Single-cluster PHD filtering and smoothing for SLAM applications

Daniel E. Clark
School of Engineering and Physical Sciences
Heriot-Watt University
Riccarton, Edinburgh EH14 4AS. UK
Email: d.e.clark@hw.ac.uk

Chee Sing Lee and Sharad Nagappa
Computer Vision and Robotics Group (VICOROB)
University of Girona,
Girona, Spain
Email: cslee@eia.udg.edu, sharad.nagappa@udg.edu

***Abstract*—The random finite set formulation for simultaneous localisation and mapping (SLAM) provides a means of estimating the number of landmarks in cluttered environments with missed detections within a unified probabilistic framework. This article develops the random finite set formulation further by introducing forward-backward smoothing. The algorithms are implemented with sequential Monte Carlo and Gaussian mixture techniques, and demonstrated in simulated scenarios. It is shown that when the vehicle closes the loop, the smoother is able to improve the estimated vehicle trajectory.**

I. Introduction

The mathematical foundation for multi-sensor multi-target data fusion was proposed by Mahler as a systematic means of combining evidence in the presence of uncertainty in a unified way using random finite sets [1], [2]. The Probability Hypothesis Density (PHD) filter was proposed as a tractable approximation to the optimal multi-target Bayesian filter [3], [4]. A Sequential Monte Carlo approach to PHD filtering was proposed by Vo, Singh and Doucet [5], and further developments have examined or refined different aspects of the SMC methodology applied to PHD filtering [6]–[9]. A closed-form solution was derived using Gaussian mixture techniques by Vo and Ma [10], which has also led to a number of further developments and applications [11]–[16].

The problem of simultaneously localising the position of a vehicle and mapping its environment (known as SLAM) has received a great deal of interest in robotics since its proposal by Smith and Cheeseman [17]. Typically, this involves propagating a joint posterior distribution of the vehicle position and landmarks, using data association strategies from the target tracking literature for managing measurements fed to the filter and the number of objects. The random finite set approach to SLAM was proposed by Mullane, Vo and Adams with a Rao-Blackwellised approach [18]. Lee, Clark and Salvi [19] investigated the random finite set approach for SLAM using single-cluster Poisson point processes, with fewer approximations in the filter. The single-cluster PHD filter [20] was originally proposed as a special case of a multi-group multi-object filter which has been developed for tracking groups of targets [21], and extended objects [22]. In addition to the work on SLAM [19], the approach has been also developed as a unified approach for jointly tracking a variable number of targets and sensor registration [23].

In this paper, we develop the single-cluster point process approach further by introducing forward-backward smoothing. The PHD smoother, based on Kitagawa's forward-backward smoother [24], was first hypothesised by Nandakumaran *et al.* [25], and then confirmed with the derivation by Mahler *et al.* [26], [27] using Finite Set Statistics. A simpler derivation of the PHD smoother was proposed by Clark [28], and it has also been investigated by Hernandez [29]. Practical implementations include a sequential Monte Carlo version [27], and a closed form solution with Gaussian mixture models [30], [31]. A full Bayesian forward-backward smoother has been developed for a simpler model, where there is at most one target [32]–[34].

The paper is structured as follows. In the next section, we review some concepts in point process theory. In section 3, we describe Bayesian filtering and forward-backward smoothing for a single-cluster process and the Probability Hypothesis Density approximations. In section 4, we describe the implementation. Section 5 presents simulated results. We conclude in section 6.

II. Point processes

Point processes are used to describe multi-object systems with uncertainty in both the number of objects and the locations of the objects. These can be conveniently described with the probability generating functional, described next.

A. Probability generating functionals (p.g.fls.)

The probability generating functional (p.g.fl.) of a point process is defined with

$$G(h) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\prod_{i=1}^{n} \int dx_i h(x_i) \right) p\left(x_1, \ldots, x_n\right), \qquad (1)$$

where $p(x_1, \ldots, x_n) \geq 0$ are known as Janossy densities, and $G(1) = 1$. The first-order factorial moment density, known in the tracking community as the Probability Hypothesis Density

(PHD), is

$$D(x) = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int p(x, x_2 \ldots, x_n) dx_2 \ldots dx_n \quad (2)$$

1) Example: Poisson point process: One of the most simple point processes is the Poisson point process, since it can be uniquely characterised by its intensity, or PHD,

$$D(x) = \mu p(x), \quad (3)$$

where μ is the mean number of objects, each independently and identically distributed according to single-object spatial distribution $p(x)$. The n^{th}-order Janossy density is found with

$$p(x_1, \ldots, x_n) = \exp(-\mu) \prod_{i=1}^{n} \mu p(x_i). \quad (4)$$

The p.g.fl. of the Poisson point process is

$$G(h) = \exp\left(\mu(p[h] - 1)\right), \quad (5)$$

where we define functional

$$p[h] = \int p(x)h(x)dx. \quad (6)$$

2) Example: Single-cluster Poisson process: For SLAM applications, we are interested in estimating the joint state of the vehicle, as well as an indeterminate number of landmarks, or objects. In this case, it is more appropriate to consider a point process that models this explicitly. We consider a hierarchical point process, with a single-object process, $s(\cdot)$, describing the vehicle state, and a conditional multi-object process describing the positions of the landmarks. This can be modelled by a cluster process [35]. The single-cluster point process has probability generating functional

$$G(h) = s\left[G_d(h|\cdot)\right] \quad (7)$$

where h is a function of both inner functional G_d, and outer functional $s[\cdot]$ is defined in equation (6). For a Poisson inner functional, this becomes

$$G(h) = s\left[\exp\left(\mu(p[h|\cdot] - 1)\right)\right] \quad (8)$$

The PHD $D(c, x)$ of a single-cluster Poisson process is given by

$$D(c, x) = s(c)D(x|c), \quad (9)$$

where, in the SLAM application, $s(c)$ is the posterior of the vehicle state, $D(x|c) = \mu(c)p(x|c)$ is the PHD of the landmarks conditioned on the vehicle state, and $\mu(c)$ is the expected number of landmarks conditioned on vehicle position c, and each landmark is distributed according to $p(x|c)$.

In the following sections, for compactness of notation, we write

$$p(c, x_1, \ldots, x_n) = p(\mathbb{X}) = s(\mathbf{X})p(\mathbf{M}|\mathbf{X}), \quad (10)$$

where $\mathbb{X} = (\mathbf{X}, \mathbf{M})$, where $\mathbf{X} = c$ is the vehicle position, and $\mathbf{M}$ is the map of features, $\{x_1, \ldots, x_n\}$.

III. Single-cluster multi-object Bayesian filtering and smoothing

Let $\mathbf{X}_k$ be the random vector that represents the vehicle state, and $\mathbf{M}_k$ be the Random Finite Set (RFS) that represents the location of map features, which exist in the space $\mathcal{X} \subseteq \mathbb{R}^{n_m}$.

$$\mathbf{X}_k = [x_{k,1} \ldots x_{k,n_x}] \quad (11)$$

$$\mathbf{M}_k = \{\mathbf{m}_{k,1} \ldots \mathbf{m}_{k,\nu_k}\} \in \mathcal{F}(\mathcal{X}) \quad (12)$$

Where $\mathcal{F}(\mathcal{X})$ is the set of all finite subsets of $\mathcal{X}$. In addition, the vehicle receives measurements which are represented as an RFS $\mathbf{Z}_k$ taken from the measurement space $\mathcal{Z} \subseteq \mathbb{R}^{n_z}$.

$$\mathbf{Z}_k = \{\mathbf{z}_{k,1} \ldots \mathbf{z}_{k,\mu_k}\} \in \mathcal{F}(\mathcal{Z}) \quad (13)$$

This RFS is the union of measurements that originate from true targets, and measurements generated by a Poisson false alarm process whose PHD is $\kappa(\mathbf{z}) = \lambda U(\mathcal{Z})$, where λ is the Poisson rate parameter and $U(\mathcal{Z})$ is the uniform distribution over $\mathcal{Z}$. The measurement model $\mathbf{z} = h(\mathbf{m}, \mathbf{X})$ relates the measurements to landmark locations and the vehicle pose.

Let $p_k(\mathbb{X}_k) = p_k(\mathbf{X}_k, \mathbf{M}_k)$ be the joint posterior probability distribution of the vehicle state $\mathbf{X}_k$ and multi-object map state $\mathbf{M}_k$ at time step k. From a methodological perspective, the estimation problem to solve remains the same as other SLAM formulations: the sequential Bayesian estimation of $p_k(\mathbb{X}_k)$. For brevity, let $\mathbb{X} = \mathbb{X}_k$ and $\mathbb{X}' = \mathbb{X}_{k-1}$. In the following sections we describe the time-evolution, data update, and smoothing equations for the PHD.

A. Time-update

The Chapman-Kolmogorov equation describes the evolution of stochastic processes over time, i.e.

$$p_{k|k-1}(\mathbb{X}|\mathbf{Z}_{1:k-1}) = \int \pi_{k|k-1}(\mathbb{X}|\mathbb{X}')p_{k-1}(\mathbb{X}')\,\delta\mathbb{X}' \quad (14)$$

where $\pi_{k|k-1}(x|x')$ is a Markov transition from time-step $k-1$ to time-step k, $p_{k-1}(x|z_{1:k-1})$ is the posterior at time-step $k-1$, and $p_{k|k-1}(x|z_{1:k-1})$ is the predicted process to time-step k. Lemma 1 describes the evolution of the parent state and conditional daughter state posteriors. This is then used to determine the evolution of the single-cluster PHD equations in Lemma 2.

Lemma 1: Suppose that the Markov transition for the cluster process factorises as

$$\pi_{k|k-1}(\mathbb{X}|\mathbb{X}') = \pi_{k|k-1}(\mathbf{X}|\mathbf{X}')\pi_{k|k-1}(\mathbf{M}|\mathbf{M}') \quad (15)$$

Then the Chapman-Kolmogorov equations for the parent and conditional daughter states are

$$p_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1}) = \int \pi_{k|k-1}(\mathbf{X}|\mathbf{X}')p_{k-1}(\mathbf{X}')dX' \quad (16)$$

$$p_{k|k-1}(\mathbf{M}|\mathbf{X}, \mathbf{Z}_{1:k-1}) = \int \pi_{k|k-1}(\mathbf{M}|\mathbf{M}')p_{k-1}(\mathbf{M}'|\mathbf{X})\delta\mathbf{M}' \quad (17)$$

Proof: Substituting the Markov transition into the Chapman-Kolmogorov equation gives

$$p_{k|k-1}(\mathbb{X}|\mathbf{Z}_{1:k-1}) \tag{18}$$
$$= \int \pi_{k|k-1}(\mathbf{X}|\mathbf{X}')p_{k-1}(\mathbf{X}')p_{k|k-1}(\mathbf{M}|\mathbf{X}',\mathbf{Z}_{1:k-1})d\mathbf{X}',$$

where the conditional daughter process $p_{k|k-1}(\mathbf{M}|\mathbf{X},\mathbf{Z}_{1:k-1})$ evolves according to

$$p_{k|k-1}(\mathbf{M}|\mathbf{X},\mathbf{Z}_{1:k-1}) = \int \pi_{k|k-1}(\mathbf{M}|\mathbf{M}')p_{k-1}(\mathbf{M}'|\mathbf{X},\mathbf{Z}_{1:k-1})\,\delta\mathbf{M}'. \tag{19}$$

Using the fact that

$$p_{k|k-1}(\mathbb{X}|\mathbf{Z}_{1:k-1}) = p_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1})p_{k|k-1}(\mathbf{M}|\mathbf{X},\mathbf{Z}_{1:k-1}), \tag{20}$$

we get the desired result. ■

Lemma 2: The predicted parent and daughter processes are found with

$$s_{k|k-1}(\mathbf{X}) = \int s_{k-1}(\mathbf{X}')\pi_{k|k-1}(\mathbf{X}|\mathbf{X}')d\mathbf{X}' \tag{21}$$

$$\tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X}') = \gamma_{k|k-1}(\mathbf{m}|\mathbf{X}) + \int D_{k-1}(\mathbf{m}'|\mathbf{X}')\tilde{\pi}_{k|k-1}(\mathbf{m}|\mathbf{m}')\,d\mathbf{m}' \tag{22}$$

$\pi_{k|k-1}(\mathbf{X}|\mathbf{X}')$ is the Markov transition density for the parent process, and $\tilde{\pi}_{k|k-1}(\mathbf{m}|\mathbf{m}')$ is the conditional Markov transition density for the daughter process; $\gamma_{k|k-1}(\mathbf{m}|\mathbf{X})$ is the PHD for the daughter birth process.

Proof: The proof is a special case of the multi-group multi-object prediction presented in [21], where there is at exactly one parent point. ■

B. Measurement update

When new measurements are received, the predicted distribution can be updated with Bayes' rule, i.e.

$$p_k(\mathbb{X}|\mathbf{Z}_{1:k}) = \frac{g_k(\mathbf{Z}_k|\mathbb{X})p_{k|k-1}(\mathbb{X}|\mathbf{Z}_{1:k-1})}{\int g_k(\mathbf{Z}_k|\mathbb{X})p_{k|k-1}(\mathbb{X}|\mathbf{Z}_{1:k-1})\,\delta\mathbb{X}} \tag{23}$$

Lemma 3 describes the updated parent state and conditional daughter state. This is then used in Lemma 4 to determine the updated single-cluster PHD equations.

Lemma 3: The parent state and conditional daughter process are updated with

$$p_k(\mathbf{X}|\mathbf{Z}_{1:k}) = \frac{L_{\mathbf{Z}_k}(\mathbf{X})p_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1})}{\int L_{\mathbf{Z}_k}(\mathbf{X})p_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1})\,d\mathbf{X}} \tag{24}$$

$$p_k(\mathbf{M}|\mathbf{X},\mathbf{Z}_{1:k}) = \frac{g_k(\mathbf{Z}_k|\mathbf{M})p_{k|k-1}(\mathbf{M}|\mathbf{X},\mathbf{Z}_{1:k-1})}{L_{\mathbf{Z}_k}(\mathbf{X})} \tag{25}$$

where

$$L_{\mathbf{Z}_k}(\mathbf{X}) = \int g_k(\mathbf{Z}_k|\mathbf{M})p_{k|k-1}(\mathbf{M}|\mathbf{X},\mathbf{Z}_{1:k-1})\,\delta\mathbf{M}. \tag{26}$$

Proof: The proof follows by noting that

$$p_k(\mathbb{X}|\mathbb{Z}_{1:k}) = p_k(\mathbf{X}|\mathbf{Z}_{1:k})p_k(\mathbf{M}|\mathbf{X},\mathbf{Z}_{1:k}). \tag{27}$$

■

Lemma 4: Let us assume that the predicted posterior is a single-cluster Poisson process. Then the single-cluster PHD update can be separated into a parent update and a daughter update as follows.

$$s_k(\mathbf{X}) = \frac{L_{\mathbf{Z}_k}(\mathbf{X})s_{k|k-1}(\mathbf{X})}{\int L_{\mathbf{Z}_k}(\mathbf{X})s_{k|k-1}(\mathbf{X})\,d\mathbf{X}} \tag{28}$$

$$\tilde{D}_k(\mathbf{m}|\mathbf{X}) = (1-p_D(\mathbf{m}|\mathbf{X}))\tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X}) + \sum_{\mathbf{z}\in\mathbf{Z}_k}\frac{p_D(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X})g_k(\mathbf{z}|\mathbf{m})}{\kappa_k(\mathbf{z}) + \int p_D(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X})g_k(\mathbf{z}|\mathbf{m})\,d\mathbf{m}} \tag{29}$$

Where $g_k(\mathbf{z}|\mathbf{m})$ is the single-object measurement likelihood, and $L_{\mathbf{Z}_k}(\mathbf{X})$ is the multi-object measurement likelihood, both conditional on the vehicle state. The multi-object likelihood is defined as

$$L_{\mathbf{Z}_k}(\mathbf{X}) = \exp\left\{-\int p_D(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X})\,d\mathbf{m}\right\}\times \prod_{\mathbf{z}\in\mathbf{Z}_k}\left(\kappa_k(\mathbf{z}) + \int p_D(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X})g_k(\mathbf{z}|\mathbf{m})\,d\mathbf{m}\right) \tag{30}$$

Proof: The multi-object likelihood $L_{\mathbf{Z}_k}(\mathbf{X})$ follows from equation (121) in [3]. The conditional PHD follows from equation (123) in [3]. ■

C. Forward-backward smoother

The forward-backward smoother [24] is used to refine state estimates in the past based on current measurements. The posterior at time-step k based on measurement sets $Z_{1:k'} = Z_1,\ldots,Z_k'$, where $k' > k$ is given by equations

$$p_{k|k'}(\mathbb{X}|\mathbf{Z}_{1:k'}) = \int p_{k|k+1}(\mathbb{X}|\mathbb{Y},\mathbf{Z}_{1:k})p_{k+1|k'}(\mathbb{Y}|\mathbf{Z}_{1:k'})\,\delta\mathbb{Y} \tag{31}$$

$$p_{k|k+1}(\mathbb{X}|\mathbb{Y},\mathbf{Z}_{1:k}) = \frac{\pi_{k+1|k}(\mathbb{Y}|\mathbb{X})p_{k|k}(\mathbb{X}|\mathbf{Z}_{1:k})}{\int \pi_{k+1|k}(\mathbb{Y}|\mathbb{X}')p_{k|k}(\mathbb{X}'|\mathbf{Z}_{1:k})\,\delta\mathbb{X}'} \tag{32}$$

Lemma 5 factorises equation (31) into parent and conditional daughter posteriors. Making Poisson approximation, this is then applied in Theorem 1 to find the smoothed single-cluster PHD equations.

Lemma 5: The parent state and conditional daughter process are updated with

$$p_{k|k+1}(\mathbf{X}|\mathbb{Y},\mathbf{Z}_{1:k}) = \frac{\pi_{k+1|k}(\mathbf{Y}|\mathbf{X})L_{\mathbf{M_Y}}(\mathbf{X})p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})}{\int \pi_{k+1|k}(\mathbf{Y}|\mathbf{X})L_{\mathbf{M_Y}}(\mathbf{X})p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})\,d\mathbf{X}} \tag{33}$$

and

$$p_{k|k+1}(\mathbf{M_X}|\mathbf{X},\mathbb{Y},\mathbf{Z}_{1:k}) = \frac{\pi_{k+1|k}(\mathbf{M_Y}|\mathbf{M_X})p_{k|k}(\mathbf{M_X}|\mathbf{X},\mathbf{Z}_{1:k})}{L_{\mathbf{M_Y}}(\mathbf{X})}, \tag{34}$$

where

$$L_{\mathbf{M_Y}}(\mathbf{X}) = \int \pi_{k+1|k}(\mathbf{M_Y}|\mathbf{M_X}) p_{k|k}(\mathbf{M_X}|\mathbf{X}, \mathbf{Z}_{1:k})\, \delta\mathbf{M_X}. \tag{35}$$

Proof: The result follows from noting that

$$p_{k|k+1}(\mathbb{X}|\mathbb{Y}, \mathbf{Z}_{1:k}) = p_{k|k+1}(\mathbf{X}|\mathbb{Y}, \mathbf{Z}_{1:k}) p_{k|k+1}(\mathbf{M_X}|\mathbf{X}, \mathbb{Y}, \mathbf{Z}_{1:k}), \tag{36}$$

and applying Lemma 3. ■

Theorem 1: Suppose that $p_{k|k}(\mathbf{M_X}|\mathbf{X}, \mathbf{Z}_{1:k})$ is Poisson and that equation (33) can be approximated with

$$p_{k|k+1}(\mathbf{X}|\mathbb{Y}, \mathbf{Z}_{1:k}) = \frac{\pi_{k+1|k}(\mathbf{Y}|\mathbf{X}) p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})}{\int \pi_{k+1|k}(\mathbf{Y}|\mathbf{X}) p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})\, d\mathbf{X}}. \tag{37}$$

Then the smoothed single-cluster PHD approximation to equation (31) becomes

$$s_{k|k'}(\mathbf{X}) = \int s_{k|k+1}(\mathbf{X}|\mathbf{Y}) s_{k+1|k'}(\mathbf{Y}) d\mathbf{Y} \tag{38}$$

$$D_{k|k'}(\mathbf{m}|\mathbf{X}) = (1 - p_S(\mathbf{m}|\mathbf{X}))\tilde{D}_{k|k}(\mathbf{m}|\mathbf{X}) + \int \frac{p_S(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k}(\mathbf{m}|\mathbf{X})\pi_{k+1|k}(\mathbf{Y}|\mathbf{m})}{\gamma_{k+1}(\mathbf{Y}) + \int p_S(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k}(\mathbf{m}|\mathbf{X})\pi_{k+1|k}(\mathbf{Y}|\mathbf{m})\, d\mathbf{m}} d\mathbf{Y} \tag{39}$$

Proof: The proof follows by applying Lemma 4, and then Campbell's theorem (see, for example, equation (11) in [28]) in equation (31). So, applying Lemma 4, to equation (33), we have the single-cluster PHD update separated into a parent update and a daughter update as follows.

$$s_{k|k+1}(\mathbf{X}|\mathbf{Y}) = \frac{\pi_{k+1|k}(\mathbf{Y}|\mathbf{X}) s_{k|k}(\mathbf{X})}{\int \pi_{k+1|k}(\mathbf{Y}|\mathbf{X}) s_{k|k}(\mathbf{Y})\, d\mathbf{Y}} \tag{40}$$

$$\tilde{D}_{k|k+1}(\mathbf{m}|\mathbf{X}) = (1 - p_S(\mathbf{m}|\mathbf{X}))\tilde{D}_{k|k}(\mathbf{m}|\mathbf{X}) + \sum_{\mathbf{y}\in\mathbf{M_Y}} \frac{p_S(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k}(\mathbf{m}|\mathbf{X})\pi_{k+1|k}(\mathbf{y}|\mathbf{m})}{\gamma_{k+1}(\mathbf{y}) + \int p_S(\mathbf{m}|\mathbf{X})\tilde{D}_{k|k}(\mathbf{m}|\mathbf{X})\pi_{k+1|k}(\mathbf{y}|\mathbf{m})\, d\mathbf{m}} \tag{41}$$

Applying Campbell's theorem gives the required result. ■

IV. Implementation

We implement the Single-Cluster PHD filter using a Dirac mixture model for the PHD of the parent. Each component of the parent mixture model is associated with a Gaussian mixture model which represents the PHD of the daughter process conditioned on that particular parent component. At each iteration of the filter, we begin with the following prior PHDs:

$$s_{k-1}(\mathbf{X}) = \sum_{i=1}^{N_{k-1}} \eta_{k-1}^{(i)} \delta(\mathbf{X} - \mathbf{X}_{k-1}^{(i)}) \tag{42}$$

$$D_{k-1}^{(i)}(\mathbf{m}|\mathbf{X}) = \sum_{j=1}^{J_{k-1}^{(i)}} w_{k-1}^{(j|i)} \mathcal{N}(\mathbf{m}; \mu_{k-1}^{(j|i)}, \mathbf{P}_{k-1}^{(j|i)}) \tag{43}$$

The notation $\mathcal{N}(m; \mu, \mathbf{P})$ is used to denote a Gaussian distribution with mean vector μ and covariance matrix $\mathbf{P}$, and $\delta(\mathbf{X} - \alpha)$ denotes the Dirac delta distribution centred at α. Mixture models may be conveniently represented by the set of their parameters, so we may alternatively express the prior PHDs like so:

$$s_{k-1}(\mathbf{X}) = \{\eta_{k-1}^{(i)}, \mathbf{X}_{k-1}^{(i)}\}_{i=1}^{N_{k-1}} \tag{44}$$

$$D_{k-1}^{(i)}(\mathbf{m}|\mathbf{X}) = \{w_{k-1}^{(j|i)}, \mu_{k-1}^{(j|i)}, \mathbf{P}_{k-1}^{(j|i)}\}_{j=1}^{J_{k-1}^{(i)}} \tag{45}$$

A. Map Prediction

The prediction for the daughter process is the same as that for the standard Gaussian mixture PHD filter. However, in accordance to the measurement-driven birth density proposed in [8], [9], the addition of the birth PHD $\gamma_{k|k-1}(\mathbf{m}|\mathbf{X})$ is postponed until the update step. Hence, the predicted map PHD consists only of the propagation of features which survive from the previous time step.

$$\tilde{D}_{k|k-1}^{(i)}(\mathbf{m}|\mathbf{X}_{k-1}^{(i)}) = p_S(\mathbf{m}|\mathbf{X}_{k-1}^{(i)}) \times \sum_{j=1}^{J_{S,k|k-1}^{(i)}} w_{S,k|k-1}^{(j|i)} \mathcal{N}(\mathbf{m}; \mu_{S,k|k-1}^{(j|i)}, \mathbf{P}_{S,k|k-1}^{(j|i)}|\mathbf{X}) \tag{46}$$

$$w_{S,k|k-1}^{(j|i)} = w_{k-1}^{(j|i)} \tag{47}$$

$$\mu_{S,k|k-1}^{(j|i)} = f(\mu_{k-1}^{(j|i)}|\mathbf{X}) \tag{48}$$

$$\mathbf{P}_{S,k|k-1}^{(j|i)} = \mathbf{F}_k \mathbf{P}_{k-1}^{(j|i)} \mathbf{F}_k^T + \mathbf{W}_k \mathbf{Q}_k \mathbf{W}_k^T \tag{49}$$

$$\mathbf{F}_k^{(j|i)} = \frac{\partial}{\partial \mathbf{m}} f(\mathbf{m}, \mathbf{X}) \Big|_{\mathbf{m}=\mu_{k-1}^{(j|i)}, \mathbf{X}=\mathbf{X}_{k-1}^{(i)}} \tag{50}$$

Here, $f(\mu|\mathbf{X})$ is the function that models the evolution of map features, conditional on the vehicle state.

B. Vehicle Prediction

The sampling property of the Dirac delta function means that substitution of (42) and (43) into (21) results in the following sum:

$$D_{k|k-1}(\mathbf{X}, \mathbf{m}) = \sum_{i=0}^{N_{k-1}} \pi_{k|k-1}(\mathbf{X}|\mathbf{X}_{k-1}^{(i)}) \tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X}_{k-1}^{(i)}) \tag{51}$$

The Markov transition density is then approximated by sampling M particles from it. The result is a new Dirac mixture for the parent, containing $N_{k-1} \times M$ components. Because the predicted maps $\tilde{D}_{k|k-1}(\mathbf{m}|\mathbf{X}_i')$ do not depend on the current predicted vehicle pose, each of the M parent components that originate from the same $\mathbf{X}_{k-1}^{(i)}$ can be assigned identical copies of the predicted map.

C. Measurement Update

Like the prediction, the measurement update for the daughter process mirrors that of the standard GM-PHD filter, modified by the measurement-driven birth density. First, the birth terms are constructed from the current measurements:

$$\gamma_{k|k-1}(\mathbf{m}|\mathbf{X}^{(i)}_{k|k-1}) = \sum_{j=1}^{J_{k|k-1,b}=|\mathbf{Z}_k|} w_b \mathcal{N}(\mathbf{m}; \mu^{(j|i)}_{k|k-1,b}; \mathbf{P}^{(j|i)}_{k|k-1,b}) \quad (52)$$

$$\mu^{(j|i)}_{k|k-1,b} = h^{-1}(\mathbf{z}_j, \mathbf{X}^{(i)}_{k|k-1}) \quad (53)$$

$$\mathbf{P}^{(j|i)}_{k|k-1,b} = \mathbf{J}^{(j)}_{k,b}\mathbf{R}\mathbf{J}^{(j),T}_{k,b} \quad (54)$$

$$\mathbf{J}^{(j)}_{k,b} = \frac{\partial}{\partial \mathbf{z}} h^{-1}(\mathbf{z}, \mathbf{X})\bigg|_{\mathbf{z}=\mathbf{z}_j, \mathbf{X}=\mathbf{X}^{(i)}_{k|k-1}} \quad (55)$$

Here, $h^{-1}(\mathbf{z}, \mathbf{X})$ is the inverse measurement model, relating a measurement and vehicle position to a feature state. The weight of birth components w_b is a parameter to be specified. The birth density is combined with the predicted PHD and a GM-PHD update is executed as normal, except that the birth terms have a probability of detection $p_D = 1$.

$$\tilde{D}^{(i)}_k(\mathbf{m}|\mathbf{X}^{(i)}_{k-1}) = (1 - p_D(\mathbf{m}|\mathbf{X}^{(i)}_{k|k-1}))\tilde{D}^{(i)}_{k|k-1}(\mathbf{m}|\mathbf{X}^{(i)}_{k-1}) + \sum_{\mathbf{z}\in\mathbf{Z}_k} \tilde{D}^{(i)}_{D,k}(\mathbf{m}|\mathbf{X}^{(i)}_{k-1}) \quad (56)$$

where

$$\tilde{D}^{(i)}_{D,k}(\mathbf{m}|\mathbf{X}^{(i)}_{k-1}) = \left[\sum_{j=1}^{J_{k|k-1}} p_D(\mu^{(j|i)}_{k|k-1}|\mathbf{X}^{(i)}_{k|k-1}) w^{(j|i)}_k \mathcal{N}(\mathbf{m}; \mu^{(j|i)}_k; \mathbf{P}^{(j|i)}_k) + \sum_{j=1}^{J_{k|k-1,b}} w^{(j|i)}_{k,b}\mathcal{N}(\mathbf{m}; \mu^{(j|i)}_{k,b}; \mathbf{P}^{(j|i)}_{k,b})\right] \Big/ L_{\mathbf{z}} \quad (57)$$

$$L_{\mathbf{z}} = \kappa_k(\mathbf{z}) + \sum_{l=1}^{J_{k|k-1}} p_D(\mu^{(l|i)}_{k|k-1}|\mathbf{X}^{(i)}_{k|k-1}) w^{(l|i)}_k + \sum_{l=1}^{J_{k|k-1,b}} w^{(l|i)}_{k,b} \quad (58)$$

$$w^{(j|i)}_k = g_k(z|\mu^{(j|i)}_{k|k-1}; \mathbf{X}^{(i)}_{k|k-1}) w^{(j|i)}_{k|k-1} \quad (59)$$

$$\mu^{(j|i)}_k = \mu^{(j|i)}_{k|k-1} - K^{(j|i)}_k(\mathbf{z} - \hat{\mathbf{z}}^{(j|i)}_k) \quad (60)$$

$$\mathbf{P}^{(j|i)}_k = (\mathbf{I} - \mathbf{K}^{(j|i)}_k\mathbf{J}^{(j|i)}_k)\mathbf{P}^{(j|i)}_{k|k-1} \quad (61)$$

$$\mathbf{J}^{(j|i)}_k = \frac{\partial}{\partial \mathbf{m}} h(\mathbf{m}, \mathbf{X})\bigg|_{\mathbf{m}=\mu^{(j|i)}_{k|k-1}, \mathbf{X}=\mathbf{X}^{(i)}_{k|k-1}} \quad (62)$$

$$\mathbf{K}^{(j|i)}_k = \mathbf{P}^{(j|i)}_{k|k-1}\mathbf{J}^{(j|i)}_k\mathbf{S}^{(j|i),-1}_k \quad (63)$$

$$\mathbf{S}^{j|i}_k = \mathbf{J}^{(j|i)}_k\mathbf{P}^{(j|i)}_{k|k-1}\mathbf{J}^{(j|i),T}_k + \mathbf{R}_k \quad (64)$$

$$\hat{\mathbf{z}}^{(j|i)}_k = h(\mu^{(j|i)}_{k|k-1}, \mathbf{X}^{(i)}_{k|k-1}) \quad (65)$$

Let $FOV_k(\mathbf{X}_k) \in \mathcal{X}$ be the vehicle's sensor field of view at time k, dependent on the current vehicle location. Assuming a constant probability of detection p_D, we have:

$$p_D(\mathbf{m}|\mathbf{X}^{(i)}_{k|k-1}) = \begin{cases} p_D & \text{if } \mathbf{m} \in FOV_k(\mathbf{X}^{(i)}_{k|k-1}) \\ 0 & \text{otherwise} \end{cases} \quad (66)$$

Consequently, for landmarks outside of the field of view, the updated feature will be identical to the predicted one because only the first term of (56) will be non-zero. This means that only the features within $FOV_k(\mathbf{X}^{(i)}_{k|k-1})$ need to be updated, and remaining feature estimates can be propagated forward untouched. In order to perform the measurement update for the parent process, we must first compute the multi-object measurement likelihood.

$$L_{\mathbf{Z}_k}(\mathbf{X}^{(i)}) = \exp\left\{-\sum_{j=1}^{J_{k|k-1}} p_D(\mu^{(j|i)}_{k|k-1}|\mathbf{X}^{(i)}_{k|k-1}) w^{(j|i)}_{k|k-1}\right\} \times \prod_{\mathbf{z}\in\mathbf{Z}_k} L_{\mathbf{z}} \quad (67)$$

With this likelihood in hand, the weights of the Dirac mixture can be updated:

$$\eta^{(i)}_k = \frac{L_{\mathbf{Z}_k}(\mathbf{X}^{(i)})}{\sum_{l=1}^{N_{k|k-1}} L_{\mathbf{Z}_k}(\mathbf{X}^{(l)})}\eta^{(i)}_{k-1} \quad (68)$$

At this point we have the updated posterior parent and daughter PHDs. However, some steps need to be taken to manage the computational complexity of the filter. During the prediction for the parent, each component in the Dirac mixture is "shotgunned" into M new components, resulting in a new mixture containing $M \times N_{k-1}$ components. Left unchecked, the size of the parent process mixture would grow exponentially with every time step. To curb this growth, we prune the mixture to the N_{k-1} components with the highest weights. The Gaussian mixtures for the daughter process also have the potential for this exponential growth, as the measurement update generates $|\mathbf{Z}|+1$ new Gaussians for each component in the predicted mixture. Many of these come from low-likelihood measurement associations and contribute little to the updated PHD.

D. Forward Backward Smoother

In this section, we consider smoothing the parent process which is approximated using a set of weighted particles. The expression for the smoothed distribution is given by equations (31) and (37). In smoothing the parent process, we re-evaluate the weights of the particles according to the backward recursion without adding additional particles.

The smoothed distribution at time $k+1$ from $k' > k$ is approximated by a set of N_{k+1} weighted particles as

$$p_{k+1|k'}(\mathbf{X}|\mathbf{Z}_{1:k'}) = \sum_{i=1}^{N_{k+1}} w^{(i)}_{k+1|k'}\delta(\mathbf{X} - \mathbf{X}^{(i)}_{k+1|k'}) \ . \quad (69)$$

The updated distribution $p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})$ at time k is approximated using N_k particles as

$$p_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k}) = \sum_{i=1}^{N_k} w^{(i)}_{k|k}\delta(\mathbf{X} - \mathbf{X}^{(i)}_{k|k}) \ . \quad (70)$$

Then, using equations (31) and (37), the smoothed distribution at time k is given by

$$p_{k|k'}(\mathbf{X}|\mathbf{Z}_{1:k'}) = \sum_{i=1}^{N_k} w_{k|k'}^{(i)} \delta(\mathbf{X} - \mathbf{X}_{k|k'}^{(i)}) \tag{71}$$

where the smoothed weights are evaluated according to the expression

$$w_{k|k'}^{(i)} = \sum_{j=1}^{N_{k+1}} \frac{w_{k+1|k'}^{(j)} w_{k|k}^{(i)} f_{k+1|k}(x_{k+1|k'}^{(j)}|x_{k|k}^{(i)})}{\sum_{l=1}^{N_k} w_{k|k}^{(l)} f_{k+1|k}(x_{k+1|k'}^{(j)}|x_{k|k}^{(l)})} \tag{72}$$

and $f(\cdot|\cdot)$ represents the Markov transition density on the parent process.

V. Simulated Results

In this section, we demonstrate smoothing of the parent process using the forward-backward particle smoother. Figure 1 illustrates the true trajectory followed by the vehicle as well as the estimated path from the filter and a fixed lag smoother (for a lag of 16 time steps). Additionally, the figure shows an estimate of the uncertainty on the estimated landmarks from the filter.

In this scenario, the uncertainty in the distribution on the vehicle position grows as the vehicle traverses the path until the loop is closed, and the uncertainty decreases at this point. By smoothing backwards at times just prior to the loop closure, we can achieve a significant reduction in the estimated vehicle position.

In the experiment conducted here, the vehicle uses the Ackerman steering motion model [36]. A set of measurements corresponding to detected landmarks is obtained from the sensor. The set of measurements consists of detected point features represented by range and bearing. The odometry noise is given by a zero mean Gaussian with standard deviation 1 m/s for velocity and $0.1°$ for steering angle. The observation noise for the range and bearing sensor measurements is given by a zero mean Gaussian with standard deviation 1 m for range and $2°$ for bearing. The probability of detection of landmarks is $p_D = 0.95$ and an average of $\lambda = 2$ false alarms are detected per scan.

100 particles are used to model the vehicle trajectory and resampling is performed when the effective number of particles reduces to 70. This allows the particle distribution to retain a sufficient number of samples in the tails of the distribution which is essential for the smoother.

Since we are smoothing the path only in the region just prior to the loop closure, we illustrate the smoothing only on the second half of the vehicle path. We apply a fixed lag smoother with a lag of 5, 8, 12 and 16 time steps. The squared error for the filter and fixed lag smoothers is shown in Figure 2. From the figure, it is clear that the correction in the filter at time $t = 261$ from the loop closure event propagates backwards in the smoother and reduces the error at those corresponding time instances.

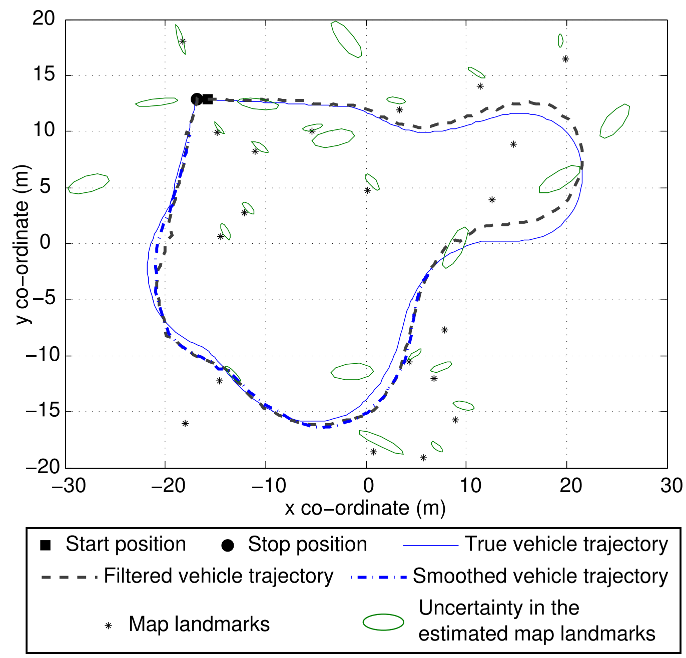

Fig. 1. True vehicle trajectory along with estimated path from the filter and fixed lag smoother with a lag of 16 time steps.

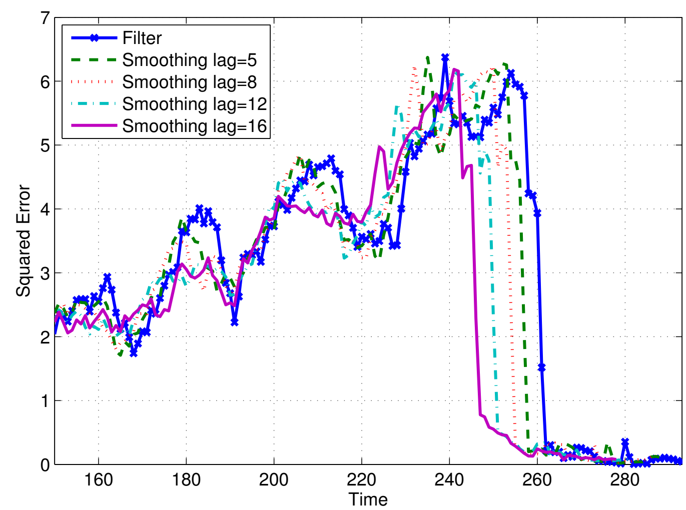

Fig. 2. Reduction in uncertainty due to loop closure in the filter ($t = 261$) is propagated backwards through the smoother resulting in lower error in the smoothed vehicle position.

VI. Conclusions

This paper has developed the random finite set approach to simultaneous localisation and mapping by introducing forward-backward smoothing to refine the vehicle trajectory. The results demonstrate that the technique is able to improve the estimate of the vehicle position after closing a loop. Future work will involve investigating smoothing the estimate of the map and introducing moving targets.

VII. Acknowledgements

Dr Clark is a Royal Academy of Engineering/ EPSRC Research Fellow. This work was supported by the Engineering and Physical Sciences Research Council EP/H010866/1. Chee Sing Lee is granted with a PhD FI Scholarship of the Catalan Government. This work has been supported by the FP7-ICT-2011-7 project PANDORA—Persistent Autonomy through Learning, Adaptation, Observation and Re-planning (Ref 288273) funded by the European Commission.

References

[1] I. Goodman, R. Mahler, and H. Nguyen. *Mathematics of Data Fusion*. Kluwer Academic Press, Boston, 1997.

[2] R. Mahler. Statistical Multisource Multitarget Information Fusion. *Artech House*, 2007.

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

[4] R. Mahler. PHD Filters of Higher Order in Target Number. *IEEE Transactions on Aerospace and Electronic Systems*, 43, No. 4:1523–1543, 2007.

[5] 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.

[6] D. E. Clark and J. Bell. Convergence Results for the Particle PHD Filter. *IEEE Trans. Sig. Proc., Vol 54 No 7: 2652–2661*, 2006.

[7] N. Whiteley, S. S. Singh, and S. Godsill. Auxiliary Particle Implementation of Probability Hypothesis Density Filter. *IEEE Transactions on Aerospace and Electronic Systems*, 46(3):1437–1454, 2010.

[8] B. Ristic, D. Clark, and B.-N. Vo. Improved SMC implementation of the PHD filter. *Proc. conf. Information Fusion (FUSION)*, 2010.

[9] B. Ristic, D. Clark, B.-N. Vo, and B.-T. Vo. Adaptive Target Birth Intensity for PHD and CPHD Filters. *IEEE Transactions on Aerospace and Electronic Systems*, 48:1656 – 1668, 2012.

[10] B. Vo and W. K. Ma. The Gaussian Mixture Probability Hypothesis Density Filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, 2006.

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

[12] J. Houssineau and D. Laneuville. PHD filter with diffuse spatial prior on the birth process with applications to GM-PHD filter. In *Information Fusion (FUSION), 2010 13th Conference on*, 2010.

[13] D. Clark, B. T. Vo, and S. J. Godsill. Gaussian mixture implementations of phd filters for non-linear dynamical models. *IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications*, 2008.

[14] D. Clark, B. T. Vo, and B.-N. Vo. Gaussian particle implementations of probability hypothesis density filters. *IEEE Aerospace Conference*, 2007.

[15] D. Clark, B. Ristic, B.-N. Vo, and B. T. Vo. Bayesian multi-object filtering with amplitude feature likelihood for unknown object SNR. *IEEE Trans. Sig. Proc.*, 58 (1):26–37, 2010.

[16] D. Clark, A. T. Cemgil, P. Peeling, and S. Godsill. Multi-object tracking of sinusoidal components in audio with the Gaussian mixture Probability Hypothesis Density filter. *Workshop on Applications of Signal Processing to Audio and Acoustics*, 2007.

[17] R. C. Smith and P. Cheeseman. On the Representation and Estimation of Spatial Uncertainty. *Int. J. Robotics Research*, 5 (4):56–68, 1986.

[18] J. Mullane, B.-N. Vo, M. D. Adams, and B.-T. Vo. A random-finite-set approach to Bayesian SLAM. *IEEE Trans. Robotics*, 99:1–15, 2011.

[19] C.-S. Lee and D. E. Clark and J. Salvi. SLAM with Single Cluster PHD Filters. *International Conference on Robotics and Automation*, 2012.

[20] A. Swain and D. Clark. The single-group PHD filter: an analytic solution. *Proc. conf. Information Fusion*, 2011.

[21] A. Swain and D. E. Clark. First-moment filters for spatial independent cluster processes. *Proceedings of SPIE 7697*, 2010.

[22] A. Swain and D. Clark. Extended object filtering using spatial independent cluster processes. *Proc. conf. Information Fusion (FUSION)*, 2010.

[23] B. Ristic and D. Clark. Particle filter for joint estimation of multi-object dynamic state and multi-sensor bias. *Proc. ICASSP*, 2012.

[24] G. Kitagawa. Non-Gaussian State-Space Modeling of Nonstationary Time Series. *Journal of the American Statistical Association*, 82(400):1032–1041, 1987.

[25] N. Nandakumaran, K. Punithakumar, and T. Kirubarajan. Improved Multi-Target Tracking Using Probability Hypothesis Density Smoothing. *Proc. SPIE Conf. of Signal and Data Proc. of Small Targets Vol. 6699*, 2007.

[26] R. P. S. Mahler, B.-N. Vo, and B.-T. Vo. The forward-backward Probability Hypothesis Density smoother. *Int. Conf. on Information Fusion*, 2010.

[27] R.P.S. Mahler, Ba-Tuong Vo, and Ba-Ngu Vo. Forward-backward probability hypothesis density smoothing. *Aerospace and Electronic Systems, IEEE Transactions on*, 48(1):707 –728, jan. 2012.

[28] D. E. Clark. First-Moment Multi-Object Forward-Backward Smoothing. *Int. Conf. on Information Fusion*, 2010.

[29] S. Hernandez. State Estimation and Smoothing for the Probability Hypothesis Density Filter. *PhD thesis, University of Wellington*, 2011.

[30] B.-N. Vo, B.-T. Vo, and R. P. S. Mahler. A closed form solution to the Probability Hypothesis Density Smoother. *Int. Conf. on Information Fusion*, 2010.

[31] B.-N. Vo, B.-T. Vo, and R. P. S. Mahler. Closed form solutions to forward-backward smoothing. *IEEE Trans. Sig. Proc., accepted for publication*, 2011.

[32] B.-T. Vo, D. Clark, B.-N. Vo, and B. Ristic. Bernoulli Forward-Backward Smoothing for Joint Target Detection and Tracking. *IEEE Transactions on Signal Processing*, 59(9):4473–4477, 2011.

[33] D. E. Clark, B. T. Vo, and B.-N. Vo. Forward-backward sequential Monte Carlo smoothing for joint target detection and tracking. *Proc. conf. Information Fusion*, 2009.

[34] D. Clark. Joint target-detection and tracking smoothers. *Proceedings of SPIE 7336*, 2009.

[35] J. E. Moyal. The General Theory of Stochastic Population Processes. *Acta Mathematica*, 108:1–31, 1962.

[36] José E. Guivant and Eduardo Mario Nebot. Optimization of the simultaneous localization and map-building algorithm for real-time implementation. *IEEE Transactions on Robotics*, 17(3):242–257, 2001.