PREPRINT: IEEE TRANS. AEROSPACE & ELECTRONIC SYS., VOL. 48, NO. 1, pp. 707-728, JAN. 2012

# Forward Backward Probability Hypothesis Density Smoothing

Ronald. P. S. Mahler*, Ba Tuong Vo†, Ba Ngu Vo‡

**Abstract**

A forward-backward Probability Hypothesis Density (PHD) smoother involving forward filtering followed by backward smoothing is proposed. 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. In addition, assuming the previous iterate is Poisson, the cardinality distribution and all moments of the backward-smoothed multi-target density are derived. It is also shown that PHD smoothing alone does not necessarily improve cardinality estimation. Using an appropriate particle implementation we present a number of experiments to investigate the ability of the proposed multi-target smoother to correct state as well as cardinality errors.

**Index Terms**

Filtering and Smoothing, tracking, random sets, point processes, finite set statistics, SMC-PHD.

**SP-EDICS: MLR-BAYL** Bayesian signal processing, **SSP-FILT** Filtering, **SSP-TRAC** Tracking algorithms, **SSP-APPL** Applications of statistical signal processing techniques

---

Acknowledgement: This work is supported by the Australian Research Council.
* Address: Advanced Technology Group, Lockheed Martin MS2 Tactical Systems, Eagan, Minnesota E-mail: ronald.p.mahler@lmco.com Tel: +1 651 456 4819 Fax: +1 651 456 3098
† Corresponding author. Address: School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Crawley, WA 6009, Australia. E-mail: ba-tuong.vo@uwa.edu.au Tel: +61 8 6488 1767 Fax: +61 8 6488 1065
‡ Address: Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia. E-mail: bnvo@unimelb.edu.au Tel: +61 3 8344 6693 Fax: +61 3 8344 6678

# I. INTRODUCTION

Filtering, smoothing and prediction are three important interrelated problems in stochastic estimation, which essentially amount to calculating

$$p_{k|l}(x_k|z_{1:l}) \tag{1}$$

the probability density of the state $x_k$ at time $k$ given the observation history $z_{1:l} = (z_1, \ldots, z_l)$ up to time $l$. Smoothing, filtering and prediction, respectively, refer to the cases $l > k$, $l = k$, and $l < k$. In filtering the objective is to recursively estimate the current state given the observation history up to the current time $k$. Smoothing can yield significantly better estimates than filtering by delaying the decision time ($k$) and using data at a later time ($l > k$) [24], [13].

The text [1] provides a comprehensive coverage of closed-form smoothing solutions for linear Gaussian models. For non-linear non-Gaussian models [17] proposed the first Sequential Monte Carlo (SMC) implementation of the *smoothing while filtering* scheme or the *filter-smoother* by extending the standard particle filter. In the *forward-backward smoothing* scheme, involving forward filtering followed by backward smoothing, SMC implementations are proposed in [15], [7], [12]. A block-based particle smoothing method was proposed in [11]. In [3] the two-filter smoother of [2], [16] was generalized and efficient SMC implementations were proposed. A SMC smoother that is linear in complexity and does not suffer from particle depletion as in [17] has recently been proposed in [10].

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 [19] 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 [14] 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 [18] to improve the tracking of maneuvering targets. These techniques are proposed for linear Gaussian models, but can be extended to non-linear non-Gaussian models via techniques such as SMC, linearization, and unscented transforms.

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 [21], [34]. Indeed, the computational intractability is more severe in multi-target smoothing than filtering. The PHD filter [20], [21] 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. Its efficiency and performance suggest that smoothing with the PHD offers a tractable solution to non-linear non-Gaussian multi-target smoothing. The first effort to solve the multi-target smoothing problem via the PHD framework was reported in [25], where a forward-backward smoothing scheme is employed. Mahler's PHD filter [20] is the natural choice for the forward filter. Based on the bin occupancy argument of [9], an approximate backward PHD smoother under Poisson assumptions was proposed in [25]. An SMC implementation adapted from [30] and [7] was presented and simulation results were given.

Inspired by the effort in [25], we derive, using rigorous mathematical arguments, the first correct backward PHD smoother. Moreover, our backward PHD recursion is exact and does not require that the smoothed PHD at the previous iteration to be Poisson. In addition, under Poisson assumptions, all moments as well as the cardinality distribution of the backward-smoothed multi-target density are derived. It is also shown that PHD smoothing alone does not necessarily improve cardinality estimation. The mathematical tools used in our derivations are Mahler's Finite Set Statistics (FISST) [20], [21] and the celebrated Campbell's theorem from point process theory [6], [29]. Using an SMC implementation adapted from [30] and [7], we present a number of experiments to investigate the ability of the proposed PHD smoother to correct for state errors and cardinality errors.

In section II, we present the necessary mathematical tools from random set theory for the development of the main results. A rigorous derivation of the PHD backward smoothing recursion is presented in Section III while the derivation of the smoothed cardinality distribution and moments are presented in Section IV. Details of the sequential Monte Carlo implementation and case studies are presented in Sections V and VI.

## II. 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 [21] from a FISST perspective or in [34] from a point process perspective. For simplicity the following notation is adopted throughout the paper:

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

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

### A. Random finite set and the Bayes multi-target filter

Suppose at time $k$ there are $M(k)$ targets with states $x_{k,1}, \dots, x_{k,M(k)}$ each taking values in a state space $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, and $N(k)$ measurements $z_{k,1}, \dots, z_{k,N(k)}$ each taking values in an observation space

$\mathcal{Z} \subseteq \mathbb{R}^{n_z}$, where $\mathbb{R}^n$ denotes the $n$th Cartesian product of $\mathbb{R}$. Then, the *multi-target state* $X_k$ and the *multi-target measurement* $Z_k$, at time $k$, are defined as

$$X_k = \{x_{k,1}, \dots, x_{k,N(k)}\} \in \mathcal{F}(\mathcal{X}),$$

$$Z_k = \{z_{k,1}, \dots, z_{k,M(k)}\} \in \mathcal{F}(\mathcal{Z}),$$

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 [20], [21], 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 [20], [21]

$$p_{k+1|k}(X_{k+1}|Z_{1:k}) = \int f_{k+1|k}(X_{k+1}|X) p_{k|k}(X|Z_{1:k}) \delta X, \tag{2}$$

$$p_{k+1|k+1}(X_{k+1}|Z_{1:k+1}) = \frac{g_{k+1}(Z_{k+1}|X_{k+1}) p_{k+1|k}(X_{k+1}|Z_{1:k})}{\int g_{k+1}(Z_{k+1}|X) p_{k+1|k}(X|Z_{1:k}) \delta X}, \tag{3}$$

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

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

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

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 [21].

The multi-target Bayes filter is generally intractable [20], [21], [34] and it is necessary to resort to more tractable approximations. The Probability Hypothesis Density (PHD) filter [20] is a first moment approximation to the full multi-target Bayes filter (2)-(3), which operates on the (single-target) state space $\mathcal{X}$.

### *B. 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 [6], [29], [20]. 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.

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

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

$$\mathbb{E}\left[f(X)\right] = \int f(X)p(X)\delta X.$$

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 [23].

Campbell's theorem relates certain types of expectation of an RFS to its PHD, and is an important result in point process theory [29]. Campbell's theorem (see [6], [29]) states that for an RFS $X$ on $\mathcal{X}$ with PHD (or intensity) $v$

$$\mathbb{E}\left[\sum_{x\in X} \zeta(x)\right] = \langle \zeta, v\rangle\,. \tag{4}$$

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

### *C. 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 (2) and update (3). Let

$$v_{k|k} = \text{filtered (updated) PHD at time } k$$

$$v_{k+1|k} = \text{predicted PHD from time } k \text{ to } k+1$$

$$\gamma = \gamma_{k+1|k} = \text{PHD of birth at time } k+1$$

$$f = f_{k+1|k} = \text{single target transition from time } k \text{ to } k+1$$

$$p_S = p_{S,k+1|k} = \text{probability of survival from } k \text{ to } k+1$$

$$g = g_{k+1} = \text{single target likelihood at time } k+1$$

$$p_D = p_{D,k+1} = \text{probability of detection at time } k+1$$

$$\kappa_{k+1} = \text{intensity function of clutter at time } k+1$$

(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

$$v_{k+1|k}(x) = \gamma(x) + \left\langle v_{k|k} p_S, f(x|\cdot) \right\rangle, \tag{5}$$

$$v_{k+1|k+1}(x) = [1 - p_D(x)] v_{k+1|k}(x) + \sum_{z \in Z_{k+1}} \frac{p_D(x) g(z|x) v_{k+1|k}(x)}{\kappa_{k+1}(z) + \left\langle p_D g(z|\cdot), v_{k+1|k} \right\rangle}. \tag{6}$$

Sequential Monte Carlo (SMC) implementations of the PHD recursion [27], [30], [35], closed form solutions [31], as well as generalizations [22], [32] have opened the door to numerous novel extensions and applications.

### D. Probability generating functionals (PGFl)

Apart from the probability density, the probability generating functional (PGFl) is another fundamental descriptor of an RFS. Following [6], [29], the *probability generating functional* (PGFl) $G[\cdot]$ of an RFS $X$ on $\mathcal{X}$ is defined by

$$G[h] \equiv \mathbb{E}[h^X], \tag{7}$$

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

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

The $i$th functional derivative of the PGFl is linear in the each of the directions $\zeta_1, ..., \zeta_i$. For more technical details on functional derivatives of the PGFl, we refer the reader to [28] and the references therein.

It was noted in [34] that a linear functional in each of the variables $\zeta_1, ..., \zeta_i$ can be identified with a measure $\mu$ 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 $\mu$ as a functional that takes the functions $\zeta_1, ..., \zeta_i$ to the reals. If the measure $\mu$ admits a density $f$ then,

$$\mu[\zeta_1, ..., \zeta_i] = \int ... \int \zeta_1(x_1), ..., \zeta_i(x_i) f(x_1, .., x_i) dx_1 ... dx_i \tag{8}$$

and the rather suggestive notation $\mu[\delta_{x_1}, ..., \delta_{x_i}] \equiv f(x_1, ..., x_i)$ can be used, where $\delta_x$ 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 [20]. It follows from [6] that the multi-target density $p$ and the PHD $v$ (if they exist) can be recovered from the PGFl by set differentiation

$$p(X) = \left.\frac{\partial}{\partial X}\right|_{h=0} G[h], \tag{9}$$

$$v(x) = \left.\frac{\partial}{\partial x}\right|_{h=1} G[h]. \tag{10}$$

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 $\rho$ (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 $v$ [6], [29]. 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

$$G[h] = \exp(\langle v, h - 1\rangle). \tag{11}$$

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 $\mathcal{X}$), [21]. The PGFl of a multi-Bernoulli RFS is given by

$$G[h] = \prod_{i=1}^{M} \left(1 - r^{(i)} + r^{(i)} \langle p^{(i)}, h \rangle \right). \tag{12}$$

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 $i$th object while $p^{(i)}$ is the probability density of the state conditional on its existence.

Examples involving Poisson and multi-Bernoulli RFSs are those associated with multi-target Markov transitions and multi-target likelihood functions. 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 $\gamma$ 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

$$\begin{aligned} G_{k+1|k}[h|X] &= \exp(\langle \gamma, h-1 \rangle) \prod_{x \in X} (1 - p_S(x) + p_S(x) \langle f(:|x), h(:) \rangle) \\ &= \exp(\langle \gamma, h-1 \rangle) \left(1 - p_S + p_S \langle f(:|\cdot), h(:) \rangle \right)^X . \end{aligned} \tag{13}$$

Similarly, the multi-target measurement is modelled by the union of a Poisson clutter RFS with intensity $\kappa$ 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

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

### *E. Factorial moment measures*

Factorial moment measures are a useful generalization of the PHD to higher order moments of an RFS. To define the factorial moment measures, we treat a measure $\mu$ as a functional that takes a function $g$ on $\mathcal{X}^i$ to the real line via:

$$\mu[g] = \int g(y_1, ..., y_i) \mu(dy_1, ..., dy_i).$$

In this sense, the PHD $v$ can be treated as a measure on $\mathcal{X}$ via $v[g] = \langle g, v \rangle$

The $i$th *factorial moment measure* $\phi^{(i)}$ of an RFS $Y$ on $\mathcal{X}$ is a measure on $\mathcal{X}^i$ defined by (see [29] pp. 111)

$$\phi^{(i)}[g] = \mathbb{E}\left[ \sum_{y_1 \neq y_2 \neq ... \neq y_i \in Y} g(y_1, ..., y_i) \right].$$

Note that the PHD measure $v[\cdot]$ is the first factorial moment measure. The *product density* $\varrho^{(i)}$ of a factorial moment measure $\phi^{(i)}$ is the density (if exists) w.r.t. the Lebesgue measure, i.e.

$$\phi^{(i)}[g] = \int g(y_1, ..., y_i)\varrho^{(i)}(y_1, ..., y_i)dy_1...dy_i.$$

Note that the PHD (function) $v(\cdot)$ is the product density of the first factorial moment measure. For a Poisson RFS with PHD (or intensity function) $v$, (see [ [29] pp. 44])

$$\varrho^{(i)}(y_1, ..., y_i) = v(y_1)...v(y_i)$$

and so

$$\begin{aligned}\int \sum_{y_1 \neq y_2 \neq ... \neq y_i \in Y} g(y_1, ..., y_i)p(Y)\delta Y &= \mathbb{E}\left[\sum_{y_1 \neq y_2 \neq ... \neq y_i \in Y} g(y_1, ..., y_i)\right] \\ &= \int g(y_1, ..., y_i)v(y_1)...v(y_i)dy_1...dy_i. \end{aligned} \tag{14}$$

We will call on the above formula for the derivations of the smoothed moments and cardinality.

## III. 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 [16]). In the multi-target case, the multi-target posterior is propagated forward to time $k$ via the multi-target Bayes recursion (2)-(3) and the smoothed multi-target density is propagated backward, from time $k$ to time $k' < k$, via the multi-target backward smoothing recursion

$$p_{k'|k}(X) = p_{k'|k'}(X) \int f_{k'+1|k'}(Y|X) \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y. \tag{15}$$

While the FISST multi-target density is not a probability density [21], the recursion (15) has the same form as the standard backward smoother expressed in terms of probability densities [16]. A simple way to derive (15) 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 [30].

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 PHD forward propagation is accomplished by the PHD recursion (5)-(6). 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*

$$v_{k'|k}(x) = v_{k'|k'}(x)\left(1 - p_S(x) + p_S(x)\left\langle \frac{f(\cdot|x)}{v_{k'+1|k'}}, v_{k'+1|k}\right\rangle\right). \tag{16}$$

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

$$\left.\frac{\partial}{\partial Y}\right|_{g=0} \exp(\langle \alpha, g\rangle)\left(\langle \beta, g\rangle + c\right) = \left(c + \sum_{y\in Y}\frac{\beta(y)}{\alpha(y)}\right)\alpha^Y. \tag{17}$$

This is a special case of Lemma 2 in Section IV.

**Proof of Proposition 1:** For the smoothed multi-target state with density $p_{k'|k}$ given by (15), the PGFl is

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

Substituting (13) for $G_{k'+1|k'}[g|X]$, gives

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

Using the assumption that the filtered multi-target state is Poisson i.e. $G_{k'|k'}[h] = \exp(\langle v_{k'|k'}, h-1\rangle)$ gives (18).

$$\begin{aligned}
G_{k'|k}[h] = \int \left.\frac{\partial}{\partial Y}\right|_{g=0} &\exp(\langle v_{k'|k'}, h\left(1 - p_S + p_S\left\langle f(:|\cdot), g(:)\right\rangle\right) - 1\rangle) \\
&\times \exp(\langle \gamma, g-1\rangle)\frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y.
\end{aligned} \tag{18}$$

The smoothed PHD can be obtained by differentiating the PGFl in (18):

$$
\begin{aligned}
v_{k'|k}(x) =& \left.\frac{\partial}{\partial x}\right|_{h=1} G_{k'|k}[h] \\
=& \left.\frac{\partial}{\partial x}\right|_{h=1} \int \left.\frac{\partial}{\partial Y}\right|_{g=0} \exp(\langle \gamma, g-1 \rangle) \\
&\times \exp(\left\langle v_{k'|k'}, h\left(1-p_S+p_S\left\langle f(:|\cdot), g(:)\right\rangle\right)-1 \right\rangle) \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \\
=& \int \left.\frac{\partial}{\partial Y}\right|_{g=0} \exp(\langle \gamma, g-1 \rangle) \\
&\times \left.\frac{\partial}{\partial x}\right|_{h=1} \exp(\left\langle v_{k'|k'}, h\left(1-p_S+p_S\left\langle f(:|\cdot), g(:)\right\rangle\right)-1 \right\rangle) \\
&\times \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y
\end{aligned} \tag{19}
$$

Now, consider the derivative w.r.t. $x$ in (19)

$$
\begin{aligned}
&\left.\frac{\partial}{\partial x}\right|_{h=1} \exp(\left\langle v_{k'|k'}, h\left(1-p_S+p_S\left\langle f(:|\cdot), g(:)\right\rangle\right)-1 \right\rangle) \\
&= \exp(\left\langle v_{k'|k'}, \left(1-p_S+p_S\left\langle f(:|\cdot), g(:)\right\rangle\right)-1 \right\rangle) \\
&\times v_{k'|k'}(x)\left(1-p_S(x)+p_S(x)\left\langle f(:|x), g(:)\right\rangle\right).
\end{aligned} \tag{20}
$$

The exponent in the RHS of (20) can be rearranged by changing the order of integration as follows

$$
\begin{aligned}
&\left\langle v_{k'|k'}, \left(1-p_S+p_S\left\langle f(:|\cdot), g(:)\right\rangle\right)-1 \right\rangle \\
&= \left\langle v_{k'|k'}(\cdot), p_S(\cdot)\left\langle f(:|\cdot), g(:)\right\rangle \right\rangle - \left\langle v_{k'|k'}, p_S \right\rangle \\
&= \left\langle \left\langle v_{k'|k'}(\cdot) p_S(\cdot), f(:|\cdot) \right\rangle, g(:) \right\rangle - \left\langle \left\langle v_{k'|k'}(\cdot) p_S(\cdot) f(:|\cdot) \right\rangle, 1 \right\rangle \\
&= \left\langle \left\langle v_{k'|k'}(\cdot) p_S(\cdot), f(:|\cdot) \right\rangle, g(:)-1 \right\rangle \\
&= \left\langle v_{k'+1|k'} - \gamma, g-1 \right\rangle.
\end{aligned} \tag{21}
$$

and hence

$$
\begin{aligned}
&\exp(\langle \gamma, g-1 \rangle) \left.\frac{\partial}{\partial x}\right|_{h=1} \exp(\left\langle v_{k'|k'}, h\left(1-p_S+p_S\left\langle f(:|\cdot), g(:)\right\rangle\right)-1 \right\rangle) \\
&= \exp(\langle \gamma, g-1 \rangle) \exp(\left\langle v_{k'+1|k'} - \gamma, g-1 \right\rangle) v_{k'|k'}(x)\left(1-p_S(x)+p_S(x)\left\langle f(:|x), g(:)\right\rangle\right) \\
&= \exp(\left\langle v_{k'+1|k'}, g-1 \right\rangle) v_{k'|k'}(x)\left(1-p_S(x)+p_S(x)\left\langle f(:|x), g(:)\right\rangle\right).
\end{aligned} \tag{22}
$$

Thus, substituting (22) into (19) gives

$$
\begin{aligned}
v_{k'|k}(x) = v_{k'|k'}(x) \int \left.\frac{\partial}{\partial Y}\right|_{g=0} \exp(\left\langle v_{k'+1|k'}, g-1 \right\rangle) \\
\times \left(1-p_S(x)+p_S(x)\left\langle f(:|x), g(:)\right\rangle\right)
\end{aligned}
$$

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

Taking $\exp(-\langle v_{k'+1|k'}, 1\rangle)$ 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

$$v_{k'|k}(x) = v_{k'|k'}(x)\exp(-\langle v_{k'+1|k'}, 1\rangle) \int \frac{\partial}{\partial Y}\bigg|_{g=0} \exp(\langle v_{k'+1|k'}, g\rangle)$$
$$\times (1 - p_S(x) + p_S(x)\langle f(:|x), g(:)\rangle)$$
$$\times \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y$$

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

$$v_{k'|k}(x) = v_{k'|k'}(x)\int\left(1 - p_S(x) + p_S(x)\sum_{y\in Y}\frac{f(y|x)}{v_{k'+1|k'}(y)}\right)$$
$$\times \frac{v^Y_{k'+1|k'}\exp(-\langle v_{k'+1|k'}, 1\rangle)}{p_{k'+1|k'}(Y)}p_{k'+1|k}(Y)\delta Y$$
$$= v_{k'|k'}(x)\int\left(1 - p_S(x) + p_S(x)\sum_{y\in Y}\frac{f(y|x)}{v_{k'+1|k'}(y)}\right)$$
$$\times p_{k'+1|k}(Y)\delta Y.$$

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

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

## IV. Smoothed Cardinality distribution and moments

In the previous section we made no assumption on the form of the smoothed multi-target density $p_{k'+1|k}$ to calculate the smoothed PHD $v_{k'|k}$. If we further assume that smoothed multi-target density $p_{k'+1|k}$ is Poisson, then the cardinality distribution and cardinality moments of the smoothed multi-target density $p_{k'|k}$ can also be calculated in closed form.

### *A. Smoothed cardinality distribution*

**Proposition 2**: *Under the premises of Proposition 1, if the smoothed multi-target density $p_{k'+1|k}$ is Poisson, then the smoothed cardinality $\rho_{k'|k}$ is given by*

$$\rho_{k'|k}(n) = \exp\left(\left\langle\frac{\gamma v_{k'+1|k}}{v_{k'+1|k'}} - v_{k'+1|k} + v_{k'+1|k'} - \gamma - v_{k'|k'}, 1\right\rangle\right)$$
$$\times \sum_{i=0}^{n}\frac{\langle v_{k'|k'}, 1 - p_S\rangle^{n-i}}{(n-i)!}\left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}}\right\rangle^i.$$

see appendix for proof.

### *B. Moments of the smoothed cardinality*

**Proposition 3**: *Under the premises of Proposition 1, if the smoothed multi-target density* $p_{k'+1|k}$ *is Poisson, then the* $n$*th moment of the smoothed cardinality is*

$$G^{(n)}_{k'|k} = n! \sum_{i=0}^{n} \frac{\left\langle v_{k'|k'}, 1-p_S \right\rangle^{n-i}}{(n-i)!} \left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}} \right\rangle^{i}$$

see appendix for proof.

Two special cases are the 1st and 2nd moments:

$$G^{(1)}_{k'|k} = \left\langle v_{k'|k'}, 1-p_S \right\rangle + \left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}} \right\rangle^{i},$$

$$G^{(2)}_{k'|k} = \left\langle v_{k'|k'}, 1-p_S \right\rangle^{2} + 2\left\langle v_{k'|k'}, 1-p_S \right\rangle \left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}} \right\rangle + 2\left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}} \right\rangle^{2}.$$

The smoothed cardinality variance is then

$$\begin{aligned} var(N_{k'|k}) &= G^{(2)}_{k'|k} + G^{(1)}_{k'|k} - \left(G^{(1)}_{k'|k}\right)^{2} \\ &= \left\langle v_{k'|k'}, 1-p_S \right\rangle + \left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}} \right\rangle + \left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}} \right\rangle^{2}. \end{aligned}$$

Note that since the smoothed RFS density $p_{k'+1|k}$ is Poisson, $var(\hat{N}_{k'+1|k}) = \left\langle v_{k'+1|k}, 1 \right\rangle$ and hence

$$\begin{aligned} var(N_{k'|k}) \;\; &= \;\; var(N_{k'+1|k}) + \left\langle v_{k'|k'}, 1-p_S \right\rangle \\ &+ \left\langle \frac{v_{k'+1|k}}{1 + \frac{\gamma}{\langle v_{k|k}(:)p_S(:), f(\cdot|:)\rangle}}, 1 \right\rangle^{2} - \left\langle \frac{v_{k'+1|k}}{1 + \frac{\langle v_{k|k}(:)p_S(:), f(\cdot|:)\rangle}{\gamma}}, 1 \right\rangle. \end{aligned} \quad (23)$$

## V. SEQUENTIAL MONTE CARLO IMPLEMENTATION

A review of the SMC implementation of the PHD forward filter from [30], and a corresponding SMC implementation of the backward smoother is presented. This implementation can be easily extended to the Gaussian particle approach proposed in [5]. A special resampling scheme is also proposed to mitigate the effects of particle depletion (the situation that all but one of the importance weights are close to zero [7]) specifically for use with the forward-backward smoother.

### *A. Forward Filter*

Suppose that the posterior PHD at time $k-1$ is of the form

$$v_{k-1|k-1}(x) = \sum_{i=1}^{L_{k-1}} w^{(i)}_{k-1|k-1} \delta_{x^{(i)}_{k-1|k-1}}(x),$$

then, given importance (or proposal) densities $q_{k|k}(\cdot|x_{k-1}, Z_k)$ and $b_{k|k-1}(\cdot|Z_k)$, along with samples $x^{(i)}_{P,k|k-1} \sim q_{k|k}(\cdot|x^{(i)}_{k|k-1}, Z_k)$ (for $i = 1, ..., L_{k-1}$) and $x^{(j)}_{\gamma,k|k-1} \sim b_{k|k-1}(\cdot|Z_k)$ (for $j = 1, ..., L_{\gamma,k}$), the predicted PHD $v_{k|k-1}$ can be computed as

$$v_{k|k-1}(x) = \sum_{i=1}^{L_{k-1}} w^{(i)}_{P,k|k-1}\delta_{x^{(i)}_{P,k|k-1}}(x) + \sum_{j=1}^{L_{\gamma,k}} w^{(j)}_{\gamma,k|k-1}\delta_{x^{(j)}_{\gamma,k|k-1}}(x),$$

where

$$w^{(i)}_{P,k|k-1} = \frac{w^{(i)}_{k-1|k-1} p_{S,k|k-1}(x^{(i)}_{k-1|k-1}) f_{k|k-1}(x^{(i)}_{P,k|k-1}|x^{(i)}_{k|k-1})}{q_{k|k}(x^{(i)}_{P,k|k-1}|x^{(i)}_{k-1|k-1}, Z_k)},$$

$$w^{(j)}_{\gamma,k|k-1} = \frac{\gamma_{k|k-1}(x^{(j)}_{\gamma,k|k-1})}{N_{\gamma,k} b_{k|k-1}(x^{(j)}_{\gamma,k|k-1}|Z_k)}.$$

Suppose that the predicted PHD at time $k$ is of the form

$$v_{k|k-1}(x) = \sum_{j=1}^{L_{k|k-1}} w^{(j)}_{k|k-1}\delta_{x^{(j)}_{k|k-1}}(x),$$

then the posterior PHD at time $k$ is given by

$$v_{k|k}(x) = \sum_{j=1}^{L_{k|k-1}} \left(w^{(j)}_{M,k|k} + w^{(j)}_{D,k|k}\right)\delta_{x^{(j)}_{k|k-1}}(x),$$

where

$$w^{(j)}_{M,k|k} = w^{(j)}_{k|k-1}(1 - p_{D,k}(x^{(j)}_{k|k-1})),$$

$$w^{(j)}_{D,k|k} = \sum_{z\in Z_k} \frac{w^{(j)}_{k|k-1} p_{D,k}(x^{(j)}_{k|k-1}) g_k(z|x^{(j)}_{k|k-1})}{\kappa_k(z) + \sum_{\ell=1}^{L_{k|k-1}} w^{(\ell)}_{k|k-1} p_{D,k}(x^{(\ell)}_{k|k-1}) g_k(z|x^{(\ell)}_{k|k-1})}.$$

### *B. Backward Smoother*

Suppose that the filtered PHD at time $k'$ and smoothed PHD at time $k'+1$ from time $k$ are given respectively by the weighted samples

$$v_{k'|k'}(x) = \sum_{i=1}^{L_{k'}} w^{(i)}_{k'|k'}\delta_{x^{(i)}_{k'|k'}}(x),$$

$$v_{k'+1|k}(x) = \sum_{j=1}^{L_{k'+1}} w^{(j)}_{k'+1|k}\delta_{x^{(j)}_{k'+1|k}}(x).$$

Then, it can be easily shown that the smoothed PHD at time $k'$ from time $k$ is given by the reweighted samples

$$v_{k'|k}(x) = \sum_{i=1}^{L_{k'}} w^{(i)}_{k'|k}\delta_{x^{(i)}_{k'|k'}}(x)$$

where

$$w_{k'|k}^{(i)} = w_{k'|k'}^{(i)}\Big(1 - p_{S,k'+1|k'}(x_{k'|k'}^{(i)}) + \sum_{j=1}^{L_{k'+1}} \frac{w_{k'+1|k}^{(j)} p_{S,k'+1|k'}(x_{k'|k'}^{(i)}) f_{k'+1|k'}(x_{k'+1|k}^{(j)}|x_{k'|k'}^{(i)})}{\gamma_{k'+1|k'}(x_{k'+1|k}^{(j)}) + \sum_{\ell=1}^{L_{k'}} w_{k'|k'}^{(\ell)} p_{S,k'+1|k'}(x_{k'|k'}^{(\ell)}) f_{k'+1|k'}(x_{k'+1|k}^{(j)}|x_{k'|k'}^{(\ell)})}\Big)$$

Notice that in this backward recursion, no new samples are produced; there is only a reweighting of the samples of the previously filtered result at each backward time step.

### C. Resampling

To mitigate the effects of particle depletion, resampling is also required in the forward backward smoother. Suppose that the posterior PHD at time $k$ is

$$v_{k|k}(x) = \sum_{j=1}^{L_{k|k}} w_{k|k}^{(j)} \delta_{x_{k|k}^{(j)}}(x),$$

where $w_{k|k}^{(j)} = w_{M,k|k}^{(j)} + w_{D,k|k}^{(j)}$ comprises weights corresponding to missed detections and measurement returns respectively. In the standard multinomial resampling technique, particles are resampled directly in proportion to the forward filter weights and directly from the forward filter samples. With this approach however, the particle population tends to be dominated by the weights of measurement updated terms, and consequently, particles whose weights are mainly derived from missed detection terms will rarely be selected for resampling. In practice, this can be problematic as track losses are more likely to be incurred, since the state space will be almost completely depleted of particles in regions (possibly occupied by targets) where the posterior has assigned relatively low weights (for whatever reason). To circumvent this issue, an adaptive resampling strategy is proposed in which a chosen number of particles are resampled separately from the missed detection and measurement updated contributions respectively and then recombined. Specifically, the strategy is to resample $L_{M,k}$ samples $\left\{x_{k|k}^{(m)}\right\}_{m\in I_{M,k}}$ from the population corresponding to missed detections $\left\{w_{M,k|k}^{(j)}, x_{k|k}^{(j)}\right\}_{j=1}^{L_{k|k}}$ as well as $L_{D,k}$ samples $\left\{x_{k|k}^{(\ell)}\right\}_{\ell\in I_{D,k}}$ from the population corresponding to measurement updates$\left\{w_{D,k|k}^{(j)}, x_{k|k}^{(j)}\right\}_{j=1}^{L_{k|k}}$. The resampled approximation to the posterior PHD is then given by

$$v_{k|k}(x) = w_{M,k|k} \sum_{m\in I_{M,k}} \delta_{x_{k|k}^{(m)}}(x) + w_{D,k|k} \sum_{\ell\in I_{D,k}} \delta_{x_{k|k}^{(\ell)}}(x),$$

where $w_{M,k|k} = \sum_{m=1}^{L_{k|k}} w_{M,k|k}^{(m)}/L_{M,k}$ and $w_{D,k|k} = \sum_{\ell=1}^{L_{k|k}} w_{D,k|k}^{(\ell)}/L_{D,k}$. Furthermore, as the number of targets is time varying, it is necessary to adaptively allocate the number of particles present at each time step. To ensure a sufficient numbers, the number of samples allocated at each time step should be selected in proportion to the expected number of targets, i.e. choose $L_{M,k} + L_{D,k} = \left\lfloor \rho \hat{N}_{k|k} \right\rceil$ for some $\rho > 0$.

## VI. EXPERIMENTS

This section presents numerical studies to verify the performance of the proposed PHD forward-backward smoother and to investigate its advantages and disadvantages over the PHD filter. Two examples are presented; the first is a typical multiple target tracking scenario involving various births and deaths while the second is a simplified scenario to specifically investigate the effect of missed detections and false alarms on filter and smoother performance. The previously described SMC implementations of the forward filter and backward smoother are used. Resampling is performed as per the adapted strategy described in the previous section. State extraction is performed via $k$-means based partitioning of the PHD samples to estimate the constituent population centres.

For performance evaluation, the Optimal SubPattern Assignment (OSPA) metric [26] is used to jointly capture, in a mathematically consistent yet intuitively meaningful way, the difference in the cardinalities and individual elements of two finite sets. An intuitive construction of the OSPA distance between two finite sets $X = \{x_1, \ldots, x_m\}$ and $Y = \{y_1, \ldots, y_n\}$ can be read as follows. The set $X$ with the smaller cardinality is initially chosen as a reference. Determine the assignment between the $m$ points of $X$ and points of $Y$, that minimizes the sum of the distances, subject to the constraint that distances are capped at a preselected maximum or cut-off value $c$. This minimized sum of distances can be interpreted as the "total localization error", which are assigned by giving the points in $X$ the "benefit of the doubt". All other points which remain unassigned are also penalized with a maximum error value of $c$. These errors can interpreted as "cardinality errors" which are "penalized at the maximum rate". The "total error" committed is then the sum of the "total localization error" and the "total cardinality error". Remarkably, the "per target error", obtained by normalizing "total error" by $n$, (the larger cardinality of the two given sets) is a proper metric [26]. In other words the "per target error" enjoys all the properties of the usual distance that we normally take for granted on a Euclidean space. The OSPA metric $\bar{d}_p^{(c)}$ is formally defined as follows. Let $d^{(c)}(x, y) := \min(c, \|x - y\|)$ for $x, y \in \mathcal{X}$, and $\Pi_k$ denote the set of permutations on $\{1, 2, \ldots, k\}$ for any positive integer $k$. Then, for $p \geq 1$, $c > 0$, and $X = \{x_1, \ldots, x_m\}$ and $Y = \{y_1, \ldots, y_n\}$,

$$\bar{d}_p^{(c)}(X,Y) := \left( \frac{1}{n} \left( \min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}(x_i, y_{\pi(i)})^p + c^p(n-m) \right) \right)^{\frac{1}{p}} \tag{24}$$

if $m \leq n$, and $\bar{d}_p^{(c)}(X, Y) := \bar{d}_p^{(c)}(Y, X)$ if $m > n$; and $\bar{d}_p^{(c)}(X, Y) = \bar{d}_p^{(c)}(Y, X) = 0$ if $m = n = 0$. The OSPA distance is thus interpreted as a $p$-th order per-target error, comprised of a $p$-th order per-target localization error and a $p$-th order per-target cardinality error. The order parameter $p$ determines the sensitivity of the metric to outliers, and the cut-off parameter $c$ determines the relative weighting of the penalties assigned to cardinality and localization errors.

### A. Single Target Motion and Observation Model

The following single-target non-linear motion and observation models are used in these experiments. Targets follow a nearly constant turn model and are observed via noisy bearings and range measurements. Specifically, the single target state $x_k = [\ \tilde{x}_k^T, \omega_k\ ]^T$ comprises the planar position and velocity $\tilde{x}_k^T = [\ p_{x,k}, \dot{p}_{x,k}, p_{y,k}, \dot{p}_{y,k}\ ]^T$ and the turn rate $\omega_k$ while the single target observation $x_k = [\ r_k, \theta_k\ ]^T$ comprises the range and angle of arrival. The state transition model is

$$\tilde{x}_k = F(\omega_{k-1})\tilde{x}_{k-1} + Gw_{k-1}$$

$$\omega_k = \omega_{k-1} + \Delta u_{k-1}$$

where

$$F(\omega) = \begin{bmatrix} 1 & \frac{\sin\omega\Delta}{\omega} & 0 & -\frac{1-\cos\omega\Delta}{\omega} \\ 0 & \cos\omega\Delta & 0 & -\sin\omega\Delta \\ 0 & \frac{1-\cos\omega\Delta}{\omega} & 1 & \frac{\sin\omega\Delta}{\omega} \\ 0 & \sin\omega\Delta & 0 & \cos\omega\Delta \end{bmatrix}, G = \begin{bmatrix} \frac{\Delta^2}{2} & 0 \\ T & 0 \\ 0 & \frac{\Delta^2}{2} \\ 0 & \Delta \end{bmatrix},$$

$w_{k-1} \sim \mathcal{N}(\cdot; 0, \sigma_w^2 I)$ and $u_{k-1} \sim \mathcal{N}(\cdot; 0, \sigma_u^2 I)$ with $\Delta = 1s$. The measurement model is given by

$$z_k = \begin{bmatrix} \arctan(p_{x,k}/p_{y,k}) \\ \sqrt{p_{x,k}^2 + p_{y,k}^2} \end{bmatrix} + \varepsilon_k$$

where $\varepsilon_k \sim \mathcal{N}(\cdot; 0, R_k)$ with $R_k = \operatorname{diag}([\ \sigma_\theta^2, \sigma_r^2\ ]^T)$.

### B. Experiment 1

In this experiment, a typical multiple target tracking scenario is employed to verify the performance of the proposed PHD forward-backward smoother (with a lag of 5 time steps) and to compare with that of the PHD filter. A total of 5 targets appears on the scene at any one time, with various births and deaths throughout the 100 time step scenario. The observation region is the half disc of radius $2000m$. The true trajectories are shown in Fig. 1 along with the start and stop positions of each track. In SMC implementations, the transition is used for the proposal while resampling is performed at each time step with a $50/50$ apportionment of missed/measurement particles and with an average of 1000 particles per surviving target.

The model parameters are as follows. The motion and measurement models are set to $\sigma_w = 3m/s^2$, $\sigma_u = 0.1\pi/180rad/s$, $\sigma_\theta = 0.3\pi/180rad$, $\sigma_r = 0.7m$. The probability of target survival and detection are $p_{S,k}(x) = 0.99$ and $p_{D,k}(x) = 0.98$. The birth process has PHD given by the Gaussian

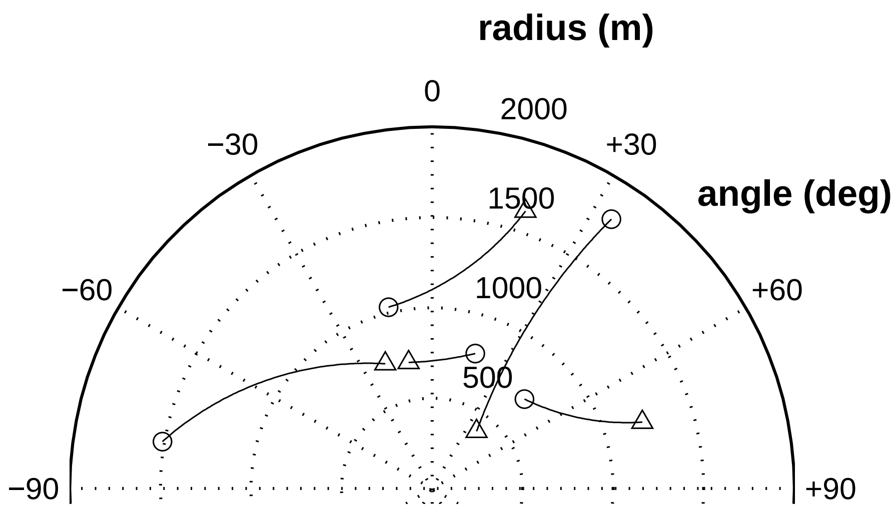


Fig. 1. Target trajectories in the $r\theta$ plane. Start/Stop positions for each track are shown with ○/△.

mixture $\gamma_k(x) = \sum_{i=1}^{5} w_b \mathcal{N}(x; m_b^{(i)}, P_b)$ where $w_b = 0.1$, $m_b^{(1)} = [\ -1500, 0, 250, 0\ \ 0\ ]^T$, $m_b^{(2)} = [\ -250, 0, 1000, 0\ \ 0\ ]^T$, $m_b^{(3)} = [\ 250, 0, 750, 0\ \ 0\ ]^T$, $m_b^{(4)} = [\ 1000, 0, 1500, 0\ \ 0\ ]^T$, $m_b^{(5)} = [\ 500, 0, 500, 0\ 0\ ]^T$, $P_b = \text{diag}([\ 10,\ 10,\ 10,\ 10,\ \pi/180\ ]^T)^2$. Clutter is Poisson with intensity $\lambda_c = 1.1 \times 10^{-3}\ (radm)^{-1}$ over the region $[0, \pi]rad \times [0, 2000]m$ (giving an average of 7 returns per scan).

The results of a sample run of the filter and smoother are shown in Figs. 2 and 3 with the true and estimated $x$ and $y$ positions versus time, and in Fig. 4 with a plot of true and estimated number of targets versus time. Both the filter and smoother are able to identify target births and deaths, and maintain target lock for the majority of each track. State estimates appear reasonable for both the filter and smoother although those for the smoother appear marginally better. Both the filter and smoother also produce reasonable estimates of the number of targets with the occasional dropped or false track.

To further investigate, 1000 Monte Carlo trials are performed for the filter and smoother on the same data. In Fig 5, the mean and standard deviation of the estimated number of targets are shown. The filtered and smoothed state estimates are evaluated using the Optimal Sub-Pattern Assignment (OSPA) metric [26] for $p = 1$ and $c = 100m$. In Fig. 6, the MC average of the estimated cardinality and OSPA miss distance for the filter and smoother is shown versus time. In Fig. 7, the MC average of the localization and cardinality components of OSPA distance are shown versus time. These results suggest that overall, the smoother slightly outperforms the filter in terms of the total miss distance. Furthermore, the smoother outperforms the filter in terms of localization error (an average improvement of $1.4m$ or 33%) but both are roughly on par in terms of cardinality error.

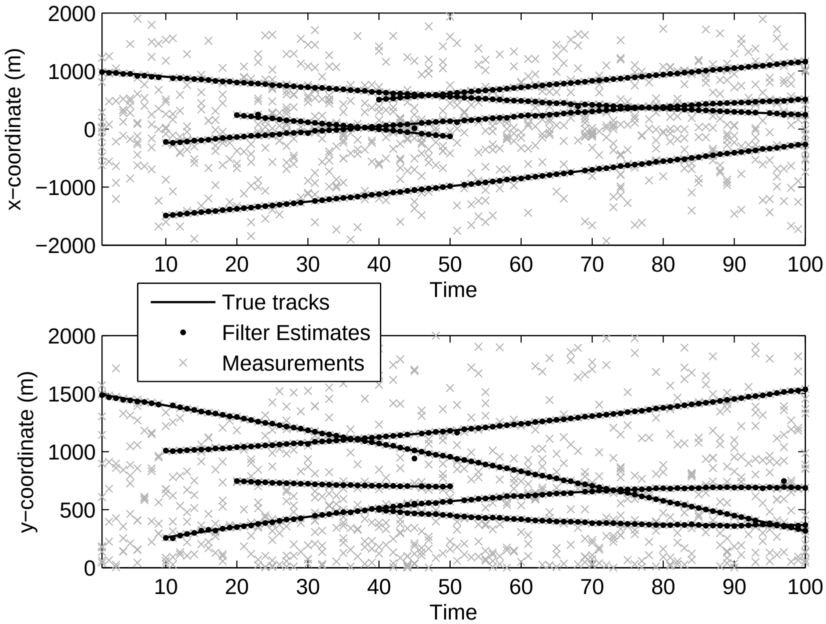


Fig. 2. Filtered state estimates versus time.

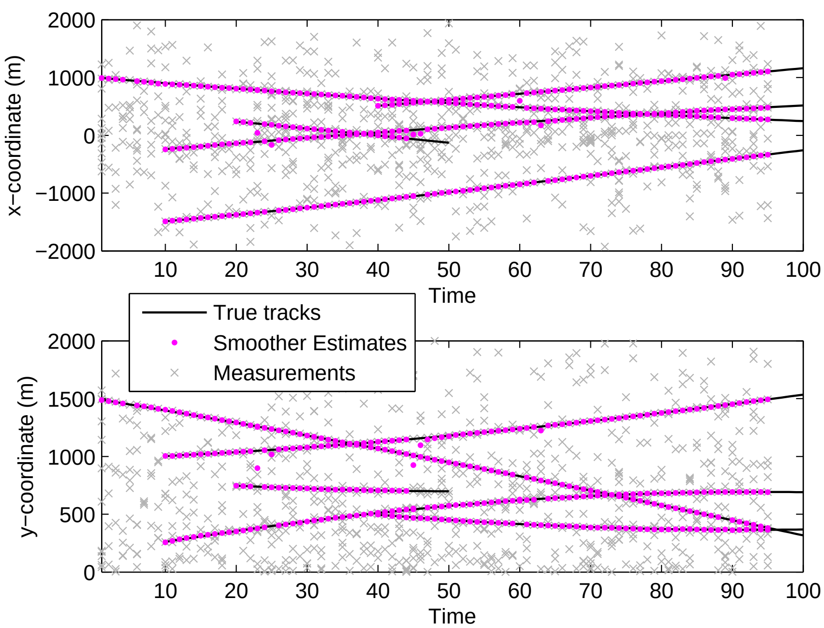


Fig. 3. Smoothed state estimates versus time.

The results for the smoother showed one apparent anomaly. There appears to be a noticeable rise in the cardinality error committed by the smoother between times $k = 46$ and $k = 50$, which as it turns out is due to the smoother prematurely dropping the track which later dies at time $k = 51$ (compare

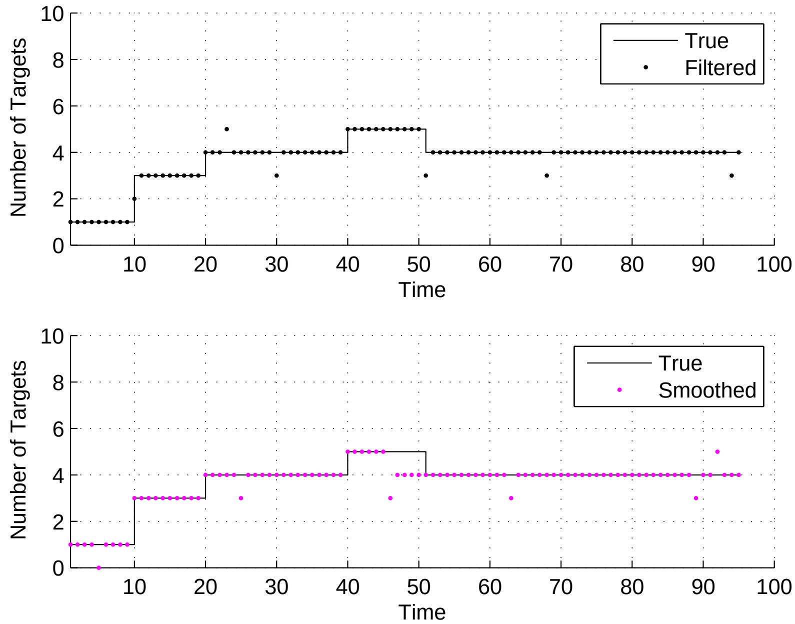


Fig. 4. Target number estimates for filter and smoother versus time.

with the single sample result in Fig 3). Moreover, the apparent coincidence of the $5$ time step difference between the actual and smoother-declared times of track death, with the smoother delay of $5$ time steps, is due to some undesirable behaviour in the smoother. This phenomenon is examined more closely in the following experiment.

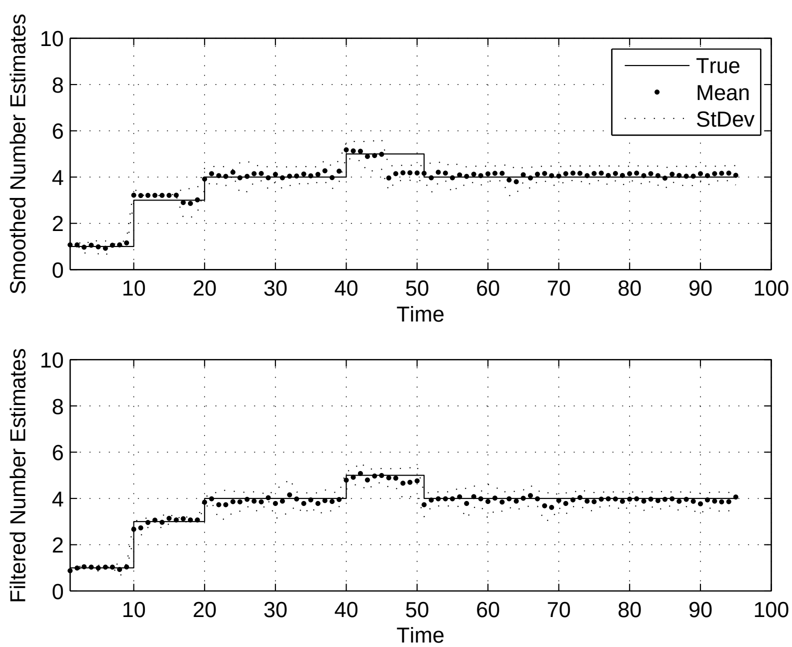


Fig. 5. Average cardinality statistics versus time for filter and smoother.

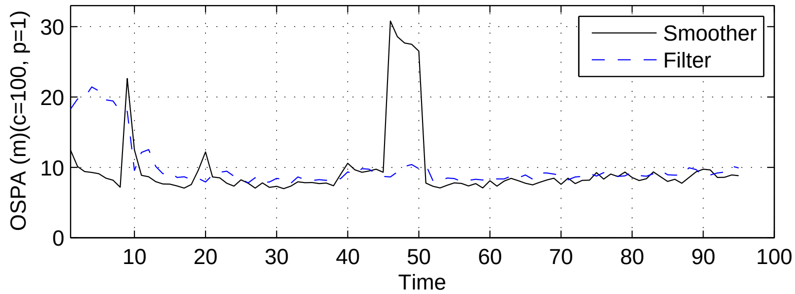


Fig. 6. OSPA miss distance versus time for filter and smoother.

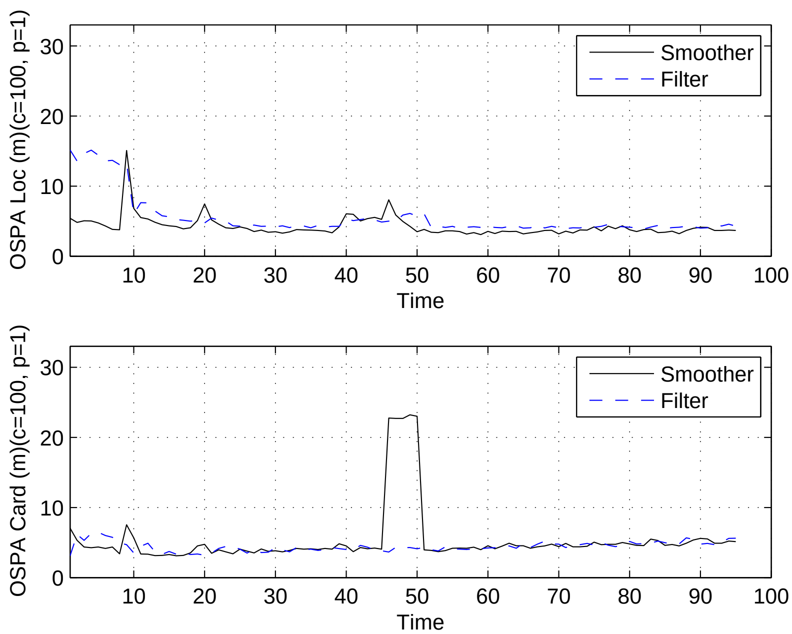


Fig. 7. OSPA localization and cardinality components versus time for filter and smoother.

### *C. Experiment 2*

This experiment investigates the performance of the filter and forward-backward smoother specifically in relation to missed detections and false alarms. The observation region is again the half disc of radius $2000m$. Exactly 2 targets are present throughout the entire scenario of $100$ time steps. Each target starts just above the horizontal axis, near the edges of the observation region, and travels diagonally upwards. Observations for the filter and smoother are generated as follows. Target detections are perfect except for measurement noise. Clutter is negligible. To investigate the effects of missed detections, some of the measurements generated by the track on the left are purposely deleted, in particular the two consecutive measurements at times $k = 10$ and $k = 11$ respectively. To investigate the effects of false alarms, two spurious measurements are artificially inserted, at times $k = 30$ and $k = 31$ respectively, in the lower corner of the observation space and mimicking the motion of an actual track. The true trajectories are shown in Fig. 8 along with the start and stop positions of each track. In implementations, the transition is used for the proposal, while resampling is performed at each time step with a $50/50$ apportionment of missed/measurement particles and with an average of $1000$ particles per surviving target.

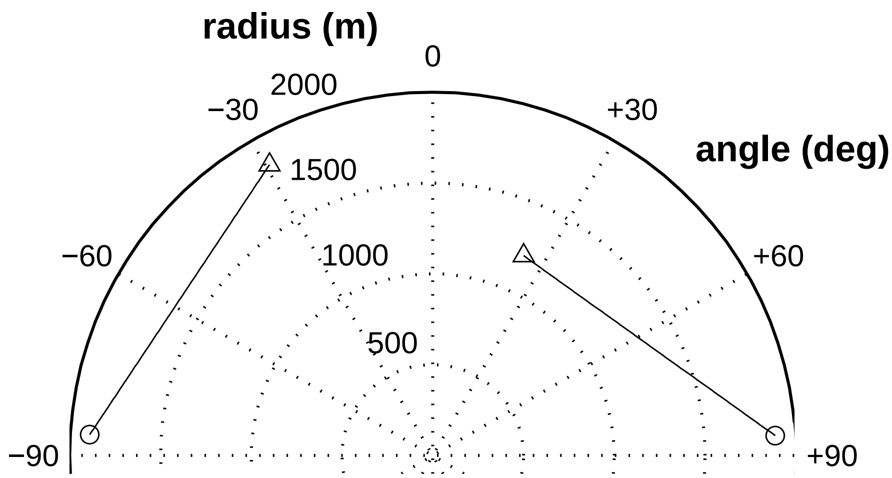


Fig. 8. Target trajectories in the $r\theta$ plane. Start/Stop positions for each track are shown with ○/△.

The parameters for the single target motion and measurement models are $\sigma_w = 2m/s^2$ and $\sigma_u = 0.2\pi/180rad/s$, and $\sigma_\theta = \pi/180rad$ and $\sigma_r = 3m$. The probability of survival is a flat $p_{S,k}(x) = 0.99$. The probability of detection is a flat $p_{D,k}(x) = 0.98$, except for two steep Gaussian notches, each centred on the locations of the two missed detections where the detection probability drops to $p_{D,k}(x) = 0.08$ (i.e. the modelled field of view is flat but drops sharply at the locations of the target where two consecutive measurements were deleted). The birth process has PHD given by a Gaussian mixture with two components, each of weight $0.01$ and centred on the initial states of the two tracks. Clutter follows a Poisson RFS with uniform spatial distribution but almost negligible mean number of returns.

The estimated $x$ and $y$ target positions for the filter and the smoother with a lag of $5$ time steps, are plotted in Figs. 9 and 10. A plot of the estimated number of targets versus time is shown in Fig. 11 for the filter and smoother. Both the filter and smoother are able to identify the target births and track their motions. It can further be seen that the filter temporarily drops one of the target tracks at times $k = 10$ and $k = 11$, due to the two consecutive missed detections at these times, whereas the smoother is able to correct for the lost measurements and hence lost track. It can also be seen that the filter declares a false track at times $k = 30$ and $k = 31$, due to false alarms, whereas the smoother is able to disregard the extraneous tracks. Surprisingly however, at times $k = 5$ and $k = 6$, the smoother drops one track which the filter does not, the same track that the filter later drops at times $k = 10$ and $k = 11$. Note again the coincidence between the chosen smoother lag of 5 time steps and the 5 time step difference between the instants when the filter and smoother suffer track losses.

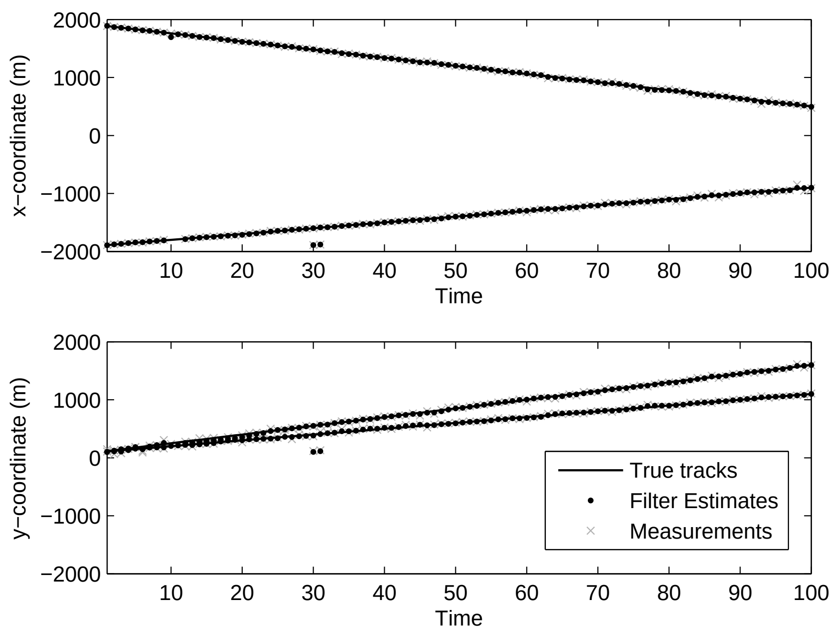


Fig. 9. Filtered state estimates versus time.

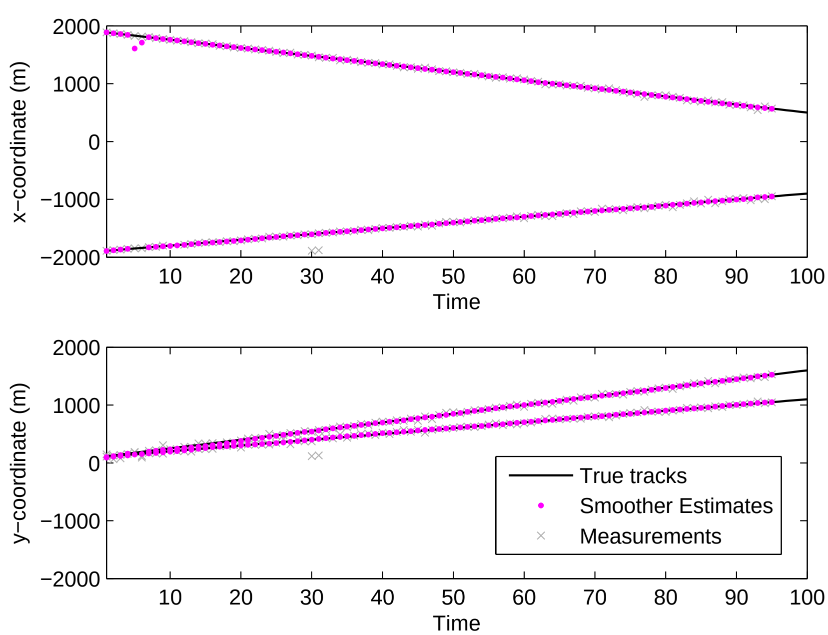


Fig. 10. Smoothed state estimates versus time.

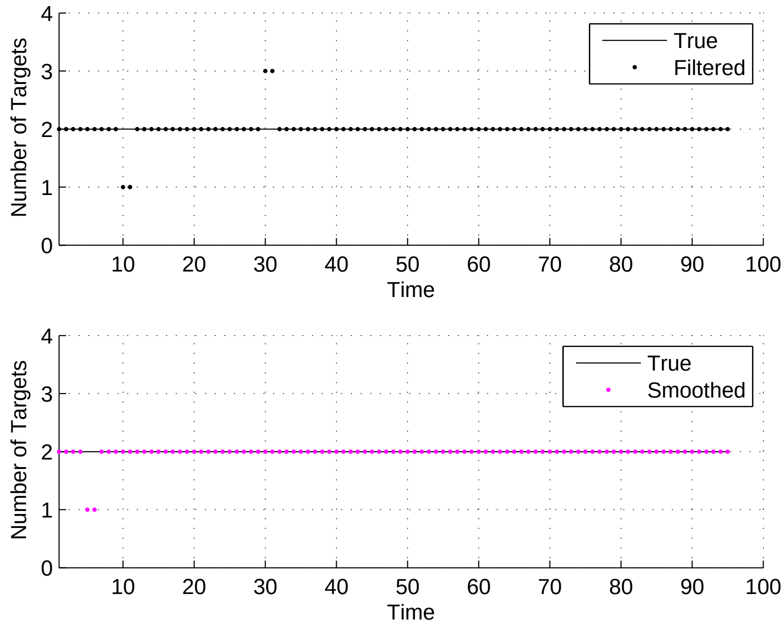


Fig. 11. Target number estimates for filter and smoother versus time.

On further investigation, it appears that such behaviour in the forward-backward smoother is not a coincidence, whereby the backwards propagation of the smoother is generally unable to restore tracks that are already lost by the filter. That is, when the filter has dropped a track for whatever reason, the corresponding backward smoother is generally unable to recover it. It is conjectured that the cause may be related to the strictly first order nature of the PHD backward smoother approximation to the full smoothed posterior density. More specifically, it may be that the information pertaining to the cardinality distribution is contained in only one parameter, the mass of the smoothed PHD, and hence may not provide sufficient confidence to facilitate proper smoothing of ambiguous tracks. It is noted however that the observed behaviour is not likely due to particle depletion, and hence unlikely to be an implementation issue, since in the current implementation, the smoother merely reweights the particles generated by the filter.

*Remark:* This behaviour may not be completely unexpected. Recall from (23) the expression for the smoothed cardinality variance. Observe that $var(N_{k'|k})$ is not necessarily smaller than $var(N_{k'+1|k})$, especially when the birth PHD $\gamma$ is small compared to the PHD of surviving targets $\left\langle v_{k|k}(:)p_S(:), f(\cdot|\,:)\right\rangle$ and when the probability of target death $1 - p_S$ is high. Hence, there is no guarantee that the forward-backward smoother improves cardinality estimates.

The results of 100 Monte Carlo runs further confirms the observations from the single sample results. In Fig 12, the average estimated number of targets for the filter and smoother are shown. In Fig. 13, the total OSPA distance for $c = 1$ and $p = 100m$ is shown, and in Fig. 14, the corresponding localization and cardinality components are shown. It can be readily seen that the smoother generally outperforms the filter. Moreover, the smoother generally achieves $50\%$ improvement in localization error, but is still on-par with the filter in cardinality error. The behaviour of the filter in regards in missed detections and false alarms is also confirmed, with peaks in the total miss distance and cardinality error component curves, at times $k = 10, 11$, and $k = 30, 31$. The corresponding behaviour of the smoother with a $5$ step delay is also confirmed, with only one peak in the total miss distance and cardinality error component curves, at times $k = 5, 6$, due to its inability to recover the missed tracks from the filter at times $k = 10, 11$.

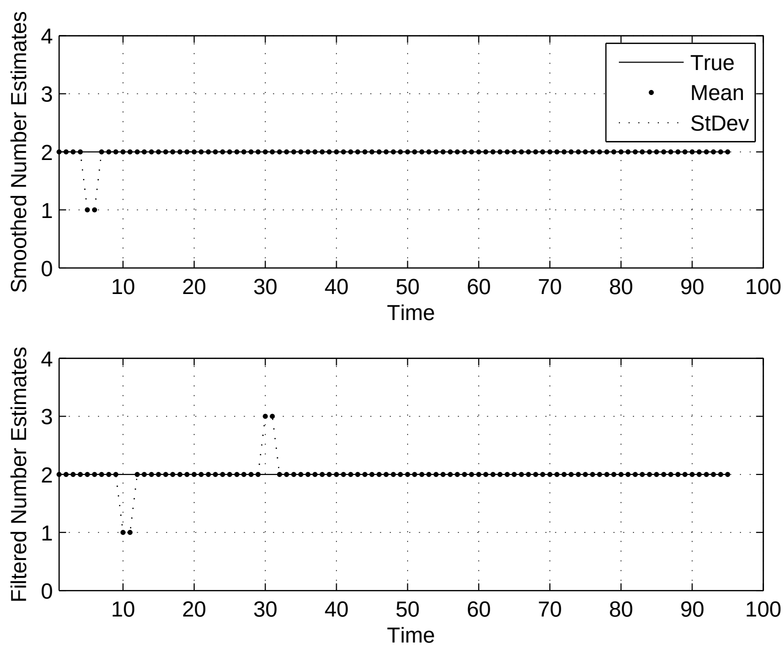


Fig. 12. Average cardinality statistics versus time for filter and smoother.

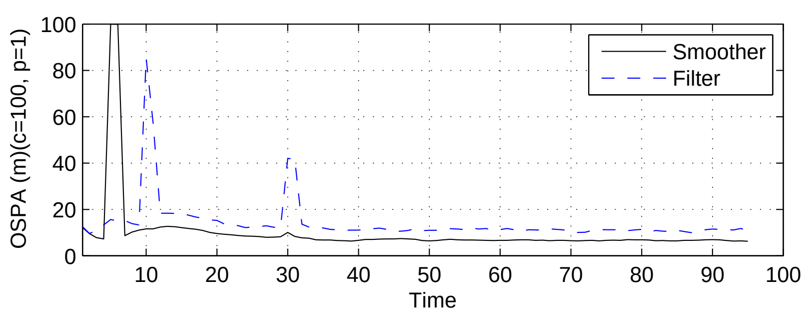


Fig. 13. OSPA miss distance versus time for filter and smoother.

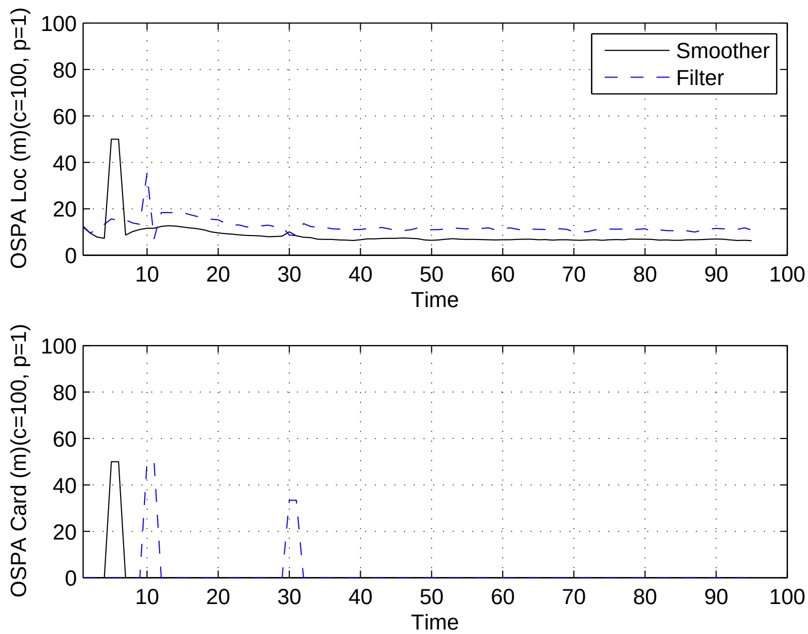


Fig. 14. OSPA localization and cardinality components versus time for filter and smoother.

## VII. Conclusions

Using FISST and Campbell's theorem, we have derived a backward PHD smoothing recursion along with recursions for the smoother cardinality distribution and moments. While the backward PHD smoothing recursion does not require the previous smoothed iterate to be Poisson, the recursion for the cardinality distribution and moments does. Case studies have shown that PHD forward-backward smoothing improves state error but does not necessarily improve cardinality error. The smoother does correct for false estimates, but does not respond well to missed detections or decreases in the number of targets. This seems to agree with theoretical analysis based on the derived backward recursion for the moments. It is possible that the cause is related to the strictly first order nature of the PHD which propagates cardinality information with a single parameter. Consequently, a PHD smoother may lack the necessary degrees of freedom to suitably control the variance of the estimated cardinality. Future works will investigate the feasibility of a Cardinalized PHD (CPHD) [22], [32] or a Multi-Bernoulli [21], [33] based forward-backward smoother, both of which may have sufficient flexibility, but will most likely incur added computational expense, to deliver improved cardinality estimates than over filtering alone.

## VIII. Appendix

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

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

*Then*

$$\frac{\partial}{\partial Y}\bigg| H_n[g] = n!\alpha^Y \sum_{i=0}^{n} e_i\left(Y;\frac{\beta}{\alpha}\right)\frac{H_{n-i}[g]}{(n-i)!}, \tag{25}$$

$$\frac{\partial}{\partial Y}\bigg|_{g=0} H_n[g] = n!\alpha^Y \sum_{i=0}^{n} e_i\left(Y;\frac{\beta}{\alpha}\right)\frac{c^{n-i}}{(n-i)!}, \tag{26}$$

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

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

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

Note that Lemma 1 used in the proof of Proposition 1 is the special case $(n = 1)$ of this Lemma.

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

$$e_i(y_1,...,y_m,y_{m+1};\phi) = e_i\left(y_1,...,y_m;\phi\right) + e_{i-1}(y_1,...,y_m;\phi)\phi(y_{m+1}) \tag{27}$$

Abbreviate: $\alpha(y_i)$ by $\alpha_i$, $\beta(y_i)$ by $\beta_i$, and note that

$$\frac{\partial}{\partial\emptyset}\bigg| H_n[g] = H_n[g] \tag{28}$$

$$\begin{aligned}\frac{\partial}{\partial y_i}\bigg| H_n[g] &= \alpha_i e^{\langle\alpha,g\rangle}\left(\langle\beta,g\rangle + c\right)^n + n\beta_i e^{\langle\alpha,g\rangle}\left(\langle\beta,g\rangle + c\right)^{n-1} \\ &= \alpha_i H_n[g] + n\beta_i H_{n-1}[g] \\ &= \alpha_i\left(H_n[g] + n\frac{\beta_i}{\alpha_i}H_{n-1}[g]\right)\end{aligned} \tag{29}$$

From (28) and (29), it is clear that (25) holds for $Y=\emptyset$ and $Y=\{y_1\}$. Suppose that (25) is true for $Y=\{y_1,...,y_m\}$, then

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

Using (29) gives

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

Set $i$ to $j+1$ gives

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

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

Using (27)

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

Hence (25) holds by the principle of induction and (26) follows as $H_n[0] = c$ .

**Proof of Proposition 2:** Recall the PGFl from (18) and substitute $h = z$ to obtain the PGF of the smoothed cardinality

$$\begin{aligned} G_{k'|k}(z) &= \int\left.\frac{\partial}{\partial Y}\right|_{g=0}\exp(\langle\gamma,g-1\rangle)\exp(\langle v_{k'|k'},z\,(1-p_S+p_S\,\langle f(:|\cdot),g(:)\rangle)-1\rangle)\frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y\\ &= \int\left.\frac{\partial}{\partial Y}\right|_{g=0}\exp(\langle\gamma,g-1\rangle)\exp(z\,\langle v_{k'|k'},1-p_S+p_S\,\langle f(:|\cdot),g(:)\rangle\rangle-\langle v_{k'|k'},1\rangle)\frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y\\ &= \int\left.\frac{\partial}{\partial Y}\right|_{g=0}\exp(\langle\gamma,g-1\rangle)\exp(z(\langle v_{k'+1|k'}-\gamma,g\rangle+\langle v_{k'|k'},1-p_S\rangle)-\langle v_{k'|k'},1\rangle)\frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)}\delta Y. \end{aligned} \tag{30}$$

where (30) follows by changing the order of integration and rearranging the term:

$$\begin{aligned}\langle v_{k'|k'},1-p_S+p_S\,\langle f(:|\cdot),g(:)\rangle\rangle &= \langle v_{k'|k'},1-p_S\rangle+\langle v_{k'|k'},p_S\,\langle f(:|\cdot),g(:)\rangle\rangle\\ &= \langle v_{k'|k'},1-p_S\rangle+\langle\langle v_{k'|k'},p_S,f(:|\cdot)\rangle,g(:)\rangle\\ &= \langle v_{k'+1|k'}-\gamma,g\rangle+\langle v_{k'|k'},1-p_S\rangle.\end{aligned}$$

Note also that

$$\begin{aligned}&\left.\frac{d^n}{dz^n}\right|\exp(z\,(\langle v_{k'+1|k'}-\gamma,g\rangle+\langle v_{k'|k'},1-p_S\rangle)-\langle v_{k'|k'},1\rangle)\\ &=\exp(z\,(\langle v_{k'+1|k'}-\gamma,g\rangle+\langle v_{k'|k'},1-p_S\rangle)-\langle v_{k'|k'},1\rangle)\,(\langle v_{k'+1|k'}-\gamma,g\rangle+\langle v_{k'|k'},1-p_S\rangle)^n\end{aligned} \tag{31}$$

$$\begin{aligned}&\left.\frac{d^n}{dz^n}\right|_{z=0}\exp(z\,(\langle v_{k'+1|k'}-\gamma,g\rangle+\langle v_{k'|k'},1-p_S\rangle)-\langle v_{k'|k'},1\rangle)\\ &=\exp(-\langle v_{k'|k'},1\rangle)\,(\langle v_{k'+1|k'}-\gamma,g\rangle+\langle v_{k'|k'},1-p_S\rangle)^n\end{aligned}$$

The cardinality distribution is given by

$$\rho_{k'|k}(n) = \frac{1}{n!}\left.\frac{d^n}{dz^n}\right|_{z=0} G_{k'|k}(z)$$

$$
\begin{aligned}
&= \frac{1}{n!} \int \frac{\partial}{\partial Y}\bigg|_{g=0} \exp(\langle \gamma, g-1 \rangle) \\
&\times \frac{d^n}{dz^n}\bigg|_{z=0} \exp(z\left(\left\langle v_{k'+1|k'} - \gamma, g \right\rangle + \left\langle v_{k'|k'}, 1-p_S \right\rangle\right) - \left\langle v_{k'|k'}, 1 \right\rangle) \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \\
&= \frac{1}{n!} \int \frac{\partial}{\partial Y}\bigg|_{g=0} \exp(\langle \gamma, g-1 \rangle) \\
&\times \exp(-\left\langle v_{k'|k'}, 1 \right\rangle) \left(\left\langle v_{k'+1|k'} - \gamma, g \right\rangle + \left\langle v_{k'|k'}, 1-p_S \right\rangle\right)^n \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \\
&= \frac{1}{n!} \exp(-\left\langle \gamma + v_{k'|k'}, 1 \right\rangle) \\
&\times \int \frac{\partial}{\partial Y}\bigg|_{g=0} \exp(\langle \gamma, g \rangle) \left(\left\langle v_{k'+1|k'} - \gamma, g \right\rangle + \left\langle v_{k'|k'}, 1-p_S \right\rangle\right)^n \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y. \qquad (32)
\end{aligned}
$$

Let $\alpha = \gamma$, $\beta = v_{k'+1|k'} - \gamma$, $c = \left\langle v_{k'|k'}, 1-p_S \right\rangle$, and applying Lemma 2 gives

$$
\begin{aligned}
\rho_{k'|k}(n) &= \exp(-\left\langle \gamma + v_{k'|k'}, 1 \right\rangle) \int \alpha^Y \sum_{i=0}^{n} e_i\left(Y; \frac{\beta}{\alpha}\right) \frac{c^{n-i}}{(n-i)!} \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \qquad (33) \\
&= \exp(-\left\langle \gamma + v_{k'|k'}, 1 \right\rangle) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int \alpha^Y e_i\left(Y; \frac{\beta}{\alpha}\right) \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \qquad (34) \\
&= \exp(-\left\langle \gamma + v_{k'|k'}, 1 \right\rangle) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) \frac{\alpha^Y}{\exp(-\left\langle v_{k'+1|k'}, 1 \right\rangle) v_{k'+1|k'}^Y} p_{k'+1|k}(Y) \delta Y \qquad (35) \\
&= \exp(\left\langle v_{k'+1|k'} - \gamma - v_{k'|k'}, 1 \right\rangle) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) \left(\frac{\alpha}{v_{k'+1|k'}}\right)^Y p_{k'+1|k}(Y) \delta Y \qquad (36) \\
&= \exp(\left\langle v_{k'+1|k'} - \gamma - v_{k'|k'}, 1 \right\rangle) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) \left(\frac{\alpha}{v_{k'+1|k'}}\right)^Y \\
&\times \exp(-\left\langle v_{k'+1|k}, 1 \right\rangle) v_{k'+1|k}^Y \delta Y \qquad (37) \\
&= \exp(\left\langle -v_{k'+1|k} + v_{k'+1|k'} - \gamma - v_{k'|k'}, 1 \right\rangle) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) \left(\frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}\right)^Y \delta Y \qquad (38) \\
&= \exp(\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}} - v_{k'+1|k} + v_{k'+1|k'} - \gamma - v_{k'|k'}, 1 \right\rangle) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) \\
&\times \exp\left(-\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}, 1 \right\rangle\right) \left(\frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}\right)^Y \delta Y. \qquad (39)
\end{aligned}
$$

where (35) and (37) follow from the Poisson assumption on $p_{k'+1|k'}$ and $p_{k'+1|k}$.

The set integral in (39) can be rewritten as follows:

$$
\int e_i\left(Y; \frac{\beta}{\alpha}\right) \exp\left(-\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}, 1 \right\rangle\right) \left(\frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}\right)^Y \delta Y \qquad (40)
$$

$$= \int \left( \sum_{S \subset Y, |S|=i} \left(\frac{\beta}{\alpha}\right)^S \right) \exp\left(-\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}, 1\right\rangle\right) \left(\frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}\right)^Y \delta Y$$

$$= \int \left( \sum_{y_1 \neq y_2 \neq \ldots \neq y_i \in Y} \frac{\beta(y_1)}{\alpha(y_1)} \ldots \frac{\beta(y_i)}{\alpha(y_i)} \right) \exp\left(-\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}, 1\right\rangle\right) \left(\frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}\right)^Y \delta Y \quad (41)$$

Note that (41) is the $i$th factorial moment, evaluated at $g$ defined by $g(y_1, ..., y_i) = \frac{\beta}{\alpha}(y_1)...\frac{\beta}{\alpha}(y_i)$, of a Poisson RFS with (FISST) density $\exp\left(-\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}, 1\right\rangle\right) \left(\frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}\right)^Y$. Since the product densities of a Poisson RFS is the product of the PHDs, using (14) we have

$$\int \left( \sum_{y_1 \neq y_2 \neq \ldots \neq y_i \in Y} \frac{\beta(y_1)}{\alpha(y_1)} \ldots \frac{\beta(y_i)}{\alpha(y_i)} \right) \exp\left(-\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}, 1\right\rangle\right) \left(\frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}}\right)^Y \delta Y$$

$$= \int \frac{\beta(y_1)}{\alpha(y_1)} \ldots \frac{\beta(y_i)}{\alpha(y_i)} \frac{\alpha}{v_{k'+1|k'}}(y_1) v_{k'+1|k}(y_1) \ldots \frac{\alpha}{v_{k'+1|k'}}(y_i) v_{k'+1|k}(y_i) dy_1 \ldots dy_i$$

$$= \left\langle \frac{\beta}{\alpha}, \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}} \right\rangle^i .$$

Therefore

$$\rho_{k'|k}(n) = \exp\left(\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}} - v_{k'+1|k} + v_{k'+1|k'} - \gamma - v_{k'|k'}, 1\right\rangle\right) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \left\langle \frac{\beta}{\alpha}, \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}} \right\rangle^i$$

$$= \exp\left(\left\langle \frac{\alpha v_{k'+1|k}}{v_{k'+1|k'}} - v_{k'+1|k} + v_{k'+1|k'} - \gamma - v_{k'|k'}, 1\right\rangle\right) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \left\langle \frac{\beta}{v_{k'+1|k'}}, v_{k'+1|k} \right\rangle^i$$

$$= \exp\left(\left\langle \frac{\gamma v_{k'+1|k}}{v_{k'+1|k'}} - v_{k'+1|k} + v_{k'+1|k'} - \gamma - v_{k'|k'}, 1\right\rangle\right)$$

$$\times \sum_{i=0}^{n} \frac{\langle v_{k'|k'}, 1 - p_S\rangle^{n-i}}{(n-i)!} \left\langle \frac{v_{k'+1|k'} - \gamma}{v_{k'+1|k'}}, v_{k'+1|k} \right\rangle^i .$$

**Proof of Proposition 3:** The $n$th moment of the smoothed cardinality is obtained by differentiating (30) at $z = 1$

$$G^{(n)}_{k'|k} = \left.\frac{d^n}{dz^n}\right|_{z=1} G_{k'|k}(z)$$

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

$$\times \left.\frac{d^n}{dz^n}\right|_{z=1} \exp\left(z\left(\langle v_{k'+1|k'} - \gamma, g\rangle + \langle v_{k'|k'}, 1 - p_S\rangle\right) - \langle v_{k'|k'}, 1\rangle\right) \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \quad (42)$$

but from (31), setting $z = 1$ gives

$$\left.\frac{d^n}{dz^n}\right|_{z=1} \exp\left(z\left(\langle v_{k'+1|k'} - \gamma, g\rangle + \langle v_{k'|k'}, 1 - p_S\rangle\right) - \langle v_{k'|k'}, 1\rangle\right)$$

$$= \exp\left(\langle v_{k'+1|k'} - \gamma, g\rangle + \langle v_{k'|k'}, 1 - p_S\rangle - \langle v_{k'|k'}, 1\rangle\right) \left(\langle v_{k'+1|k'} - \gamma, g\rangle + \langle v_{k'|k'}, 1 - p_S\rangle\right)^n$$

$$= \exp\left(\langle v_{k'+1|k'} - \gamma, g\rangle - \langle v_{k'|k'}, p_S\rangle\right) \left(\langle v_{k'+1|k'} - \gamma, g\rangle + \langle v_{k'|k'}, 1 - p_S\rangle\right)^n .$$

Hence, substituting this into (42) we have

$$\begin{aligned}
G^{(n)}_{k'|k} &= \int \frac{\partial}{\partial Y}\bigg|_{g=0} \exp\left(\langle \gamma, g-1\rangle\right) \exp\left(\left\langle v_{k'+1|k'} - \gamma, g\right\rangle - \left\langle v_{k'|k'}, p_S\right\rangle\right) \\
&\quad \times \left(\left\langle v_{k'+1|k'} - \gamma, g\right\rangle + \left\langle v_{k'|k'}, 1-p_S\right\rangle\right)^n \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \\
&= \exp\left(-\langle \gamma, 1\rangle - \left\langle v_{k'|k'}, p_S\right\rangle\right) \\
&\quad \times \int \frac{\partial}{\partial Y}\bigg|_{g=0} \exp\left(\left\langle v_{k'+1|k'}, g\right\rangle\right) \left(\left\langle v_{k'+1|k'} - \gamma, g\right\rangle + \left\langle v_{k'|k'}, 1-p_S\right\rangle\right)^n \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \\
&= \exp\left(-\left\langle v_{k'+1|k'}, 1\right\rangle\right) \\
&\quad \times \int \frac{\partial}{\partial Y}\bigg|_{g=0} \exp\left(\left\langle v_{k'+1|k'}, g\right\rangle\right) \left(\left\langle v_{k'+1|k'} - \gamma, g\right\rangle + \left\langle v_{k'|k'}, 1-p_S\right\rangle\right)^n \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y.
\end{aligned}$$

Let $\alpha = v_{k'+1|k'}$, $\beta = v_{k'+1|k'} - \gamma$, $c = \left\langle v_{k'|k'}, 1-p_S\right\rangle$, and using Lemma 2 gives

$$\begin{aligned}
G^{(n)}_{k'|k} &= \exp\left(-\left\langle v_{k'+1|k'}, 1\right\rangle\right) \int n!\alpha^Y \sum_{i=0}^{n} e_i\left(Y; \frac{\beta}{\alpha}\right) \frac{c^{n-i}}{(n-i)!} \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \\
&= n! \exp\left(-\left\langle v_{k'+1|k'}, 1\right\rangle\right) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) \alpha^Y \frac{p_{k'+1|k}(Y)}{p_{k'+1|k'}(Y)} \delta Y \\
&= n! \exp\left(-\left\langle v_{k'+1|k'}, 1\right\rangle\right) \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) \frac{\alpha^Y}{\exp\left(-\left\langle v_{k'+1|k'}, 1\right\rangle\right) v^Y_{k'+1|k'}} p_{k'+1|k}(Y) \delta Y \\
&= n! \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int e_i\left(Y; \frac{\beta}{\alpha}\right) p_{k'+1|k}(Y) \delta Y \\
&= n! \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int \left( \sum_{S \subset Y, |S|=i} \left(\frac{\beta}{\alpha}\right)^S \right) p_{k'+1|k}(Y) \delta Y \\
&= n! \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \int \left( \sum_{y_1 \neq y_2 \neq \ldots \neq y_i \in Y} \frac{\beta(y_1)}{\alpha(y_1)} \cdots \frac{\beta(y_i)}{\alpha(y_i)} \right) p_{k'+1|k}(Y) \delta Y
\end{aligned}$$

Since $p_{k'+1|k}$ is Poisson, using the same arguments as those in Proposition 2, the set integral can be simplified as follows:

$$\begin{aligned}
\int \left( \sum_{y_1 \neq y_2 \neq \ldots \neq y_i \in Y} \frac{\beta(y_1)}{\alpha(y_1)} \cdots \frac{\beta(y_i)}{\alpha(y_i)} \right) p_{k'+1|k}(Y) \delta Y &= \int \frac{\beta(y_1)}{\alpha(y_1)} \cdots \frac{\beta(y_i)}{\alpha(y_i)} v_{k'+1|k}(y_1) \ldots v_{k'+1|k}(y_i) dy_1 \ldots dy_i \\
&= \left\langle \frac{\beta}{\alpha}, v_{k'+1|k} \right\rangle^i.
\end{aligned}$$

Hence,

$$G^{(n)}_{k'|k} = n! \sum_{i=0}^{n} \frac{c^{n-i}}{(n-i)!} \left\langle \frac{\beta}{\alpha}, v_{k'+1|k} \right\rangle^i$$

$$= n! \sum_{i=0}^{n} \frac{\left\langle v_{k'|k'}, 1 - p_S \right\rangle^{n-i}}{(n-i)!} \left\langle \frac{v_{k'+1|k'} - \gamma}{v_{k'+1|k'}}, v_{k'+1|k} \right\rangle^{i}$$

$$= n! \sum_{i=0}^{n} \frac{\left\langle v_{k'|k'}, 1 - p_S \right\rangle^{n-i}}{(n-i)!} \left\langle v_{k'+1|k}, 1 - \frac{\gamma}{v_{k'+1|k'}} \right\rangle^{i}.$$

## REFERENCES


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

[2] Y. Bresler, ”Two- filter formula for discrete-time non-linear Bayesian smoothing,” *International Journal of Control*, Vol. 43, No. 2, 629–641, 1986.

[3] M. Briers, A. Doucet, and S. Maskell, ”Smoothing algorithms for state-space models,” *Annals of the Institute of Statistical Mathematics*, to appear 2008.

[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] D. Clark, B.-T. Vo, B.-N. Vo, and S. Godsill, “Gaussian Mixture Implementations of PHD Filters for Nonlinear Dynamical Models,” *IET Seminar on Target Tracking*, Birmingham, Apr. 2008.

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

[7] A. Doucet, S. Godsill, and C. Andrieu, “On sequential Monte Carlo sampling methods for Bayesian filtering,” *Statistics and Computing*, Vol. 10, No. 3, pp. 197–208, 2000.

[8] A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.

[9] O. Erdinc, P. Willett, and Y; Bar-Shalom*,* ”The Bin-Occupancy Filter and Its Connection to the PHD Filters,” *IEEE Trans. Signal Processing,* vol.57, no.11, pp.4232-4246, Nov. 2009.

[10] P. Fearnhead, D. Wyncoll, and J. Tawn, ”A sequential smoothing algorithm with linear computational cost,” *Technical report*, Department of Mathematics and Statistics, Lancaster University, 2008

[11] W. Fong, S. J. Godsill, A. Doucet, and M. West, “Monte Carlo smoothing with application to audio signal enhancement,” *IEEE Trans. on Signal Processing*, Vol. 50, No. 2, pp. 438–449, Feb. 2002.

[12] S. J. Godsill, A. Doucet, and M. West, “Monte Carlo smoothing for nonlinear time series,” *Jnl. American Statistical Association*, Vol. 99, pp. 156–168, Mar. 2004.

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

[14] 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.

[15] M. Hürzeler, and H. R. Künsch, “Monte Carlo approximations for general state space models,” *Journal of Computational and Graphical Statistics*, Vol. 7, No. 2, pp. 175–193, Jun. 1998.

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

[17] G. Kitagawa, “Monte Carlo filter and smoother for nonlinear non-Gaussian state-space models,” *Journal of Computational Graphical Statistics*, Vol. 5, No. 1, pp. 1–25, Mar. 1996.

[18] 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.

[19] 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.

[20] 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.

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

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

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

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

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

[26] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Processing*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

[27] H. Sidenbladh, "Multi-target particle filtering for the Probability Hypothesis Density", *Proc. Int'l Conf. on Information Fusion*, Cairns, Australia, pp. 800-806, 2003.

[28] S. S. Singh, B.-N. Vo, A. Baddeley, and S. Zuyev, "Filters for spatial point processes", SIAM J. Control and Optimization, Volume 48, Issue 4, pp. 2275-2295, Jun. 2009

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

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] T. Zajic, R. and Mahler, "A Particle-Systems implementation of the PHD multi-target tracking filter," in I. Kadar (ed.), *Signal Processing, Sensor Fusion and Target Recognition XII,* SPIE Proc., Vol. 5096, 2003.