

Convergence Analysis of the Gaussian Mixture PHD Filter

Daniel Clark* and Ba-Ngu Vo†

Abstract—The Gaussian mixture Probability Hypothesis Density (PHD) filter was proposed recently for jointly estimating the time-varying number of targets and their states from a sequence of sets of observations without the need for measurement-to-track data association. It was shown that, under linear-Gaussian assumptions, the posterior intensity at any point in time is a Gaussian mixture. This article proves uniform convergence of the errors in the algorithm and provides error bounds for the pruning and merging stages. In addition, uniform convergence results for the extended Kalman PHD Filter are given, and the unscented Kalman PHD Filter implementation is discussed.

Index Terms—Multi-target tracking, optimal filtering, point processes, random sets, PHD filter.

SP-EDICS: SSP-a Filtering, **SSP-d** Tracking algorithms.

I. INTRODUCTION

The closed-form solution to the PHD (Probability Hypothesis Density) filter was recently derived to provide a solution for multiple target tracking with linear/Gaussian models without the need for measurement-to-track data association [1], [2]. It was shown that when the initial prior intensity of the random-set of targets is a Gaussian mixture, the posterior intensity at any time step is also a Gaussian mixture.

This article demonstrates the uniform convergence of the errors for each of the stages of the Gaussian Mixture PHD Filter [1], [2] using results already established for the particle implementation of the PHD filter [3] and Wiener's Theory of Approximation [4]. Error bounds are provided in L_1 for the pruning and merging stages of the algorithm, based on those established for the Gaussian Sum filter [5].

Extensions of the Gaussian Mixture PHD filter proposed in [2], namely the Extended Kalman PHD filter and the Unscented Kalman PHD filter, are also discussed. Convergence results for the Extended Kalman PHD filter are given based on the Gaussian Sum filter developed by Sorenson and Alspach [5], the L_1 convergence properties discussed in

Anderson and Moore [6], and the fact that densities can be represented by a linear combination of Gaussians in L_1 [4]. Taken with the convergence results of the L_1 error, the Gaussian mixture approximation then converges to the true posterior intensity.

The results show that, under linear Gaussian assumptions of the dynamic model, the Gaussian Mixture posterior intensity can approximate the true posterior intensity to any desired degree of accuracy. In addition, error bounds have been established for the pruning and merging stages of the algorithm which ensure that the accuracy of these stages can be controlled.

II. BACKGROUND

The random-set framework for multiple target tracking developed by Mahler [7] using Finite Set Statistics offers a distinct alternative to the traditional approach to multiple target tracking, which involves single target stochastic filters assigned to each target which are managed by a data association technique, by treating the collection of individual targets as a set-valued state and the collection of individual observations as a set-valued observation. The set-valued state is predicted and updated at each time-step based on the set-valued observation. The multiple target posterior can be estimated using a generalisation of the single target Bayesian filtering equations to a multiple target scenario. This model can also incorporate clutter, or false measurements, into the framework.

The complexity of computing this recursion grows exponentially with the number of targets and so the optimal filter must be approximated. To alleviate the complexity of computing the multi-target posterior, a recursion was derived for the first order moment of the multi-target posterior distribution, known as the PHD filter [7].

The Sequential Monte Carlo implementation of the PHD filter [8] provided a solution for multiple target tracking with non-linear/non-Gaussian target models. Convergence of this algorithm has been shown [8] [3], [9], and applications have been demonstrated on realistic synthetic data [10] and real data [11], [12].

The closed-form version of the PHD filter for linear-Gaussian target dynamics was developed recently to provide a multi-target tracker without the complexity of the particle filtering approach [1]. This algorithm provided a means of estimating the set of targets at each time-step but did not provide continuity of the individual targets. The Gaussian mixture multi-target tracker [13] was developed recently and

This work is supported in part by a research grant awarded by the Australian Research Council.

* Corresponding author. Address: Department of Electrical, Electronic and Computer Engineering, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, UK. E-mail: decl1@hw.ac.uk Tel.: +44 131 451 3357 Fax: +44 131 451 4155

† Address: Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, Victoria 3010, Australia. E-mail: bv@ee.unimelb.edu.au Tel.: +61 3 8344 6693 Fax: +61 3 8344 6678

it was shown that individual Gaussians within the mixture are able to track targets successfully. This has also been demonstrated successfully for tracking objects in forward-scan sonar in clutter [14]. We show here the convergence properties of the closed-form PHD filter, also known as the Gaussian Mixture PHD filter.

III. RANDOM-SET MULTI-TARGET FILTERING

In this section, the random-set approach to multiple target filtering is described. The multiple target tracking model is formulated using random-sets for the observations, target states and target estimates. The number of targets is also a parameter that needs to be estimated. The posterior PHD or posterior intensity is given as the first-order statistical moment of the posterior random-set of targets. The PHD represents the expectation of a multi-target posterior and is multimodal, where each mode represents an expected target location. Finally, the PHD filter recursion for predicting and updating the intensity is described.

A. Multiple Target Tracking Model

Let (Ω, F, P) be a probability space on which we have defined two finite set-valued stochastic processes $X = \{X_k : k \in \mathbb{N}\}$ and $Z = \{Z_k : k \in \mathbb{N} \setminus \{0\}\}$. Process X is called the state process and process Z is called the observation process, and these processes are used to formulate the multiple target Bayesian Filtering equations. The multiple target inference model adopted here uses Finite Set Statistics [7] as a means of directly extending the single target Bayesian recursive state estimation to a multiple target scenario. Instead of using a random vector to represent a target state, a random finite set of vectors is used to represent a variable number of target states.

The set of objects tracked at time k is modelled by the Random Finite Set (RFS) X_k , which includes the set of targets survived at time k from the previous time step, the set of targets spawned from X_{k-1} and the set of targets that appear spontaneously at time k . The RFS X_k is the set of target states, $X_k = \{x_{k,1}, \dots, x_{k,T_k}\}$, where $x_{k,i}$ represents the state of an individual target and T_k is the number of targets at time k .

The measurements at time k are modelled by RFS $Z_k = \{z_{k,1}, \dots, z_{k,m_k}\}$, where $z_{k,j}$ represents a single-target measurement or false alarm and m_k is the number of observations at time k . This RFS includes measurements from target states in X_k and also measurements due to clutter. The Bayesian recursion for the multiple target model is determined from the following prior and posterior calculations:

$$p_{k|k-1}(X_k|Z_{1:k-1}) = \int f_{k|k-1}(X_k|X_{k-1}, Z_{1:k-1}) p_{k-1}(X_{k-1}|Z_{1:k-1}) \mu(dX_{k-1}), \quad (1)$$

$$p_k(X_k|Z_{1:k}) \propto g_k(Z_k|X_k) p_{k|k-1}(X_k|Z_{1:k-1}), \quad (2)$$

where $p_{k|k-1}(X_k|Z_{1:k-1})$, $p_k(X_k|Z_{1:k})$, $f_{k|k-1}(X_k|X_{k-1}, Z_{1:k-1})$, $g_k(Z_k|X_k)$, represent the multitarget prior, posterior, transition and likelihood respectively. The dominating measure μ is constructed from the Lebesgue measure as described in [15].

The multi-target filtering problem is to estimate the unobserved signal process $X_{0:k} = \{X_0, \dots, X_k\}$ based on observations $Z_{1:k} = \{Z_1, \dots, Z_k\}$, i.e. to obtain $\hat{X}_k = \{\hat{x}_{k,1}, \dots, \hat{x}_{k,T_k}\}$, where $\hat{x}_{k,i}$ are the individual target estimates and \hat{T}_k is the estimate of the number of targets at time k .

B. The Probability Hypothesis Density (PHD) filter

The PHD, also known as the intensity in the point process literature, is defined as the density, $v_k(x_k|Z_{1:k})$, whose integral, $\int_S v_k(x_k|Z_{1:k}) dx_k$, on any region S of the state space is the expected number of targets in S . The estimated object states can be detected as the peaks of this distribution. The posterior intensity $v_k(x)$ is propagated in time via the PHD prediction and measurement update equations:

$$v_{k|k-1}(x) = \int \phi_{k|k-1}(x, \zeta) v_{k-1}(\zeta) d\zeta + \gamma_k(x), \quad (3)$$

$$v_k(x) = [1 - p_{D,k}(x)] v_{k|k-1}(x) + \sum_{z \in Z_k} \frac{\Psi_{k,z}(x) v_{k|k-1}(x)}{\kappa_k(z) + \int \Psi_{k,z}(\xi) v_{k|k-1}(\xi) d\xi}, \quad (4)$$

where $\phi_{k|k-1}(x, \xi) = p_{S,k}(\xi) f_{k|k-1}(x|\xi) + \beta_{k|k-1}(x|\xi)$, $\kappa_k(z) = \lambda_k c_k(z)$ and $\Psi_{k,z} = p_{D,k}(x) g(z|x)$. The conditioning on measurements $Z_{1:k}$ is removed in v_k for notational simplicity. In the prediction equation, $\beta_{k|k-1}$ is the PHD for spawned target birth of a new target at time k , $p_{S,k}$ is the probability of target survival and $f_{k|k-1}(x|x_{k-1})$ is the single target motion distribution¹. In the data update equation, $g(z|x)$ is the single target likelihood, $p_{D,k}$ is the probability of detection, λ_k is the Poisson parameter specifying the expected number of false alarms and c_k is the probability distribution over the observation space of clutter points.

The PHD filter model assumes that each target evolves and generates observations independently of one another, that the clutter is independent of target-induced measurements, and that the predicted multi-target RFS governed by $p_{k|k-1}$ is Poisson.

IV. THE GAUSSIAN MIXTURE PHD FILTER

In this section, we describe the linear-Gaussian multiple target model and the recently developed Gaussian Mixture PHD filter.

A. The Linear Gaussian Multiple Target Model

The multiple target model for the PHD recursion is described here. Each target follows a linear Gaussian dynamical model,

$$f_{k|k-1}(x|\zeta) = \mathcal{N}(x; F_{k-1}\zeta, Q_{k-1}), \quad (5)$$

$$g_k(z|x) = \mathcal{N}(z; H_k x, R_k), \quad (6)$$

where $\mathcal{N}(\cdot; m, P)$ denotes a Gaussian density with mean m and covariance P , F_{k-1} is the state transition matrix, Q_{k-1} is the process noise covariance, H_k is the observation matrix, and R_k is the observation noise covariance.

¹The same notation is used for multi-target and single-target densities. There is no danger of confusion since in the single-target case the arguments are vectors whereas in the multi-target case the arguments are finite sets.

The survival and detection probabilities are state independent, $p_{S,k}(x) = p_{S,k}$, and $p_{D,k}(x) = p_{D,k}$. The intensities of the spontaneous birth and spawned targets are Gaussian mixtures,

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}), \quad (7)$$

$$\beta_{k|k-1}(x|\zeta) = \sum_{j=1}^{J_{\beta,k}} w_{\beta,k}^{(j)} \mathcal{N}(x; F_{\beta,k-1}^{(j)} \zeta + d_{\beta,k-1}^{(j)}, Q_{\beta,k-1}^{(j)}), \quad (8)$$

where $J_{\gamma,k}$, $w_{\gamma,k}^{(i)}$, $m_{\gamma,k}^{(i)}$, $P_{\gamma,k}^{(i)}$, $i = 1, \dots, J_{\gamma,k}$, are given model parameters that determine the shape of the birth intensity. Similarly, $J_{\beta,k}$, $w_{\beta,k}^{(j)}$, $F_{\beta,k-1}^{(j)}$, $d_{\beta,k-1}^{(j)}$, and $Q_{\beta,k-1}^{(j)}$, $j = 1, \dots, J_{\beta,k}$, determine the shape of the spawning intensity of a target with previous state ζ . Note that the sum of the weights in the above summations represent the expected number of spontaneous birth and spawned targets respectively, since the Gaussian mixture is an intensity function and not a probability distribution.

B. Prediction

Assume that each target follows a linear Gaussian dynamical model, that the survival and detection probabilities are constant, that the intensities of the birth and spawned targets are Gaussian mixtures, and that the posterior intensity at time $k-1$ is a Gaussian mixture of the form

$$v_{k-1}(x) = \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}(x; m_{k-1}^{(i)}, P_{k-1}^{(i)}), \quad (9)$$

then the predicted intensity to time k is also a Gaussian mixture and is given by

$$v_{k|k-1}(x) = v_{S,k|k-1}(x) + v_{\beta,k|k-1}(x) + \gamma_k(x), \quad (10)$$

where

$$v_{S,k|k-1}(x) = p_{S,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} \mathcal{N}(x; m_{S,k|k-1}^{(j)}, P_{S,k|k-1}^{(j)}), \quad (11)$$

$$m_{S,k|k-1}^{(j)} = F_{k-1} m_{k-1}^{(j)}, \quad (12)$$

$$P_{S,k|k-1}^{(j)} = Q_{k-1} + F_{k-1} P_{k-1}^{(j)} F_{k-1}^T, \quad (13)$$

$$v_{\beta,k|k-1}(x) = \sum_{j=1}^{J_{k-1}} \sum_{\ell=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(\ell)} \mathcal{N}(x; m_{\beta,k|k-1}^{(j,\ell)}, P_{\beta,k|k-1}^{(j,\ell)}), \quad (14)$$

$$m_{\beta,k|k-1}^{(j,\ell)} = F_{\beta,k-1}^{(\ell)} m_{k-1}^{(j)} + d_{\beta,k-1}^{(\ell)}, \quad (15)$$

$$P_{\beta,k|k-1}^{(j,\ell)} = Q_{\beta,k-1}^{(\ell)} + F_{\beta,k-1}^{(\ell)} P_{k-1}^{(j)} (F_{\beta,k-1}^{(\ell)})^T. \quad (16)$$

The proof of this is given in [2].

C. Measurement Update

Under the above assumptions, and that the predicted intensity to time k is a Gaussian mixture of the form

$$v_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}), \quad (17)$$

then the posterior intensity at time k is also a Gaussian mixture, and is given by

$$v_k(x) = (1 - p_{D,k}) v_{k|k-1}(x) + \sum_{z \in Z_k} v_{D,k}(x; z) \quad (18)$$

where

$$v_{D,k}(x; z) = \sum_{j=1}^{J_{k|k-1}} \frac{p_{D,k} w_{k|k-1}^{(j)} q_k^{(j)}(z)}{\kappa_k(z) + p_{D,k} \sum_{\ell=1}^{J_{k|k-1}} w_{k|k-1}^{(\ell)} q_k^{(\ell)}(z)} \mathcal{N}(x; m_{k|k}^{(j)}(z), P_{k|k}^{(j)}), \quad (19)$$

$$q_k^{(j)}(z) = \mathcal{N}(z; H_k m_{k|k-1}^{(j)}, H_k P_{k|k-1}^{(j)} H_k^T + R_k), \quad (20)$$

$$m_{k|k}^{(j)}(z) = m_{k|k-1}^{(j)} + K_k^{(j)} (z - H_k m_{k|k-1}^{(j)}), \quad (21)$$

$$P_{k|k}^{(j)} = [I - K_k^{(j)} H_k] P_{k|k-1}^{(j)}, \quad (22)$$

$$K_k^{(j)} = P_{k|k-1}^{(j)} H_k^T (H_k P_{k|k-1}^{(j)} H_k^T + R_k)^{-1}. \quad (23)$$

The proof of this is given in [2].

V. CONVERGENCE OF THE L_1 ERROR

This section shows the L_1 convergence of the Gaussian mixture PHD filter; in other words, proving that each step in time of the PHD filter will maintain a suitable approximation error that converges to zero as the number of Gaussians in the mixture tends to infinity. This is achieved through the successive application of triangle inequalities and Hölder's inequality. Finally the observation update is shown to converge using an adaptation of the previous result on particle PHD convergence [3].

Results for the convergence properties of the Gaussian mixture PHD Filter are now established. Convergence of the L_1 error is first shown, $\lim_{J_k \rightarrow \infty} \langle v_k^{J_k} - v_k, \phi \rangle = 0$, for any function ϕ , where $v_k^{J_k}$ is the Gaussian mixture approximation to v_k with J_k Gaussian components. The $\langle \cdot, \cdot \rangle$ notation defines the usual inner product

$$\langle v_k, \phi \rangle = \int v_k(x_k | Z_{1:k}) \phi(x_k) dx_k. \quad (24)$$

and the operator notations $f_{k|k-1} \phi$, $v f_{k|k-1}$ are defined by

$$(f_{k|k-1} \phi)(x_{k-1}) = \int f_{k|k-1}(x_k | x_{k-1}) \phi(x_k) dx_k, \quad (25)$$

$$(v f_{k|k-1})(x_k) = \int v(x_{k-1}) f_{k|k-1}(x_k | x_{k-1}) dx_{k-1}. \quad (26)$$

Note that $\langle v f_{k|k-1}, \phi \rangle = \langle v, f_{k|k-1} \phi \rangle$. Also the PHD prediction equation (3) can be written as

$$v_{k|k-1} = (p_{S,k} v_{k-1}) f_{k|k-1} + v_{k-1} \beta_{k|k-1} + \gamma_k. \quad (27)$$

In the proofs, we use an instance of Hölder's Inequality (see, for example pp. 27 [16]),

$$|\langle v, \phi \rangle| \leq \|v\|_1 \|\phi\|_\infty. \quad (28)$$

The data update equation assumes a Poisson model and, hence, is only an approximation. The clutter parameters need to be determined from the data and cannot be inferred from the recursion. For the purpose of these proofs, it has been assumed that the correct density c_k and average number of Poisson clutter points λ_k are known.

THEOREM 1 Any density on \mathbb{R}^d can be approximated as closely as desired in L_1 by a linear combination of Gaussian densities,

$$v(x) = \lim_{n \rightarrow \infty} \sum_{i=1}^n w^{(i)} \mathcal{N}(x; \mu_i, P_i) \quad (29)$$

Proof: This result is due to Wiener's theorem on approximation [4].

This means that given any $\varepsilon > 0$, a positive integer N can be found such that

$$\int |v(x) - \sum_{i=1}^n w^{(i)} \mathcal{N}(x; \mu_i, P_i)| dx \leq \varepsilon, \quad (30)$$

for $n \geq N$. This result shall be used to establish bounds for the error in the Gaussian approximation to the posterior intensity.

A. Initialisation

It is assumed that the initial intensity is known. By Theorem 1, this initial intensity can be approximated to any arbitrary degree of accuracy, so that, for any bounded measurable function φ and any given $\varepsilon_0 > 0$, there is a positive integer J such that

$$|\langle v_0 - v_0^{J_0}, \varphi \rangle| \leq \varepsilon_0 \|\varphi\|_\infty, \quad (31)$$

for any $J_0 > J$, using Hölder's Inequality where

$$\|v_0 - v_0^{J_0}\|_1 \leq \varepsilon_0. \quad (32)$$

The notation v^J is used to denote the Gaussian mixture approximation to the density v , where J is the number of Gaussians in the mixture.

B. Prediction Equation

Let us assume that the approximation of the posterior intensity, $v_{k-1}^{J_{k-1}}$, by a sum of Gaussians converges uniformly to the true posterior intensity v_{k-1} . Then, given any $\varepsilon_{k-1} > 0$, an integer J can be found such that

$$|\langle v_{k-1} - v_{k-1}^{J_{k-1}}, \varphi \rangle| \leq \varepsilon_{k-1} \|\varphi\|_\infty, \quad (33)$$

for $J_{k-1} \geq J$, using Hölder's Inequality.

LEMMA 1 After the prediction step, there exist real numbers $b_{k|k-1}$, d_k and $e_{k|k-1}$ such that

$$|\langle v_{k|k-1}^{J_{k|k-1}} - v_{k|k-1}, \varphi \rangle| \leq (b_{k|k-1} \varepsilon_{k-1} + d_k + e_{k|k-1}) \|\varphi\|_\infty, \quad (34)$$

where d_k and $e_{k|k-1}$ are dependent on the models for the spontaneous birth and spawned target models.

Proof: Expanding the prediction density using equation (27) and using the triangle inequality,

$$\begin{aligned} |\langle v_{k|k-1}^{J_{k|k-1}} - v_{k|k-1}, \varphi \rangle| &\leq | \langle (p_{S,k} v_{k-1} f_{k|k-1})^{J_{k-1}} - p_{S,k} v_{k-1} f_{k|k-1}, \varphi \rangle | \\ &\quad + | \langle (v_{k-1} \beta_{k|k-1})^{J_{k-1} J_{\beta,k}} - v_{k-1} \beta_{k|k-1}, \varphi \rangle | + | \langle \gamma_k^{J_{\gamma,k}} - \gamma_k, \varphi \rangle | \end{aligned} \quad (35)$$

Taking the first term on the right hand side, which concerns the predicted intensity for existing targets, adding and subtracting

$\langle p_{S,k} v_{k-1}^{J_{k-1}}, f_{k|k-1} \varphi \rangle$, and using the triangle inequality again we get

$$| \langle (p_{S,k} v_{k-1} f_{k|k-1})^{J_{k-1}} - p_{S,k} v_{k-1} f_{k|k-1}, \varphi \rangle | \quad (36)$$

$$\begin{aligned} &\leq | \langle (p_{S,k} v_{k-1} f_{k|k-1})^{J_{k-1}}, \varphi \rangle - \langle p_{S,k} v_{k-1}^{J_{k-1}}, f_{k|k-1} \varphi \rangle | \\ &\quad + p_{S,k} | \langle v_{k-1}^{J_{k-1}} - v_{k-1}, f_{k|k-1} \varphi \rangle |, \end{aligned} \quad (37)$$

the first term on the right hand side is zero due to the linear Gaussian prediction model. Moreover,

$$\begin{aligned} (f_{k|k-1} \varphi)(x_{k-1}) &= \int f_{k|k-1}(x_k | x_{k-1}) \varphi(x_k) dx_k \quad (38) \\ &\leq \|\varphi\|_\infty \int f_{k|k-1}(x_k | x_{k-1}) dx_k \\ &= \|\varphi\|_\infty \end{aligned}$$

where the last equation follows from the fact that $f_{k|k-1}(x_k | x_{k-1})$ is a transition density. Hence,

$$| \langle v_{S,k|k-1}^{J_{k-1}} - p_{S,k} v_{k-1} f_{k|k-1}, \varphi \rangle | \leq p_{S,k} \|\varphi\|_\infty \varepsilon_{k-1}, \quad (39)$$

for $J_{k-1} \geq J$.

Now consider the birth model; there exists a constant d_k and integer J such that

$$| \langle \gamma_k^{J_{\gamma,k}} - \gamma_k, \varphi \rangle | \leq d_k \|\varphi\|_\infty, \quad (40)$$

for $J_{\gamma,k} \geq J$, since we assume that we can model this exactly.

Finally, for the spawned target model, adding and subtracting $\langle v_{k-1}^{J_{k-1}}, \beta_{k|k-1}^{J_{\beta,k}} \varphi \rangle$ and applying the triangle inequality gives

$$\begin{aligned} &| \langle (v_{k-1} \beta_{k|k-1})^{J_{k-1} J_{\beta,k}} - v_{k-1} \beta_{k|k-1}, \varphi \rangle | \quad (41) \\ &\leq | \langle (v_{k-1} \beta_{k|k-1})^{J_{k-1} J_{\beta,k}}, \varphi \rangle - \langle v_{k-1}^{J_{k-1}}, \beta_{k|k-1}^{J_{\beta,k}} \varphi \rangle | \\ &\quad + | \langle v_{k-1}^{J_{k-1}}, \beta_{k|k-1}^{J_{\beta,k}} \varphi \rangle - \langle v_{k-1}, \beta_{k|k-1} \varphi \rangle |, \end{aligned} \quad (42)$$

the first term on the right is zero due to the linear Gaussian spawned target model. Using an argument similar to the prediction for existing targets, equations (37-39), there exists a number $e_{k|k-1}$, such that the second term is less than or equal to $e_{k|k-1} \|\varphi\|_\infty$, for $J_{\beta,k} J_{k|k-1} \geq J$. This number, $e_{k|k-1}$, is dependent on the L_1 norm of the spawned target intensity, $\|\beta_{k|k-1}\|_1$. The lemma is proved by combining the three results above and setting $b_{k|k-1} = p_{S,k}$.

C. Measurement Equation

Let us assume that the approximation of the prediction intensity, $v_{k|k-1}^{J_{k|k-1}}$, by a sum of Gaussians converges uniformly to the true prediction intensity $v_{k|k-1}$. Then, using the same arguments as in (31), we have for any $\varepsilon_{k|k-1} > 0$, an integer J can be found such that

$$|\langle v_{k|k-1}^{J_{k|k-1}} - v_{k|k-1}, \varphi \rangle| \leq \varepsilon_{k|k-1} \|\varphi\|_\infty, \quad (43)$$

LEMMA 2 After the measurement update step, there exists a real number b_k , dependent on the number of measurements such that

$$|\langle v_k^{J_k} - v_k, \varphi \rangle| \leq b_k \varepsilon_{k|k-1} \|\varphi\|_\infty. \quad (44)$$

We assume that the predicted intensity $v_{k|k-1}$ is non-zero. This is a reasonable assumption as there would be no intensity to update when the measurements are received if it were zero.

Proof: Using the convergence result for the particle PHD filter (Lemma 2, Clark [3]), we have the inequality

$$\begin{aligned} |\langle v_k^{J_k} - v_k, \Phi \rangle| &\leq (1 - p_{D,k}) |\langle v_{k|k-1}^{J_{k|k-1}} - v_{k|k-1}, \Phi \rangle| \\ &+ \sum_{z \in Z_k} \left(\frac{1}{\langle v_{k|k-1}, \Psi_{k,z} \rangle} \left(\|\Phi\|_\infty \left| \langle v_{k|k-1}^{J_{k|k-1}} - v_{k|k-1}, \Psi_{k,z} \rangle \right| \right. \right. \\ &\quad \left. \left. + \left| \langle v_{k|k-1}^{J_{k|k-1}} - v_{k|k-1}, \Phi \Psi_{k,z} \rangle \right| \right) \right), \end{aligned} \quad (45)$$

using the assumption, we find that it is less than or equal to

$$\varepsilon_{k|k-1} \|\Phi\|_\infty \left((1 - p_{D,k}) + \sum_{z \in Z_k} \left(\frac{(2 \|\Psi_{k,z}\|_\infty)}{\langle v_{k|k-1}, \Psi_{k,z} \rangle} \right) \right), \quad (46)$$

so that the lemma is proved with

$$b_k = \left((1 - p_{D,k}) + \sum_{z \in Z_k} \left(\frac{(2 \|\Psi_{k,z}\|_\infty)}{\langle v_{k|k-1}, \Psi_{k,z} \rangle} \right) \right). \quad (47)$$

VI. PRUNING AND MERGING OF GAUSSIAN COMPONENTS

Since the number of Gaussians used to represent the Gaussian mixture increases at each time step, methods are required to ensure that the complexity of the algorithm is controlled. This is achieved through pruning, to eliminate the Gaussians with low weights, and merging, to combine Gaussians with similar means [2]. This section considers the errors introduced in these stages and shows that they can be controlled. The first approximation shows that a bound can be placed on the error introduced by eliminating terms with negligible weights, using a result for the Gaussian sum filter [5]. The second approximation arises from the tendency of many terms to converge to the same result so that they can be combined by adding their weights. When two terms are approximately equal, a bound on the error can be introduced so that the errors introduced in the merging stage are within tolerable limits, using another result for the Gaussian sum filter [17].

The number of Gaussian components used to represent the Gaussian mixture increases without bound; at time k , the posterior intensity requires

$$(J_{k-1}(1 + J_{\beta,k}) + J_{\gamma,k})(1 + |Z_k|) = O(J_{k-1}|Z_k|) \quad (48)$$

Gaussian components, where $|Z_k|$ is the number of measurements at time k and $O(\cdot)$ represents the asymptotic complexity. Clearly this has implications for the complexity of the algorithm, so it would be useful to reduce the total number of components required to represent the PHD. To alleviate these problems, components with small weights, $w_k^{(i)}$, are pruned, and components with similar means, $m_k^{(i)} \approx m_k^{(j)}$, are merged. The full procedure is given in Vo and Ma [2]. It is shown here that bounds can be put on the L_1 error when these methods are used.

A. Pruning

The pruning stage of the algorithm allows us to drop terms with negligible weights. It is shown here that the error introduced in this stage can be bounded. Suppose that the posterior intensity at time k is given by the sum of Gaussians,

$$v_k(x) = \sum_{i=1}^{J_k} w_k^{(i)} \mathcal{N}(x; m_k^{(i)}, P_k^{(i)}). \quad (49)$$

Assume, without loss of generality, that the components with indices $i = 1, \dots, N_p$ are those with weights, $w_k^{(i)}$, less than some specified threshold δ_1 . Prune these components and replace $v_k(x)$ by $v_k^P(x)$,

$$v_k^P(x) = \frac{\sum_{i=1}^{J_k} w_k^{(i)}}{\sum_{i=N_p+1}^{J_k} w_k^{(i)}} \sum_{i=N_p+1}^{J_k} w_k^{(i)} \mathcal{N}(x; m_k^{(i)}, P_k^{(i)}), \quad (50)$$

where components with indices $i = N_p + 1, \dots, J_k$ are the surviving components. The following bound can be established (from Sorenson and Alspach [5]),

$$\|v_k - v_k^P\|_1 \leq 2 \sum_{i=1}^{N_p} w_k^{(i)} \leq 2N_p \delta_1. \quad (51)$$

This shows that the L_1 error can be selected to fall within specified bounds for the pruning stage of the algorithm.

B. Merging

Several methods for Gaussian mixture reduction using merging techniques have been proposed for Gaussian sum filters. The first of which was derived by Alspach who provided an L_1 error for approximating two Gaussian components with the same covariance as follows [17]. Suppose that two components have the same covariance $P_k := P_k^{(1)} = P_k^{(2)}$, and similar means, $m_k^{(1)} \approx m_k^{(2)}$, so that for some threshold δ_2 ,

$$(m_k^{(1)} - m_k^{(2)})^T P_k^{-1} (m_k^{(1)} - m_k^{(2)}) \leq (\delta_2)^2. \quad (52)$$

Consider approximating $v_k(x)$ by

$$v_k^M(x) = \sum_{i=3}^{J_k} w_k^{(i)} \mathcal{N}(x; m_k^{(i)}, P_k^{(i)}) + \tilde{w}_k^{(l)} \mathcal{N}(x; \tilde{m}_k^{(l)}, P_k), \quad (53)$$

where the weight and mean of the new component are given by

$$\tilde{w}_k^{(l)} = w_k^{(1)} + w_k^{(2)}, \quad (54)$$

$$\tilde{m}_k^{(l)} = \frac{1}{\tilde{w}_k^{(l)}} (w_k^{(1)} m_k^{(1)} + w_k^{(2)} m_k^{(2)}), \quad (55)$$

then the following bound holds,

$$\|v_k - v_k^M\|_1 \leq \frac{2w_k^{(1)}w_k^{(2)}}{w_k^{(1)} + w_k^{(2)}} \delta_2. \quad (56)$$

Note that as the covariance decreases, the distance between the terms must also decrease to retain the same bound. Unfortunately, this requires that both of the covariance matrices are the same which may be an unrealistic assumption.

Salmond proposed two techniques for merging Gaussian components named Joining and Clustering algorithms [18].

In the Joining algorithm, the two components, i and j , which are closest using the distance measure

$$\delta_3^2 = \frac{w_k^{(i)} w_k^{(j)}}{w_k^{(i)} + w_k^{(j)}} (m_k^{(i)} - m_k^{(j)})^T P_k^{-1} (m_k^{(i)} - m_k^{(j)}) \quad (57)$$

are merged, where P_k is the covariance of the entire mixture. It was shown that the minimum distance increases monotonically as the reduction proceeds, and that it is bounded by the dimension of the state space where a threshold is chosen to be a constant fraction of this. In the Clustering algorithm, the Gaussians with the largest weights are chosen as principal components which define cluster centres. The covariance in equation (57) is replaced with the covariance of the principal component, and components in set L within a specified threshold can be merged with the following calculations to preserve the overall covariance of the cluster.

$$\tilde{w}_k^{(\ell)} = \sum_{i \in L} w_k^{(i)}, \quad (58)$$

$$\tilde{m}_k^{(\ell)} = \frac{1}{\tilde{w}_k^{(\ell)}} \sum_{i \in L} w_k^{(i)} m_k^{(i)}, \quad (59)$$

$$\tilde{P}_k^{(\ell)} = \frac{1}{\tilde{w}_k^{(\ell)}} \sum_{i \in L} w_k^{(i)} (P_k^{(i)} + (\tilde{m}_k^{(\ell)} - m_k^{(i)})(\tilde{m}_k^{(\ell)} - m_k^{(i)})^T). \quad (60)$$

This procedure was used in the original formulation of the Gaussian mixture PHD filter [2] and is appropriate since the intensity is multi-modal, where the principal components represent the expected target states.

Williams developed a reduction algorithm which considered the overall change in the probability distribution by evaluating the cost of each possible action and selecting the one which has the minimum effect on the entire mixture in an L_2 sense [19]. The components are merged with equations (58-60) above, which preserves the mixture mean and covariance. This is good for probability distributions but it may not be desirable for intensities as this has the effect of smearing out the modes.

VII. NON-LINEAR TARGET DYNAMIC MODELS

This section considers the convergence for the nonlinear extensions of the Gaussian mixture PHD filter proposed in [2]. As with the linear case, the survival and detection probabilities are assumed constant and the intensities of the birth and spawned target intensities are Gaussian but the state and observation processes can be relaxed to the nonlinear model:

$$x_k = \varphi_k(x_{k-1}, \mathbf{v}_{k-1}), \quad (61)$$

$$z_k = h_k(x_k, \mathbf{e}_k), \quad (62)$$

where φ_k and h_k are known nonlinear functions, \mathbf{v}_{k-1} and \mathbf{e}_k are zero-mean Gaussian process noise and measurement noise with covariances Q_{k-1} and R_k respectively. Due to the nonlinearity of φ_k and h_k , the posterior intensity can no longer be represented as a Gaussian mixture. However, the proposed Gaussian mixture PHD filter can be adapted to accommodate models with mild nonlinearities.

The results here show that the intensity function can be approximated by a set of extended Kalman filters where the covariance of each separate Gaussian component is sufficiently small for the time evolution of its mean and covariance

to be calculated accurately. These are based on the results established for the Gaussian sum filter [6]. In a low noise environment, the EK PHD filter can be nearly optimal. In a high noise environment, it may be necessary to reinitialise the algorithm such that the error covariance of each Gaussian is sufficiently small. If these conditions can not be met, then it may be more appropriate to use the particle PHD filter [8], which can use non-linear dynamic models and non-Gaussian state and observation noises, although this will result in a higher computational complexity.

We now establish the conditions for uniform convergence of the extended Kalman PHD filter. It is shown that, as the covariance term tends to zero, the approximation is optimal. In addition, convergence for the Unscented Kalman PHD filter is discussed.

A. Extended Kalman Prediction Update

Using the PHD prediction equation,

$$v_{k|k-1}(x) = \int \phi_{k|k-1}(x, \zeta) v_{k-1}(\zeta) d\zeta + \gamma_k(x), \quad (63)$$

we show that the predicted intensity for the EK PHD filter can be given by a sum of Gaussians. The extended Kalman prediction tools for existing targets are given by

$$w_{S,k|k-1}^{(j)} = p_{S,k} w_{k-1}^{(j)}, \quad (64)$$

$$m_{S,k|k-1}^{(j)} = \varphi_k(m_{k-1}^{(j)}, 0), \quad (65)$$

$$P_{S,k|k-1}^{(j)} = G_{k-1}^{(j)} Q_{k-1} [G_{k-1}^{(j)}]^T + F_{k-1}^{(j)} P_{k-1}^{(j)} [F_{k-1}^{(j)}]^T, \quad (66)$$

where

$$F_{k-1}^{(j)} = \left. \frac{\partial \varphi_k(x_{k-1}, 0)}{\partial x_{k-1}} \right|_{x_{k-1}=m_{k-1}^{(j)}}, G_{k-1}^{(j)} = \left. \frac{\partial \varphi_k(m_{k-1}^{(j)}, \mathbf{v}_{k-1})}{\partial \mathbf{v}_{k-1}} \right|_{\mathbf{v}_{k-1}=0}. \quad (67)$$

LEMMA 3 If we know the dynamic model, and the posterior intensity at time $k-1$ is given by the sum of Gaussians,

$$v_{k-1}(x) = \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}(x; m_{k-1}^{(i)}, P_{k-1}^{(i)}), \quad (68)$$

then the predicted intensity approaches a sum of Gaussians in L_1 ,

$$v_{k|k-1}^{EK}(x) \rightarrow v_{S,k|k-1}(x) + v_{\beta,k|k-1}(x) + \gamma_k(x), \quad (69)$$

as $P_{k-1}^{(i)} \rightarrow 0$.²

Proof: We assume that we know the birth intensity, γ_k , so that by Theorem 1, we can represent this by a sum of Gaussians as closely as we wish in L_1 ,

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}). \quad (70)$$

For the existing targets, using the intensity at time $k-1$, $v_{k-1}(x)$, and the extended Kalman filter prediction equations, we obtain an approximate expression for the predicted estimate of each Gaussian component, $\mathcal{N}(x; m_{k-1}^{(i)}, P_{k-1}^{(i)})$, to a new

²The *EK* superscript refers to the extended Kalman approximation.

Gaussian component, $\mathcal{N}(x; m_{S,k|k-1}^{(i)}, P_{S,k|k-1}^{(i)})$. Then using the result for the EK Gaussian Sum filter [6], we find that

$$v_{S,k|k-1}^{EK}(x) \rightarrow p_{S,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} \mathcal{N}(x; m_{S,k|k-1}^{(j)}, P_{S,k|k-1}^{(j)}), \quad (71)$$

uniformly in x as $P_{k-1}^{(i)} \rightarrow 0$ for $i = 1, \dots, J_{k-1}$, where

$$m_{S,k|k-1}^{(j)} = \Phi_k(m_{k-1}^{(j)}, 0), \quad (72)$$

$$P_{S,k|k-1}^{(j)} = G_{k-1}^{(j)} Q_{k-1} [G_{k-1}^{(j)}]^T + F_{k-1}^{(j)} P_{k-1}^{(j)} [F_{k-1}^{(j)}]^T. \quad (73)$$

Finally, we come to the predicted intensity for spawned targets, $\beta_{k|k-1}(x|\zeta)$. Using the PHD prediction equations for the EK PHD filter, each of the Gaussian components at time $k-1$ produces $J_{\beta,k}$ Gaussian components,

$$\beta_{k|k-1}(x|m_{k-1}^{(i)}) = \sum_{l=1}^{J_{\beta,k}} w_{\beta,k}^{(l)} \mathcal{N}(x; F_{\beta,k-1}^{(l)} m_{k-1}^{(i)} + d_{\beta,k-1}^{(l)}, Q_{\beta,k-1}^{(l)}). \quad (74)$$

Similar to the result used for the prediction of existing targets, the sum over the J_{k-1} components approaches a Gaussian sum

$$v_{\beta,k|k-1}^{EK}(x) \rightarrow \sum_{j=1}^{J_{k-1}} \sum_{\ell=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(\ell)} \mathcal{N}(x; m_{\beta,k|k-1}^{(j,\ell)}, P_{\beta,k|k-1}^{(j,\ell)}), \quad (75)$$

where

$$m_{\beta,k|k-1}^{(j,\ell)} = F_{\beta,k-1}^{(\ell)} m_{k-1}^{(j)} + d_{\beta,k-1}^{(\ell)}, \quad (76)$$

$$P_{\beta,k|k-1}^{(j,\ell)} = G_{\beta,k-1}^{(j)} Q_{\beta,k-1} [G_{\beta,k-1}^{(j)}]^T + F_{\beta,k-1}^{(j)} P_{\beta,k-1}^{(j)} [F_{\beta,k-1}^{(j)}]^T. \quad (77)$$

B. Extended Kalman Measurement Update

Using the EK PHD filter measurement update equation,

$$v_k(x) = [1 - p_{D,k}(x)] v_{k|k-1}(x) + \sum_{z \in Z_k} \frac{\Psi_{k,z}(x) v_{k|k-1}(z)}{\kappa_k(z) + \int \Psi_{k,z}(\xi) v_{k|k-1}(\xi) d\xi}, \quad (78)$$

we show that the posterior intensity converges to a sum of Gaussians uniformly in L_1 . The PHD update components are given by

$$S_k^{(j)} = U_k^{(j)} R_k [U_k^{(j)}]^T + H_k^{(j)} P_{k|k-1}^{(j)} [H_k^{(j)}]^T, \quad (79)$$

$$K_k^{(j)} = P_{k|k-1}^{(j)} [H_k^{(j)}]^T [S_k^{(j)}]^{-1}, \quad (80)$$

$$P_{k|k}^{(j)} = [I - K_k^{(j)} H_k^{(j)}] P_{k|k-1}^{(j)}, \quad (81)$$

where

$$H_k^{(j)} = \left. \frac{\partial h_k(x_k, 0)}{\partial x_k} \right|_{x_k = m_{k|k-1}^{(j)}}, U_k^{(j)} = \left. \frac{\partial h_k(m_{k|k-1}^{(j)}, \epsilon_k)}{\partial \epsilon_k} \right|_{\epsilon_k = 0}. \quad (82)$$

LEMMA 4 With the non-linear measurement equation $z_k = h_k(x_k, \epsilon_k)$ and the predicted intensity given by the sum of Gaussians,

$$v_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}), \quad (83)$$

the updated density approaches the Gaussian sum

$$v_k^{EK}(x) \rightarrow (1 - p_{D,k}) v_{k|k-1}(x) + \sum_{z \in Z_k} v_{D,k}(x; z), \quad (84)$$

uniformly in x_k and Z_k as $P_{k|k-1}^{(i)} \rightarrow 0$ for $i = 1, \dots, J_{k|k-1}$.

Proof: Clearly the term on the left of the measurement equation, $[1 - p_{D,k}(x)] v_{k|k-1}(x)$, is a Gaussian sum, since the probability of detection $p_{D,k}(x) = p_{D,k}$ is assumed to be a constant and $v_{k|k-1}(x)$ is given by the predicted intensity.

Taking the numerator of the term on the right inside the summation and using the predicted intensity (equation (69)),

$$\Psi_{k,z}(x) v_{k|k-1}(x) = p_{D,k} g_k(z|x) v_{k|k-1}^{EK}(x) \quad (85)$$

$$= p_{D,k} \mathcal{N}(z; h_k(x), R_k) \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}), \quad (86)$$

(which, by Anderson and Moore [6] pp 215-216)

$$\rightarrow p_{D,k} \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(z; h_k(m_{k|k-1}^{(i)}), R_k + H_k^{(i)T} P_{k|k-1}^{(i)} H_k^{(i)}) \mathcal{N}(x; m_{k|k}^{(i)}, P_{k|k}^{(i)}), \quad (87)$$

uniformly as $P_{k|k-1}^{(i)} \rightarrow 0$ for all $i = 1, \dots, J_{k|k-1}$.

Now consider the denominator,

$$\kappa_k(z) + \int \Psi_{k,z}(\xi) v_{k|k-1}(\xi) d\xi. \quad (88)$$

Taking the integral and the L_1 convergence result discussed above, by equation (69),

$$\int \Psi_{k,z}(\xi) v_{k|k-1}(\xi) d\xi \quad (89)$$

$$= p_{D,k} \int \mathcal{N}(z; h_k(\xi), R_k) \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(\xi; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}) d\xi \quad (90)$$

$$\rightarrow p_{D,k} \int \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} q_k^{(j)}(z) \mathcal{N}(\xi; m_{k|k}^{(i)}(\xi), P_{k|k}^{(i)}) d\xi, \quad (91)$$

where

$$q_k^{(j)}(z) = \mathcal{N}(z; h_k(m_{k|k-1}^{(i)}), H_k^{(i)T} P_{k|k-1}^{(i)} H_k^{(i)} + R_k). \quad (92)$$

Changing the order of the summation and integral, this is equal to

$$p_{D,k} \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \int q_k^{(j)}(z) \mathcal{N}(\xi; m_{k|k}^{(i)}(\xi), P_{k|k}^{(i)}) d\xi \quad (93)$$

$$= p_{D,k} \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} q_k^{(j)}(z) \mathcal{N}(x; m_{k|k}^{(i)}, P_{k|k}^{(i)}), \quad (94)$$

so that,

$$v_k(x) \rightarrow (1 - p_{D,k}) v_{k|k-1}(x) \quad (95)$$

$$+ p_{D,k} \sum_{i=1}^{J_{k|k-1}} \frac{w_{k|k-1}^{(i)} q_k^{(j)}(z)}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} q_k^{(l)}(z)} \mathcal{N}(x; m_{k|k}^{(i)}, P_{k|k}^{(i)}),$$

uniformly as $P_{k|k-1}^{(i)} \rightarrow 0$ for all $i = 1, \dots, J_{k|k-1}$, where

$$m_{k|k}^{(j)}(z) = m_{k|k-1}^{(j)} + K_k^{(j)}(z - h_k(m_{k|k-1}^{(j)})). \quad (96)$$

C. The Unscented Kalman PHD Filter

Instead of linearising the model, as is the case with the extended Kalman filter, the unscented Kalman filter [20] approximates the mean and covariance with a set of sigma points using the unscented transform. It can be shown that the predicted mean converges to an estimate which is accurate to a second order, which is more accurate than the estimate given by the extended Kalman filter, and that the predicted covariance converges to the same as that estimated through linearisation using the EKF. In the unscented PHD filter, the unscented transform is applied in the prediction step to each term in the Gaussian mixture and the update step is the same as the Gaussian mixture PHD filter update. The convergence analysis of the UK PHD filter is omitted here, and the interested reader is referred to the work by Julier and Uhlmann [21] for an analysis of the convergence of the unscented Kalman filter.

VIII. CONCLUSIONS

A consequence of Wiener's Theory of Approximation is that density functions can be approximated uniformly with a sum of Gaussians. This result has been used to show that the error for the recently proposed Gaussian mixture PHD filter converges uniformly for each of the steps in the algorithm. Error bounds have been provided for the pruning and merging stages, which are used to reduce the number of Gaussian components, based on those established for the Gaussian sum filter. These results give further theoretical justification for the use of the Gaussian mixture PHD filter in multiple target tracking problems.

Proofs of uniform convergence are also derived for the extended Kalman PHD filter. The accuracy of the unscented Kalman PHD filter is discussed as an extension to the results already established for the unscented Kalman filter.

REFERENCES

- [1] B. Vo and W. K. Ma. A closed-form solution to the Probability Hypothesis Density filter. in *Proc. Int'l Conf. on Information Fusion*, Philadelphia, 2005.
- [2] B. Vo and W. K. Ma. The Gaussian Mixture Probability Hypothesis Density Filter. *IEEE Transactions on Signal Processing*, to appear, 2006.
- [3] D. E. Clark and J. Bell. Convergence Results for the Particle PHD Filter. *IEEE Transactions on Signal Processing*, to appear, 2006.
- [4] J. T.-H. Lo. Finite-dimensional sensor orbits and optimal non-linear filtering. *IEEE Trans. IT*, IT-18(5):583–588, 1972.
- [5] H. W. Sorenson and D. L. Alspach. Recursive Bayesian estimation using Gaussian sum. *Automatica*, 7:465–479, 1971.
- [6] B. D. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, New Jersey, 1979.
- [7] R. Mahler. Multi-target Bayes filtering via first-order multi-target moments. *IEEE Trans. AES*, 39(4):1152–1178, 2003.
- [8] B.-N. Vo, S. Singh, and A. Doucet. Sequential Monte Carlo methods for Multi-target Filtering with Random Finite Sets. *IEEE Trans. Aerospace Elec. Systems*, 41, No.4:1224–1245, 2005.
- [9] A. M. Johansen, S. S. Singh, A. Doucet, and B.-N. Vo. Convergence of the SMC implementation of the PHD filter. *Methodology and Computing in Applied Probability*, to appear., 2006.
- [10] M. Tobias and A. D. Lanterman. Probability Hypothesis Density-based multi-target tracking with bistatic range and Doppler observations. *IEE Radar, Sonar and Navigation, Volume 152, Issue 3*, p. 195-205, 2005.
- [11] D. E. Clark and J. Bell. Bayesian Multiple Target Tracking in Forward Scan Sonar Images Using the PHD Filter. *IEE Radar, Sonar and Navigation, October 2005. Volume 152, Issue 5*, p. 327-334.
- [12] N. Ikoma, T. Uchino, and T. Maeda. Tracking of feature points in image sequence by SMC implementation of PHD filter. *ICE 2004 Annual Conference, 4-6 Aug, 2004. p 1696 - 1701 vol. 2.*
- [13] D. Clark, K. Panta, and B. Vo. The GM-PHD Filter Multiple Target Tracker. *Proc. International Conference on Information Fusion. Florence.*, July 2006.
- [14] D. Clark, B. Vo, and J. Bell. GM-PHD Filter Multi-target Tracking in Sonar Images. *Proc. SPIE Defense and Security Symposium. Orlando, Florida*, April 2006.
- [15] B.N. Vo, S. Singh, and A. Doucet. Sequential Monte Carlo Implementation of the PHD filter for Multi-target Tracking. *Proc. FUSION 2003*, pages 792–799, 2003.
- [16] B. Rynne and M. Youngson. *Linear Functional Analysis*. Springer-Verlag, 2000.
- [17] D. L. Alspach. *A Bayesian Approximation Technique for Estimation and Control of Discrete Time Systems*. PhD thesis, University of California, San Diego, 1970.
- [18] D. Salmond. *Tracking in Uncertain Environments*. PhD thesis, University of Sussex, 1989.
- [19] J. L. Williams. Gaussian mixture reduction for tracking multiple maneuvering targets in clutter. Master's thesis, Air Force Institute of Technology, 2003.
- [20] S. Julier and J. Uhlmann. A new extension of the kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, Orlando, FL.*, 1997.
- [21] S. J. Julier and J. K. Uhlmann. A General Method for Approximating Nonlinear Transformations of Probability Distributions. *Technical Report, RRG, Dept. of Engineering Science, University of Oxford.*, 1996.