Notes on the Probabilistic Data Association Filter

İlker Bayram
ilker.bayram@ieee.org

1 Introduction

The probabilistic data association filter (PDAF) is an extension of the Kalman filter (or possibly the extended Kalman filter) to address scenarios where

- there might be spurious detections, referred to as clutter,
- the detector might miss the actual target (returning only clutter, or nothing at all).

The input to the algorithm is a set of detections obtained by a detector. The output is a Gaussian approximation to the pdf of the state given all of the observations.

In this note, we derive PDAF, along with a track existence probability, following [1, 2]. We also discuss a version with simplified assumptions for handling occasional outliers, rather than multiple clutter detections.

2 Model

Let us consider the underlying model and introduce some notation.

2.1 State Update

We consider a state space model of the form

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

where x denotes the state, F, the state update matrix, and v, noise. We will assume that $v(k) \sim \mathcal{N}(0, Q)$ for all k, and that $v(i)$ and $v(j)$ are independent for $i \neq j$.

More general versions of the algorithm can be derived by allowing F, Q to vary with k.

2.2 Observation Model

Given the state $x(k)$, the observation model is

$$z(k) = H\,x(k) + w(k), \tag{2}$$

where H is a known observation matrix, and $w(k) \sim \mathcal{N}(0, R)$, again with independent occurrences for different k.

The algorithm can be generalized by allowing H, R, to vary with k.

2.3 Probability of Detection

For this note, for simplicity we assume that the probability of detection is constant, P_D. A more general version of the algorithm can be derived by allowing P_D to depend on location $x(k)$.

2.4 Clutter Model

The number of clutter detections is modeled as a Poisson random variable determined by a spatial density parameter λ. If the region over which the detector is active has volume V, the expected number of clutter measurements is λV. In the following, the parameter V will not appear in the algorithm steps, so we forgo a further discussion for now.

The parameter λ has the following interpretation : Consider an infinitesimal volume of measure δ in any region that the detector is inspecting. Then, the probability of making a spurious detection in that region is $\delta\lambda$.

A more general algorithm can be derived by allowing λ to vary throughout the considered volume.

3 Derivation

Let us introduce some additional notation.

At each instant k, the detector provides a set of detections, namely $z(k) = [z_1(k), z_2(k), \ldots, z_{d_k}(k)]$, where $z_i(k) \in \mathbb{R}^m$. We also denote the set of detections up to time k as Z^k.

We write $x(k|l)$ to refer to the random variable $x(k)$ conditioned on Z^l. We will assume that the distribution of this variable is Gaussian with mean $\hat{x}(k|l)$, covariance $P(k|l)$. We write $x(k|l) \sim \mathcal{N}\big(\hat{x}(k|l), P(k|l)\big)$ to denote that.

Given the distribution of $x(k-1|k-1)$, and the set of detections at k, namely $z(k)$, PDAF is essentially an update procedure to obtain $x(k|k)$.

Given this notation, pseudo-code for PDAF is presented in Algorithm 1.

3.1 Update Equations

To simplify derivation, we will rely on some hypotheses which form a partition of the sample space. Consider the following events concerning the detections at the k^{th} stage (we drop the index k to simplify notation) :

$$H_i(m) = E_m \cap \begin{cases} \text{none of the detections are associated with the object,} & \text{if } i = 0, \\ \text{the } i^{\text{th}} \text{ detection is associated with the object,} & \text{if } i = 1, 2, \ldots, m. \end{cases} \tag{3}$$

where $E_m = \{m \text{ detections are made}\}$.

If $(m, j) \neq (n, k)$, then $H_i(m) \cap H_n(k) = \emptyset$. Also,

$$\bigcup_{m=0}^{\infty} \bigcup_{i=0}^{m} H_i(m) = \Omega. \tag{4}$$

Therefore, $\{H_i(m)\}_{i,m}$ form a partition of the sample space.

Algorithm 1 The Probabilistic Data Association Filter (PDAF)

Input: $x(k-1)|Z^{k-1} \sim \mathcal{N}(\hat{x}(k-1|k-1), P(k-1|k-1))$, detections $z_1, \ldots, z_m$, parameters λ, P_D

Output: Distribution of $x(k)|Z^k$ approximated as $\mathcal{N}(\hat{x}(k|k), P(k|k))$

$\hat{x} \leftarrow F\, \hat{x}(k-1|k-1)$ ▷ predicted mean

$P \leftarrow F\, P(k-1|k-1)\, F^T + Q$ ▷ predicted covariance

$\hat{z} \leftarrow H\, \hat{x}$ ▷ predicted mean for measurement

$S \leftarrow H\, P\, H^T + R$ ▷ predicted covariance for measurement

$\tilde{\beta}_0 \leftarrow 1 - P_D$ ▷ likelihood of missed detection

$\tilde{\beta}_i \leftarrow P_D \dfrac{\mathcal{N}(z_i; \hat{z}, S)}{\lambda}$ for $i = 1, 2, \ldots, m$ ▷ likelihood that i^{th} obs. belongs to object

$\beta_i \leftarrow \dfrac{\tilde{\beta}_i}{\sum_{i=0}^m \tilde{\beta}_i}$ for $i = 0, 1, \ldots, m$ ▷ probability for the i^{th} hypothesis

$\nu_i \leftarrow z_i - \hat{z}$ for $i = 1, 2, \ldots, m$ ▷ innovation for the i^{th} hypothesis

$\nu \leftarrow \sum_{i=1}^m \beta_i\, \nu_i$ ▷ average innovation

$W \leftarrow P\, H\, S^{-1}$ ▷ Kalman gain

$\mathbf{\hat{x}(k|k)} \leftarrow \hat{x} + W\, \nu$ ▷ sought mean

$P^c \leftarrow P - P\, H\, S^{-1}\, H^T\, P$ ▷ covariance for each hypothesis, except missed detection

$\tilde{P} \leftarrow W\, \left(\sum_{i=1}^m \beta_i \nu_i\, \nu_i^T - \nu\, \nu^T\right)\, W^T$ ▷ excess covariance due to merging all components into one

$\mathbf{P(k|k)} \leftarrow \beta_0\, P + (1-\beta_0)\, P^c + \tilde{P}$ ▷ approximate covariance for the Gaussian mixture to output

Consider an application of the total probability theorem, using this partition. Suppose $z(k) = [z_1(k), \ldots z_d(k)]$, so we have d detections at the k^{th} instant. We can write, for some set $S \subset \mathbb{R}^n$,

$$\begin{aligned} P(x(k) \in S | x(k-1|k-1), z(k)) &= \sum_{m=0}^{\infty} \sum_{i=0}^{m} P\big((x(k) \in S) \cap H_i(m) \,\big|\, x(k-1|k-1), z(k)\big) \\ &= \sum_{i=0}^{d} P\big((x(k) \in S) \cap H_i(d) \,\big|\, x(k-1|k-1), z(k)\big) \\ &= \sum_{i=0}^{d} P\big((x(k) \in S) \big|\, H_i(d), x(k-1|k-1), z(k)\big)\; \underbrace{P\big(H_i(d) \big|\, x(k-1|k-1), z(k)\big)}_{\beta_i} \end{aligned} \tag{5}$$

where equality in the second line follows because we already know from $z(k)$ that there are d detections, so $P(H_i(m)|z(k)) = 0$ if $m \neq d$.

From the last equation, we find the distribution of $x(k)$ conditioned on $x(k-1|k-1), z(k)$ is given by

$$x(k)\big|x(k-1|k-1), z(k) \sim \sum_{i=0}^{d} \beta_i\, p_i(\cdot) \tag{6}$$

where $p_i(\cdot)$ denotes the distribution of $x(k)\big|\, H_i(d), x(k-1|k-1), z(k)$. In what follows, we will first derive p_i, and proceed to find an expression for β_i.

Distribution of $x(k)$ Conditioned on H_i

Conditioned on H_i, with $i > 0$, estimation of the pdf of $x(k|k)$ by updating the distribution of $x(k-1|k-1)$ can be performed with a regular Kalman filter. With an abuse of notation we write the update rule as

$$x(k|k)|H_i := x(k|k, H_i) \sim \mathcal{N}\big(\hat{x}_i(k|k), P(k|k)\big), \tag{7}$$

where

$$x_i(k|k) = \hat{x}(k|k-1) + W\,\big(z_i(k) - H\,\hat{x}(k|k-1)\big), \tag{8}$$

and

$$P(k|k) = P(k|k-1) - P(k|k-1)\,H^T(H\,P(k|k-1)\,H^T + R)^{-1}\,H\,P(k|k-1). \tag{9}$$

Here, $\hat{x}(k|k-1)$ denotes the 'predicted' mean of $x(k)$, given Z^{k-1}, that is

$$\hat{x}(k|k-1) = \mathbb{E}(x(k)|Z^{k-1}) = F\,x(k-1|k-1), \tag{10}$$

$P(k|k-1)$ denotes the covariance of $x(k)$ given Z^{k-1}, which can be written as,

$$P(k|k-1) = \mathrm{cov}(x(k)|Z^{k-1}) = F\,P(k-1|k-1)\,F^T + Q, \tag{11}$$

and W denotes the Kalman gain,

$$W = P(k|k-1)\,H^T(H\,P(k|k-1)\,H^T + R)^{-1}. \tag{12}$$

If we condition on H_0 (missed detection), then we have

$$x(k|k, H_0) \sim \mathcal{N}\big(\hat{x}(k|k-1), P(k|k-1)\big). \tag{13}$$

Putting it all together, we find that $x(k|k)$ is a Gaussian mixture of the form

$$x(k|k) \sim \beta_0\,\mathcal{N}\big(\hat{x}(k|k-1), P(k|k-1)\big) + \sum_{i=1}^{m} \beta_i\,\mathcal{N}\big(\hat{x}_i(k|k), P(k|k)\big). \tag{14}$$

In order to derive a recursive algorithm, we need to make sure that the density of $x(k|k)$ matches in form to the density of $x(k-1|k-1)$ – i.e., a Gaussian. Otherwise, we would end up with a Gaussian mixture with an ever increasing number of components. PDAF tackles this issue by approximating the Gaussian mixture with a single Gaussian distribution. This approximation is realized through moment matching. That is, we compute the mean and covariance of the mixture (say μ and P), and declare that $x(k|k) \sim \mathcal{N}(\mu, P)$. Therefore, all we need are expressions for the mean and covariance of the distribution in (14).

Using $\sum_{i=0}^{m} \beta_i = 1$, and the expressions for $\hat{x}_i$, we find that

$$\mathbb{E}(x(k|k)) = \hat{x}(k|k-1) + W\,\underbrace{\sum_{i=1}^{m} \beta_i\,\nu_i}_{\nu}, \tag{15}$$

where $\nu_i = z_i(k) - H\,\hat{x}(k|k-1)$ is the "innovation" in the i^{th} detection. Here ν can be interpreted as the combined innovation given the full set of detections.

In order to find the expression for the covariance, let us write $\mathbb{E}(x(k|k)) \cdot \mathbb{E}(x(k|k))^T$ and $\mathbb{E}(x(k|k)\, x(k|k)^T)$, using the total probability theorem with the partition $\{H_i\}_{i=0}^m$.

$$\begin{aligned}\mathbb{E}(x(k|k)) \cdot \mathbb{E}(x(k|k))^T &= (\hat{x}(k|k-1) + W\nu) \cdot (\hat{x}(k|k-1) + W\,\nu)^T \\ &= \hat{x}(k|k-1)\,\hat{x}(k|k-1)^T + \hat{x}(k|k-1)\,\nu^T\,W^T + W\,\nu\,\hat{x}(k|k-1)^T \\ &\quad + W\,\nu\,\nu^T\,W^T \end{aligned} \tag{16}$$

$$\begin{aligned}\mathbb{E}(x(k|k)\,x(k|k)^T) &= \beta_0\,\left(P(k|k-1) + \hat{x}(k|k-1)\,\hat{x}(k|k-1)^T\right) \\ &\quad + \sum_{i=1}^m \beta_i\,\left(P(k|k) + \hat{x}_i(k|k)\,\hat{x}_i(k|k)^T\right) \\ &= \beta_0\,P(k|k-1) + (1-\beta_0)\,P(k|k) + \beta_0\,\hat{x}(k|k-1)\,\hat{x}(k|k-1)^T \\ &\quad + \sum_{i=1}^m \beta_i\hat{x}(k|k-1)\,\hat{x}(k|k-1)^T + \sum_{i=1}^m \beta_i\hat{x}(k|k-1)\,\nu_i^T\,W^T \\ &\quad + \sum_{i=1}^m \beta_i W\nu_i\,\hat{x}(k|k-1) + \sum_{i=1}^m \beta_i\,W\nu_i\,\nu_i^T\,W^T \\ &= \beta_0\,P(k|k-1) + (1-\beta_0)\,P(k|k) + \hat{x}(k|k-1)\,\hat{x}(k|k-1)^T \\ &\quad + \hat{x}(k|k-1)\,\nu^T\,W^T + W\,\nu\,\hat{x}(k|k-1)^T + W\,\left(\sum_{i=1}^m \beta_i\,\nu_i\,\nu_i^T\right)\,W^T \end{aligned} \tag{17}$$

We finally obtain $\mathrm{cov}(x(k|k))$ via the difference as

$$\begin{aligned}\mathrm{cov}(x(k|k)) &= \mathbb{E}(x(k|k)\,x(k|k)^T) - \mathbb{E}(x(k|k)) \cdot \mathbb{E}(x(k|k))^T \\ &= \beta_0\,P(k|k-1) + (1-\beta_0)\,P(k|k) + W\,\left(\left(\sum_{i=1}^m \beta_i\nu_i\,\nu_i^T\right) - \nu\,\nu^T\right)\,W^T. \end{aligned} \tag{18}$$

For an alternative expression, let us take $\nu_0 = 0$. Then, we have

$$\sum_{i=0}^m \beta_i\,\nu_i = \sum_{i=1}^m \beta_i\,\nu_i = \nu. \tag{19}$$

Using this observation, we can write

$$\sum_{i=0}^m \beta_i\,(\nu_i - \nu)\,(\nu_i - \nu)^T = \sum_{i=0}^m \beta_i\,\nu_i\,\nu_i^T - \sum_{i=0}^m \beta_i\,\nu\,\nu_i^T - \sum_{i=0}^m \beta_i\,\nu_i\,\nu^T + \sum_{i=0}^m \beta_i\,\nu\,\nu^T \tag{20}$$

$$= \sum_{i=1}^m \beta_i\,\nu_i\,\nu_i^T - \nu\,\nu^T - \nu\,\nu^T + \nu\,\nu^T \tag{21}$$

$$= \left(\sum_{i=1}^m \beta_i\,\nu_i\,\nu_i^T\right) - \nu\,\nu^T. \tag{22}$$

Therefore,

$$\mathrm{cov}(x(k|k)) = \beta_0\,P(k|k-1) + (1-\beta_0)\,P(k|k) + W\,\left(\sum_{i=0}^m \beta_i(\nu_i - \nu)\,(\nu_i - \nu)^T\right) \tag{23}$$

3.2 Probability of Hypotheses

What remains is to find expressions for $\beta_i = P\big(H_i(d)\big|\, x(k-1|k-1), z(k)\big)$. Observe that

$$P\big(H_i(d)\big|\, x(k-1|k-1), z(k)\big) \propto P\big(z(k)\big|H_i(d),\, x(k-1|k-1)\big) \cdot P\big(H_i(d)\big|\, x(k-1|k-1)\big). \quad (24)$$

For the first term, given $H_i(d)$, we already know that there are c false detections, or clutter, where

$$c = \begin{cases} d, & \text{if } i = 0, \\ d-1, & \text{if } i \neq 0. \end{cases} \quad (25)$$

By the assumed clutter model, conditioned on the number of clutter observations c, clutter measurements are independent and uniformly distributed over the observation field. Therefore, any set of clutter measurements will have likelihood $1/V^c$, where V is the total volume over which we make detections. Using $\mathcal{N}_k$ to denote the pdf of $z(k)$ given $x(k|k-1)$, we thus find (invoking the independence of the observations),

$$P\big(z(k)\big|H_i(d),\, x(k-1|k-1)\big) = \begin{cases} \dfrac{1}{V^d}, & \text{if } i = 0, \\ \mathcal{N}_k(z_i(k))\, \dfrac{1}{V^{d-1}}, & \text{if } i \neq 0. \end{cases} \quad (26)$$

For the second term, observe that knowing $x(k-1|k-1)$ does not provide any additional information about the event $H_i(m)$ for any i, or m – note that $H_i(m)$ does not contain any information about the location of the detections z. In other words, the event $H_i(m)$ is independent of $x(k-1|k-1)$. Therefore,

$$P\big(H_i(d)\big|\, x(k-1|k-1)\big) = P\big(H_i(d)\big). \quad (27)$$

$$P(H_i(d)) = \begin{cases} P\{d \text{ clutter measurements and no true detection}\}, & \text{if } i = 0, \\ P\{d-1 \text{ clutter measurements and true detection at the } i^{\text{th}} \text{ location}\}, & \text{if } i \neq 0. \end{cases} \quad (28)$$

$$= \begin{cases} (1-P_D)\, \dfrac{e^{-\lambda V}(\lambda V)^d}{d!}, & \text{if } i = 0, \\ \dfrac{P_D}{d}\, \dfrac{e^{-\lambda V}(\lambda V)^{d-1}}{(d-1)!}, & \text{if } i \neq 0, \end{cases} \quad (29)$$

$$= \frac{e^{-\lambda V}}{d!} \cdot \begin{cases} (1-P_D)\,(\lambda V)^d, & \text{if } i = 0, \\ P_D\,(\lambda V)^{d-1}, & \text{if } i \neq 0, \end{cases} \quad (30)$$

Putting it together, we find that

$$\beta_i \propto \begin{cases} 1 - P_D, & \text{if } i = 0, \\ \dfrac{P_D}{\lambda}\, \mathcal{N}_k(z_i(k)), & \text{if } i \neq 0. \end{cases} \quad (31)$$

4 Probability of Existence

In this section, we consider an extension of PDAF that takes into account a probability of existence. In particular, suppose the underlying track may cease to exist at some instances. More precisely, we define an event E_k as the event that there is truly a track at instant k. Following [2], we assume E_k is a Markov chain, and

$$\begin{bmatrix} P(E_k|Z^{k-1}) \\ P(E_k^c|Z^{k-1}) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} P(E_{k-1}|Z^{k-1}) \\ P(E_{k-1}^c|Z^{k-1}) \end{bmatrix}. \tag{32}$$

We can now write

$$P(E_k|Z^k) = \frac{P(z(k)|E_k, Z^{k-1}) \cdot P(E_k|Z^{k-1})}{P(z(k)|Z^{k-1})} \tag{33}$$

and

$$P(E_k^c|Z^k) = \frac{P(z(k)|E_k^c, Z^{k-1}) \cdot P(E_k^c|Z^{k-1})}{P(z(k)|Z^{k-1})} \tag{34}$$

$$= \frac{P(z(k)|E_k^c, Z^{k-1}) \cdot \big(1 - P(E_k|Z^{k-1})\big)}{P(z(k)|Z^{k-1})} \tag{35}$$

Since the denominators in (33), (35) are common, and the two events partition the sample space, it suffices to compute the numerators to derive an expression for the sought probabilities.

Using the results from the previous section, we find (recall that $z(k)$ has d elements).

$$P(z(k)|E_k, Z^{k-1}) = \sum_{i=0}^{d} P\big(z(k)\big|H_i(d),\, x(k-1|k-1)\big) \cdot P\big(H_i(d)\big|\, x(k-1|k-1)\big) \tag{36}$$

$$= (1-P_D)\frac{e^{-\lambda V}\lambda^d}{d!} + P_D \frac{e^{-\lambda V}\lambda^{d-1}}{d!} \sum_{i=1}^{d} \mathcal{N}_k(z_i(k)) \tag{37}$$

$$= \frac{e^{-\lambda V}\lambda^d}{d!} \left((1-P_D) + \frac{P_D}{\lambda} \sum_{i=1}^{d} \mathcal{N}_k(z_i(k)) \right) \tag{38}$$

On the other hand,

$$P(z(k)|E_k^c, Z^{k-1}) = \frac{e^{-\lambda V}\lambda^d}{d!}. \tag{39}$$

Note that,

$$\frac{P(E_k|Z^k)}{P(E_k^c|Z^k)} = \frac{\left((1-P_D) + \frac{P_D}{\lambda} \sum_{i=1}^{d} \mathcal{N}_k(z_i(k)) \right) \cdot P(E_k|Z^{k-1})}{1 - P(E_k|Z^{k-1})} = c \tag{40}$$

Using $P(E_k|Z^k) + P(E_k^c|Z^k) = 1$, we solve for $P(E_k|Z^k)$ to find that

$$P(E_k|Z^k) = \frac{c}{1+c} \tag{41}$$

$$= \frac{1 - P_D\left(1 - \frac{1}{\lambda}\sum_{i=1}^{d} \mathcal{N}_k(z_i(k))\right)}{1 - P_D\left(1 - \frac{1}{\lambda}\sum_{i=1}^{d} \mathcal{N}_k(z_i(k))\right) \cdot P(E_k|Z^{k-1})} \cdot P(E_k|Z^{k-1}) \tag{42}$$

5 A Simple Special Case

We now consider a scenario where instead of clutter, we have occasional outliers. That is, at each instant, our filter receives a single observation. But that observation can come from the true target of interest or be an outlier. In keeping with the notation, we take the probability that we get to observe the true target to be P_D. We further assume that the outliers are sampled from a uniform distribution with pdf equal to U, where $1/U$ can be interpreted as the volume of the support of the uniform distribution.

Under this assumption, we only have two hypothesis at each instant that partition the sample space:

- H_0 : Observation is an outlier
- H_1 : Observation comes from the true target

Following the derivation in Section 3.2, we find that, for a given observation z at instant k (note this is a single observation), and $x(k|k-1) \sim \mathcal{N}_k$

$$\beta_i := P\big(H_i(d)\big|\, x(k-1|k-1), z(k)\big) \tag{43}$$

$$\propto \begin{cases} U \cdot (1-P_D) & \text{if } i = 0, \\ \mathcal{N}_k(z) \cdot P_D & \text{if } i = 1. \end{cases} \tag{44}$$

In addition, we assume that the $F = I$ in (1), and $H = I$ in (2). That is, the state update observation models are

$$x(k) = x(k-1) + v(k-1), \tag{45}$$

$$z(k) = x(k) + w(k), \tag{46}$$

respectively.

Algorithm 2 PDAF for the simplified setup in Sec. 5

Input: $x(k-1)|Z^{k-1} \sim \mathcal{N}(\hat{x}(k-1|k-1), P(k-1|k-1))$, detection z, parameters U, P_D

Output: Distribution of $x(k)|Z^k$ approximated as $\mathcal{N}(\hat{x}(k|k), P(k|k))$

$\hat{x} \leftarrow \hat{x}(k-1|k-1)$ ▷ predicted mean

$P \leftarrow P(k-1|k-1) + Q$ ▷ predicted covariance

$\hat{z} \leftarrow \hat{x}$ ▷ predicted mean for measurement

$S \leftarrow P + R$ ▷ predicted covariance for measurement

$\tilde{\beta}_0 \leftarrow U \cdot (1-P_D)$ ▷ likelihood of outlier

$\tilde{\beta}_1 \leftarrow P_D \dfrac{\mathcal{N}(z; \hat{z}, S)}{\lambda}$ ▷ likelihood of true observation

$\beta_i \leftarrow \dfrac{\tilde{\beta}_i}{\tilde{\beta}_0 + \tilde{\beta}_1}$ for $i = 0, 1$ ▷ probability for the i^{th} hypothesis

$\nu_1 \leftarrow (z - \hat{z})$ ▷ innovation

$W \leftarrow P\, S^{-1}$ ▷ Kalman gain

$\hat{\mathbf{x}}(\mathbf{k}|\mathbf{k}) \leftarrow \hat{x} + \beta_1\, W\, \nu_1$ ▷ sought mean

$P^c \leftarrow P - P\, S^{-1}\, P$ ▷ covariance for H_1

$\tilde{P} \leftarrow \beta_1\, \beta_0\, W\, \nu_1\, \nu_1^T\, W^T$ ▷ excess covariance due to merging all components into one

$\mathbf{P}(\mathbf{k}|\mathbf{k}) \leftarrow \beta_0\, P + (1-\beta_0)\, P^c + \tilde{P}$ ▷ approximate covariance for the Gaussian mixture to output

Under these changes, the algorithm update steps are as in Algo. 2.

References

[1] Y. Bar-Shalom, F. Daum, and J. Huang. The probabilistic data association filter. *IEEE Control Systems Magazine*, 29(6):82–100, Dec 2009.

[2] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *IEEE Transactions on Automatic Control*, 39(6):1237–1241, 1994.