The Forward-Backward Probability Hypothesis Density Smoother

Ronald. P. S. Mahler
Advanced Technology Group,
Lockheed Martin MS2 Tactical Systems
Eagan, Minnesota
ronald.p.mahler@lmco.com

Ba-Ngu Vo & Ba-Tuong Vo
School of Electrical, Electronic & Computer Engineering
The University of Western Australia
35 Stirling Hwy, Crawley WA 6009
{ba-ngu.vo, ba-tuong.vo}@uwa.edu.au

Abstract − *A forward-backward Probability Hypothesis Density (PHD) smoother involving forward filtering followed by backward smoothing is derived. The forward filtering is performed by Mahler's PHD recursion. The PHD backward smoothing recursion is derived using Finite Set Statistics (FISST) and standard point process theory. Unlike the forward PHD recursion, the proposed backward PHD recursion is exact and does not require the previous iterate to be Poisson.*

Keywords: PHD, Filtering, Smoothing, tracking, random sets, point processes, finite set statistics.

1 Introduction

Filtering, smoothing and prediction are three important interrelated problems in stochastic estimation. Smoothing can yield significantly better estimates than filtering by delaying the decision and using data at a later time [15], [5]. The text [1] provides a comprehensive coverage of closed-form smoothing solutions for linear Gaussian models.

In a multi-target scenario the number of states and the states themselves vary in time in a random fashion. This is compounded by false measurements, detection uncertainty and data association uncertainty. Consequently, filtering and smoothing in the multi-target realm is extremely challenging. In [9] a Probabilistic Data Association (PDA) multi-target smoothing algorithm was proposed to improve tracking performance in clutter. An Interacting Multiple Model (IMM) smoothing method was proposed in [6] to improve the tracking of maneuvering targets. In [4] a fixed lag smoothing scheme with IMM-PDA was proposed to improve the tracking of agile targets in clutter. The use of fixed-interval smoothing in IMM-MHT was proposed in [8] to improve the tracking of maneuvering targets.

As with the multi-target filtering problem, the challenge in multi-target smoothing is the high dimensionality of the distributions on the multi-target state space [11],[22]. Indeed, the computational intractability is more severe in multi-target smoothing than filtering. The PHD filter [10], [11] is a recent multi-target filter that operates on the single-target state space and, consequently, avoids the high dimensionality that results from multiple targets. In recent years developments of the Probability Hypothesis Density (PHD) and Cardinalized PHD filters [10], [18], [19], [12], [20] has attracted substantial interest. Its efficiency and performance suggest that smoothing with the PHD offers a tractable solution to non-linear non-Gaussian multi-target smoothing. The first attempt to solve the multi-target smoothing problem via the PHD framework was reported in [16], where a forward-backward smoothing scheme is employed. Mahler's PHD filter [10] is the natural choice for the forward filter. Based on physical intuition, an approximate backward PHD smoother under Poisson assumptions was proposed [16].

Inspired by the attempt in [16], we derive, using rigorous mathematical arguments rather than intuition, the first correct backward PHD smoother. It turns out that our backward PHD recursion is slightly different from the result of [16]. Moreover, our backward PHD recursion is exact and does not require that the smoothed PHD at the previous iteration to be Poisson. The mathematical tools used in our derivations are Mahler's Finite Set Statistics (FISST) [10], [11] and

the celebrated Campbell's theorem from point process theory [2], [17].

2 Background

This section presents relevant background required for the derivation of the main results, including Bayesian multi-target filtering, the PHD filter, tools from FISST and point process theory such as probability generating functionals, Campbell's theorem and factorial moments. Further background material can be found in [11] from a FISST perspective or in [22] from a point process perspective. For simplicity the following notation is adopted throughout the paper:

$$\langle f, g \rangle = \int f(x)g(x)dx$$

$$\langle \langle f(:|\cdot), g(:) \rangle, h(\cdot) \rangle = \int \left(\int f(x|y)g(x)dx \right) h(y)dy$$

2.1 Random finite set and the Bayes multi-target filter

Suppose at time k there are $M(k)$ targets with states $x_{k,1}, \ldots, x_{k,M(k)}$ each taking values in a state space $\mathcal{X} \subseteq R^{n_x}$, and $N(k)$ measurements $z_{k,1}, \ldots, z_{k,N(k)}$ each taking values in an observation space $\mathcal{Z} \subseteq R^{n_z}$, where R^n denotes the nth Cartesian product of R. Then, the *multi-target state* X_k and the *multi-target measurement* Z_k, at time k, are defined as

$$X_k = \{x_{k,1}, \ldots, x_{k,N(k)}\} \in \mathcal{F}(\mathcal{X}),$$
$$Z_k = \{z_{k,1}, \ldots, z_{k,M(k)}\} \in \mathcal{F}(\mathcal{Z}),$$

where $\mathcal{F}(\mathcal{X})$ and $\mathcal{F}(\mathcal{Z})$ denote the spaces of all finite subsets of $\mathcal{X}$ and $\mathcal{Z}$, respectively. In the Bayesian estimation paradigm, the state and measurement are treated as realizations of random variables. Since the (multi-target) state X_k and measurement Z_k are finite sets, the concept of a *random finite set* is required.

In essence, a *random finite set* (RFS) X on $\mathcal{X}$, is simply a finite-set-valued random variable or a random variable taking values in $\mathcal{F}(\mathcal{X})$. As with random vectors, the *probability density* of an RFS (if it exists) is a very useful descriptor in filtering and estimation. However, standard tools for random vectors are not appropriate for RFSs since the space $\mathcal{F}(\mathcal{X})$ does not inherit the usual Euclidean notion of integration and density. Mahler's Finite Set Statistics (FISST) provides practical mathematical tools for dealing with RFSs [10], [11], including a consistent notion of integration and density.

Using the FISST notion of integration and density, the multi-target Bayes filter that propagates the *multi-target posterior density* $p_{k|k}(\cdot|Z_{1:k})$ in time is given by

[10, 11]

$$p_{k|k-1}(X_k|Z_{1:k-1}) =$$
$$\int f_{k|k-1}(X_k|X)p_{k-1|k-1}(X|Z_{1:k-1})\delta X, \quad (1)$$
$$p_{k|k}(X_k|Z_{1:k}) = \frac{g_k(Z_k|X_k)p_{k|k-1}(X_k|Z_{1:k-1})}{\int g_k(Z_k|X)p_{k|k-1}(X|Z_{1:k-1})\delta X}, \quad (2)$$

where $p_{k|k-1}$ denotes the predicted multi-target density, $f_{k|k-1}$ is the *multi-target transition density*, g_k is the *multi-target likelihood* and

$$\int f(X)\delta X = \sum_{i=0}^{\infty} \frac{1}{i!} \int f(\{x_1, \ldots, x_i\})dx_1 \cdots dx_i.$$

is the set integral of a function $f : \mathcal{F}(\mathcal{X}) \to R$.

Like the standard (vector) posterior, the multi-target posterior captures all information about the multi-target state. However, optimal Bayes estimators for random vectors, such as expected a posteriori or maximum a posteriori, are not applicable to RFSs. Suitable Bayes optimal estimators for RFSs have been established in [11].

The multi-target Bayes filter is generally intractable and it is necessary to resort to more tractable approximations [10], [12][11], [22], [21]. The Probability Hypothesis Density (PHD) filter [10] is a first moment approximation to the full multi-target Bayes filter (1)-(2), which operates on the (single-target) state space $\mathcal{X}$.

2.2 The PHD and Campbell's theorem

The PHD, commonly known in point process theory as an *intensity function*, is a first-order statistical moment of an RFS [2], [17], [10]. The PHD of an RFS X on $\mathcal{X}$, is a non-negative function v on $\mathcal{X}$ such that its integral over any region S gives the expected number of elements of X that are in S, i.e.

$$E\left[\sum_{x \in X} 1_S(x)\right] = \langle 1_S, v \rangle,$$

where 1_S is the indicator function of the set S, and E denotes the expectation operator. Note that given a (FISST) multi-target density p, E can be expressed in terms of a set integral as follows

$$E[f(X)] = \int f(X)p(X)\delta X.$$

The local maxima of the PHD are points in $\mathcal{X}$ with the highest local concentration of expected number of elements, and can be used to generate estimates for the elements of X. A simple multi-target estimator can be obtained by, first, estimating the number of states, $\hat{N}$ by rounding the PHD mass $\langle 1, v \rangle$ and, second, choosing the $\hat{N}$ highest maxima of the PHD v. The Bayes optimality of this estimator has been discussed in [13].

Campbell's theorem relates certain types of expectation of an RFS to its PHD, and is an important result in point process theory [17]. Campbell's theorem (see [2], [17]) states that for an RFS X on $\mathcal{X}$ with PHD (or intensity) v

$$E\left[\sum_{x \in X} \zeta(x)\right] = \langle \zeta, v \rangle. \tag{3}$$

We will call on Campbell's theorem for the derivation of our main results.

2.3 The PHD filter

The PHD filter recursively propagates the PHD of the multi-target state in time based on the following assumptions:

- Each target evolves and generates measurements independently of one another

- The surviving and birth RFSs are independent of each other

- The clutter RFS is Poisson and independent of the target generated measurements

- The predicted multi-target RFS is Poisson

The PHD recursion consists of a prediction step and an update step that respectively approximate the Bayes multi-target prediction (1) and update (2). Let

$$v_{k|k} = \text{filtered (updated) PHD at time } k$$
$$v_{k+1|k} = \text{predicted PHD from time } k \text{ to } k+1$$
$$\gamma = \gamma_{k+1|k} = \text{PHD of birth at time } k+1$$
$$f = f_{k+1|k} = \text{single-target transition to } k+1$$
$$p_S = p_{S,k+1|k} = \text{probability of survival to } k+1$$
$$g = g_{k+1} = \text{single-target likelihood at time } k+1$$
$$p_D = p_{D,k+1} = \text{probability of detection at } k+1$$

(Note that for simplicity we have dropped the time indices from the model parameters $\gamma_{k+1|k}, f_{k+1|k}$, $p_{S,k+1|k}, g_{k+1}, p_{D,k+1}$.) Then the PHD prediction and update are respectively given by

$$v_{k+1|k}(x) = \gamma(x) + \langle v_{k|k} p_S, f(x|\cdot) \rangle, \tag{4}$$
$$v_{k+1|k+1}(x) = [1 - p_D(x)] v_{k+1|k}(x)$$
$$+ \sum_{z \in Z_{k+1}} \frac{p_D(x) g(z|x) v_{k+1|k}(x)}{\kappa_{k+1}(z) + \langle p_D g(z|\cdot), v_{k+1|k} \rangle}. \tag{5}$$

Sequential Monte Carlo (SMC) and Gaussian mixture implementations of the PHD recursion [18], [19], as well as generalizations [12], [20] have opened the door to numerous novel extensions and applications.

2.4 Probability generating functionals (PGFl)

Apart from the probability density, the probability generating functional (PGFl) is another fundamental descriptor of an RFS. Following [2, 17], the *probability generating functional* (PGFl) $G[\cdot]$ of an RFS X on $\mathcal{X}$ is defined by

$$G[h] \equiv E[h^X], \tag{6}$$

where h is any real-valued function on $\mathcal{X}$ such that $0 \leq h(x) \leq 1$, and

$$h^X = \prod_{x \in X} h(x), \text{ with } h^\emptyset = 1$$

The functional derivative of the PGFl can be defined, if the limit exists, as follows

$$G^{(1)}[g; \zeta] = \lim_{\varepsilon \to 0} \frac{G[g + \varepsilon \zeta] - G[g]}{\varepsilon}$$
$$G^{(i)}[g; \zeta_1, ..., \zeta_i] = \left(G^{(i-1)}[g; \zeta_1, ..., \zeta_{i-1}] \right)^{(1)}[\zeta_i]$$

The ith functional derivative of the PGFl is linear in the each of the directions $\zeta_1, ..., \zeta_i$. Moreover, it is a Fréchet derivative if it is also continuous in each of the directions.

It was noted in [22] that a linear functional in each of the variables $\zeta_1, ..., \zeta_i$ can be identified with a measure μ on $\mathcal{X}^i$ via

$$\mu[\zeta_1, ..., \zeta_i] = \int \zeta_1(x_1), ..., \zeta_i(x_i) \mu(dx_1, ..., dx_i).$$

That is, we treat the measure μ as a functional that takes the functions $\zeta_1, ..., \zeta_i$ to the reals. If the measure μ admits a density f then,

$$\mu[\zeta_1, ..., \zeta_i] = \int ... \int \zeta_1(x_1) ... \zeta_i(x_i) f(x_1, .., x_i) dx_1 ... dx_i \tag{7}$$

and the rather suggestive notation $\mu[\delta_{x_1}, ..., \delta_{x_i}] \equiv f(x_1, ..., x_i)$ can be used, where δ_x can be interpreted as a Dirac delta centered at x.

Treating $G^{(i)}[g; \cdot, ..., \cdot]$ as a measure, and $G^{(i)}[g; \delta_{x_1}, ..., \delta_{x_i}]$ as its density, we use the set derivative notation

$$\left. \frac{\partial}{\partial \{x_1, ..., x_i\}} \right|_h G[\cdot] = G^{(i)}[h; \delta_{x_1}, ..., \delta_{x_i}]$$

since this is suggestive of ordinary derivatives. The rules for this type of differentiation are established in [10]. It follows from [2] that the multi-target density p and the PHD v (if they exist) can be recovered from the PGFl by set differentiation

$$p(X) = \left. \frac{\partial}{\partial X} \right|_{h=0} G[h], \tag{8}$$

$$v(x) = \left. \frac{\partial}{\partial x} \right|_{h=1} G[h]. \tag{9}$$

The *cardinality* (number of elements) of X, denoted as $|X|$, is a discrete random variable whose *probability generating function* (PGF) $G(\cdot)$ can be obtained by setting the function h in the PGFl $G[\cdot]$ to a constant z. Note the distinction between the PGF and PGFl by the round and square brackets on the argument. The cardinality distribution ρ (the probability distribution of the cardinality $|X|$) and the PGF $G(\cdot)$ are Z-transform pairs.

A *Poisson* RFS X on $\mathcal{X}$ is one that is completely characterized by its PHD function v [2, 17]. The cardinality of a Poisson RFS is Poisson with mean $\langle v, 1\rangle$, and for a given cardinality the elements of X are each independent and identically distributed with probability density $v/\langle v, 1\rangle$. The PGFl of a Poisson RFS is

$$G[h] = e^{\langle v, h-1\rangle}. \tag{10}$$

A *multi-Bernoulli* RFS X on $\mathcal{X}$ is a union $\bigcup_{i=1}^{M} X^{(i)}$ of independent RFSs $X^{(i)}$ that has probability $1 - r^{(i)}$ of being empty, and probability $r^{(i)} \in (0,1)$ of being a singleton whose (only) element is distributed according to a probability density $p^{(i)}$ (defined on $\mathcal{X}$), [11]. The PGFl of a multi-Bernoulli RFS is given by

$$G[h] = \prod_{i=1}^{M} \left(1 - r^{(i)} + r^{(i)}\langle p^{(i)}, h\rangle\right). \tag{11}$$

A multi-Bernoulli RFS is thus completely described by the multi-Bernoulli parameters $\{(r^{(i)}, p^{(i)})\}_{i=1}^{M}$. The parameter $r^{(i)}$ is the existence probability of the ith object while $p^{(i)}$ is the probability density of the state conditional on its existence. Multi-Bernoulli approximations have been successfully applied to multi-target filtering, see for example [21], [23].

Given a multi-target state X, at time k, the multi-target at the next time step is modeled by the union of a Poisson birth RFS with intensity γ and a multi-Bernoulli surviving RFS with parameter set $\{(p_S(x), f(\cdot|x) : x \in X\}$. If the birth RFS and the surviving RFS are independent, then the PGFl $G_{k+1|k}[\cdot|X]$ of the multi-target transition density $f_{k+1|k}(\cdot|X)$ is given by

$$G_{k+1|k}[h|X] = e^{\langle \gamma, h-1\rangle} \left(1 - p_S + p_S \langle f(:|\cdot), h(:)\rangle\right)^X. \tag{12}$$

Similarly, the multi-target measurement is modelled by the union of a Poisson clutter RFS with intensity κ and a multi-Bernoulli detection RFS with parameters $\{(p_D(x), g(\cdot|x) : x \in X\}$. If the clutter RFS and the detection RFS are independent, then the PGFl $G_{k+1}[\cdot|X]$ of the multi-target likelihood $g_{k+1}(\cdot|X)$ is given by

$$G_{k+1}[h|X] = e^{\langle \kappa, h-1\rangle} \left(1 - p_D + p_D \langle g(:|\cdot), h(:)\rangle\right)^X.$$

3 The PHD Smoother

Forward-backward smoothing consists of forward filtering followed by backward smoothing. In the forward filtering, the posterior density is propagated forward to time k via the Bayes recursion. In the backward smoothing step, the smoothed density is propagated backward, from time k to time $k' < k$, via the backward smoothing recursion (see for example [7]). In the multi-target case, the multi-target posterior is propagated forward to time k via the multi-target Bayes recursion (1)-(2) and the smoothed multi-target density is propagated backward, from time k to time $k' < k$, via the multi-target backward smoothing recursion

$$p_{k'|k}(X) = p_{k'|k'}(X) \int f_{k'+1|k'}(Y|X) \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y. \tag{13}$$

While the FISST multi-target density is not a probability density [11], the recursion (13) has the same form as the standard backward smoother expressed in terms of probability densities [7]. A simple way to derive (13) is to first apply the same argument as per the standard backward smoother to relevant RFS probability densities, then invoke the relationship between FISST density/integration with measure theoretic density/integration in [18].

As with the multi-target Bayes filter, the multi-target forward-backward smoother is computationally intractable in general. We consider in this paper a first order moment approximation that propagates the PHD forward and backward. The first moment captures the first order statistic of the underlying multi-target state. The PHD forward propagation is accomplished by the PHD recursion (4)-(5). The PHD backward propagation is given by the following result.

Proposition 1: *If the filtered and the predicted multi-target RFSs are Poisson, then the smoothed PHD $v_{k'|k}$ can be computed recursively by*

$$v_{k'|k}(x) = v_{k'|k'}(x)\left(1 - p_S(x) + p_S(x)\left\langle \frac{f(\cdot|x)}{v_{k'+1|k'}}, v_{k'+1|k}\right\rangle\right). \tag{14}$$

Note that the smoothed PHD from the previous time step, $v_{k'+1|k}$, need not be Poisson. To prove this result, we need the following mathematical aid.

Lemma 1: *Given $\alpha, \beta : \mathcal{X} \to R$, and $c \in R$,*

$$\left.\frac{\partial}{\partial Y}\right|_{g=0} e^{\langle \alpha, g\rangle} \left(\langle \beta, g\rangle + c\right) = \left(c + \sum_{y \in Y} \frac{\beta(y)}{\alpha(y)}\right) \alpha^Y. \tag{15}$$

This is a special case of Lemma 2 in the appendix.

Proof of Proposition 1: For the smoothed multi-target state with density $p_{k'|k}$ given by (13), the PGFl is

$$G_{k'|k}[h] = \int h^X p_{k'|k}(X)\delta X$$

$$= \int \int h^X f_{k'+1|k'}(Y|X)p_{k'|k'}(X)\delta X$$

$$\times \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

$$= \int \int h^X \left.\frac{\partial}{\partial Y}\right|_{g=0} G_{k'+1|k'}[g|X]p_{k'|k'}(X)\delta X$$

$$\times \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

$$= \int \left.\frac{\partial}{\partial Y}\right|_{g=0} \int h^X G_{k'+1|k'}[g|X]p_{k'|k'}(X)\delta X$$

$$\times \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

Substituting (12) for $G_{k'+1|k'}[g|X]$, gives

$$G_{k'|k}[h] = \int \left.\frac{\partial}{\partial Y}\right|_{g=0} \int h^X (1 - p_S + p_S \langle f(:|\cdot),g(:)\rangle)^X$$

$$\times e^{\langle\gamma,g-1\rangle}p_{k'|k'}(X)\delta X \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

$$= \int \left.\frac{\partial}{\partial Y}\right|_{g=0} G_{k'|k'}[h(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)]$$

$$\times e^{\langle\gamma,g-1\rangle}\frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

Using the assumption that the filtered multi-target state is Poisson i.e. $G_{k'|k'}[h] = e^{\langle v_{k'|k'},h-1\rangle}$ gives (16).

$$G_{k'|k}[h] = \int \left.\frac{\partial}{\partial Y}\right|_{g=0} e^{\langle v_{k'|k'},h(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)-1\rangle}$$

$$\times e^{\langle\gamma,g-1\rangle}\frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y. \qquad (16)$$

The smoothed PHD can be obtained by differentiating the PGFl in (16):

$$v_{k'|k}(x) = \left.\frac{\partial}{\partial x}\right|_{h=1} G_{k'|k}[h]$$

$$= \left.\frac{\partial}{\partial x}\right|_{h=1} \int \left.\frac{\partial}{\partial Y}\right|_{g=0} e^{\langle\gamma,g-1\rangle}$$

$$\times e^{\langle v_{k'|k'},h(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)-1\rangle}\frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

$$= \int \left.\frac{\partial}{\partial Y}\right|_{g=0} e^{\langle\gamma,g-1\rangle}$$

$$\times \left.\frac{\partial}{\partial x}\right|_{h=1} e^{\langle v_{k'|k'},h(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)-1\rangle}$$

$$\times \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y \qquad (17)$$

Now, consider the derivative w.r.t. x in (17)

$$\left.\frac{\partial}{\partial x}\right|_{h=1} e^{\langle v_{k'|k'},h(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)-1\rangle}$$

$$= e^{\langle v_{k'|k'},(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)-1\rangle}$$

$$\times v_{k'|k'}(x)\left(1 - p_S(x) + p_S(x)\langle f(:|x),g(:)\rangle\right). \qquad (18)$$

The exponent in the RHS of (18) can be rearranged by changing the order of integration as follows

$$\langle v_{k'|k'},(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)-1\rangle$$

$$= \langle v_{k'|k'}(\cdot),p_S(\cdot)\langle f(:|\cdot),g(:)\rangle\rangle - \langle v_{k'|k'},p_S\rangle$$

$$= \langle\langle v_{k'|k'}(\cdot)p_S(\cdot),f(:|\cdot)\rangle,g(:)\rangle$$

$$- \langle\langle v_{k'|k'}(\cdot)p_S(\cdot)f(:|\cdot)\rangle,1\rangle$$

$$= \langle\langle v_{k'|k'}(\cdot)p_S(\cdot),f(:|\cdot)\rangle,g(:)-1\rangle$$

$$= \langle v_{k'+1|k'} - \gamma,g-1\rangle. \qquad (19)$$

and hence

$$e^{\langle\gamma,g-1\rangle}\left.\frac{\partial}{\partial x}\right|_{h=1} e^{\langle v_{k'|k'},h(1-p_S+p_S\langle f(:|\cdot),g(:)\rangle)-1\rangle}$$

$$= e^{\langle\gamma,g-1\rangle}e^{\langle v_{k'+1|k'}-\gamma,g-1\rangle}v_{k'|k'}(x)$$

$$\times (1 - p_S(x) + p_S(x)\langle f(:|x),g(:)\rangle)$$

$$= e^{\langle v_{k'+1|k'},g-1\rangle}v_{k'|k'}(x)$$

$$\times (1 - p_S(x) + p_S(x)\langle f(:|x),g(:)\rangle). \qquad (20)$$

Thus, substituting (20) into (17) gives

$$v_{k'|k}(x) = v_{k'|k'}(x)\int \left.\frac{\partial}{\partial Y}\right|_{g=0} e^{\langle v_{k'+1|k'},g-1\rangle}$$

$$\times (1 - p_S(x) + p_S(x)\langle f(:|x),g(:)\rangle)$$

$$\times \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

Taking $e^{-\langle v_{k'+1|k'},1\rangle}$ outside the derivative, and applying Lemma 1 with $\alpha = v_{k'+1|k'}$, $\beta = p_S(x)f(\cdot|x)$, $c = 1 - p_S(x)$ yields

$$v_{k'|k}(x) = v_{k'|k'}(x)\int\left(1-p_S(x)+p_S(x)\sum_{y\in Y}\frac{f(y|x)}{v_{k'+1|k'}(y)}\right)$$

$$\times \frac{v_{k'+1|k'}^Y e^{-\langle v_{k'+1|k'},1\rangle}}{p_{k'+1|k'}(Y)}p_{k'+1|k}(Y)\delta Y \qquad (21)$$

Using the Poisson assumption on $p_{k'+1|k'}$ gives

$$v_{k'|k}(x) = v_{k'|k'}(x)\int\left(1-p_S(x)+p_S(x)\sum_{y\in Y}\frac{f(y|x)}{v_{k'+1|k'}(y)}\right)$$

$$\times p_{k'+1|k}(Y)\delta Y.$$

Taking the integral inside the bracket and applying Campbell's theorem gives

$$v_{k'|k}(x) = v_{k'|k'}(x)\left(1-p_S(x)+p_S(x)\left\langle\frac{f(\cdot|x)}{v_{k'+1|k'}},v_{k'+1|k}\right\rangle\right). \qquad (22)$$

4 Conclusions

Using FISST and Campbell's theorem, we have derived a backward PHD smoothing recursion that does not require the previous smoothed iterate to be Poisson. This smoothing recursion also admits a closed form solution under multi-target linear Gaussian assumptions as shown in the companion paper [24]. It is also possible to derive the recursion for the cardinality distribution and moments using this framework (for more details see [14]). Our derivation consists of two key arguments. The first argument uses FISST differentiation techniques to derive (21) and the second argument uses Campbell's theorem to arrive at the end result. An alternative, and simpler, way of obtaining (21) is given in [3] by reusing the PHD update (2).

5 Appendix

Lemma 2: *Given* $\alpha, \beta : \mathcal{X} \to R$, *and* $c \in R$, *define for each integer* $n \geq 0$,

$$H_n[g] = e^{\langle \alpha, g \rangle} \left(\langle \beta, g \rangle + c \right)^n.$$

Then

$$\frac{\partial}{\partial Y}\bigg| H_n[g] = n!\alpha^Y \sum_{i=0}^{n} e_i\left(Y; \frac{\beta}{\alpha}\right) \frac{H_{n-i}[g]}{(n-i)!}, \quad (23)$$

$$\frac{\partial}{\partial Y}\bigg|_{g=0} H_n[g] = n!\alpha^Y \sum_{i=0}^{n} e_i\left(Y; \frac{\beta}{\alpha}\right) \frac{c^{n-i}}{(n-i)!}, \quad (24)$$

where $\frac{\beta}{\alpha}$ *denotes the point wise quotient of the functions* β, α *and* $e_i(Y; \phi)$ *denotes the ith elementary symmetric function evaluated at* $[\phi(y)]_{y \in Y}$, *i.e.*

$$e_i(Y; \phi) = \sum_{S \subset Y, |S|=i} \phi^S,$$

with the standard convention that $e_i(Y; \phi) = 0$ *for* $|Y| < i$, *so that the sum effectively contains* $|Y| + 1$ *terms, when* $|Y| < n$.

Proof: The proof makes use of the elementary symmetric function identity

$$e_i(y_1, ..., y_m, y_{m+1}; \phi) = e_i(y_1, ..., y_m; \phi)$$
$$+ e_{i-1}(y_1, ..., y_m; \phi)\phi(y_{m+1}) \quad (25)$$

Abbreviate: $\alpha(y_i)$ by α_i, $\beta(y_i)$ by β_i, and note that

$$\frac{\partial}{\partial \emptyset}\bigg| H_n[g] = H_n[g] \quad (26)$$

$$\frac{\partial}{\partial y_i}\bigg| H_n[g] = \alpha_i e^{\langle \alpha, g \rangle} \left(\langle \beta, g \rangle + c \right)^n$$
$$+ n\beta_i e^{\langle \alpha, g \rangle} \left(\langle \beta, g \rangle + c \right)^{n-1}$$
$$= \alpha_i H_n[g] + n\beta_i H_{n-1}[g]$$
$$= \alpha_i \left(H_n[g] + n\frac{\beta_i}{\alpha_i} H_{n-1}[g] \right) \quad (27)$$

From (26) and (27), it is clear that (23) holds for $Y = \emptyset$ and $Y = \{y_1\}$. Suppose that (23) is true for $Y = \{y_1, ..., y_m\}$, then

$$\frac{\partial}{\partial Y \cup \{y_{m+1}\}}\bigg| H_n[g]$$
$$= \frac{\partial}{\partial y_{m+1}}\bigg| \frac{\partial}{\partial Y}\bigg| H_n[g]$$
$$= n!\alpha^Y \frac{\partial}{\partial y_{m+1}}\bigg| \sum_{j=0}^{n} \frac{e_j\left(Y; \frac{\beta}{\alpha}\right)}{(n-j)!} H_{n-j}[g]$$
$$= n!\alpha^Y \frac{\partial}{\partial y_{m+1}}\bigg| \left(\begin{array}{c} \sum_{j=0}^{n-1} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j)!} H_{n-j}[g] \\ +e_n\left(Y; \frac{\beta}{\alpha}\right) H_0[g] \end{array} \right)$$
$$= n!\alpha^Y \left(\begin{array}{c} \sum_{j=0}^{n-1} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j)!} \frac{\partial}{\partial y_{m+1}}\bigg| H_{n-j}[g] \\ +e_n\left(Y; \frac{\beta}{\alpha}\right) \frac{\partial}{\partial y_{m+1}}\bigg| H_0[g] \end{array} \right)$$

Using (27) gives

$$\frac{\partial}{\partial Y \cup \{y_{m+1}\}}\bigg| H_n[g]$$
$$= n!\alpha^Y \left(\begin{array}{c} \sum_{j=0}^{n-1} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j)!} \alpha_{m+1} \\ \times \left(\begin{array}{c} H_{n-j}[g] \\ +(n-j)\frac{\beta_{m+1}}{\alpha_{m+1}} H_{n-j-1}[g] \end{array} \right) \\ +\alpha_{m+1} e_n\left(Y; \frac{\beta}{\alpha}\right) H_0[g] \end{array} \right)$$
$$= n!\alpha^Y \alpha_{m+1} \left(\begin{array}{c} \sum_{j=0}^{n-1} \left(\begin{array}{c} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j)!} H_{n-j}[g] + \\ \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j-1)!} \frac{\beta_{m+1}}{\alpha_{m+1}} H_{n-j-1}[g] \end{array} \right) \\ +e_n\left(Y; \frac{\beta}{\alpha}\right) H_0[g] \end{array} \right)$$
$$= n!\alpha^{Y \cup \{y_{m+1}\}} \left(\begin{array}{c} \sum_{j=0}^{n-1} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j)!} H_{n-j}[g] \\ +\sum_{j=0}^{n-1} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j-1)!} \frac{\beta_{m+1}}{\alpha_{m+1}} H_{n-j-1}[g] \\ +e_n\left(Y; \frac{\beta}{\alpha}\right) H_0[g] \end{array} \right)$$
$$= n!\alpha^{Y \cup \{y_{m+1}\}} \left(\begin{array}{c} \frac{H_n[g]}{n!} + \sum_{j=1}^{n-1} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j)!} H_{n-j}[g] \\ +\sum_{j=0}^{n-2} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j-1)!} \frac{\beta_{m+1}}{\alpha_{m+1}} H_{n-j-1}[g] \\ +e_{n-1}\left(Y; \frac{\beta}{\alpha}\right) \frac{\beta_{m+1}}{\alpha_{m+1}} H_0[g] \\ +e_n\left(Y; \frac{\beta}{\alpha}\right) H_0[g] \end{array} \right)$$

Set i to $j+1$ gives

$$\frac{\partial}{\partial Y \cup \{y_{m+1}\}}\bigg| H_n[g]$$
$$= n!\alpha^{Y \cup \{y_{m+1}\}} \left(\begin{array}{c} \frac{H_n[g]}{n!} + \sum_{j=1}^{n-1} \frac{e_j(Y; \frac{\beta}{\alpha})}{(n-j)!} H_{n-j}[g] \\ +\sum_{i=1}^{n-1} \frac{e_{i-1}(Y; \frac{\beta}{\alpha})}{(n-i)!} \frac{\beta_{m+1}}{\alpha_{m+1}} H_{n-i}[g] \\ +e_{n-1}\left(Y; \frac{\beta}{\alpha}\right) \frac{\beta_{m+1}}{\alpha_{m+1}} H_0[g] \\ +e_n\left(Y; \frac{\beta}{\alpha}\right) H_0[g] \end{array} \right)$$

$$= n!\alpha^{Y \cup \{y_{m+1}\}} \left(+ \sum_{i=1}^{n-1} \begin{pmatrix} e_i\left(Y; \frac{\beta}{\alpha}\right) + \\ e_{i-1}\left(Y; \frac{\beta}{\alpha}\right) \frac{\beta_{m+1}}{\alpha_{m+1}} \end{pmatrix} \frac{H_{n-i}[g]}{(n-i)!} + \begin{pmatrix} e_{n-1}\left(Y; \frac{\beta}{\alpha}\right) \frac{\beta_{m+1}}{\alpha_{m+1}} \\ + e_n\left(Y; \frac{\beta}{\alpha}\right) \end{pmatrix} H_0[g] \right)$$

$$= n!\alpha^{Y \cup \{y_{m+1}\}} \left(\frac{H_n[g]}{n!} + \sum_{i=1}^{n-1} \left(e_i\left(Y \cup \{y_{m+1}\}; \frac{\beta}{\alpha}\right)\right) \frac{H_{n-i}[g]}{(n-i)!} + e_n\left(Y \cup \{y_{m+1}\}; \frac{\beta}{\alpha}\right) H_0[g] \right)$$

Using (25)

$$\frac{\partial}{\partial Y \cup \{y_{m+1}\}} \bigg| H_n[g]$$

$$= n!\alpha^{Y \cup \{y_{m+1}\}} \sum_{i=0}^{n} e_i\left(Y \cup \{y_{m+1}\}; \frac{\beta}{\alpha}\right) \frac{H_{n-i}[g]}{(n-i)!}$$

Hence (23) holds by the principle of induction and (24) follows as $H_n[0] = c$.

References

[1] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, New York: Wiley, 2001.

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

[3] D. Clark, "First-moment Multi-object Forward-Backward Smoothing," *Proc Int. Conf. Information Fusion*, Edinburgh, UK, 2010.

[4] B. Chen, and J. K. Tugnait, "Multisensor tracking of a maneuvering target in clutter using IMMPDA fixed-lag smoothing," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 36, No. 3, pp. 983–991, Jul. 2000.

[5] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, New York: Cambridge University Press, 1989.

[6] R. E. Helmick, W. D. Blair, and S. A. Hoffman, "One step fixed-lag smoothers for Markovian switching systems," *IEEE Trans. Automatic Control*, Vol. 41, No. 7, pp. 1051–1056, Jul. 1996.

[7] G. Kitagawa, "Non-Gaussian state space modelling of nonstationary time series," *Journal of American Statistical Association*, Vol. 82, pp. 1032-1063, 1987.

[8] W. Koch, "Fixed-interval retrodiction approach to Bayesian IMM-MHT for maneuvering multiple targets," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 36, No. 1, pp. 2–14, Jan. 2000.

[9] A. K. Mahalanabis, B. Zhou, N. K. Bose, "Improved multi-target tracking in clutter by PDA smoothing," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 26, No. 1, pp. 113–121, Jan. 1990.

[10] R. Mahler, "Multi-target Bayes Filtering via First-Order Multi-target Moments," *IEEE Trans. Aerospace & Electronic Systems*, Vol. 39, No. 4, pp. 1152-1178, 2003.

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

[12] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerospace & Electronic Systems*, Vol. 43, No. 4, pp. 1523-1543, 2007.

[13] R. Mahler, "Implementation and Application of PHD/CPHD Filters," Tutorial Notes, *Int'l Conf. on Information Fusion, Seattle, WA, 2009*.

[14] R. Mahler, B.-N. Vo, B.-T. Vo, "Multi Target Smoothing via First Order Moments," *IEEE Trans, Aerospace & Electroni Systems*, Submitted 2009.

[15] J. Meditch, "A survey of data smoothing for linear and nonlinear dynamic systems," *Automatica,* Vol. 9, 151–162, 1973.

[16] N. Nandakumaran, K. Punithakumar and T. Kirubarajan, "Improved Multi-target Tracking Using Probability Hypothesis Density Smoothing", SPIE 2007, San Diego.

[17] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic geometry and its applications*, John Wiley & Sons, 1995.

[18] B.-N. Vo, S. Singh and A. Doucet, "Sequential Monte Carlo methods for Multi-target filtering with Random Finite Sets", *IEEE Trans. Aerospace & Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, 2005.

[19] B.-N. Vo, and W.-K. Ma, "The Gaussian mixture Probability Hypothesis Density filter," *IEEE Trans. Signal Processing, IEEE Trans. Signal Processing*, Vol. 54, No. 11, pp. 4091-4104, 2006.

[20] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the Cardinalized Probability Hypothesis Density filter," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3553–3567, Jul. 2007.

[21] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The Cardinality Balanced Multi-Target Multi-Bernoulli Filer and its Implementations," *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 409–423, Feb. 2009.

[22] B.-T. Vo, "Random Finite Sets in Multi-Object Filtering," *PhD Thesis*, School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Oct. 2008.

[23] B.-N. Vo, B.-T. Vo, N.-T. Pham and D. Suter, "Joint Detection and Estimation of Multiple Objects from Image Observations," *IEEE Trans. Signal Processing*, vol. 58, no. 10, Oct. 2010.

[24] B.-N. Vo, B.-T. Vo, and R. Mahler, "A closed form solution to the Probability Hypothesis Density Smoother," *Proc Int. Conf. Information Fusion*, Edinburgh, UK, 2010.