Gaussian Mixture Implementations of Probability Hypothesis Density Filters for Non-linear Dynamical Models

Daniel Clark [*], Ba-Tuong Vo [†], Ba-Ngu Vo [‡], and Simon Godsill [§]

Abstract

The Probability Hypothesis Density (PHD) filter is a multiple-target filter for recursively estimating the number of targets and their state vectors from sets of observations. The filter is able to operate in environments with false alarms and missed detections. Two distinct algorithmic implementations of this technique have been developed. The first of which, called the Particle PHD filter, requires clustering techniques to provide target state estimates which can lead to inaccurate estimates and is computationally expensive. The second algorithm, called the Gaussian Mixture PHD (GM-PHD) filter does not require clustering algorithms but is restricted to linear-Gaussian target dynamics, since it uses the Kalman filter to estimate the means and covariances of the Gaussians. This article provides a review of Gaussian filtering techniques for non-linear filtering and shows how these can be incorporated within the Gaussian mixture PHD filters. Finally, we show some simulated results of the different variants.

1 Introduction

The multiple-target tracking framework based on random-sets was proposed to unify the problems of detecting, identifying, classifying and tracking targets within a unified Bayesian paradigm [1]. An optimal multi-target Bayes filter, analogous to the single target Bayes filter, can be derived which propagates a multi-target posterior density in time. The complexity of computing this recursion grows exponentially with the number of targets and is thus not practical for more than a few targets. The Probability Hypothesis Density (PHD) filter was proposed as a practical suboptimal alternative to computing the full multiple-target posterior distribution by propagating the first-order moment statistic [2]. This approach has led to efficient multiple target tracking algorithms for jointly estimating the number of targets and their states from a sequence of observation sets with data association uncertainty, missed detections, false alarms and noisy measurements.

A closed-form solution to the PHD filter was derived under linear assumptions on the system and observation equations and Gaussian process and observation noises, called the Gaussian Mixture PHD filter [3]. The PHD is approximated at each stage with a mixture of Gaussians, where the means and covariances of the Gaussian components are calculated according to the Kalman filter equations, and the weights are calculated according to the PHD filter equations.

It is a common misperception that target trajectories can not be maintained with the PHD filter since in the original formulation methods for target state estimation and track continuity were not explicitly defined [2]. In the Gaussian mixture implementation, the multiple target states are estimated by taking the Gaussian components with highest weights and tracks can be maintained by labelling the Gaussian components [4]. More complex methods for dealing with target resolution uncertainty have been developed using this approach as a basis [5]. These techniques are also directly applicable to other Gaussian mixture intensity based filters such as the Cardinalized PHD filter [6] and to the methods described in this paper. However, since the main focus of the paper is on different Gaussian filtering techniques and not track continuity, we do not exploit these techniques here.

In this paper, we investigate strategies for calculating the Gaussian mixture recursion with non-linear motion and observation models. If the posterior distribution can be approximated by a Gaussian density, then the core objective is to find techniques for Gaussian quadrature. Approaches for this include sigma-point filters such as the unscented Kalman filter, central difference, and divided difference filters. Other approaches which have been studied include Gauss-Hermite quadrature, which computes the integral exactly, and Quasi-Monte filtering, which attempts to deterministically choose sample points to calculate the density more efficiently than standards Monte Carlo integration. These techniques are based on approximating the means and covariances in the Kalman filter recursion. An alternative approach uses Monte Carlo integration and importance sampling to approximate the integrals in the Bayes filter directly, called the Gaussian Particle filter [7]. Based on the Gaussian mixture framework, we compare different strategies for approximating the PHD recursion under non-linear target dynamics.

2 Gaussian Filtering

The single-target filtering problem is to estimate recursively in time, the probability distribution $p(x_k|\mathcal{Z}_k)$ of the signal, where $x_{0:t} := \{x_0, \ldots, x_k\}$ be an unobserved signal process of dimension n that we wish to estimate, and $\mathcal{Z}_k := \sigma(\{z_1, \ldots, z_k\})$ be the σ-algebra generated by noisy observations of dimension $m \leq n$

[*]Daniel Clark is in the Department of Electrical and Computing Engineering at Heriot-Watt University. dec1@hw.ac.uk

[†]Ba Tuong Vo is in the Department of Electrical and Computing Engineering at University of Western Australia. vob@

[‡]Ba-Ngu Vo is in the Department of Electrical, Electronic Engineering at the University of Melbourne. bv@ee.unimelb.edu.au

[§]Simon Godsill is in the Department of Engineering at the University of Cambridge sjg30@cam.ac.uk

related to this process.

Expressed in Bayesian terms, the problem is to estimate recursively in time the posterior distribution, $p_k(x_k|Z_k)$. The posterior distribution is predicted according to the transition density $f_{k|k-1}(x_k|x_{k-1})$,

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

When the new measurement, z_k, has been observed, the posterior is updated with the likelihood $g_k(z_k|x_k)$ according to Bayes' rule,

$$p_k(x_k|Z_k) = \frac{g_k(z_k|x_k)p_{k|k-1}(x_k|Z_{k-1})}{\int g_k(z_k|x_k)p_{k|k-1}(x_k|Z_{k-1})dx_k}. \quad (2)$$

Gaussian filtering is a special case of the nonlinear filtering problem which is restricted to Gaussian distributions with Gaussian state and observation models, so we assume that each target follows a Gaussian dynamical and observation model i.e.

$$f_{k|k-1}(x|\zeta) = \mathcal{N}(x; \varphi_{k-1}(\zeta), Q_{k-1}), \quad (3)$$
$$g_k(z|x) = \mathcal{N}(z; h_k(x), R_k). \quad (4)$$

When functions φ_{k-1} and h_k are linear and the process and observation noises, Q_{k-1} and R_k, are Gaussian, then the optimal estimate is given by the Kalman filter [8], which can be written as the following closed form version [9] of the Bayes filter,

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

$$\mathcal{N}(x; m_k(z), P_{k|k}) = \frac{g_k(z|x)\mathcal{N}(x; m_{k|k-1}, P_{k|k-1})}{\mathcal{N}(z; \hat{z}_{k|k-1}, S_{k|k-1})}, \quad (6)$$

where $\varphi_{k-1}(x) = F_{k-1}x$ and $h_k(x) = H_kx$ here, $\hat{z}_{k|k-1}$ is the predicted measurement and $S_{k|k-1}$ is the innovation covariance.

Since a Gaussian is uniquely characterised by its mean and covariance, it suffices to find the recursion in terms of these two moments. The predicted state estimate and covariance to time k are

$$\hat{x}_{k|k-1} = \int \varphi_{k-1}(x_{k-1})\mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1})dx_{k-1} \quad (7)$$

$$P_{k|k-1} = \int (x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T\mathcal{N}(x_k, \varphi_{k-1}(m_{k-1}), Q_k)dx_k \quad (8)$$

$$= Q_k + \int (\varphi_{k-1}(x_{k-1}) - \hat{x}_{k|k-1}) \quad (9)$$
$$\times(\varphi_{k-1}(x_{k-1}) - \hat{x}_{k|k-1})^T\mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1})dx_{k-1}$$

When a measurement is received, the updated mean $\hat{x}_k$ and covariance P_k are given by

$$\hat{x}_k = E(x_k|Z_k) \quad (10)$$
$$= \int x_kN(x; m_k(z), P_{k|k})dx_k,$$
$$P_k = E((x_k - \hat{x}_k)(x_k - \hat{x}_k)^T|Z_k) \quad (11)$$
$$= \int (x_k - \hat{x}_k)(x_k - \hat{x}_k)^TN(x; m_k(z), P_{k|k})dx_k,$$

which can be computed with

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k(z_k - \hat{z}_{k|k-1}) \quad (12)$$
$$P_k = (I - P_{x,z,k|k-1})P_{k|k-1} \quad (13)$$

where the Kalman gain K_k, predicted measurement $\hat{z}_{k|k-1}$, innovation covariance $S_{k|k-1}$, and covariances $P_{x,z,k|k-1}$ and $P_{z,z,k}$ are given by

$$K_k = P_{x,z,k|k-1}S^{-1}_{k|k-1} \quad (14)$$
$$\hat{z}_{k|k-1} = \int h_k(x_k)\mathcal{N}(x_k\hat{x}_{k|k-1}, P_{k|k-1})dx_k \quad (15)$$
$$S_{k|k-1} = R_k + P_{z,z,k} \quad (16)$$
$$P_{x,z,k|k-1} = E\left((x_k - \hat{x}_{k|k-1})(h_k(x_k) - \hat{z}_{k|k-1})^T\right) \quad (17)$$
$$P_{z,z,k} = E\left((h_k(x_k) - \hat{z}_{k|k-1})(h_k(x_k) - \hat{z}_{k|k-1})^T\right). \quad (18)$$

In the next few subsections, we present different methods of approximating the Gaussian filter with nonlinear dynamical and observation models. The next three subsections are based on methods for calculating the predicted mean and covariance, in equations (7) and (8), and then computing (14-18) in order to compute the mean and covariance update, in equations (12) and (13). The technique presented in the final subsection takes a different approach in that it attempts to approximate the Gaussian Bayes recursion, in equations (5) and (6), directly through Monte Carlo integration and importance sampling.

2.1 Taylor Series Extensions

The extended Kalman filter linearises about the current mean and covariance using Taylor series expansions. The nonlinear functions, φ_{k-1} and h_k can then be expanded in terms of their Taylor series,

$$\varphi_{k-1}(x_k) = \varphi_{k-1}(\hat{x}_{k|k}) + F_k(x_k - \hat{x}_{k|k}) + \ldots \quad (19)$$
$$h_k(x_k) = h_k(\hat{x}_{k|k}) + H_k(x_k - \hat{x}_{k|k}) + \ldots, \quad (20)$$

where F_{k-1} and G_{k-1} are the following partial derivatives required for the state equation,

$$F_{k-1} = \left.\frac{\partial\varphi_{k-1}(x_{k-1})}{\partial x_{k-1}}\right|_{x_{k-1}=\hat{x}_{k-1}}, G_{k-1} = \left.\frac{\partial\varphi_{k-1}(\hat{x}_{k-1}, w_{k-1})}{\partial w_{k-1}}\right|_{w_{k-1}=0}, \quad (21)$$

and similarly H_k and U_k are the partial derivatives required for the measurement equation,

$$H_k = \frac{\partial h_k(x)}{\partial x}_{x=\hat{x}_{k|k-1}}, U_k = \left.\frac{\partial h_k(\hat{x}_{k|k-1}, \varepsilon_k)}{\partial\varepsilon_k}\right|_{\varepsilon_k=0}. \quad (22)$$

Higher order expansions can also be found, such as the truncated or modified second order filter [10], or iterated extended Kalman filters [11].

2.2 Stirling's Interpolation

The divided difference filter is based on polynomial approximation of a non-linear function using Stirling's interpolation [12]. The advantage of this expansion over the Taylor series expansion used in the extended Kalman filter is that it does not require calculation of Jacobians. The first order version is also known as the Central Difference filter [13].

Define $P^{1/2}$ to be the (lower) Cholesky square root matrix of covariance matrix P, so that $P = P^{1/2}(P^{1/2})^T$, and let $h^2 = 3$. The non-linear function φ_{k-1} can be written in the form

$$\varphi_{k-1}(x) = \varphi_{k-1}(P^{1/2}z) =: \tilde{\varphi}_{k-1}(z), \tag{23}$$

which can be approximated using Stirling's interpolation formula,

$$\tilde{\varphi}_{k-1}(\hat{z}) + \tilde{D}_{\Delta z}\tilde{\varphi}_{k-1}(z) + \frac{1}{2}\tilde{D}^2_{\Delta z}\tilde{\varphi}_{k-1}(z) + \ldots, \tag{24}$$

where $\tilde{D}_{\Delta z}$ and $\tilde{D}^2_{\Delta z}$ are the first and second order divided difference operators defined by

$$\tilde{D}_{\Delta z}\tilde{\varphi}_{k-1}(z) = \frac{1}{2h}\sum_{i=1}^{n_x}\Delta z_i\left(\Delta\varphi_{k-1}(\hat{z}+h[P^{1/2}]_i) - \Delta\varphi_{k-1}(\hat{z}-h[P^{1/2}]_i)\right) \tag{25}$$

and

$$\tilde{D}^2_{\Delta z}\tilde{\varphi}_{k-1}(z) = \frac{1}{h^2}\times \sum_{i=1}^{n_x}(\Delta z_i)^2\left(\Delta\varphi_{k-1}(\hat{z}+h[P^{1/2}]_i) - \Delta\varphi_{k-1}(\hat{z}-h[P^{1/2}]_i) - 2\tilde{\varphi}_{k-1}(\hat{z})\right), \tag{26}$$

where $[P^{1/2}]_i$ is the i^{th} column of matrix $P^{1/2}$.

In order to compute equations (14-18), we are required to calculate the divided difference matrices $S^{(1)}_{x_k,\hat{x}_{k-1}}$ and $S^{(1)}_{z_k,\hat{x}_{k|k-1}}$ for the first order divided difference filter (DD1), and matrices $S^{(2)}_{x_k,\hat{x}_{k-1}}$ and $S^{(2)}_{z_k,\hat{x}_{k|k-1}}$ for DD2.

$$[S^{(1)}_{x_k,\hat{x}_{k-1}}]_j = \frac{1}{2h}\left(\varphi_k\left(\hat{x}_k + h[P^{1/2}_{k-1}]_j\right) - \varphi_k\left(\hat{x}_k - h[P^{1/2}_{k-1}]_j\right)\right). \tag{27}$$

We first compute the first order divided difference matrix,

$$[S^{(1)}_{z_k,\hat{x}_{k|k-1}}]_j = \frac{1}{2h}\left(h_k\left(\hat{x}_k + h[P^{1/2}_{k|k-1}]_j\right) - h_k\left(\hat{x}_k - h[P^{1/2}_{k|k-1}]_j\right)\right). \tag{28}$$

and the secord-order one,

$$\begin{aligned} S^{(2)}_{x_k,\hat{x}_{k-1}} =& \frac{\sqrt{h^2-1}}{4h^2}\left(-2\varphi_k(\hat{x}_k)\right. \\ &\left. + \varphi_k\left(\hat{x}_k + h[P^{1/2}_{k-1}]_j\right) - \varphi_k\left(\hat{x}_k - h[P^{1/2}_{k-1}]_j\right)\right) \end{aligned} \tag{29}$$

The second order divided difference matrix needed for the update is

$$\begin{aligned} S^{(2)}_{z_k,\hat{x}_{k|k-1}} =& \frac{\sqrt{h^2-1}}{4h^2}\left(-2h_k\left(\hat{x}_{k|k-1}\right)\right. \\ &\left. + h_k\left(\hat{x}_k + h[P^{1/2}_{k|k-1}]_j\right) - h_k\left(\hat{x}_k - h[P^{1/2}_{k|k-1}]_j\right)\right). \end{aligned} \tag{30}$$

The mean and covariance can be computed with

$$\begin{aligned} \hat{x}_{k|k-1} &= f_k(\hat{x}_{k-1}) + \frac{1}{2h^2}\left(f(\hat{x}_{k-1} + h[P^{1/2}_{k-1}]_j)\right. \\ &\quad \left. + f(\hat{x}_{k-1} - h[P^{1/2}_{k-1}]_j)\right) \end{aligned} \tag{31}$$

$$P_{k|k-1} = S^{(1)}_{x_k,\hat{x}_{k-1}}(S^{(1)}_{x_k,\hat{x}_{k-1}})^T + S^{(2)}_{x_k,\hat{x}_{k-1}}(S^{(2)}_{x_k,\hat{x}_{k-1}})^T + Q_k \tag{32}$$

2.3 Gauss-Hermite Quadrature

The Gauss-Hermite filter [13] uses the Gauss-Hermite quadrature rule to compute the integral

$$\int \varphi_{k-1}(x)\mathcal{N}(x;m,P)dx = \sum_{i=1}^{m} w_i\varphi_{k-1}(x_i), \tag{33}$$

which holds for all polynomials up to degree $2m-1$ with quadrature points x_i. Let J be the symmetric tridiagonal matrix with zero diagonals and $J_{i,i+1} = \sqrt{i/2}$ for $1 \le i \le m-1$. Then x_i are the quadrature points and the weights are $w_i = ([v_i]_1)^2$, where v_i is the i^{th} normalised eigenvector of J. The integral is approximated with

$$I_m = \sum_{i_1}^{m}\cdots\sum_{i_n}^{m} w_{i_1}\cdots w_{i_n}\varphi_{k-1}(q_{i_1},\ldots,q_{i_n}) \tag{34}$$

where $q_i = x_1/\sqrt{2}$ for $1 \le i \le m-1$., which is exact for all polynomials up to degree $2m-1$.

The Unscented Kalman filter [14] is a specific instance of the Gauss-Hermite filter which is exact for all quadratic polynomials. A set of sigma-points are chosen so that their mean and covariance are $\hat{x}_{k-1}$ and P_{k-1}. The non-linear transform is applied to each of the points to obtain a set of transformed points with mean $\hat{x}_{k|k-1}$ and covariance $P_{k|k-1}$. If the state dimension is n, then $M := 2n+1$ sample points are chosen deterministically with

$$x^{(0)}_{k-1} = \hat{x}_{k-1}, \qquad w^{(0)}_{k-1} = \kappa/(n+\kappa), \tag{35}$$

$$x^{(i)}_{k-1} = \hat{x}_{k-1} + \left(\sqrt{(n+\kappa)P_{k-1}}\right)_i, \qquad w^{(i)}_{k-1} = \kappa/2(n+\kappa), \tag{36}$$

$$x^{(i+n)}_{k-1} = \hat{x}_{k-1} - \left(\sqrt{(n+\kappa)P_{k-1}}\right)_i, \qquad w^{(i+n)}_{k-1} = \kappa/2(n+\kappa), \tag{37}$$

where $\kappa \in \mathbb{R}$ and $w^{(i)}_{k-1}$ is the weight associated with the i^{th} sigma point at time $k-1$. Similarly, a set of points are computed for the observation equation h_k.

2.4 Monte Carlo Integration

The Quasi Monte Carlo Kalman filter (QMC-KF) [15] uses Quasi Monte Carlo integration [16] to compute the integral of a non-linear function multiplied by a Gaussian distribution. A deterministic set points is found by finding a mapping ϕ that projects the integration domain to the unit hypercube. Suppose that $\{u^{(j)}_{k-1}, j = 1,\ldots,M\}$ is a quasi-random sequence of vectors on the unit hypercube $[0,1)$. These can be transformed into a quasi-Gaussian sequence $\{x^{(j)}_{k-1}, j = 1,\ldots,M\}$ with mean m_{k-1} and covariance P_{k-1} by performing a Cholesky decomposition, so that $P_{k-1} = S^T_{k-1}S_{k-1}$, transforming sequence $\{u^{(j)}_{k-1}\}$ to $\{y^{(j)}_{k-1}\}$ via $(y^{(j)}_{k-1})_i = \phi^{-1}((u^{(j)}_{k-1})_i)$ for each dimension i, and finally computing $x^{(j)}_{k-1} = m_{k-1} + S_{k-1}y^{(j)}_{k-1}$ for $j = 1,\ldots,M$.

The integral can then be approximated with

$$\int \varphi_{k-1}(x_{k-1})\mathcal{N}(x, m_{k-1}, P_{k-1})dx_{k-1} \approx \frac{1}{M}\sum_{j=1}^{M} \varphi_{k-1}(x_{k-1}^{(j)}). \tag{38}$$

A non-deterministic alternative to the QMC-KF can be found by drawing a set of sample points $\{x_{k-1}^{(j)}, j = 1, \ldots, N\}$ randomly from the Gaussian distribution, so that $x_{k-1}^{(i)} \sim \mathcal{N}(x; m_{k-1}, P_{k-1})$. The integral can be computed with equation (38), where the deteministic points are replaced with the randomly chosen samples.

2.5 Gaussian Particle Filtering

The Gaussian Particle filter [7], takes a different approach to the above Gaussian filters in that it does not approximate all the means and covariances in equations (14-18) but instead uses Monte Carlo integration and importance sampling to approximate the Bayes filter directly. This approach has been extended for more general probability distributions [17] within a Gaussian sum framework, and we have shown that this approach can be extended to intensity functions [18]. Suppose that the Gaussian posterior density, $\mathcal{N}(\xi; m_{k-1}, P_{k-1})$, is projected through nonlinear state function $\varphi_{k-1}(\cdot)$ with process noise Q_{k-1}. Then we are required to calculate

$$\int \mathcal{N}(x; \varphi_{k-1}(\xi), Q_{k-1})\mathcal{N}(\xi; m_{k-1}, P_{k-1})d\xi. \tag{39}$$

Since the state function is nonlinear, there does not exist a closed form solution, so we resort to Monte Carlo integration as follows, For $j = 1, \ldots, M$, sample particles $x_{k-1}^{(j)}$ from $\mathcal{N}(\cdot; m_{k-1}, P_{k-1})$ and $x_k^{(j)}$ from $\mathcal{N}(\cdot; \varphi_{k-1}(x_{k-1}^{(j)}), Q_{k-1})$. Then the integral can be approximated with

$$\frac{1}{M}\sum_{j=1}^{M} \mathcal{N}(\cdot; \varphi_{k-1}(x_{k-1}^{(j)}), Q_{k-1}), \tag{40}$$

which, by the strong law of large numbers, converges almost surely to (39) as $M \to \infty$. Kotecha and Djuric [7] showed that the sample means and covariances converge to the MMSE of the means and covariances.

Importance sampling is performed for the Gaussian prediction density. For $j = 1, ..., M$ draw sample particles $x_k^{(j)}$ from an importance function $\pi_k^{(i)}(\cdot|\mathcal{Z}_{k-1}, z)$ where $j = 1, ..., M$. Compute the weight $\xi_k^{(j)}(z)$ of each sample $x_k^{(j)}$ according to

$$\xi_k^{(j)}(z) = \frac{\mathcal{N}(z; h_k(x_k^{(j)}), R_k)\mathcal{N}(x_k^{(j)}; m_{k|k-1}, P_{k|k-1})}{\pi_k\left(x_k^{(j)}|\mathcal{Z}_{k-1}, z\right)}.$$

Choices for this importance function could include the associated prediction component, $\mathcal{N}(x_{k|k-1}; m_{k|k-1}, P_{k|k-1})$, or use one of the Gaussian filter updated densities given above Gaussian with measurement z_k. The sum of these weights converges almost surely to

$$\frac{1}{M}\sum_{j=1}^{M} \xi_k^{(j)}(z) \to \int \mathcal{N}(z; h_k(x_k), R_k)\mathcal{N}(x_k; m_{k|k-1}, P_{k|k-1})dx_k, \tag{41}$$

as $M \to \infty$, which is used as an approximation to the Gaussian

$$\mathcal{N}(z; \hat{z}_{k|k-1}, S_{k|k-1}). \tag{42}$$

The next section describes the PHD filter methodology and how the Gaussian filters can be implemented within a mixture framework.

3 Random Set Filtering

3.1 Multiple Target Bayes Filtering

The multiple target tracking framework based on random-sets was proposed by Mahler [2] as a mathematically rigorous attempt to unify the problems of detection, classification and tracking. The approach uses a Bayesian paradigm for recursively estimating and updating a multi-target density function based on measurements received at each time-step.

The set of objects tracked at time k is modelled by the point process or Random Finite Set (RFS)

$$X_k = \left(\bigcup_{x\in X_{k-1}} Y_{k|k-1}(x)\right) \cup \left(\bigcup_{x\in X_{k-1}} B_{k|k-1}(x)\right) \cup \Gamma_k. \tag{43}$$

where $Y_{k|k-1}$ is the RFS of targets survived at time t from multi-target state X_{k-1} [1] at time $k-1$, $B_{k|k-1}$ is the RFS of targets spawned from X_{k-1} and Γ_k is the RFS of targets that appear spontaneously at time t. The multi-target measurement at time t is modelled by RFS

$$Z_k = K_k \cup \left(\bigcup_{x\in X_k} \Theta_k(x)\right), \tag{44}$$

where $\Theta_k(X_k)$ is the RFS of measurements from multi-target state X_k and K_k is the RFS of measurements due to clutter.

The optimal multi-target Bayes filter propagates the multi-target posterior density $p_k(\cdot|Z_{1:k})$ conditioned on the sets of observations up to time k, $Z_{1:k}$, with the following recursion

$$p_{k|k-1}(X_k|Z_{1:k-1}) = \int f_{k|k-1}(X_k|X)p_{k-1}(X|Z_{1:k-1})\mu_s(dX), \tag{45}$$

$$p_k(X_k|Z_{1:k}) = \frac{g_k(Z_k|X_k)p_{k|k-1}(X_k|Z_{1:k-1})}{\int g_k(Z_k|X)p_{k|k-1}(X|Z_{1:k-1})\mu_s(dX)}, \tag{46}$$

where the dynamic model is governed by the transition density $f_{k|k-1}(X_k|X_{k-1})$ and multi-target likelihood $g_k(Z_k|X_k)$ and μ_s takes the place of the Lebesgue measure, as described in [19].

The function $g_{k|k}(Z_k|X_k)$ is the joint multi-target likelihood function, or global density, of observing the set of measurements, Z, given the set of target states, X, which is the total probability density of association between measurements in Z and parameters in X. The parameters for this density are the set of observations, $Z = \{z_1, ..., z_k\}$, the unknown set of target

[1]Note that the same notation is used for a random variable and its realization.

states, $X = \{x_1, ..., x_k\}$, the sensor noise distribution, or observation noise, clutter models, and the detection profile of sensor or field of view.

The PHD filter is an approximation developed to alleviate the computational intractability in the multi-target Bayes filter. Instead of propagating the multi-target posterior density in time, the PHD filter propagates the posterior intensity, a first-order statistical moment of the posterior multi-target state [2]. This strategy is reminiscent of the constant gain Kalman filter, which propagates the first moment (the mean) of the single-target state.

For a RFS X on $\mathcal{X}$ with probability distribution P, its first-order moment is a non-negative function v on $\mathcal{X}$, called the *intensity*, such that for each region $S \subseteq \mathcal{X}$ [20]

$$\int |X \cap S| P(dX) = \int_S v(x)dx. \tag{47}$$

In other words, the integral of v over any region S gives the expected number of elements of X that are in S. Hence, the total mass $\hat{N} = \int v(x)dx$ gives the expected number of elements of X. The local maxima of the intensity v are points in $\mathcal{X}$ with the highest local concentration of expected number of elements, and hence can be used to generate estimates for the elements of X. The simplest approach is to round $\hat{N}$ and choose the resulting number of highest peaks from the intensity. The intensity is also known in the tracking literature as the Probability Hypothesis Density (PHD) [2].

3.2 The PHD Recursion

Let v_k and $v_{k|k-1}$ denote the respective intensities associated with the multi-target posterior density p_k and the multi-target predicted density $p_{k|k-1}$ in the recursion (45)-(46).

The prediction equation is given by

$$v_{k|k-1}(x) = \int p_{S,k}(\zeta) f_{k|k-1}(x|\zeta) v_{k-1}(\zeta) d\zeta + \gamma_k(x), \tag{48}$$

where

$$\begin{aligned} f_{k|k-1}(\cdot|\zeta) &= \text{single target transition density at time } k, \\ p_{S,k}(\zeta) &= \text{probability of target existence at time } k, \\ \gamma_k(\cdot) &= \text{intensity of spontaneous births at time } k, \end{aligned}$$

and the update equation is given by

$$\begin{aligned} v_k(x) &= [1 - p_{D,k}(x)] v_{k|k-1}(x) \\ &+ \sum_{z \in Z_k} \frac{p_{D,k}(x) g_k(z|x) v_{k|k-1}(x)}{\kappa_k(z) + \int p_{D,k}(\xi) g_k(z|\xi) v_{k|k-1}(\xi) d\xi} \end{aligned} \tag{49}$$

where

$$\begin{aligned} Z_k &= \text{measurement set at time } k, \\ g_k(\cdot|x) &= \text{single target measurement likelihood at time} k \\ p_{D,k}(x) &= \text{probability of target detection at time } k \\ \kappa_k(\cdot) &= \text{intensity of clutter measurements at time } k, \end{aligned}$$

For simplicity we do not consider spawning in this paper. It is clear from (48)-(49) that the PHD filter completely avoids the combinatorial problem that arises from the unknown association of measurements with appropriate targets. Furthermore, since the posterior intensity is a function on the single-target state space $\mathcal{X}$, the PHD recursion requires much less computational power than the multi-target recursion (45)-(46).

3.3 Gaussian Mixture Implementations

Along with the Gaussian model for individual targets, the multi-target model includes certain assumptions on the birth, death and detection of targets. We omit the labelling of the Gaussian components to maintain target continuity [4] for simplicity of presentation. These are summarized below:

•. The survival and detection probabilities are state independent, i.e.

$$p_{S,k}(x) = p_{S,k}, \tag{50}$$

$$p_{D,k}(x) = p_{D,k}. \tag{51}$$

•. The intensities of the birth RFSs is a Gaussian mixture of the form

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}), \tag{52}$$

where $J_{\gamma,k}$, $w_{\gamma,k}^{(i)}$, $m_{\gamma,k}^{(i)}$, $P_{\gamma,k}^{(i)}$, $i = 1, \ldots, J_{\gamma,k}$, are given model parameters that determine the shape of the birth intensity.

GM-PHD Prediction: Suppose that the posterior intensity at time $k-1$ is a Gaussian mixture of the form

$$v_{k-1}(x) = \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}(x; m_{k-1}^{(i)}, P_{k-1}^{(i)}). \tag{53}$$

Substituting linear transition density $f_{k|k-1}$, posterior intensity v_{k-1} and and birth intensity γ_k into the PHD prediction, we have

$$\begin{aligned} &v_{k|k-1}(x) \\ &= p_{S,k} \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}) + \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \mathcal{N}(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}), \end{aligned} \tag{54}$$

where the means $m_{S,k|k-1}^{(i)}$ and covariances $P_{S,k|k-1}^{(i)}$ are computed with a Gaussian filter prediction.

GM-PHD Update: Suppose that the predicted intensity to time k is a Gaussian mixture of the form

$$v_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}). \tag{55}$$

Substituting the prediction intensity $v_{k|k-1}$ and linear Gaussian observation g_k into the PHD update equation, this simplifies to

$$\begin{aligned} v_k(x) &= (1 - p_{D,k}) v_{k|k-1}(x) \\ &+ \sum_{z \in Z_k} \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) \mathcal{N}(x; m_{k|k}^{(j)}(z), P_{k|k}^{(j)}) \end{aligned} \tag{56}$$

where the weights of the components are given by (57) and the means and covariances are calculated with a Gaussian filter update.

$$w_k^{(j)}(z) = \frac{p_{D,k} w_{k|k-1}^{(j)} q_k^{(j)}(z)}{\kappa_k(z) + p_{D,k} \sum_{\ell=1}^{J_{k|k-1}} w_{k|k-1}^{(\ell)} q_k^{(\ell)}(z)} \tag{57}$$

The updated means $m_k(z)^{(i)}$ and covariances $P_k^{(i)}$ are calculated with the chosen Gaussian filter and $q_k^{(i)}$ is either calculated by finding the predicted measurement, $\hat{z}_{k|k-1}^{(i)}$ and innovation covariance, $S_{k|k-1}^{(i)}$, for each component by approximating equations (15) and (16) or, in the case of the Gaussian Particle filter, by computing the sum of the particle weights in equation (42).

4 Simulations

In this section, we present simulated results for the different Gaussian filters. We consider a standard state space model for a single-speaker Time Difference Of Arrival (TDOA) problem. The notation α_k is used to represent the speaker location at the kth time frame. By defining a state vector $\mathbf{x}_k = [\ \alpha_k^T, \phi_k^T\ ]^T \in \mathbb{R}^n$ where n is the state dimension and ϕ_k contains some kinematic variables for the speaker motion (e.g., velocity), we model $\mathbf{x}_k$ by a dynamic process:

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{w}_k \tag{58}$$

where $\mathbf{A}$ and $\mathbf{B}$ are some pre-specified matrices, and $\mathbf{w}_k$ is a time-uncorrelated random Gaussian vector with zero mean and covariance $\mathbf{I}$. In speaker location tracking, it is popular to employ the Langevin model in which ϕ_k consists of the (x, y) velocities. The state space equations for the Langevin model are given by

$$\alpha_k = \alpha_{k-1} + T\phi_{k-1} \tag{59}$$

$$\phi_k = e^{-\beta T}\phi_{k-1} + \nu\sqrt{1 - e^{-2\beta T}}\mathbf{w}_k \tag{60}$$

Here, β and ν are model parameters called the rate constant and the steady-state root-mean-square velocity, respectively.

Next, we consider the TDOA measurements. We denote by $z_k^{[q]}$ the TDOA measured from the qth microphone pair at time frame k. The measured TDOAs are modelled by:

$$z_k^{[q]} = \tau_q(\mathbf{C}\mathbf{x}_k) + v_k^{[q]}, \quad q = 1, \ldots, Q. \tag{61}$$

Here, $\mathbf{C} = [\ \mathbf{I}\ \mathbf{0}\ ]$ so that $\mathbf{C}\mathbf{x}_k = \alpha_k$,

$$\tau_q(\alpha_k) = \frac{1}{c}(\|\alpha_k - \mathbf{u}_{2,q}\| - \|\alpha_k - \mathbf{u}_{1,q}\|) \tag{62}$$

is the true TDOA value, $\{\mathbf{u}_{1,q}, \mathbf{u}_{2,q}\}$ are the position vectors of the qth microphone pair, and $v_k^{[q]}$ is time-uncorrelated noise. We assume that $v_k^{[q]}$ is independent of $v_k^{[p]}$ for any $q \neq p$, and that each $v_k^{[q]}$ follows a Gaussian distribution with zero mean and variance σ_v^2.

Let $f(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \mathbf{A}\mathbf{x}_{k-1}, \mathbf{B}\mathbf{B}^T)$ and $g_q(z_k^{[q]}|\mathbf{x}_k) = \mathcal{N}(z_k^{[q]}; \tau_q(\mathbf{C}\mathbf{x}_k), \sigma_v^2)$, be the state transition density and the likelihood for the qth speaker respectively. Let $z_{1:k}^{[1:Q]}$ define the sequence containing $z_i^{[q]}$ for $i = 1, \ldots, k$ and for $q = 1, \ldots, Q$.

The room has dimensions $3m \times 3m \times 2.5m$, with four microphone pairs with spacing of $0.5m$. The data has been partitioned into 60 frames of length $256ms$ to measure the TDOA. Two speakers enter and leave the scene at different times. The process and observation noises are $w_k = N([0;0]; diag([0:01;0:01]))$. and $v_k = N(0; 4109)$ respectively. The probabilities of survival and detection are $p_{S,k} = 0.95$ and $p_{D,k} = 0.7$.

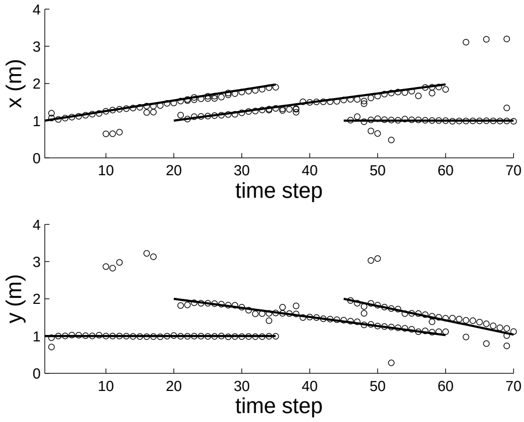

Figure 1: Unscented Kalman GM-PHD filter results

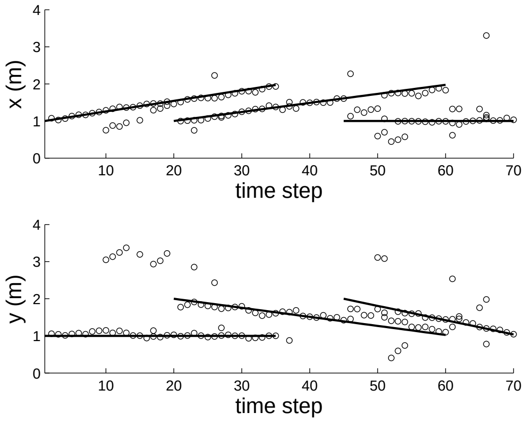

Figure 2: DD2 GM-PHD filter results

We have implemented four of the Gaussian filters for comparison on this non-linear tracking example, the unscented filter,the second-order divided difference filter, the Gaussian particle filter and the Monte Carlo Kalman filter. In the Gaussian particle filter and Monte Carlo Kalman filter, we use 100 samples for each component to ensure that the complexity is not too great. Not all of the Gaussian filters described previously will be appropriate for this model; in particular the Kalman filter is unsuitable and the extended Kalman filter generally has poorer performance than the unscented filter. Figures 1 to 4 show the results of the filters on the same target measurements where circles indicate target state estimates and the solid lines are the true trajectories. By inspection, the unscented filter identified and followed the correct target trajectories better than the others, followed by the Gaussian particle filter, DD2 filter and Monte Carlo filter respectively. Whilst these results may be indicative of performance in this example, it should be stressed that some of the other filters may perform better in different non-linear environments. It should further be noted that the Gaussian mixture implementation of the Cardinalized PHD filter [6] would be expected to outperform the PHD filter due to its improved estimate of the target number.

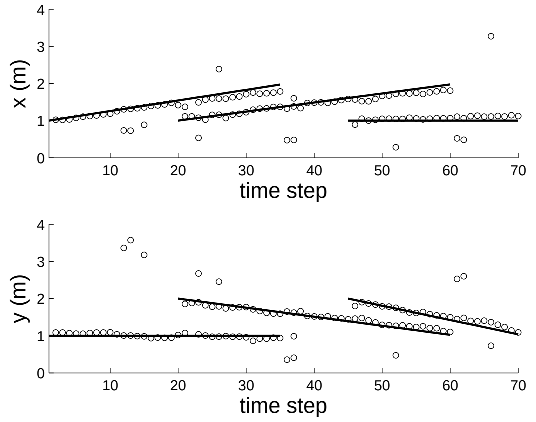

Figure 3: Gaussian Particle GM-PHD filter results

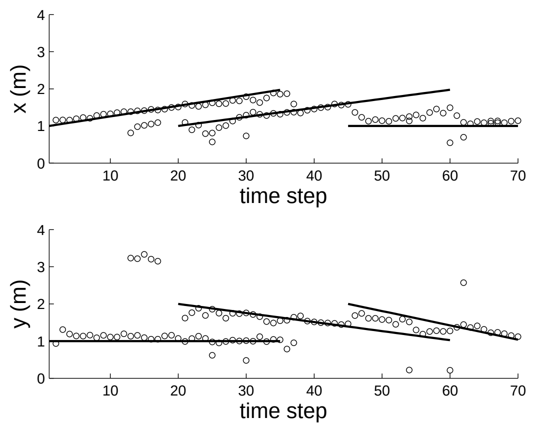

Figure 4: Monte Carlo Kalman GM-PHD filter results

5 Conclusion

This article presents a review of the Gaussian mixture implementations of the Probability Hypothesis Density filters. Although a closed form solution to the Gaussian mixture Probability Hypothesis Density filter only exists for linear-Gaussian dynamical models, some non-linear dynamical models can be accommodated by using Gaussian quadrature techniques. We present a review of these Gaussian filtering techniques, including recent developments in this area. We then show how these can be incorporated within the GM-PHD filter framework and present some simulated results.

Acknowledgments

Daniel Clark wishes to thank the Royal Society of Edinburgh, the Carnegie Trust for the Universities of Scotland and the Royal Academy of Engineering for research travel grants which enabled him to visit Australia to pursue this work. Simon Godsill was supported by the DIF-DTC Tracking Project.

References

[1] R. P. S. Mahler. Statistical Multisource Multitarget Information Fusion. *Artech House*, 2007.

[2] R. Mahler. Multitarget Bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic Systems*, 39, No.4:1152–1178, 2003.

[3] B. Vo and W. K. Ma. The Gaussian Mixture Probability Hypothesis Density Filter. *IEEE Transactions on Signal Processing, Vol 54, No 11 pp. 4091-4104.*, 2006.

[4] D. Clark, K. Panta, and B. Vo. The GM-PHD Filter Multiple Target Tracker. *Proc. International Conference on Information Fusion. Florence.*, July 2006.

[5] K. Panta, D. Clark, and B-N. Vo. Data Association and Track Management for the Gaussian Mixture Probability Hypothesis Density Filter. *Aerospace and Electronic Systems, IEEE Transactions on (to appear)*, 2008.

[6] B. T. Vo, B-N. Vo, and A. Cantoni. Analytic Implementations of Probability Hypothesis Density Filters. *IEEE Transactions on Signal Processing, Vol 55 No 7 Part 2*, pages 3553–3567, 2007.

[7] J. Kotecha and P. Djuric. Gaussian particle filtering. *IEEE Transactions on Signal Processing*, 51, 10:2592 – 2602, 2003.

[8] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[9] Y. C. Ho and R. C. K. Lee. A Bayesian approach to problems in stochastic estimation and control. *IEEE Trans. AC*, AC-9:333–339, 1964.

[10] A. Jazwinski. Stochastic processes and filtering theory. *Academic Press*, 1970.

[11] B. D. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, New Jersey, 1979.

[12] M. Norgaard, N. K. Poulson, and O. Ravn. Advances in Derivative Free State Estimation. *Technical Report: IMM-REP-1998-15, Technical University of Denmark*, 2000.

[13] K. Ito and K. Xiong. Gaussian Filters for Nonlinear Filtering Problems. *IEEE Transactions on Automatic Control, Vol 45 No 5*, pages 910–927, 2000.

[14] 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.

[15] D. Gao and X. Wang. Quasi-Monte Carlo Filtering in Nonlinear Dynamical Systems. *IEEE Transactions on Signal Processing, Vol 54 No 6*, pages 2087–2098, 2006.

[16] P. Fearnhead. Using random quasi-monte carlo within particle filters, with application to financial time series. *J. Computational Graphical Statistics, Vol 14*, pages 751–769, 2005.

[17] J. Kotecha and P. Djuric. Gaussian sum particle filtering. *IEEE Transactions on Signal Processing*, 51, 10:2603 – 2613, 2003.

[18] D. Clark, B-T. Vo, and B-N. Vo. Gaussian Particle Implementations of Probability Hypothesis Density Filters. *IEEE Aerospace Conference, Big Sky, Montana, USA*, 2007.

[19] B. 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.

[20] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes*. Springer, 1988.