process and measurement noises are independent, white, and Gaussian and have

$$E\{\mathbf{v}_k\mathbf{v}_k^T\} = \mathbf{Q} \tag{2}$$

$$E\{\mathbf{w}_k\mathbf{w}_k^T\} = \mathbf{R} \tag{3}$$

as their associated covariance matrices. Extension to time-varying systems is obvious and will be avoided here for clarity. Based on $\{\mathbf{y}_k\}$, the optimal estimator would be a Kalman filter, but target tracking is made interesting by the data association problem that at time (scan) k, no single $\mathbf{y}_k$ is available, but instead, a set of candidate observations $\mathcal{Z}_k = \{\mathbf{z}_k(l)\}_{l=1}^{n_k}$ are available. In practice, $\mathcal{Z}_k$ are threshold exceedances of matched filter outputs, but for the purposes of PDAF specification, we have the following.

Assumption 1: In the development of the PDAF, it is assumed that $\mathcal{Z}_k$, which are the observations at time k, comprise n_k constituents.

1) The observation $\mathcal{Z}_k$ takes of the form of n_k vectors $\{\mathbf{z}_k(l)\}_{l=1}^{n_k}$ of the same dimension as $\{\mathbf{y}_k\}$. These are the locations of whatever threshold exceedances have been observed.
2) With probability P_d, the true measurement $\{\mathbf{y}_k\}$ [from (1)] may be among the $\{\mathbf{z}_k(l)\}_{l=1}^{n_k}$, and with probability $1-P_d$, it may be absent, corresponding in this latter case to a missed detection.
3) The number of *false-alarm* constituents (either n_k or n_k-1, depending on whether the true measurement is present) of $\mathcal{Z}_k$ is accorded a Poisson distribution with mean λV, in which λ is referred to as the spatial clutter density, and V is the observation volume.
4) False-alarm measurements are uniformly distributed within the observation volume V and are independent.

The ordering of the measurements conveys no information and may be considered a random permutation. Thus, determination of which (if any) of the n_k constituents of $\mathcal{Z}_k$ is target generated is the *association problem*. □

[1]Portions of this paper have appeared as [16]; however, this paper contains considerable modifications and extensions.

It is possible to take issue with any of the above assumptions, but they are standard and we use them. There will be a need for modifications for development of the PDAF-AI, PDAF-BD, and PDAF-BDAI, and we will give these as needed.

B. Original PDAF

At the outset , let us note the informing feature of the PDAF; it is entirely optimal, *except* that after each scan, its posterior track probability density function—ideally a mixture of Gaussian pdfs—is converted to a single Gaussian mode having the same mean and variance. Thus, at each scan, estimation is built upon a Gaussian prior and converted to a Gaussian mixture posterior, which is then forced back to Gaussianity for the succeeding scan.

Assume that the target location up to time $k-1$ is estimated as $\hat{\mathbf{x}}_{k-1|k-1}$ with associated covariance $\mathbf{P}_{k-1|k-1}$. The notation in the subscripts indicates that the estimate is *conditioned* on $\{\mathcal{Z}_i\}_{i=-\infty}^{k-1}$. Operation of the PDAF based on one scan of data (the kth) can be summarized as [2], [3] follows.

1) Predict the target location at scan k from prior at scan $k-1$:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}. \tag{4}$$

2) From $\mathcal{Z}_k$, form the "innovations" (ν's)

$$\nu_k(l) = \mathbf{z}_k(l) - \mathbf{H}\hat{\mathbf{x}}_k \tag{5}$$

of all n_k candidate measurements. Compute

$$\mathbf{S}_k = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R} \tag{6}$$

for the innovations covariance of the true measurement in which

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T + \mathbf{Q} \tag{7}$$

is the prediction covariance and is calculated separately since it is also used in a later step.

3) Assuming that $n_k > 0$, calculate the association probabilities

$$\beta(\theta) = \begin{cases} c\dfrac{(1-P_d)\lambda}{P_d}\sqrt{|2\pi\mathbf{S}_k|}, & \theta = 0 \\ ce^{-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)}, & 1 \le \theta \le n_k \end{cases} \tag{8}$$

in which c ensures that $\sum_{\theta=0}^{n_k}\beta(\theta) = 1$. Here, $\beta(\theta)$ is the posterior probability that measurement θ is from the true target; $\beta(0)$ is the posterior probability that all measurements at this scan are spurious.

4) Use these βs to form a synthetic "innovation," and update the track according to

$$\mathbf{x}_{k|k} = \mathbf{F}\mathbf{x}_{k-1|k-1} + \mathbf{W}_k\nu_k \tag{9}$$

in which

$$\mathbf{W}_k = \mathbf{P}_{k|k-1}\mathbf{H}^T\mathbf{S}_k^{-1} \tag{10}$$

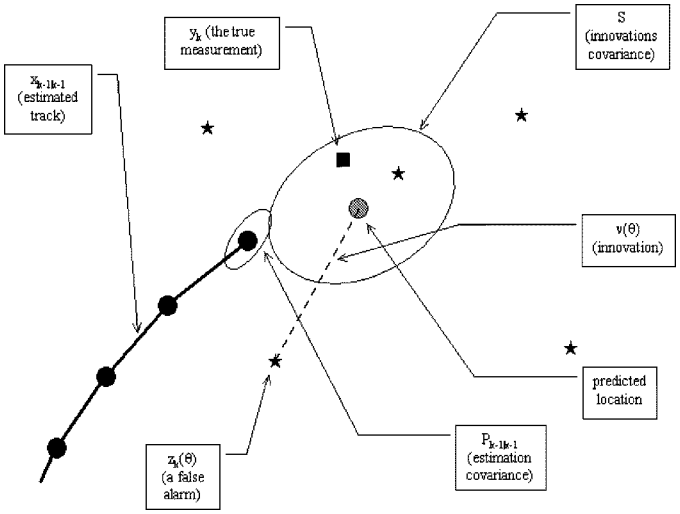

Fig. 3. Representation of the one-step tracking as performed by the PDAF.

is the Kalman gain, and

$$\nu_k = \sum_{\theta=1}^{n_k}\beta(\theta)\nu_k(\theta) \tag{11}$$

is the "aggregate" innovation.

5) Update

$$\mathbf{P}_{k|k} = \beta(0)\mathbf{P}_{k|k-1} + \sum_{\theta=1}^{n_k}\beta(\theta)[\mathbf{P}_{k|k-1} - \mathbf{W}_k\mathbf{S}_k\mathbf{W}_k^T] + \mathbf{W}_k\left[\sum_{\theta=1}^{n_k}\beta(\theta)\nu_k(\theta)\nu_k(\theta)^T - \nu_k\nu_k^T\right]\mathbf{W}_k^T \tag{12}$$

for the estimation covariance. The third term in (12) is often referred to as the "spread of the innovations."

Reference to Fig. 3 may be helpful. The above sequence may be unfamiliar, and the reader is encouraged examine the derivation in [2] and [3]. It should be noted that in practice, the predicted measurement is often enclosed by a "gate" whose volume is proportional to $|\mathbf{S}_k|$ and whose function is to reduce computation by ignoring any $\mathbf{z}_k(l)$ for which $\beta(l) \sim 0$. In theory, no gate is necessary, and the modifications to the above if a gate is used are fairly straightforward.

C. PDAF-AI

As discussed earlier, it has recently been shown [3], [8] that the use of amplitude information can be of significant benefit to the PDAF.

Assumption 2: In the PDAF-AI, we have the following.

1) The observation $\mathcal{Z}_k = \{\mathbf{z}_k(l), a_k(l)\}_{l=1}^{n_k}$. The $\mathbf{z}_k(l)$s are as in Assumption 1; now, $a_k(l)$ is the amplitude associated with the lth threshold exceedance at scan k.

This replaces item 1 of Assumption 1. □

That is, instead of "measurements" consisting simply of the locations of threshold exceedances, these are augmented by information as to how much the threshold was exceeded. Thus, it may be expected that a strong target return would be more recognizable as such than if this confidence information were thrown away by the detector, and in fact, this is so. It is interesting that the PDAF structure is little altered by the presence of

amplitude information; the only change is to the calculation of the βs. In fact, we have that (8) is replaced by

$$\beta(\theta)=\begin{cases} c\,\dfrac{(1-P_d)\lambda}{P_d}\sqrt{|2\pi\mathbf{S}_k|}, & \theta=0 \\ ce^{-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)}\left(\dfrac{f_1(a_k(\theta))}{f_0(a_k(\theta))}\right), & 1\le\theta\le n_k \end{cases} \tag{13}$$

in which $f_0(\cdot)$ and $f_1(\cdot)$ are the probability distributions of measured amplitude a_θ, respectively, under false-alarm and true target hypotheses and conditioned on the event that the threshold has been exceeded. Because the locations and the amplitudes are independent, the posterior probability $\beta(\theta)$ should be multiplied by $f_0(\cdot)$ (for $\theta=0$) or $f_1(\cdot)$ (for $1\le\theta\le n_k$). It is often convenient to express quantities in terms of dimensionless likelihood ratios, and hence, we divide all probabilities by f_0 and renormalize to get the (f_1/f_0) factor in (13). Replacement of (8) by (13) is the only operational difference between the PDAF and PDAF-AI. This extends to the PDAF-BD and PDAF-BDAI; although the signal processing (meaning the thresholding) is different from either PDAF or PDAF-AI, the only variation with regard to tracking is again in the calculation of the βs. In the following, we show how to form these.

III. Development of the PDAF-BD and PDAF-BDAI

A. Statistical Testing

We assume that a test of absence or presence of a target at location $\mathbf{z}_k(l)$ is to be performed. Hypothesis $\mathbf{H}$ is that there is no target at location $\mathbf{z}_k(l)$ and, hence, that the measured return is due simply to noise. Hypothesis $\mathbf{K}$ is that there is indeed a target at location $\mathbf{z}_k(l)$ and, hence, that the return is due to a combination of noise and signal energy. That is, we write

$$\mathbf{H}: \quad f(a_k(l))=e^{-a_k(l)} \tag{14}$$

$$\mathbf{K}: \quad f(a_k(l))=\frac{1}{1+\rho}\,e^{-a_k(l)/(1+\rho)} \tag{15}$$

in which $a_k(l)$ is the corresponding amplitude (magnitude-square output of a matched filter, with a Swerling I target fluctuation model implicit), and ρ is the SNR.[2] The usual implementation is according to the Neyman–Pearson criterion [13] that the probability of detection be maximized subject to a constraint on the false alarm rate, and the resulting test can easily be shown to be a comparison of $a_k(l)$ to a fixed threshold. From the Bayesian viewpoint, the appropriate test is

$$\left(\frac{1}{1+\rho}\,e^{a_k(l)\rho/(1+\rho)}=\right)\frac{f_K(a_k(l))}{f_H(a_k(l))}\underset{H}{\overset{K}{\gtrless}}\frac{\Pr(H)[c_{KH}-c_{HH}]}{\Pr(K)[c_{HK}-c_{KK}]} \tag{16}$$

in which

$\Pr(j)$	hypothesis $j(\in\{H,K\})$;
$f_j(\cdot)$	pdf given hypothesis j;
c_{ij}	cost of making decision i when j is true.

We note that these costs are not easily available.

[2]In the formulation given, it is apparent that the returns are assumed perfectly prenormalized such that the target-absent mean is unity. If some other target-model—such as a CA-CFAR distribution—is desirable, then the succeeding development must be modified. This modification is straightforward.

The "prior" probabilities $\Pr(H)$ and $\Pr(K)$ are not well-posed; the latter amounts to the probability that a target is located *exactly* at the test's coordinates $\mathbf{z}_k(l)$ given the prior tracking information, and this is zero. If a sampling grid of resolution cells is available, then the quantity can be calculated, but since the answer is configuration-specific, we prefer to avoid this and simply note that $\Pr(K)\propto\left(\frac{1}{\sqrt{|2\pi\mathbf{S}_k|}}\right)e^{-(1/2)\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)}$ and $\Pr(H)\propto(1/V)$, where V is the volume of the validation region. Therefore, we have

$$\frac{\Pr(K)}{\Pr(H)}\propto e^{-(1/2)\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)} \tag{17}$$

in which

$$\nu_k(l)=\mathbf{z}_k(l)-\mathbf{HF}\hat{\mathbf{x}}_{k-1|k-1} \tag{18}$$

is as before the *spatial* innovation of the lth measurement at time k, and $\hat{\mathbf{x}}_{k-1|k-1}$ is the estimate of the state given data up to scan $k-1$. It should be noted that the prior probabilities are of *isolated* tests, meaning that each observation is tested separately, which is fair.

At any rate, from (14)–(17), we have, with reference to Fig. 1, the test

$$a_k(l)\underset{H}{\overset{K}{\gtrless}}\frac{\rho+1}{2\rho}\,\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)+\eta. \tag{19}$$

In (19) η is a tunable parameter, meaning that the proportionality in (17) has been absorbed within it.

B. Probabilities of Detection and False Alarm

In the case of the PDAF and PDAF-AI, the probabilities of detection and false alarm are straightforward and do not depend on location. The PDAF-BD has a location-dependent threshold, and hence, these probabilities are easy only as conditioned by ν—both are more probable for smaller ν, meaning that it is more likely to see a threshold exceedance close to $\mathbf{HF}\hat{\mathbf{x}}_{k-1|k-1}$ than one further away. At any rate, it will be useful to have the unconditioned quantities, and these can be calculated as follows.

1) Detection: Via the test and threshold of (19), it is possible to calculate the overall probability of detection as

$$\begin{aligned} P_d&=\int\Pr(detection|\nu)f(\nu)\,d\nu \\ &=\int e^{-(1/(\rho+1))[(\rho+1/2\rho)\nu^T\mathbf{S}^{-1}\nu+\eta]}\,\frac{1}{\sqrt{|2\pi\mathbf{S}|}} \\ &\quad\cdot e^{-(1/2)\nu^T\mathbf{S}^{-1}\nu}\,d\nu \\ &=\frac{1}{\sqrt{|2\pi\mathbf{S}|}}\,e^{-\eta/(1+\rho)}\int e^{-(1/2)\nu^T[(\rho/(\rho+1))\mathbf{S}]^{-1}\nu}\,d\nu \\ &=\left(\frac{\rho}{\rho+1}\right)^{n_z/2}e^{-\eta/(\rho+1)} \end{aligned} \tag{20}$$

by averaging over all possible *true* innovations ν. Here, n_z is the dimension of the measurement [i.e., of $\mathbf{y}_k$; see (1)].

2) False Alarm: In the PDAF (and PDAF-AI), false alarms are assumed to be generated by an underlying Poisson point process, and hence, we have the probability mass function (pmf)

of the number of false alarms in a measurement space volume V as

$$\mu(m) = \frac{(\lambda V)^m}{m!}\, e^{-\lambda V} \tag{21}$$

in which λ is the average number of false alarms per unit volume for the (fixed) thresholding used. The expression for μ is necessary in the evaluation of the βs. The above is so simple that it may seem strange to devote much space to it, but in the case of the PDAF-BD and the PDAF-BDAI, the number of false alarms is controlled by the detection thresholding, and hence, the answer is not straightforward.

For the PDAF-BD, we cannot rely on the same homogeneous Poisson point process since the spatially varying PDAF-BD threshold may (for small ν) be below the spatially invariant threshold assumed by the PDAF and PDAF-AI. As such, we propose the following.

Assumption 3: There exists an underlying homogeneous Poisson point process with spatial density $\tilde{\lambda}$. To each event in this new process we attach an amplitude with a unit exponential distribution, and each amplitude is independent. □

Let us suppose that amplitudes from the events from this new Poisson point process (with spatial density $\tilde{\lambda}$) are tested against the threshold τ with only those which exceed τ being kept and those whose amplitudes lie below τ discarded. It is easy to see that the result is also a Poisson point process, this time with spatial density $\lambda = \tilde{\lambda} e^{-\tau}$. This thresholded Poisson point process (with spatial density λ) is that which generates false alarms for the PDAF and PDAF-AI; the unthresholded Poisson point process (with spatial density $\tilde{\lambda}$) models the input to the PDAF-BD and PDAF-BDAI, but it is important to realize that the *same* underlying Poisson point process is assumed. The hierarchy of clutter-generating Poisson point processes is sketched in Fig. 4.

Now, for a given event from the underlying Poisson point process (with spatial density $\tilde{\lambda}$), the probability that it is reported to the PDAF or PDAF-AI, meaning that it exceeds the threshold τ, is $P_{te} = e^{-\tau}$. For the PDAF-BD or PDAF-BDAI, the corresponding probability is

$$\begin{aligned} P_{te} &= \int_V \Pr(\text{threshold exceedance}|\nu) f(\nu)\, d\nu \\ &= \int_V e^{-(\rho+1/2\rho)\nu^T \mathbf{S}^{-1}\nu - \eta}\, \frac{1}{V}\, d\nu \\ &= \frac{1}{V}\, e^{-\eta} \sqrt{\left|2\pi \frac{\rho}{\rho+1}\mathbf{S}\right|} \int_V \frac{1}{\sqrt{\left|2\pi \frac{\rho}{\rho+1}\mathbf{S}\right|}} \\ &\quad \cdot e^{-(1/2)\nu^T((\rho/(\rho+1))\mathbf{S})^{-1}\nu}\, d\nu \\ &\simeq \frac{1}{V}\, e^{-\eta} \sqrt{\left|2\pi \frac{\rho}{\rho+1}\mathbf{S}\right|} \end{aligned} \tag{22}$$

if the volume V is sufficiently big. [The parameter η is from (19).] The last step in (22) follows from the identification of the integrand as a Gaussian density function whose restriction

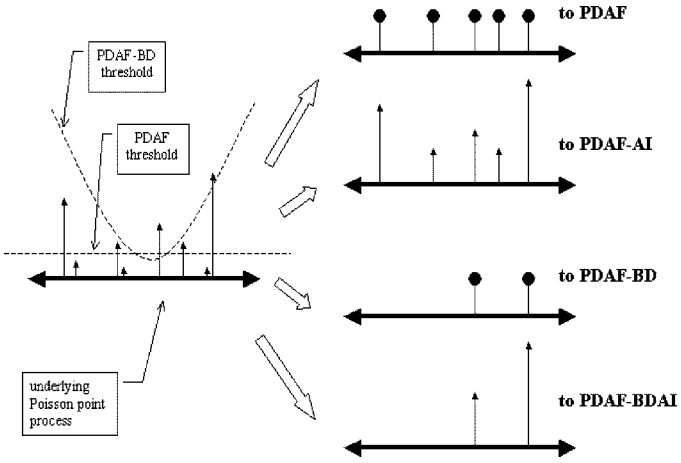

Fig. 4. Illustration of the underlying Poisson point process (with spatial density $\tilde{\lambda}$) as it applies to the four tracking algorithms. Input to the PDAF consists only of locations of exceedances of a spatially invariant threshold; input to the PDAF-AI is similar, but amplitudes also are reported. Locations of exceedances of a spatially varying threshold are passed to the PDAF-BD, and the PDAF-BDAI has, additionally, the corresponding amplitudes.

to a volume V integrates approximately to unity, provided V is large. Now, assume that the underlying Poisson process has generated n points in the volume V. The probability that there are m threshold exceedances (false alarms) is binomial with mean nP_{te}. That is, we have for the probability that there are m false alarms

$$\begin{aligned} \mu(m) &= \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} (P_{te})^m (1-P_{te})^{n-m} \frac{(\tilde{\lambda}V)^n}{n!}\, e^{-\tilde{\lambda}V} \\ &= \frac{(\tilde{\lambda}VP_{te})^m}{m!}\, e^{-\tilde{\lambda}V} \sum_{l=0}^{\infty} \frac{(\tilde{\lambda}V)^l}{l!} (1-P_{te})^l \\ &= \frac{(\tilde{\lambda}VP_{te})^m}{m!}\, e^{-\tilde{\lambda}VP_{te}} \end{aligned} \tag{23}$$

in which P_{te} is given in (22). Comparing $\mu(m)$ of (21) to $\mu(m)$ of (23), we see that for the PDAF-BD, the number of false alarms in a volume V is again Poisson but with a different mean: λV for the PDAF and

$$\tilde{\lambda}VP_{Te} = \lambda e^{\tau-\eta} \sqrt{\left|2\pi \frac{\rho}{\rho+1}\mathbf{S}\right|} \tag{24}$$

for the PDAF-BD. We note that the above expressions relate to the special case of the nonuniform false alarm process generated by the PDAF-BD threshold. There is a more general treatment of these in [10] and [11] from which an alternative derivation of (22)–(24) is possible.

C. PDAF-BD

The impact of the spatially varying detection threshold on PDAF-BD operation is only through the posterior association probabilities. In accordance with the standard PDAF nomenclature, we refer to these as the βs, in which $\beta(\theta)$ is the probability that measurement $\mathbf{z}_k(\theta)$ is target generated (and all others are clutter), whereas $\beta(0)$ is the probability that all measurements

are clutter. It is shown in Appendix B that we have

$$\beta(\theta)=\begin{cases} c\lambda e^{-\eta}\sqrt{\left|2\pi\,\dfrac{\rho}{\rho+1}\,\mathbf{S}_k\right|} \\ \quad\cdot\left[\left(\dfrac{\rho+1}{\rho}\right)^{n_z/2} e^{\eta/(1+\rho)}-1\right] & \theta=0 \\ c & 1\le\theta\le n_k \end{cases} \tag{25}$$

in which c is a constant such that $\sum_{\theta=0}^{n_k}\beta(\theta)=1$. This is remarkable and simple: $\beta(\theta)$ *is a constant*. The anomaly is presumably due to the fact that a spatially "surprising" measurement must have had a larger amplitude to exceed its threshold. Here, the effects of large amplitude and large distance from predicted position (prior information) cancel out each other and make a constant weighting.

D. PDAF-BDAI

Here, both locations and amplitudes of events that exceed the (spatially varying) threshold are reported. As for the PDAF-AI and PDAF-BD, the only algorithmic impact is through the βs. Assuming the innovation of measurement θ is $\nu_k(\theta)$ and that the corresponding amplitude is $a_k(\theta)$, it is derived in Appendix A that we get

$$\beta(\theta)=\begin{cases} c(1+\rho)\lambda\sqrt{\left|2\pi\dfrac{\rho}{\rho+1}\mathbf{S}_k\right|} \\ \quad\cdot\left[\left(\dfrac{\rho+1}{\rho}\right)^{n_z/2}-e^{-\eta/(1+\rho)}\right] & \theta=0 \\ c\cdot e^{(\rho/(\rho+1))a_k(\theta)-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)} & 1\le\theta\le n_k \end{cases} \tag{26}$$

where, as before, c is a normalizing constant such that $\sum_{\theta=0}^{n_k}\beta(\theta)=1$.

IV. Comparison

In this section, we compare the PDAF, the PDAF-AI, the PDAF-BD, and the PDAF-BDAI; there is little alternative but that the basis be simulation. For these simulations, we choose the common two-dimensional (2-D) kinematic model with direct discrete-time process noise [2]. Accordingly, we have

$$\mathbf{F}=\begin{Bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{Bmatrix} \quad \mathbf{H}=\begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{Bmatrix}$$

$$\mathbf{Q}=\sigma_p^2\begin{Bmatrix} \dfrac{\Delta t^4}{4} & \dfrac{\Delta t^3}{2} & 0 & 0 \\ \dfrac{\Delta t^3}{2} & \Delta t^2 & 0 & 0 \\ 0 & 0 & \dfrac{\Delta t^4}{4} & \dfrac{\Delta t^3}{2} \\ 0 & 0 & \dfrac{\Delta t^3}{2} & \Delta t^2 \end{Bmatrix} \quad \mathbf{R}=\sigma_m^2\begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} \tag{27}$$

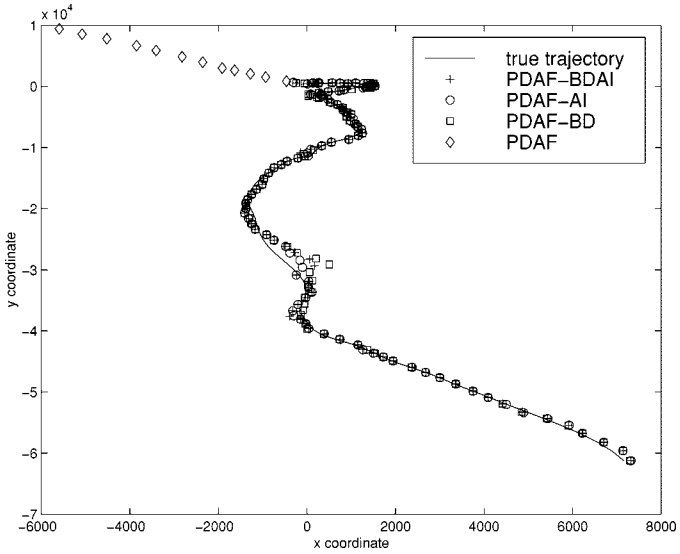

Fig. 5. Example of track with $\rho = 12$ dB, $\lambda = 10^{-6.5}\,\mathrm{m}^{-2}$, $\sigma_p = 0.07$, $\sigma_m = 100$. The PDAF loses track early; the PDAF-BDAI, PDAF-AI, and PDAF-BD hold track for the full 100 scans.

in (1) and (3). Measurements are of position only, and the model is in all respects linear/Gaussian. True detections are generated along with an associated amplitude; thresholding of this amplitude determines whether or not there is a miss. We then have some notes and parameter values:

- We have chosen $\Delta t = 30$ seconds, which is a fast but not-unreasonable scan rate for active sonar.
- We choose a track length $T = 100$ scans of data.
- Tracks are initialized by two-point differencing. This common track-initiation technique means that tracks are begun with two points for which there is no association uncertainty: Position is initialized for the second and velocity by the difference, with the associated initial uncertainties easy to derive and given (for example) in [2] and [4].

Targets begin their trajectories at scan $k = 0$ with position coordinates (0, 0) and velocity coordinates (5, 5) m/s, corresponding to 13.8 knots. A typical—but somewhat self-serving—tracking situation is given in Fig. 5.

Most studies of tracking performance are parameterized by P_d and by the clutter return density λ. In this case, we cannot use the former since for the new approach, P_d is not constant; hence, we use the SNR ρ instead. Each of the schemes takes as a parameter the detection threshold, given simply by τ for the PDAF and PDAF-AI, but in a more implicit fashion by η in the PDAF-BD and the PDAF-BDAI. We have no particular insight at present as to how η should be chosen,[3] hence, we adopt the simple and presumably fair expedient that the aggregate probabilities of detection for all four schemes be the same. This means that we have

$$\eta=-(1+\rho)\log[P_d((\rho+1)/\rho)^{n_z/2}] \tag{28}$$

[3]This would be the subject of further research. There is corresponding insight for the PDAF, but results appear to date to be more theoretical than applied, and there are no results for the PDAF-AI.

from (20), in which $P_d = e^{-\tau/(1+\rho)}$ for the PDAF and PDAF-AI, as given by the Swerling I assumption.[4] The explicit appearance of τ here amplifies the fact that independent specification of λ may be incompatible with σ_m, which is the standard deviation of the measurement error for each dimension. There is, in fact, through σ_m, an implied resolution cell grid at which threshold exceedances are interrogated, and, for example, $\lambda = 10^{-5}$ m^{-2} (a false alarm every 10^5 m^2, on average) and $\sigma_m = 100$ m (a 1.44×10^6 m^{2}resolution cell) makes very little sense indeed. Thus, we have adopted the convention that

$$\tau = -\log[12\sigma_m^2\lambda] \tag{29}$$

with the intuition that resolution cells be square and of side $\sqrt{12}\sigma_m$ with the "12" arising from the implied uniform distribution of a target within a resolution cell registering a hit.

For the theoretical development of the PDAF-BD and PDAF-BDAI, it was convenient to posit an underlying false-alarm Poisson point process having spatial density $\tilde{\lambda}$ and associated amplitudes that have a unit-exponential distribution; see Assumption 3. Now, since $\tilde{\lambda} = \lambda e^{\tau}$ (λ is the clutter spatial density for the PDAF), the implication is that $\tilde{\lambda}$ can be quite large. There is no mathematical problem with this, but the implementation can be very slow indeed. Thus, since the lowest detection threshold value for the PDAF-BD is η ($<\tau$), we adopt the equivalent expedient that the underlying Poisson point process has spatial density $\hat{\lambda} = \lambda e^{\tau-\eta}$. For each clutter point so generated, we also form an amplitude variate with distribution

$$f(a) = \begin{cases} e^{\eta-a}, & a > \eta \\ 0, & a \le \eta \end{cases} \tag{30}$$

which is thresholded using τ for the PDAF and PDAF-AI and using (19) for the PDAF-BD and the PDAF-BDAI.

First, we explore the relationship between the in-track percentage and SNR. A simulation is judged "in-track" if at the end of 100 scans, the true and estimated positions are less than $\sqrt{2}(10\sigma_m)$ apart.[5] It is clearly shown in Figs. 6 and 7 that the PDAF-BD offers considerable improvement with respect to the original PDAF, and its performance is close to those of PDAF-AI and PDAF-BDAI. Further, the PDAF-BDAI outperforms the PDAF-AI, at least at the operating point chosen. The ordering is particularly apparent in Fig. 7, in which the situation is of a more maneuvering (σ_p) but better observed (σ_m) target than in Fig. 6.

From Figs. 8 and 9, we find that the performance of the PDAF degrades significantly when λ or σ_m increases, with less of a problem for the other algorithms. Presumably, this is due to the relatively high SNR (=12 dB), meaning that all three other algorithms can maintain track via their respective uses of amplitude information.

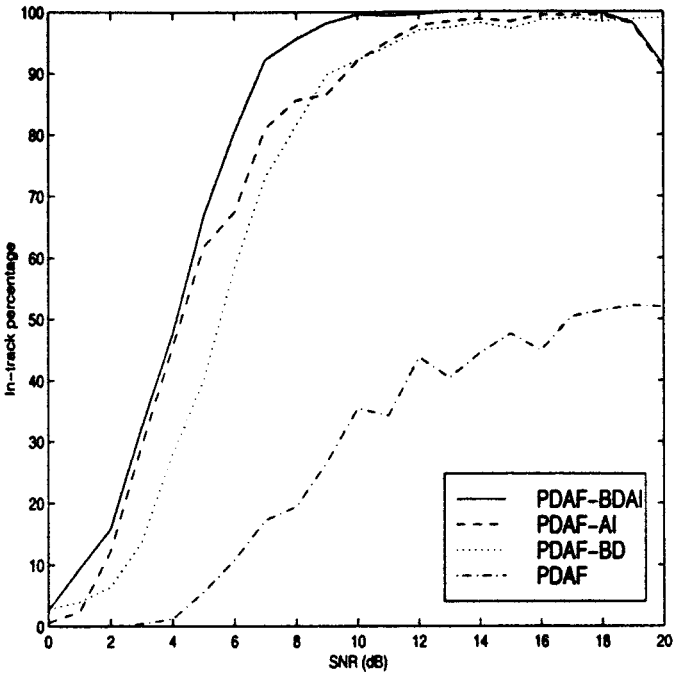

Fig. 6. In-track percentage (out of 500 Monte Carlo runs) as a function of SNR, with $\lambda = 10^{-6}$ m^{-2} (for the PDAF), $\sigma_p = 0.01$, $\sigma_m = 100$. The behavior of the two amplitude-dependent algorithms for high SNR appears to be a Monte Carlo artifact.

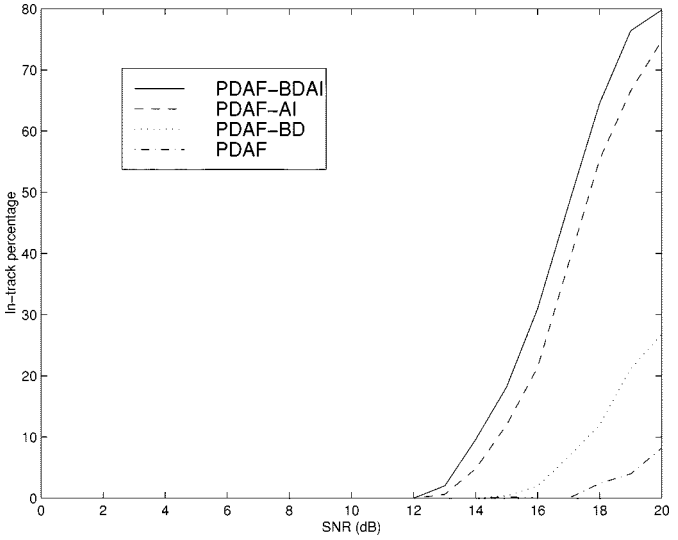

Fig. 7. In-track percentage (out of 500 Monte Carlo runs) as a function of SNR, with $\lambda = 10^{-6}$ m^{-2}, $\sigma_p = 0.1$, $\sigma_m = 10$.

Now, we compare the performance of four algorithms in different situations. The results are given in Tables I and II, respectively (of the *in-track* percentage and tracking RMSE, the former is generally accorded higher importance). From Table I, we observe that the PDAF-BD has tracking performance between that of the PDAF and PDAF-AI, which is generally and gratifyingly closer to the latter than to the former. The PDAF-BDAI has the best performance.

[4]Although there is guidance—for example, see the concept of the tracker operating characteristic (TOC) in [3]—for the choice of detection threshold τ at a particular SNR in the PDAF, there are no generally accepted prescriptive rules. Consequently, for the PDAF, and indeed in the PDAF-AI, τ may be thought of as a design parameter. In the case of the PDAF-BD (and PDAF-BDAI), the quantity η may be assigned the same role. The selection in (28) is just one possible choice: one that we hope is fair.

[5]There are many possible criteria for judging a track "lost," but since a track loss amounts to a burgeoning instability in the algorithm, for the most part, their decisions are the same. In this case, a track is considered lost if the estimation error is greater than ten times the measurement standard deviation per dimension.

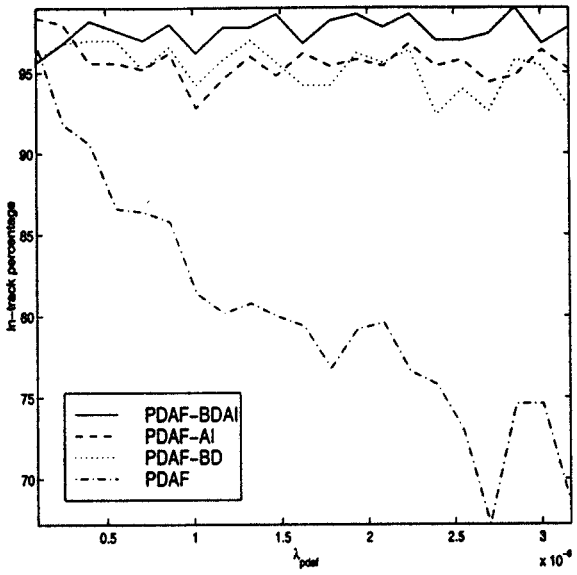

Fig. 8. In-track percentage (out of 500 Monte Carlo runs) as a function of λ, with $SNR = 12$ dB, $\sigma_p = 0.01$, and $\sigma_m = 10$.

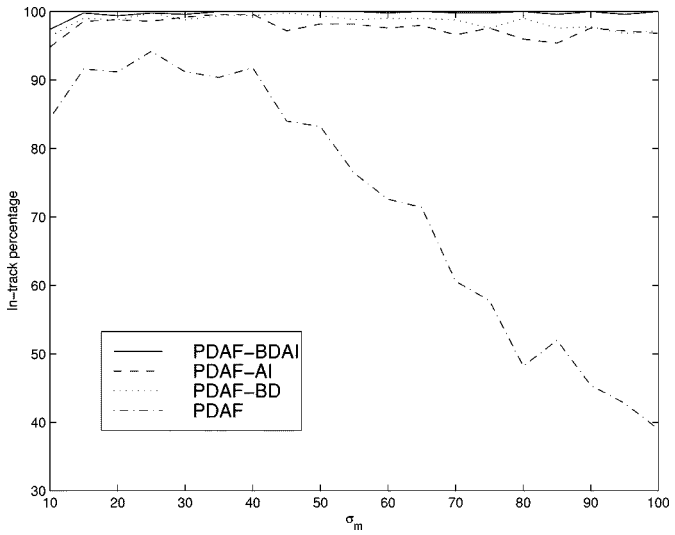

Fig. 9. In-track percentage (out of 500 Monte Carlo runs) as a function of σ_m, with $\lambda = 10^{-6}$ m^{-2}, $\sigma_p = 0.01$, $SNR = 12$ dB.

The tracking error is the RMSE *over the whole track* for those simulations in which *all* the algorithms keep in-track at their conclusion; naturally, inclusion of those tracks that become lost skews the results well beyond interpretability. From Table II, we can see that the PDAF-BD has lower RMSE than PDAF in most situations. The PDAF-AI and PDAF-BDAI are comparable in this regard and have the best RMSE.

V. Summary

The usual target tracking model is of separation between detection and tracking subsystems. In the absence of information from the latter, the former has little choice but to do the best job it can. It provides Neyman–Pearson optimal performance: the most powerful test subject to a constraint on false-alarm rate. If there *is* some information flow from tracker to detector, particularly in terms of predicted measurement location and association confidence (innovations covariance), then a *Bayesian* detector is appropriate. The difference is not in the statistic tested, but rather in the threshold. In fact, assuming that the prior probability is Gaussian (which fits with the PDAF assumptions, hence, our use of this model), the threshold is proportional to the normalized innovation and, hence, is lowest near where a detection is expected (at the predicted measurement).

In this paper, the threshold shape has been derived, and appropriate modification to the PDAF—we call it the "PDAF-BD," for *B*ayes *D*etector—is made. Simulation has revealed that the performance of PDAF-BD is considerably better than that of the PDAF and is often only slightly degraded relative to the PDAF-AI, which is that version of the PDAF appropriate to transmission from detector to tracker of full amplitude information for *all* returns. One perhaps remarkable feature of the PDAF-BD is that the posterior association probabilities (that

TABLE I
In-Track Percentage for Various Situations. The Last Four Columns Refer to the PDAF, the PDAF-BD, the PDAF-AI, and PDAF-BDAI, Respectively

σ_m	σ_p	ρ (dB)	λ	PD	BD	AI	BDAI
10	0.1	6	$10^{-5.5}$	0	0	0	0
10	0.1	6	10^{-6}	0	0	0	0
10	0.1	6	$10^{-6.5}$	0	0	0	0
10	0.1	6	10^{-7}	0	0	0	0
10	0.1	12	$10^{-5.5}$	0	0	0	0
10	0.1	12	10^{-6}	0	0	0	0
10	0.1	12	$10^{-6.5}$	0	0	0	0
10	0.1	12	10^{-7}	0	0	0	0
10	0.01	6	$10^{-5.5}$	0	0	1	2
10	0.01	6	10^{-6}	0	0	0	1
10	0.01	6	$10^{-6.5}$	0	0	0	0
10	0.01	6	10^{-7}	0	0	0	0
10	0.01	12	$10^{-5.5}$	66	89	93	98
10	0.01	12	10^{-6}	80	93	96	98
10	0.01	12	$10^{-6.5}$	95	98	98	100
10	0.01	12	10^{-7}	97	94	98	94
100	0.1	6	$10^{-5.5}$	0	0	5	15
100	0.1	6	10^{-6}	0	0	0	10
100	0.1	6	$10^{-6.5}$	0	0	2	8
100	0.1	6	10^{-7}	0	0	1	3
100	0.1	12	$10^{-5.5}$	0	0	89	99
100	0.1	12	10^{-6}	0	0	90	99
100	0.1	12	$10^{-6.5}$	1	59	88	100
100	0.1	12	10^{-7}	19	83	81	97
100	0.01	6	$10^{-5.5}$	1	38	69	78
100	0.01	6	10^{-6}	11	67	70	86
100	0.01	6	$10^{-6.5}$	22	70	70	80
100	0.01	6	10^{-7}	35	82	71	88
100	0.01	12	$10^{-5.5}$	6	61	97	100
100	0.01	12	10^{-6}	42	96	97	100
100	0.01	12	$10^{-6.5}$	61	98	97	100
100	0.01	12	10^{-7}	67	100	92	100

TABLE II
RMSE FOR VARIOUS SITUATIONS. RMSE IS CALCULATED OVER THE WHOLE TRACK BUT ONLY FOR THOSE SITUATIONS FOR WHICH THE TRACK IS MAINTAINED UNTIL THE FINAL SCAN. A BLANK MEANS THAT NO RELIABLE TRACKS WERE REPORTED. THE LAST FOUR COLUMNS REFER TO THE PDAF, THE PDAF-BD, THE PDAF-AI, AND THE PDAF-BDAI, RESPECTIVELY

σ_m	σ_p	ρ (dB)	λ	PD	BD	AI	BDAI
10	0.1	6	$10^{-5.5}$	—	—	—	—
10	0.1	6	10^{-6}	—	—	—	—
10	0.1	6	$10^{-6.5}$	—	—	—	—
10	0.1	6	10^{-7}	—	—	—	—
10	0.1	12	$10^{-5.5}$	—	—	—	—
10	0.1	12	10^{-6}	—	—	—	—
10	0.1	12	$10^{-6.5}$	—	—	—	—
10	0.1	12	10^{-7}	—	—	—	—
10	0.01	6	$10^{-5.5}$	—	—	—	—
10	0.01	6	10^{-6}	—	—	—	—
10	0.01	6	$10^{-6.5}$	—	—	—	—
10	0.01	6	10^{-7}	—	—	—	—
10	0.01	12	$10^{-5.5}$	26	26	21	24
10	0.01	12	10^{-6}	28	27	24	26
10	0.01	12	$10^{-6.5}$	29	33	29	32
10	0.01	12	10^{-7}	32	42	31	42
100	0.1	6	$10^{-5.5}$	—	—	—	—
100	0.1	6	10^{-6}	—	—	—	—
100	0.1	6	$10^{-6.5}$	—	—	—	—
100	0.1	6	10^{-7}	—	—	—	—
100	0.1	12	$10^{-5.5}$	—	—	—	—
100	0.1	12	10^{-6}	—	—	—	—
100	0.1	12	$10^{-6.5}$	252	272	155	155
100	0.1	12	10^{-7}	274	242	191	206
100	0.01	6	$10^{-5.5}$	—	—	—	—
100	0.01	6	10^{-6}	382	230	137	177
100	0.01	6	$10^{-6.5}$	197	176	142	166
100	0.01	6	10^{-7}	287	310	228	263
100	0.01	12	$10^{-5.5}$	192	168	95	97
100	0.01	12	10^{-6}	141	129	94	96
100	0.01	12	$10^{-6.5}$	137	105	99	101
100	0.01	12	10^{-7}	125	108	104	106

a given threshold exceedance is target-generated rather than a false-alarm) are **uniform** and independent of location. This is in distinction to the PDAF behavior in which closer (to the expected target location) returns are accorded a higher posterior association probability, and it presumably arises from the spatially varying detection threshold, whose implication is that a more-distant measurement must have had a higher amplitude in order that an exceedance occurred.

Through its use of a spatially varying threshold, the PDAF-BD can be thought of as exploiting amplitude information, as does the PDAF-AI, with this confidence information implicit and never requiring transmission. The PDAF-AI outperforms the new PDAF-BD, but it appeared appropriate to develop the PDAF-BDAI, which is the extension of the PDAF-BD in which amplitude information (that is, the amount by which the detection threshold was exceeded) is explicitly transmitted from detector to tracker (an extension parallel to that from the PDAF to PDAF-AI). Accordingly, the PDAF-BDAI outperforms the PDAF-AI in almost all situations studied.

We offer these further notes.

- This study has focused on extension of the PDAF to incorporate Bayesian detection. The Gaussian distributions assumed by the PDAF mesh well with Bayes detection, as has been seen. However, there exist other tracking algorithms, and there is no reason why the Bayes detection idea could not be applied to them.
- Since, in effect, only detections close to the predicted measurement are allowed, the PDAF-BD is less of a computational load than the others. This might be considered an intelligent "gating" mechanism, and certainly relatively few threshold exceedances are registered. Thus, the Bayesian detector idea may fit nicely with data fusion in which data transmission requirements (from detectors to tracker) may be strict.
- As far as we are aware, there is at present no detection system that allows nonconstant thresholding, at least not on the scale proposed here. Thus, this work is several generations ahead of its platform.
- While there is guidance in this regard (see [3]), the choice of the detection threshold τ for the PDAF is something of an art. In the PDAF-BD, there is no fixed detection threshold, but, perhaps unfortunately, there is a tunable parameter η whose role is similar. Exploration of the effect of η on tracking performance, and indeed of its robustness with respect to η, has not been explored in this paper and should be considered a suggestion for further research.

Future work on the PDAF-BD and its underlying ideas should include the incorporation of target maneuver, its use for multiple targets, and its development for multiple sensors and data fusion. An additional concern is that the feedback between tracking and detection may render the decision process of track acceptance/rejection more difficult than for the PDAF and PDAF-BD since in those cases, the quantity and locations of false alarms are arguably independent of the track estimation process.

APPENDIX A
PDAF-BDAI

Modification on the PDAF to any of the other algorithms (PDAF-AI, PDAF-BD, and PDAF-BDAI) discussed in this paper requires only a reformulation of the posterior association probabilities—these are the βs of (8). No other filtering steps need to be modified. In this appendix, we derive the βs for the PDAF-BDAI. Results for the PDAF-BD are found by stripping away the amplitude information and are presented in the succeeding Appendix.

Assuming that at the present scan we have $n_k \in \{0, 1, \cdots\}$ threshold exceedances, then according to [2], we define the events $\theta \in \{0, 1, \cdots, n_k\}$ such that $\theta = l$ means that the lth measurement is target generated and that the others are false alarms, and $\theta = 0$ means that all measurements are false. We denote by $\beta(\theta)$ the probability, conditioned on all measurements in the current scan (and naturally, implicitly, on n_k), that θ is true.

We first require the conditional observation probability densities

$$f(\mathbf{z}_k(l)|\theta=l,\, n_k=m)=\frac{1}{\sqrt{|2\pi\mathbf{S}_k|}}\,e^{-(1/2)\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)} \tag{31}$$

and for $p\neq l$, we have (32), shown at the bottom of the page, where (22) is used as $\Pr(\text{threshold exceedance at } \mathbf{z}_k(l))$. This is particularly interesting; under the new detection model, any randomly chosen false alarm has a spatial Gaussian distribution. For the standard PDAF model, this probability is, naturally, uniform.

Let us denote $\vec{\nu}_k=\{\nu_k(1),\nu_k(2),\cdots,\nu_k(n_k)\}$ and $\vec{a}_k=\{a_k(1),a_k(2),\cdots,a_k(n_k)\}$ as the full respective innovation and amplitude information available at the kth scan. Let us also define ν_k^* and a_k^* to be, respectively, the innovation and return amplitude for the *true* target; note that it is not known to the tracker which return is true, and in fact, the true return may be missed (below threshold). For the case $\theta=0$, we have the joint probability

$$\begin{aligned}p(\theta=0,\,\vec{\nu}_k,\,\vec{a}_k,\,n_k=m)&\\ =\iint f(\vec{\nu}_k,\,\vec{a}_k|\theta=0,\,n_k=m,\,\nu_k^*,\,a_k^*)&\\ \times\Pr(\theta=0|n_k=m,\,\nu_k^*,\,a_k^*)\Pr(n_k=m|\nu_k^*,\,a_k^*)&\\ \times f(\nu_k^*,\,a_k^*)\,d\nu_k^*\,da_k^*&\end{aligned} \tag{33}$$

which is a joint probability (density) of mixed type, involving both real and discrete random variables. One ingredient to (33) is

$$\begin{aligned}f(\vec{\nu}_k,\,\vec{a}_k|\theta=0,\,n_k=m,\,\nu_k^*,\,a_k^*)&\\ =f(\vec{\nu}_k,\,\vec{a}_k|\theta=0,\,n_k=m)&\\ =\left(\frac{1}{\sqrt{\left|2\pi\,\frac{\rho}{\rho+1}\,\mathbf{S}_k\right|}}\right)^m&\\ \times\exp\left(-((\rho+1)/2\rho)\sum_{l=1}^{m}\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)\right)&\\ \times\exp\left(-\sum_{l=1}^{m}(a_k(l)-((\rho+1)/2\rho)\nu_l^T\mathbf{S}_k^{-1}\nu_l-\eta)\right)&\\ =\left(\frac{1}{\sqrt{\left|2\pi\,\frac{\rho}{\rho+1}\,\mathbf{S}_k\right|}}\right)^m\exp\left(-\sum_{l=1}^{m}(a_k(l)-\eta)\right)&\end{aligned} \tag{34}$$

from (32) and the fact that if a is unit exponential, then

$$\begin{aligned}f(a|a>\tau)&=\frac{f(a)\cdot u(a-\tau)}{\Pr(a>\tau)}\\ &=\frac{e^{-a}\cdot u(a-\tau)}{e^{-\tau}}\\ &=e^{-(a-\tau)}u(a-\tau).\end{aligned}$$

Two other ingredients are the straightforward binary expressions

$$\begin{aligned}\Pr(\theta=0|n_k=m,\,\nu_k^*,\,a_k^*)&\\ =\mathbf{I}\left(a_k^*<\frac{\rho+1}{2\rho}(\nu_k^*)^T\mathbf{S}_k^{-1}\nu_k^*+\eta\right)&\end{aligned} \tag{35}$$

and

$$\begin{aligned}\Pr(n_k=m|\nu_k^*,\,a_k^*)&\\ =\mu(m)\mathbf{I}\left(a_k^*<\frac{\rho+1}{2\rho}(\nu_k^*)^T\mathbf{S}_k^{-1}\nu_k^*+\eta\right)&\\ +\mu(m-1)\mathbf{I}\left(a_k^*>\frac{\rho+1}{2\rho}(\nu_k^*)^T\mathbf{S}_k^{-1}\nu_k^*+\eta\right)&\end{aligned} \tag{36}$$

in which $\mathbf{I}(\cdot)$ is the indicator function, unity if its argument is true, and zero otherwise. The last ingredient is the product for the explicit pdf of the innovation and return amplitude from the true target:

$$f(\nu_k^*,\,a_k^*)=\frac{1}{\sqrt{|2\pi\mathbf{S}_k|}}\,e^{-(1/2)(\nu_k^*)^T\mathbf{S}_k^{-1}\nu_k^*}\,\frac{1}{1+\rho}\,e^{-(a_k^*/(1+\rho))}. \tag{37}$$

$$\begin{aligned}f(\mathbf{z}_k(l)|\theta=p,\,n_k=m)&=\frac{\Pr(\text{threshold exceedance at }\mathbf{z}_k(l)|\text{event at }\mathbf{z}_k(l))f(\mathbf{z}_k(l))}{\Pr(\text{threshold exceedance at }\mathbf{z}_k(l))}\\ &=\frac{e^{-((\rho+1)/2\rho)\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)-\eta}\cdot\frac{1}{V}}{\frac{1}{V}\,e^{-\eta}\sqrt{\left|2\pi\,\frac{\rho}{\rho+1}\,\mathbf{S}_k\right|}}\\ &=\frac{1}{\sqrt{\left|2\pi\,\frac{\rho}{\rho+1}\,\mathbf{S}_k\right|}}\,e^{-(1/2)\nu_k(l)^T[(\rho/(\rho+1))\mathbf{S}_k]^{-1}\nu_k(l)}\end{aligned} \tag{32}$$

Substituting (34)–(37) into (33), we get

$$
\begin{aligned}
&p(\theta = 0, \vec{\nu}_k, \vec{a}_k, n_k = m) \\
&\quad = \left(\frac{1}{\sqrt{\left| 2\pi \frac{\rho}{\rho+1} \mathbf{S}_k \right|}} \right)^m \exp\left(-\sum_{l=1}^{m} (a_k(l) - \eta) \right) \\
&\qquad \times \iint \mu(m) \mathbf{I}\left(a_k^* < \frac{\rho+1}{2\rho} (\nu_k^*)^T \mathbf{S}_k^{-1} \nu_k^* + \eta \right) \\
&\qquad \times \frac{1}{\sqrt{|2\pi \mathbf{S}|}} e^{-(1/2)(\nu_k^*)^T \mathbf{S}_k^{-1} \nu_k^*} \frac{1}{1+\rho} e^{-(a_k^*/(1+\rho))} \, da_k^* \, d\nu_k^* \\
&\quad = \left(\frac{1}{\sqrt{|2\pi \frac{\rho}{\rho+1} \mathbf{S}|}} \right)^m \exp\left(-\sum_{l=1}^{m} (a_k(l) - \eta) \right) \mu(m) \\
&\qquad \times \int \frac{1}{\sqrt{|2\pi \mathbf{S}|}} e^{-(1/2)(\nu_k^*)^T \mathbf{S}^{-1} \nu_k^*} \\
&\qquad \times \left(1 - e^{-(1/(1+\rho))[((\rho+1)/2\rho)(\nu_k^*)^T \mathbf{S}^{-1} \nu_k^* + \eta]} \right) d\nu_k^* \\
&\quad = \left(\frac{1}{\sqrt{\left| 2\pi \frac{\rho}{\rho+1} \mathbf{S} \right|}} \right)^m \exp\left(-\sum_{l=1}^{m} (a_k(l) - \eta) \right) \mu(m) \\
&\qquad \times \left[1 - e^{-\eta/(1+\rho)} \int \frac{1}{\sqrt{|2\pi \mathbf{S}|}} e^{-((1+\rho)/2\rho)(\nu_k^*)^T \mathbf{S}^{-1} \nu_k^*} \, d\nu_k^* \right] \\
&\quad = \left(\frac{1}{\sqrt{\left| 2\pi \frac{\rho}{\rho+1} \mathbf{S}_k \right|}} \right)^m \exp\left(-\sum_{l=1}^{m} (a_k(l) - \eta) \right) \mu(m) \\
&\qquad \times \left[1 - \left(\frac{\rho}{\rho+1} \right)^{n_z/2} e^{-\eta/(1+\rho)} \right]. \qquad (38)
\end{aligned}
$$

Consider now the case that $\theta \neq 0$. Let us define $\tilde{\nu}_k$ and $\tilde{a}_k$ such that $(\tilde{\nu}_k, \nu_k^*) = \vec{\nu}_k$, and $(\tilde{a}_k, a_k^*) = \vec{a}_k$; that is, $\tilde{\nu}_k$ and $\tilde{a}_k$ are from the false alarms. We again seek

$$
\begin{aligned}
&p(\theta, \vec{\nu}_k, \vec{a}_k, n_k = m) \\
&\quad = p(\theta, \tilde{\nu}_k, \nu_k(\theta), \tilde{a}_k, a_k(\theta), n_k = m) \\
&\quad = f(\tilde{\nu}_k, \tilde{a}_k | \theta, n_k = m, \nu_k(\theta), a_k(\theta)) \\
&\qquad \times \Pr(\theta | \theta \neq 0, n_k = m, \nu_k(\theta), a_k(\theta)) \\
&\qquad \times \Pr(\theta \neq 0 | \nu_k(\theta), a_k(\theta), n_k = m) \\
&\qquad \times \Pr(n_k = m | \nu_k(\theta), a_k(\theta)) f(\nu_k(\theta), a_k(\theta)). \qquad (39)
\end{aligned}
$$

As before, we have the ingredients

$$
\begin{aligned}
&f(\tilde{\nu}_k, \tilde{a}_k | \theta, n_k = m, \nu_k(\theta), a_k(\theta)) \\
&\quad = \left(\frac{1}{\sqrt{\left| 2\pi \frac{\rho}{\rho+1} \mathbf{S}_k \right|}} \right)^{m-1} \\
&\qquad \times \exp\left(-((\rho+1)/2\rho) \sum_{l=1, \, l \neq \theta}^{m} \nu_k(l)^T \mathbf{S}_k^{-1} \nu_k(l) \right) \\
&\qquad \times \exp\left(-\sum_{l=1, \, l \neq \theta}^{m} (a_k(l) - ((\rho+1)/2\rho) \right. \\
&\qquad\qquad \left. \cdot \nu_k(l)^T \mathbf{S}_k^{-1} \nu_k(l) - \eta) \right) \\
&\quad = \left(\frac{1}{\sqrt{\left| 2\pi \frac{\rho}{\rho+1} \mathbf{S}_k \right|}} \right)^{m-1} e^{-\sum_{l=1, \, l \neq \theta}^{m} [a_k(l) - \eta]} \qquad (40)
\end{aligned}
$$

and

$$
\Pr(\theta | \theta \neq 0, n_k = m, \nu(\theta), a_k(\theta)) = \frac{1}{m} \qquad (41)
$$

and

$$
\begin{aligned}
&\Pr(\theta \neq 0 | \nu_k(\theta), a_k(\theta), n_k = m) \\
&\qquad = \mathbf{I}\left(a_k(\theta) > \frac{\rho+1}{2\rho} \nu_k(\theta)^T \mathbf{S}_k^{-1} \nu_k(\theta) + \eta \right). \qquad (42)
\end{aligned}
$$

Repeated from (36) and (37), we get $\Pr(n_k = m | \nu_k(\theta), a_k(\theta))$ and $f(\nu_k(\theta), a_k(\theta))$. There is no need to integrate, so we have

$$
\begin{aligned}
&p(\theta, \vec{\nu}_k, \vec{a}_k, n_k = m) \\
&\quad = \left(\frac{1}{\sqrt{\left| 2\pi \frac{\rho}{\rho+1} \mathbf{S}_k \right|}} \right)^{m-1} \\
&\qquad \times \exp\left(-\sum_{l=1, l \neq \theta}^{m} [a_k(l) - \eta] \right) \frac{1}{m} \mu(m-1) \\
&\qquad \times \frac{1}{\sqrt{|2\pi \mathbf{S}_k|}} e^{-(1/2)\nu_k(\theta)^T \mathbf{S}_k^{-1} \nu_k(\theta)} \frac{1}{1+\rho} e^{-(a_k(\theta)/(1+\rho))}. \qquad (43)
\end{aligned}
$$

Putting (38) and (43) together, and using formula (23) for the false-alarm pmf $\mu(\cdot)$, we have the expression shown at the bottom of the next page.

Thus, we get (26); that is

$$\beta(\theta)=\begin{cases} c(1+\rho)\lambda\sqrt{\left|2\pi\,\dfrac{\rho}{\rho+1}\mathbf{S}_k\right|} \\ \quad\cdot\left[\left(\dfrac{\rho+1}{\rho}\right)^{n_z/2}-e^{-\eta/(1+\rho)}\right] & \theta=0 \\ c\cdot e^{(\rho/(\rho+1))a_k(\theta)-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)} & 1\le\theta\le n_k \end{cases}$$

in which c is a constant such that $\sum_{\theta=0}^{n_k}\beta(\theta)=1$.

Appendix B
PDAF-BD

In the previous Appendix, we derived the posterior association probabilities (the βs) for the PDAF-BDAI. It is a simple matter, in this Appendix, to remove the amplitude information (AI). We have

$$\begin{aligned} &p(\theta=0,\vec{\nu}_k,m)\\ &\quad=\iint f(\vec{\nu}_k|\theta=0,n_k=m,\nu_k^*,a_k^*)\\ &\qquad\times\Pr(\theta=0|n_k=m,\nu_k^*,a_k^*)\\ &\qquad\times\Pr(n_k=m|\nu_k^*,a_k^*)f(\nu_k^*,a_k^*)\,d\nu_k^*\,da_k^*. \end{aligned}\tag{44}$$

There is little difference here from the previous Appendix, except the lack of the amplitudes "$\vec{a}_k$." Hence

$$\begin{aligned} &p(\theta=0,\vec{\nu}_k,n_k=m)\\ &\quad=\left(\frac{1}{\sqrt{\left|2\pi\,\dfrac{\rho}{\rho+1}\,\mathbf{S}_k\right|}}\right)^m\\ &\qquad\times\mu(m)\left[1-\left(\frac{\rho}{\rho+1}\right)^{n_z/2}e^{-\eta/(1+\rho)}\right]\\ &\qquad\times\exp\left(-((\rho+1)/2\rho)\sum_{l=1}^{m}\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)\right). \end{aligned}\tag{45}$$

Now, for $\theta\neq 0$, we get

$$\begin{aligned} &p(\theta,\vec{\nu}_k,n_k=m)\\ &\quad=\int f(\theta,\tilde{\nu}_k,\nu_k(\theta),a_k(\theta),n_k=m)\,da_k(\theta)\\ &\quad=\int f(\tilde{\nu}_k|\theta,n_k=m,\nu_k(\theta),a_k(\theta))\\ &\qquad\times\Pr(\theta|\theta\neq 0,n_k=m,\nu_k(\theta),a_k(\theta))\\ &\qquad\times\Pr(\theta\neq 0|\nu_k(\theta),a_k(\theta),n_k=m)\\ &\qquad\times\Pr(n_k=m|\nu_k(\theta),a_k(\theta))f(\nu_k(\theta),a_k(\theta))\,da_k(\theta). \end{aligned}\tag{46}$$

The above terms are similar to those in the previous Appendix. Thus, we have

$$\begin{aligned} &p(\theta,\vec{\nu},n_k=m)\\ &\quad=\int\left(\frac{1}{\sqrt{\left|2\pi\,\dfrac{\rho}{\rho+1}\,\mathbf{S}_k\right|}}\right)^{m-1}\\ &\qquad\times\exp\left(-((\rho+1)/2\rho)\sum_{l=1,\,l\neq\theta}^{m}\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)\right)\\ &\qquad\times\frac{1}{m}\,\mu(m-1)\mathbf{I}\left(a_k(\theta)>\frac{\rho+1}{2\rho}\,\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)+\eta\right)\\ &\qquad\times\frac{1}{\sqrt{|2\pi\mathbf{S}_k|}}\,e^{-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)}\,\frac{1}{1+\rho}\\ &\qquad\times e^{-(a_k(\theta)/(1+\rho))}\,da_k(\theta)\\ &\quad=\frac{1}{m}\,\mu(m-1)\left(\frac{1}{\sqrt{\left|2\pi\,\dfrac{\rho}{\rho+1}\,\mathbf{S}_k\right|}}\right)^{m-1}\\ &\qquad\times\exp\left(-((\rho+1)/2\rho)\sum_{l=1,\,l\neq\theta}^{m}\nu_k(l)^T\mathbf{S}_k^{-1}\nu_k(l)\right)\\ &\qquad\times\frac{1}{\sqrt{|2\pi\mathbf{S}_k|}}\,e^{-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)}\\ &\qquad\times e^{-(1/(1+\rho))[((\rho+1)/2\rho)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)+\eta]}. \end{aligned}\tag{47}$$

$$\begin{aligned} \frac{p(\theta=0,\vec{\nu}_k,\vec{a}_k,n_k=m)}{p(\theta,\vec{\nu}_k,\vec{a}_k,n_k=m)} &= \frac{e^{-[a_k(\theta)-\eta]}m\mu(m)\left[1-\left(\dfrac{\rho}{\rho+1}\right)^{n_z/2}e^{-\eta/(1+\rho)}\right]}{\sqrt{\left|2\pi\,\dfrac{\rho}{\rho+1}\,\mathbf{S}_k\right|}\,\mu(m-1)\dfrac{1}{\sqrt{|2\pi\mathbf{S}_k|}}\,e^{-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta)}\,\dfrac{1}{1+\rho}\,e^{-(a_k(\theta)/(1+\rho))}}\\ &= \frac{(1+\rho)\lambda\sqrt{|2\pi\mathbf{S}_k|}\left[1-\left(\dfrac{\rho}{\rho+1}\right)^{n_z/2}e^{-\eta/(1+\rho)}\right]}{e^{(\rho/(\rho+1))a_k(\theta)-(1/2)\nu_k(\theta)^T\mathbf{S}_k^{-1}\nu_k(\theta).}} \end{aligned}$$

Putting (45), (47), and (26) together, we have

$$\beta(0) = c\lambda e^{-\eta}\sqrt{\left|2\pi\,\frac{\rho}{\rho+1}\,\mathbf{S}_k\right|}\left[\left(\frac{\rho+1}{\rho}\right)^{n_z/2} e^{\eta/(1+\rho)} - 1\right]$$

and for $\theta \neq 0$

$$\beta(\theta) = c$$

as in (25), in which c normalizes the sum of the probabilities to unity.

REFERENCES

[1] D. Avitzour, "A maximum likelihood approach to data association," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, pp. 560–565, Apr. 1992.

[2] Y. Bar-Shalom and X. R. Li, *Estimation and Tracking: Principles, Techniques and Software*. Boston, MA: Artech House, 1993.

[3] ——, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS, 1995.

[4] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Boston, MA: Artech House, 1999.

[5] T. Fortmann, Y. Bar-Shalom, M. Scheffé, and S. Gelfand, "Detection thresholds for tracking in clutter—A connection between estimation and signal processing," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 221–229, 1985.

[6] T. Kirubarajaran and Y. Bar-Shalom, "Low observable target motion analysis using amplitude information," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, pp. 1367–1384, Oct. 1996.

[7] H. Gauvrit, J.-P. LeCadre, and C. Jauffret, "A formulation of multitarget tracking as an incomplete data problem," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, pp. 1242–1257, Oct. 1997.

[8] D. Lerro and Y. Bar-Shalom, "Interacting multiple model tracking with target amplitude feature," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, pp. 494–509, Apr. 1993.

[9] K. Molnar and J. Modestino, "Application of the EM algorithm for multi-target/multi-sensor tracking problem," *IEEE Trans. Signal Processing*, vol. 47, pp. 115–129, Jan. 1998.

[10] D. Musicki, "Tracking in clutter using clutter map information," in *Proc. Int. Symp. Inform. Theory Applicat. (ISITA) Conf.*, pp. 1237–1241.

[11] D. Musicki and R. Evans, "Data association using clutter map information," in *Proc. Workshop Signal Process. Applicat. (WOSPA)*, Brisbane, Australia, Dec. 1997, pp. 227–230.

[12] K. R. Pattipati, S. Deb, Y. Bar Shalom, and R. B. Washburn, "A new relaxation algorithm and passive sensor data association," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 198–213, Feb. 1992.

[13] H. V. Poor, *An Introduction to Signal Detection and Estimation*: Springer-Verlag, 1988.

[14] A. Poore, "Multidimensional assignment formulation of data association problems arising from multitarget and multisensor tracking," *Comput. Optim. Applicat.*, vol. 3, pp. 27–57, 1994.

[15] R. L. Streit and T. E. Luginbuhl, "Probabilistic multi-hypothesis tracking,", NUWC-NPT Tech. Rep. 10 428, Feb. 1995.

[16] P. Willett and R. Niu, "A PDAF with a Bayesian detector," in *Proc. FUSION Conf.*, July 1999.

Peter Willett (SM'97) received the B.A.Sc. degree from the University of Toronto, Toronto, ON, Canada, in 1982 and the Ph.D. degree from Princeton University, Princeton, NJ, in 1986.

He is a Professor with the University of Connecticut, Storrs, where he has been since 1986. His interests are generally in the areas of detection theory, target tracking, and signal processing.

Dr. Willett is an Associate Editor for IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS and for IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS.

Ruixin Niu was born on August 29, 1972. He received the B.S. degree from Xi'an Jiaotong University, Xi'an, China, in 1994 and the M.S. degree from Institute of Electronics, Academica Sinica, Beijing, China, in 1997, both in electrical engineering.

He is a Graduate Student/Research Assistant at University of Connecticut, Storrs, where he is pursuing the Ph.D. degree. His research interests lie in signal processing, detection, and estimation theory.

Yaakov Bar-Shalom (S'63–M'66–SM'80–F'84) was born on May 11, 1941. He received the B.S. and M.S. degrees from the Technion—Israel Institute of Technology, Haifa, in 1963 and 1967, respectively, and the Ph.D. degree from Princeton University, Princeton, NJ, in 1970, all in electrical engineering.

From 1970 to 1976, he was with Systems Control, Inc., Palo Alto, CA. Currently, he is the School of Engineering Distinguished Professor with the Department of Electrical and Systems Engineering and Director of the Estimation and Signal Processing (ESP) Laboratory at the University of Connecticut, Storrs. His research interests are in estimation theory and stochastic adaptive control. He has published over 250 papers and book chapters in these areas. He coauthored the monograph *Tracking and Data Association* (New York: Academic, 1988), the graduate text *Estimation and Tracking: Principles, Techniques and Software* (Boston, MA: Artech House, 1993), the text *Multitarget-Multisensor Tracking: Principles and Techniques* (Storrs, CT: YBS, 1995), and edited the books *Multitarget-Multisensor Tracking: Applications and Advances* (Boston, MA: Artech House, vol. I, 1990; vol. II, 1992; vol. III, 2000). He has been consulting to numerous companies and originated the series of multitarget-multisensor tracking short courses offered via UCLA Extension, at Government Laboratories, private companies, and overseas. He has also developed the commercially available interactive software packages MULTIDAT TM for automatic track formation and tracking of maneuvering or splitting targets in clutter, VARDAT TM for data association from multiple passive sensors, BEARDAT TM for target localization from bearing and frequency measurements in clutter, IMDAT TM for image segmentation and target centroid tracking, and FUSEDAT TM for fusion of possibly heterogeneous multisensor data for tracking.

Dr. Bar-Shalom served as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL during 1976 and 1977, and from 1978 to 1981, he was an Associate Editor of *Automatica*. He was Program Chairman of the 1982 American Control Conference, General Chairman of the 1985 ACC, and Co-Chairman of the 1989 IEEE International Conference on Control and Applications. From 1983 to 1987, he served as Chairman of the Conference Activities Board of the IEEE Control Systems Society, and from 1987 to 1989, he was a member of the Board of Governors of the IEEE CSS. He is a member of the Board of Directors of the International Society of Information Fusion (1999–2001) and Y2K President of ISIF. In 1987, he received the IEEE CSS Distinguished Member Award. Since 1995, he has been a Distinguished Lecturer of the IEEE AESS and has given several keynote addresses at major national and international conferences. He is co-recipient of the M. Barry Carlton Award for the best paper in the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS in 1995 and the 1998 University of Connecticut AAUP Excellence Award for Research.