First-Moment Multi-Object Forward-Backward Smoothing

Daniel E. Clark
Joint Research Institute in Signal and Image Processing
Heriot-Watt University
Edinburgh, UK.
d.e.clark@hw.ac.uk

Abstract – *The optimal solution to the problem of detecting, tracking and identifying multiple targets can be found through a direct generalisation of the Bayes filter to multi-object systems using Mahler's Finite Set Statistics. Due to the inherent complexity of the multi-object Bayes filter, Mahler proposed to propagate the first-order multi-object moment density, known as the Probability Hypothesis Density (PHD), instead of the multi-object posterior. This was derived using the concept of the probability generating functional (p.g.fl.) from point process theory. In this paper, I derive multi-object first-moment smoothers for forward-backward smoothing through a new formulation of the p.g.fl. smoother which takes advantage of the p.g.fl. Bayes update. This formulation permits the straightforward derivation of first-moment multi-object smoothers, including the PHD smoother.*

Keywords: Multi-object estimation, Finite Set Statistics (FISST), Probability Hypothesis Density (PHD) filters, forward-backward smoothing

1 Introduction

Stochastic filtering, prediction, and smoothing are fundamental concepts in the theory of estimation of dynamic systems. The system is a partially observed physical object whose behaviour over time is governed by a set of equations modelling the dynamics and the relationship between the observations and the object state. Uncertainty in the system is due to the noisy nature of the problem, either from unknown and unpredictable motion of the system, or from inaccuracy in observing measurements of the system through a noisy sensor. Filtering, prediction and smoothing are precise mathematical descriptions of the problem of estimating the state of the system based on noisy observations of its behaviour over time: Prediction is the forecasting of the state of the system at some future point in time based on measurements up to the current time. Filtering is the estimation at each point in time of the state of the system based on all of the measurements up to that point. Smoothing differs from prediction and filtering in that the estimate of the state of the system at a specific point in time can be determined from a batch of measurements, some of which may be collected later than the time that we are interested in. This means that there is inevitably a delay in producing the estimate of the state at that time, though more accurate estimates can be obtained since more information is available about the system. The most commonly used probabilistic smoother formulation is the forward-backward smoother [1]. Forward-backward smoothing processes the measurements sequentially; first by running a Bayes filter forward in time and then a filter backward in time to improve the state estimates with future measurements.

The use of Finite Set Statistics (FISST) for multi-object smoothing was proposed recently for jointly detecting whether an object is observed and the object state density function [2] in cluttered environments and tractable practical implementations were then derived using sequential Monte Carlo methodology [3]. The multi-object forward-backward Probability Hypothesis Density (PHD) smoother has been investigated recently by two independent groups: Kiruba *et al.* first investigated the PHD smoother and hypothesised a possible form of the solution [4]. A formal derivation was recently found using FISST by Mahler, Vo and Vo [5], and an analytic recursion was subsequently discovered using a Gaussian mixture representation of the PHD [6]. In this paper, I propose a new formulation of the p.g.fl. smoother which enables the simple derivation of multi-object smoothers that exploits the p.g.fl. Bayes filter. Using this formulation, I show how to derive the forward-backward multi-object PHD smoother counterpart to the PHD filter [7].

The paper is structured as follows: In Section 2, I discuss the background material required for deriving the smoothers. In Section 3, I describe the multi-object Bayes filter, its probability generating functional (p.g.fl.) form, and first-moment approximations. In Section 4, I describe the multi-object forward-backward smoother and the new p.g.fl. formulation that exploits a mathematical parallel between the p.g.fl. Bayes update. Using this new formulation, I derive the PHD smoother as a specific case. The paper concludes in Section 5.

2 Background

In this section, I summarise the mathematical material required to derive the multi-object smoothers.

Spatial point process

A spatial point process is defined as a random finite set of points,

$$X = \{x_1, \ldots, x_N\} \tag{1}$$

where the number and their locations are random [8, 9]. A multi-object probability distribution p_X defines the distribution of the points, which is a mixture of joint probability distributions for any given number of targets. The cardinality distribution $\rho(n)$ is a discrete probability distribution in the number of objects and satisfies

$$\sum_{n=0}^{\infty} \rho(n) = 1. \tag{2}$$

Probability Generating Functionals

Point processes can be uniquely characterised in terms of their probability generating functional (p.g.fl.). This concept is analogous to the probability generating function (p.g.f.) from probability theory in that it is can be used for determining moments of distributions. Mahler proposed the use of the p.g.fl. as a means of finding the first-order moment densities in the multi-object Bayes filter [7].

The probability generating functional (p.g.fl.) of a (multi-object) probability distribution p_X is defined as the expectation value of the symmetric function

$$h^X := \prod_{x \in X} h(x), \tag{3}$$

so that

$$G_X[h] := E\left(h^X\right) = \int p_X(X) \cdot h^X \delta X, \tag{4}$$

using the set integral notation from FISST,

$$\int p(X)\delta X := p(\phi) + \sum_{n=1}^{\infty} \frac{1}{n!} \int p(\{x_1, \ldots, x_n\}) dx_1 \ldots dx_n, \tag{5}$$

where $p(\{x_1, \ldots, x_n\})$ is defined as the joint distribution scaled by its cardinality,

$$p(\{x_1, \ldots, x_n\}) := n! \cdot \rho(n) \cdot p(x_1, \ldots, x_n). \tag{6}$$

The $n!$ factorial term accounts for the fact we need to consider all permutations in the joint distribution. The set derivative is defined by taking the functional derivatives with respect to each point x in point process $X = \{x_1 \ldots x_m\}$, i.e.

$$\frac{\delta G[h]}{\delta X} := \frac{\delta G[h]}{\delta x_1 \ldots \delta x_m} \tag{7}$$

Moments of multi-object posterior can be obtained by taking the set-derivative of the p.g.fl. G at $h = 1$, i.e.

$$D(X) = \left. \frac{\delta G}{\delta X}[h] \right|_{h=1}. \tag{8}$$

In particular, the first-order moment of multi-object probability density $p(X)$, more commonly known as the Probability Hypothesis Density in the target tracking literature [7], can be found by taking the functional derivative of its p.g.fl. G evaluated at $h = 1$, i.e.

$$D(x) = \frac{\delta}{\delta x} G[1] \tag{9}$$

This result was shown by Mahler in [7]. An equivalent result was shown in the point process literature by Moyal [10], in the context of stochastics population processes.

Another important result [10, 7] using set derivatives is an analogy with the fundamental theorem of calculus for multi-object distributions. We can recover the multi-object distribution from the p.g.fl. by taking the set derivative of the p.g.fl. and then setting $h = 0$, i.e.

$$p(X) = \frac{\delta}{\delta X} G[0]. \tag{10}$$

Mahler refers to this result as the Fundamental Theorem of Multi-object Calculus [11].

Campbell's Theorem

In the context of Finite Set Statistics, Campbell's Theorem states that for any non-negative measurable function f, the following holds

$$\int \left(\sum_{x \in X} f(x) \right) \cdot p(X) \delta X = \int f(x) \cdot D(x) dx, \tag{11}$$

where $p(X)$ is some multi-object distribution and $D(x)$ is its first-moment density. For a proof see, for example, Stoyan *et al.* [12], page 100. This result was first applied for FISST smoothing by Mahler, Vo and Vo [5] for deriving the PHD smoother.

Poisson point processes

Suppose that the cardinality distribution $\rho(n)$ is given by

$$\rho(n) = \frac{\exp(-\lambda) \cdot \lambda^n}{n!}, \tag{12}$$

and the multi-object distribution is

$$p(Y) := e^{-\lambda} \cdot \lambda^n \cdot \prod_{y_i \in Y} s(y_i). \tag{13}$$

The p.g.fl. of a Poisson point process is given by

$$G_P[h] = \exp\left(\mu \int h(x) \cdot s(x) dx - \mu\right) \quad (14)$$

The intensity function, or PHD, of a Poisson point process is found by taking the functional derivative of (14), evaluated at $h = 1$, i.e.

$$D(x) := \frac{\delta}{\delta x} G_P[h]\bigg|_{h=1} = \mu \cdot s(x), \quad (15)$$

where μ gives the expected number of objects that are distributed according to $s(x)$.

In the following section, I summarise the FISST approach to multi-object Bayes filtering [11] and the PHD [7] approximations.

3 Multi-Object Bayes Filter

The objective of multi-object Bayesian multi-object filtering [11] is to determine at each time-step k the posterior probability density, $p_{k|k}(X_k|Z_{1:k})$, of multi-object state pX_k, where $Z_{1:k} = (Z_1, \ldots, Z_k)$ denotes the accumulated observations up to time-step k. The multi-object posterior can be computed sequentially via the prediction and update steps. Suppose that $p_{k|k}(X_k|Z_{1:k})$ is known and that a new set of measurements Z_{k+1} corresponding to time-step $k+1$ has been received. Then the predicted and updated multi-object posterior densities are calculated with the following recursion [11]

$$p_{k+1|k}(X|Z_{1:k}) = \int f_{k+1|k}(X|X') \cdot p_{k|k}(X'|Z_{1:k}) \delta X' \quad (16)$$

$$p_{k+1|k+1}(X|Z_{1:k+1}) = \frac{g_{k+1}(Z_{k+1}|X) \cdot p_{k+1|k}(X|Z_{1:k})}{\int g_{k+1}(Z_{k+1}|X) \cdot p_{k+1|k}(X|Z_{1:k}) \delta X}, \quad (17)$$

where $f_{k+1|k}(X|X')$ is a multi-object transition density and $g_{k+1}(Z_{k+1}|X)$ is a multi-object likelihood. The prediction corresponds to the Chapman-Kolmogorov equation and the update corresponds to Bayes rule. This recursion is a non-trivial generalisation, since the transition density needs to consider the uncertainty in target number, which can change over time due to targets entering and leaving the state space, and the multi-object likelihood needs to consider detection uncertainty and false alarms. It is also clear that the integrals in the recursion are non-standard and need the notion of a set integral [11].

The Bayesian multi-object filter described here in its most general form cannot be implemented in a computationally tractable manner. Among the approximations that were recently proposed of particular importance are the PHD filter and the CPHD filter [7, 14]. Instead of propagating the full multi-target density $p_{k|k}(X_k|Z_{1:k})$, the PHD filter propagates its first order moment $D_{k|k}(x)$ called the Probability Hypothesis Density (PHD) or intensity function. The Cardinalized Probability Hypothesis Density (CPHD) filter propagates not only the PHD but also the cardinality distribution. The first implementations of the PHD filters used sequential Monte Carlo approximations [15, 16] and closed-form solutions were later discovered using Gaussian mixture approximations [17, 18].

3.1 Multi-object Prediction

In order to determine moment densities, Mahler [7] considered the p.g.fl. multi-object prediction,

$$G_{k+1|k}[h] := \int h^X \cdot p_{k+1|k}(X|Z_{1:k}) \delta X \quad (18)$$

$$= \int \int h^X \cdot f_{k+1|k}(X|X') \cdot p_{k|k}(X'|Z_{1:k}) \delta X' \delta X \quad (19)$$

$$= \int G_{k+1|k}[h|X'] \cdot p_{k|k}(X'|Z_{1:k}) \delta X', \quad (20)$$

where $G_{k+1|k}[h|X']$ is the p.g.fl. Markov transition (see equation (13.52) in [11]),

$$G_{k+1|k}[h|X'] := \int h^X \cdot f_{k+1|k}(X|X') \delta X. \quad (21)$$

PHD Prediction

I now consider the model for the PHD filter [7, 11] without target spawning. The predicted state set X at time-step $k+1$ consists of the set of objects surviving from time-step k, $S_{k+1|k}(X')$, and objects appearing spontaneously at time-step k, Γ_{k+1}, i.e.

$$X = \left(\bigcup_{i=1}^{n'} S_{k+1|k}(x'_i)\right) \cup \Gamma_{k+1}. \quad (22)$$

Taking the first-order moment of the multi-object prediction in equation (16), the predicted PHD becomes [7]

$$D_{k+1|k}(x) = \gamma_{k+1}(x) + \int p_S(x') \cdot f_{k+1|k}(x|x') \cdot D_{k|k}(x') dx', \quad (23)$$

where $f_{k+1|k}(x|x')$ is the single-object Markov transition, $\gamma_{k+1}(x)$ is the intensity of spontaneous births, and $D_{k|k}(x')$ is the posterior intensity at time-step k.

3.2 Multi-object Bayes Update

In a similar manner to the multi-object prediction, Mahler's calculation of the PHD update [7] requires the p.g.fl. Bayes update,

$$G_{k+1|k+1}[h] := \int h^X \cdot p_{k+1|k+1}(X|Z_{1:k+1}) \delta X \quad (24)$$

$$= \frac{\int h^X \cdot g_{k+1}(Z_{k+1}|X) \cdot p_{k+1|k}(X|Z_{1:k}) \delta X}{\int g_{k+1}(Z_{k+1}|X') \cdot p_{k+1|k}(X'|Z_{1:k}) \delta X'} \quad (25)$$

The numerator and denominator of the p.g.fl. Bayes update are found with set derivatives of the following joint p.g.fl.

$$F[g,h] := \int h^X \cdot G_{k+1}[g|X] \cdot p_{k+1|k}(X|Z_{1:k})\delta X, \quad (26)$$

where the p.g.fl. multi-object likelihood is defined with

$$G_{k+1}[g|X] := \int g^Z \cdot g_{k+1}(Z|X)\delta Z. \quad (27)$$

The p.g.fl. Bayes update is found with

$$G_{k+1|k+1}[h] = \frac{\frac{\delta}{\delta Z}F[0,h]}{\frac{\delta}{\delta Z}F[0,1]}, \quad (28)$$

where the set-derivative is taken with respect to g at the elements in observation set Z. Mahler then assumes that the predicted multi-object distribution, $p_{k+1|k}(X|Z_{1:k})$, is either a Poisson distribution, for the PHD update, or an i.i.d cluster distribution, for the CPHD update, to determine an updated intensity in terms of the predicted density

Taking the functional derivative of the p.g.fl. Bayes update with respect to x at $h=1$ and $g=0$ then leads to an update formula for the PHD in terms of the PHD prediction,

$$D_{k+1|k+1}(x) = L_{Z_{k+1}}(x) \cdot D_{k+1|k}(x), \quad (29)$$

where $D_{k+1|k}(x)$ is the predicted first-moment density and $L_{Z_{k+1}}(x)$ is a PHD pseudo-likelihood.

PHD Update

For a given multi-target state $X = \{x_1, \dots, x_n\}$, the random measurement set collected by the sensor is of the form

$$Z = \left(\bigcup_{i=1}^{n} \Sigma(x_i)\right) \cup \Theta, \quad (30)$$

where $\Sigma(x_i)$ is the Random Finite Set (RFS) generated by the single target state x_i, which either contains a single measurement z_i or is empty and takes on the value $\emptyset$, and Θ contains false alarms with expected number λ. Let λ be the Poisson distributed false alarms, each identically spatially distributed according to c.

When the predicted multi-object distribution and clutter distribution are assumed to be Poisson, the pseudo-likelihood in the PHD filter update [7] equation (29) is

$$L_{Z_{k+1}}(x) = (1 - p_D(x)) + \sum_{z \in Z} \frac{p_D(x) L_z(x)}{\lambda c(z) + D_{k+1|k}[p_D \cdot L_z]}, \quad (31)$$

where $D_{k+1|k}(x)$ is the predicted intensity to time-step k, $p_D(x)$ is the probability of detection, and $L_z(x)$ is the single-object likelihood. and

4 Multi-object Forward-Backward Smoother

In an analogous fashion to the multi-object Bayes filter, we can generalise the forward-backward smoother to multi-object distributions [2]. This leads to the following form of multi-object forward-backward smoother,

$$p_{k'|k}(X) = \int \frac{f_{k'+1|k'}(Y|X) \cdot p_{k'|k'}(X)}{p_{k'+1|k'}(Y)} p_{k'+1|k}(Y)\delta Y, \quad (32)$$

where X consists of a random finite set of target state vectors at time-step k, and $k' < k$. The p.g.fl. forward-backward smoother is then given by

$$G_{k'|k}[h] = \int h^X p_{k'|k}(X)\delta X \quad (33)$$
$$= \int \left(\frac{\int h^X \cdot f_{k'+1|k'}(Y|X) \cdot p_{k'|k'}(X)\delta X}{p_{k'+1|k'}(Y)} \right) p_{k'+1|k}(Y)\delta Y$$

We can expand the denominator, $p_{k'+1|k'}(Y)$, inside the brackets above with the Bayes filter prediction, so that

$$G_{k'|k}[h] = \int \left(\frac{\int h^X \cdot f_{k'+1|k'}(Y|X) \cdot p_{k'|k'}(X)\delta X}{\int f_{k'+1|k'}(Y|V) \cdot p_{k'|k'}(V)\delta V} \right) p_{k'+1|k}(Y)\delta Y. \quad (34)$$

Notice that the term inside the brackets is analogous to the p.g.fl. Bayes update except that the posterior $p_{k'|k'}(Y)$ is updated with the Markov transition $f_{k'+1|k'}(Y|X)$. I shall exploit this mathematical parallel to find the smoothed first-moment density. Hence the forward-backward smoothed PHD becomes

$$D_{k'|k}(x) = \int D_{k'+1|k'}(x|Y) \cdot p_{k'+1|k}(Y)\delta Y, \quad (35)$$

where the PHD $D_{k'+1|k'}(x|Y)$ is found by taking the first-moment of the term inside the brackets of equation (34), i.e.

$$D_{k'+1|k'}(x|Y) = \frac{\delta}{\delta x} \left\{ \frac{\int h^X \cdot f_{k'+1|k'}(Y|X) \cdot p_{k'|k'}(X)\delta X}{\int f_{k'+1|k'}(Y|V) \cdot p_{k'|k'}(V)\delta V} \right\} \Bigg|_{h=1} \quad (36)$$

The PHD in equation (36) takes the same form as the PHD update equation except that here we replace the measurement-updated likelihood with a Markov transition likelihood. This leads to a PHD update formula of the form

$$D_{k'+1|k'}(x|Y) = M(x|Y) \cdot D_{k'|k'}(x), \quad (37)$$

where $M(x|Y)$ is a PHD pseudo-likelihood Markov transition analogous to the PHD pseudo-likelihood for the PHD update. By substituting into equation (35), the smoothed PHD then becomes

$$D_{k'|k}(x) = D_{k'|k'}(x) \cdot \int M(x|Y) \cdot p_{k'+1|k}(Y)\delta Y. \quad (38)$$

Forward-Backward PHD Smoother

If we assume that the filtered multi-object density is a Poisson or an i.i.d. cluster process, we can use pseudo-likelihoods Markov transitions analogous to the PHD and CPHD pseudo-likelihoods in equations 16.109 and 16.328 in [11] to find the PHD and CPHD forward-backward smoothers. The PHD smoother has been derived using Finite Set Statistics by Mahler, Vo and Vo [5] and it was also been proposed by Kiruba *et al.* [4]. I propose a simpler approach by using the direct analogy with the PHD pseudo-likelihood: replace the likelihood L_z with Markov transition $M_y(x) := f_{k'+1|k'}(y|x)$, probability of detection p_D with probability of survival p_S, and clutter PHD κ with birth PHD $\gamma_{k'}$. Hence the pseudo-likelihood Markov transition becomes (compare with equation (31))

$$M(x|Y) = (1 - p_S(x)) + \sum_{y\in Y} \frac{p_S(x)M_y(x)}{\gamma_{k'}(y) + D_{k'|k'}[p_S \cdot M_y]}, \quad (39)$$

so that the forward-backward PHD smoother then becomes

$$\begin{aligned} D_{k'|k}(x) &= D_{k'|k'}(x) \quad (40)\\ &\times \int \left((1 - p_S(x)) + \sum_{y\in Y} \frac{p_S(x)M_y(x)}{\gamma_{k'}(y) + D_{k'|k'}[p_S \cdot M_y]} \right) p_{k'+1|k}(Y)\delta Y \\ &= \left((1 - p_S(x)) + D_{k'+1|k}\left[\frac{p_S(x)M_y(x)}{\gamma_{k'} + D_{k'|k'}[p_S \cdot M_y]} \right] \right) D_{k'|k'}(x). \end{aligned}$$

where the last line follows from Campbell's Theorem (the same argument is used in [5]).

The discovery of the mathematical parallel between the p.g.fl. Bayes update and the term inside the brackets in equation (34) enables us to exploit many of the existing results for multi-object Bayes filtering [11, 7, 14, 19]. However, when Campbell's Theorem can not be applied, such as in the CPHD filter pseudo-likelihood (Equation 16.328 in [11]), we still have a set integral which may not be tractable.

5 Summary

I derive Bayesian multi-object forward-backward smoothers via first-order multi-object densities using a new form of smoothing probability generating functional that exploits the Bayes update probability generating functional.

Acknowledgements

The author is a Royal Academy of Engineering/ EPSRC Research Fellow. This work was supported by the Engineering and Physical Sciences Research Council EP/H010866/1.

References

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

[2] D. E. Clark. Joint Target-Detection and Tracking Smoothers. *SPIE Defense and Security Programme*, 2009.

[3] D. Clark, B-T. Vo, and B-N. Vo. Sequential Monte Carlo smoothing for the Joint Target-Detection and Tracking Problem. *Proc. International Conference on Information Fusion.*, July 2009.

[4] N. Nandakumaran, K. Punithakumar, and T. Kirubarajan. Improved Multi-Target Tracking Using Probability Hypothesis Density Smoothing. *Proc. SPIE Conference of Signal and Data Processing of Small Targets, San Diego, Volume 6699, CA, August 2007.*

[5] R. P. S. Mahler and B.-N. Vo and B.-T. Vo. The Forward-Backward Probability Hypothesis Density Smoother. *Proc. ISIF Int. Conf. on Information Fusion*, 2010.

[6] B.-N. Vo, B.-T. Vo, and R. P. S. Mahler. A Closed Form Solution to the Probability Hypothesis Density Smoother. *Proc. ISIF Int. Conf. on Information Fusion*, 2010.

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

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

[9] S. K. Srinivasan. *Stochastic Point Processes and Their Applications*. Griffin's Statistical Monographs and Courses, 1973.

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

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

[12] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. New York: Wiley, 2 edition, 1995.

[13] R. Mahler, B.-N. Vo, , and B.-T. Vo. Multi-Target Forward Backward Smoothing with the Probability Hypothesis Density . *submitted to IEEE Trans. AES*, 2009.

[14] R. Mahler. PHD Filters of Higher Order in Target Number. *IEEE Trans. AES. Vol. 43 No 4*, pages 1523–1543, 2007.

[15] B. Vo, S. Singh, and A. Doucet. Sequential Monte Carlo methods for Multi-target Filtering with Random Finite Sets. *IEEE Trans. Aerospace Elec. Systems*, 41, No.4:1224–1245, 2005.

[16] T. Zajic and R. Mahler. A particle-systems implementation of the PHD multitarget tracking filter. *SPIE*

Vol. 5096 Signal Processing, Sensor Fusion and Target Recognition, pages 291–299, 2003.

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

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

[19] B. T Vo, B-N Vo, and A Cantoni. The Cardinality Balanced Multi-Target Multi-Bernoulli Filter and Its Implementations. *IEEE Trans. Signal Processing*, 57:7:409–423, 2009.