Derivation of the PHD filter based on direct Kullback-Leibler divergence minimisation

Ángel F. García-Fernández, Ba-Ngu Vo

Dept. of Electrical and Computer Engineering

Curtin University

Australia

{angel.garciafernandez, ba-ngu.vo}@curtin.edu.au

Abstract—**In this paper, we provide a novel derivation of the probability hypothesis density (PHD) filter without using probability generating functionals or functional derivatives. The PHD filter fits in the context of assumed density filtering and implicitly performs Kullback-Leibler divergence (KLD) minimisations after the prediction and update steps. The novelty of this paper is that the KLD minimisation is performed directly on the multitarget prediction and posterior densities.**

Keywords—**random finite sets; PHD filter; multiple target tracking; Kullback-Leibler divergence.**

I. INTRODUCTION

In the random finite set (RFS) approach to multiple target tracking, we are interested in estimating a set, which contains the states of the targets at the current time, based on a sequence of measurements. State estimation is based on the posterior probability density function (PDF), i.e., the PDF of the current set of targets given the sequence of measurements, as it contains all information of interest about target states at the current time. Theoretically, the posterior can be calculated recursively by the prediction and update steps. However, in general, the posterior cannot be computed in closed-form so approximations are necessary [1], [2].

A widely-used approximation to the posterior is provided by the probability hypothesis density (PHD) filter [3], which has been applied in different fields such as multitarget tracking [4], robotics [5], computer vision [6], road mapping [7] and sensor control [8], [9]. Importantly, the PHD filter recursion avoids the computational complexity of evaluating measurement-to-target association hypotheses. The PHD filter was originally derived by Mahler using probability generating functionals (PGFLs) and functional derivatives [3], [10]. The Faà di Bruno's formulae for functional derivatives [2], [11], [12] provides elegant derivations of the original, as well as more general PHD filters, based on PGFLs. Alternative derivations for the PHD filter based on measure theory [13] have been proposed in [14], [15]. An interesting interpretation of the PHD filter in terms of the probability existence of targets on infinitesimal regions of the state space is given in [16].

This paper presents a new derivation of the PHD filter without resorting to PGFLs or functional derivatives. An extended version of this work which also includes a new derivation of the cardinalised PHD (CPHD) filter [1] can be found in [17]. We think the presented derivation can make the PHD filter accessible to a wider audience. The PHD filter fits in the context of assumed density filtering (ADF) [18], [19]. In ADF, we propagate a certain type of PDF in the Bayesian filtering recursion. As the output PDF of the prediction and/or update steps might not be of the considered type, we have to approximate it by a PDF of the type under consideration to continue with the filtering recursion. Ideally, this approximation is obtained by some optimality criterion such as minimising the Kullback-Leibler divergence (KLD) w.r.t. the true PDF [20], [21]. KLD minimisations are sometimes referred to as moment projection (M-projection) or information-projection (I-projection) depending on the order of the PDFs in the KLD [20]. The PHD filter follows this scheme with M-projections using Poisson PDFs [1].

In a Poisson RFS, the number of elements in the set is Poisson distributed and its elements are independent identically distributed (IID). If the prior is Poisson, the result of Bayes' rule with the standard measurement model, is no longer Poisson so the PHD filter performs KLD minimisation to approximate the posterior [3]. In the prediction step with the usual modelling assumptions and Poisson input, the output is Poisson if there is no target spawning and the RFS of new born targets is Poisson. Otherwise, KLD minimisation is performed to obtain a Poisson PDF. The resulting recursions are illustrated in Figures 1 and 2. By applying KLD minimisation to the output PDF of the prediction and update steps, we directly obtain the PHD filter recursion.

The rest of this paper is organised as follows. In Section II, we review the Bayesian filtering recursion using RFSs. In Section III, we provide an useful theorem for KLD minimisation using Poisson RFSs. The PHD filter equations are derived in Section IV. Finally, conclusions are given in Section V.

II. BAYESIAN FILTERING WITH RANDOM FINITE SETS

In this section we review the Bayesian filtering recursion with RFSs, which consists of the usual prediction and update steps. As we only need to consider one prediction and update step, we omit the time index of the filtering recursion for notational simplicity.

In the standard RFS framework to target tracking, a single target state $x \in \mathbb{R}^{n_x}$ and the state $X \in \mathcal{F}(\mathbb{R}^{n_x})$, which denotes the space of all finite subsets of $\mathbb{R}^{n_x}$, so X is a set whose elements are single target state vectors. In the
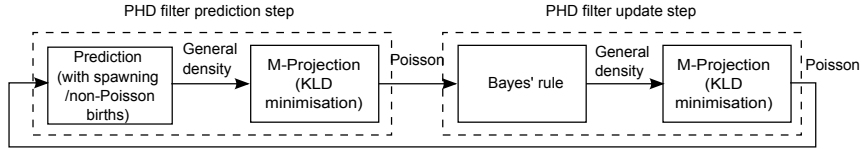
Fig. 2. PHD filter with non-Poisson births and/or target spawning diagram. The PHD filter assumes that the PDFs involved are Poisson. The output of the prediction and Bayes' rule, which are given by Eqs. (4) and (1), are no longer Poisson but, in order to be able to perform filtering, it obtains the best Poisson approximation to the corresponding density by minimising the KLD divergence (M-projection).
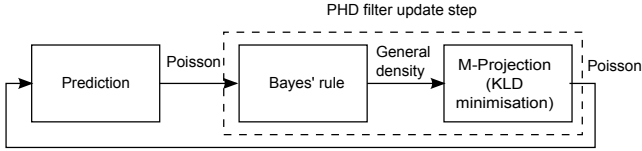

Fig. 1. PHD filter with Poisson births and no target spawning diagram. The PHD filter assumes that the PDFs involved are Poisson. The output of Bayes' rule, which is given by Eq. (1), is no longer Poisson but, in order to be able to perform filtering, it obtains the best Poisson approximation to the posterior by minimising the KLD divergence (M-projection).

update step, the state is observed by measurements that are represented as a set $Z \in \mathcal{F}(\mathbb{R}^{n_z})$. Given a prior PDF $\nu(\cdot)$ and the PDF $f(Z|X)$ of the measurement Z given the state X, the posterior PDF of X after observing Z is given by Bayes' rule [3]

$$q(X) = \frac{f(Z|X)\nu(X)}{p(Z)} \quad (1)$$

where the PDF $p(Z)$ of the measurement is given by the set integral

$$p(Z) = \int f(Z|X)\nu(X)\delta X \quad (2)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int f(Z|\{x_1,...,x_n\})$$
$$\times \nu(\{x_1,...,x_n\}) d(x_1,...,x_n). \quad (3)$$

The Bayesian filtering recursion is completed with the prediction step. Given a posterior PDF $q(\cdot)$, the prior PDF $\omega(\cdot)$ at the next time step is given by the Chapman-Kolmogorov equation

$$\omega(X') = \int \gamma(X'|X) q(X) \delta X \quad (4)$$

where $X' \in \mathcal{F}(\mathbb{R}^{n_x})$ denotes the state at the next time step and $\gamma(X'|X)$ is the PDF of the state X' given the state X.

As in single target filtering, the prediction and update steps cannot be computed in closed-form in general. In single target filtering, a well-known technique is assumed density filtering [18], [19], in which the PDF before the prediction and update step is assumed to be of a certain form. Then, we compute the output PDF via Bayes' rule or Chapman-Kolmogorov equation and project it to the same family of PDFs so that the Bayesian recursion can be performed. Ideally, the projection should be performed by minimising the Kullback-Leibler divergence

[21]. This is exactly what the PHD filter does with Poisson PDFs, respectively, see Figures 1 and 2.

We also want to remark at this point the similarity between the PDF of the measurement (2) and the prediction step (4). As we will see, unless we consider target spawning/non-Poisson births in the PHD filter, both equations are identical. Therefore, we first compute the update step, which requires the calculation of (2), and use this knowledge to derive the prediction step.

III. KULLBACK-LEIBLER MINIMISATION

As indicated in the previous section, the PHD filtering recursion involves KLD minimisations. As a result, in order to derive the filters based on this framework, we need to indicate how KLD minimisations are performed. In Section III-A, we first review the concepts of PHD and cardinality distribution of an RFS density. In Section III-B, we review a known result for KLD minimisation for Poisson RFS.

A. PHD and Cardinality Distribution

Given an RFS density $\pi(\cdot)$, its PHD is [1, Eq. (16.33)]

$$D_\pi(x) = \int \pi(\{x\} \cup X) \delta X$$
$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int \pi(\{x, x_1,...,x_n\}) d(x_1,...,x_n). \quad (5)$$

The PHD is also called intensity function in stochastic geometry [22]. The cardinality distribution of $\pi(\cdot)$ is [1, Eq. (11.115)]

$$\rho_\pi(n) = \frac{1}{n!} \int \pi(\{x_1,...,x_n\}) d(x_1,...,x_n). \quad (6)$$

B. Poisson RFS

The PHD filter propagates a Poisson PDF. If $\nu(\cdot)$ is Poisson, it can be written as [1, Eq. (11.122)]

$$\nu(\{x_1,...,x_n\}) = e^{-\lambda_\nu} \lambda_\nu^n \prod_{j=1}^{n} \breve{\nu}(x_j) \quad (7)$$

where $\breve{\nu}(\cdot)$ is a PDF on the single target state space and $\lambda_\nu \geq 0$. A Poisson PDF is characterised by its PHD $D_\nu(x) = \lambda_\nu \breve{\nu}(x)$. Applying (6), its cardinality distribution is Poisson with parameter λ_ν

$$\rho_\nu(n) = \frac{e^{-\lambda_\nu} \lambda_\nu^n}{n!}$$

Theorem 1. *Given an RFS density $\pi\left(\cdot\right)$, the Poisson density $\nu\left(\cdot\right)$ that minimises the KLD*

$$\mathrm{D}\left(\pi\left\|\nu\right.\right)=\int\pi\left(X\right)\log\frac{\pi\left(X\right)}{\nu\left(X\right)}\delta X$$

is characterised by its PHD $D_\nu\left(\cdot\right)=D_\pi\left(\cdot\right)$.

Theorem 1 is proved in [3, Theorem 4].

IV. PHD FILTER DERIVATION

In this section, we provide the derivation of the PHD filter based on direct KLD minimisation. In Section IV-A, we review the PHD filter update and the proposed proof is provided in Section IV-B. Lastly, in Section IV-C, we address the prediction step.

A. Review of the PHD Filter Update

The PHD filter update is developed under the assumptions

- U1 The set $Z = Z_1 \cup Z_2$ where Z_1 and Z_2 are the independent RFSs of measurements coming from targets and clutter, respectively.
- U2 Given $X = \{x_1, ..., x_n\}$, $Z_1 = \cup_{i=1}^n \tilde{Z}_{1,i}$ where $\tilde{Z}_{1,i} = \oslash$ with probability $1 - p_D\left(x_i\right)$, otherwise $\tilde{Z}_{1,i} = \{z_i\}$ where z_i is distributed according to a PDF $l\left(\cdot | x_i\right)$ and sets $\tilde{Z}_{1,i}$ $i = 1, ..., n$ are independent.
- U3 The set Z_2 of clutter measurements is Poisson with RFS density $c\left(\cdot\right)$.
- U4 The prior is Poisson.

Under Assumptions U1, U2 and U3, which define the standard measurement model, the PDF of the measurement given the state is [1, Eq. (12.42)]

$$
\begin{aligned}
&f\left(\{z_1, ..., z_{n_z}\} | \{x_1, ..., x_n\}\right)\\
&= e^{\lambda_c}\left[\prod_{i=1}^{n_z}\lambda_c\check{c}\left(z_i\right)\right]\left[\prod_{i=1}^n\left(1 - p_D\left(x_i\right)\right)\right]\\
&\quad \times \sum_{\sigma\in\Xi_{n,n_z}}\prod_{i:\sigma_i>0}\frac{p_D\left(x_i\right)l\left(z_{\sigma_i}|x_i\right)}{\left(1 - p_D\left(x_i\right)\right)\lambda_c\check{c}\left(z_{\sigma_i}\right)}
\end{aligned}
\tag{8}
$$

where Ξ_{n,n_z} is a set which contains all the vectors $\sigma = \left(\sigma_1, ..., \sigma_n\right)$ that indicate associations of n_z measurements to n targets, which can be either detected or undetected, taking into account that only one measurement can be associated with a given target. If $\sigma \in \Xi_{n,n_z}$, $\sigma_i \in \{1, ..., n_z\}$ indicates the measurement associated with target i and $\sigma_i = 0$ indicates that target i has not been detected.

The PHD filter update equation is [3]

$$
\begin{aligned}
D_q\left(x\right) &= \left(1 - p_D\left(x\right)\right)D_\nu\left(x\right) + p_D\left(x\right)\\
&\quad \times \sum_{z\in Z}\frac{l\left(z|x\right)D_\nu\left(x\right)}{D_c\left(z\right) + \int p_D\left(x'\right)l\left(z|x'\right)D_\nu\left(x'\right)dx'}
\end{aligned}
\tag{9}
$$

where $x' \in \mathbb{R}^{n_x}$. Note that (9) is the PHD of the posterior $q\left(\cdot\right)$, see (1), under Assumptions U1-U4. This is proved in the next subsection.

B. Proof of the PHD Filter Update

The aim is to compute the PHD of the posterior, which characterises the best Poisson approximation to the posterior in the KLD sense, see Theorem 1. Using (1) and (5), we get

$$
\begin{aligned}
D_q(x) &= \frac{1}{p\left(Z\right)}\sum_{n=0}^{\infty}\frac{1}{n!}\int f(Z| \{x, x_1, ..., x_n\})\\
&\quad \times \nu\left(\{x, x_1, ..., x_n\}\right)d\left(x_1, ..., x_n\right).
\end{aligned}
\tag{10}
$$

First, we compute the denominator in Section IV-B1 and then we complete the proof in Section IV-B2.

1) Density of the measurement: The denominator of (10) corresponds to the PDF $p\left(\cdot\right)$ of Z. This PDF can be obtained from well-known results of Poisson point processes theory [15]. According to Assumption U1, Z is the union of two independent sets Z_1 and Z_2. Under Assumption U2, set Z_1 comes from performing thinning [22] on X with a probability $p_D\left(\cdot\right)$ followed by a displacement [23] with Markov transition $l\left(\cdot | x\right)$. As X is Poisson distributed, see Assumption U4, we can apply the thinning and displacement theorems [22], [23] so that we get that Z_1 is Poisson distributed with intensity

$$\lambda_\nu\int p_D\left(x\right)l\left(z|x\right)\check{\nu}\left(x\right)dx.$$

For completeness, we also derive the previous result in Appendix A. Under Assumption U1, Z is the union of two independent Poisson RFSs. Consequently, we can apply the superposition theorem [23], which says that Z is Poisson distributed with intensity given by the sum of the intensities

$$D_p\left(z\right) = \lambda_c\check{c}\left(z\right) + \lambda_\nu\int p_D\left(x\right)l\left(z|x\right)\check{\nu}\left(x\right)dx. \tag{11}$$

Therefore, the denominator of (10) is given by

$$
\begin{aligned}
p\left(Z\right) &= e^{-\lambda_\nu\int p_D(x)\check{\nu}(x)dx - \lambda_c}\\
&\quad \prod_{z\in Z}\left[\lambda_c\check{c}\left(z\right) + \lambda_\nu\int p_D\left(x\right)l\left(z|x\right)\check{\nu}\left(x\right)dx\right].
\end{aligned}
\tag{12}
$$

2) Rest of the proof: We perform the following decomposition

$$
\begin{aligned}
&f\left(Z | \{x, x_1, ..., x_n\}\right)\\
&= \left(1 - p_D\left(x\right)\right)f\left(Z | \{x_1, ..., x_n\}\right)\\
&\quad + p_D\left(x\right)\sum_{z\in Z}l\left(z|x\right)f\left(Z\setminus\{z\} | \{x_1, ..., x_n\}\right)
\end{aligned}
\tag{13}
$$

where $B\setminus A = \{z\in B| z\notin A\}$. Note that $f\left(Z|X\right)$ goes through all the possible data association hypotheses, see (8). There are two hypotheses for target x, it can be either detected or not detected. If it is not detected, which happens with probability $\left(1 - p_D\left(x\right)\right)$, all the measurements have originated from the rest of the targets or clutter. This is represented by the first term of (13). The other hypothesis is that x is detected, which happens with probability $p_D\left(x\right)$. If it is detected, it can be associated with any of the $z \in Z$ measurements and the rest $Z\setminus\{z\}$ of the measurements have originated from the rest of the targets or clutter. This is represented by the second term of (13).

Substituting (7) and (13) into (10), we obtain

$$D_q(x) = \frac{1}{p(Z)} \sum_{n=0}^{\infty} \frac{1}{n!} (1 - p_D(x)) \int f(Z | \{x_1, ..., x_n\})$$

$$\times e^{-\lambda_\nu} \lambda_\nu^{n+1} \breve{\nu}(x) \prod_{j=1}^{n} \breve{\nu}(x_j) d(x_1, ..., x_n)$$

$$+ \frac{1}{p(Z)} \sum_{n=0}^{\infty} \frac{1}{n!} p_D(x) \sum_{z \in Z} l(z|x)$$

$$\times \int f(Z \setminus \{z\} | \{x_1, ..., x_n\})$$

$$\times e^{-\lambda} \lambda_\nu^{n+1} \breve{\nu}(x) \prod_{j=1}^{n} \breve{\nu}(x_j) d(x_1, ..., x_n). \quad (14)$$

Using (3) and (7) in (14), we obtain

$$D_q(x) = (1 - p_D(x)) \lambda_\nu \breve{\nu}(x)$$

$$+ \frac{1}{p(Z)} p_D(x) \sum_{z \in Z} l(z|x) \lambda_\nu \breve{\nu}(x) f(Z \setminus \{z\}). \quad (15)$$

Finally, we substitute (12) into (15) to get

$$D_q(x) = (1 - p_D(x)) \lambda_\nu \breve{\nu}(x) + p_D(x) \lambda_\nu \breve{\nu}(x)$$

$$\sum_{z \in Z} \frac{l(z|x)}{\lambda_c \breve{c}(z) + \int p_D(x') l(z|x') \lambda_\nu \breve{\nu}(x') dx'}.$$

which completes the proof of (9).

C. PHD Filter Prediction

The PHD prediction assumes

- P1 The set $X' = X_1' \cup X_2' \cup X_3'$ where X_1', X_2' and X_3' are the independent sets of surviving targets, newborn targets and spawned targets, respectively.
- P2 Given $X = \{x_1, ..., x_n\}$, $X_1' = \cup_{i=1}^{n} \tilde{X}_{1,i}'$ where $\tilde{X}_{1,i}' = \oslash$ with probability $1 - p_S(x_i)$, otherwise $\tilde{X}_{1,i}' = \{x_i'\}$ where x_i' has a PDF $g(\cdot|x_i)$ and sets $\tilde{X}_{1,i}'$ $i = 1, ..., n$ are independent.
- P3 The set X_2' of new born targets has an RFS density $b(\cdot)$.
- P4 Given $X = \{x_1, ..., x_n\}$, $X_3' = \cup_{i=1}^{n} X_{3,i}''$ where $X_{3,i}''$ has PHD $D_\xi(\cdot|x_i)$ and $X_{3,i}''$ $i = 1, ..., n$ are independent.
- P5 The posterior $q(\cdot)$ is Poisson.

The PHD of the prior at the next time step is [3]

$$D_\omega(x') = D_b(x')$$

$$+ \int [p_S(x) g(x'|x) + D_\xi(x'|x)] D_q(x) dx. \quad (16)$$

If there is no target spawning and the density $b(\cdot)$ is Poisson, we see that U1-U4 are equivalent to P1-P3 and P5, therefore, the PDF of the measurement (2) and the prediction step (4) are analogous. Consequently, in this case, it can be directly established that the density $\omega(\cdot)$ is Poisson, with PHD given by (16) setting $D_\xi(\cdot) = 0$, and there is no need to perform KLD minimisation, as illustrated by Figure 1. Under

Assumption P4, the set X_3' of spawned targets constitutes a cluster process with centers given by $q(\cdot)$ [24, Chap. 6]. As a result, its PHD is given by

$$D_\xi(x') = \int D_\xi(x'|x) D_q(x) dx$$

where we have used (6.3.3) in [24]. By the superposition theorem [22] and Assumption P1, the PHD of X' is given by the sum of the PHDs of X_1', X_2', X_3' so we get (16). Due to the fact that (16) is the PHD of the prior, it represents the best Poisson fit to the prior in the KLD sense, see Theorem 1, as illustrated in Figure 2.

V. Conclusions

In this paper, we have provided a derivation of the PHD filter based on KLD minimisation of the PDFs, without resorting to PGFLs or functional derivatives. This derivation uses an intuitive decomposition of the multi-target likelihood function. By assuming a Poisson prior and performing KLD minimisation, the PHD filter update avoids data association. This has important practical implications as the computational burden of the filter is significantly lowered, albeit with some drawbacks such as the spooky effect [25].

It should also be noted that the resulting minimised KLD is an indicator of the performance of the filter. The lower this KLD is, the closer the approximated posterior is to the true posterior and performance improves. This analysis for the (non-linear) Kalman filter update in single target applications was performed in [26]. Unfortunately, even though we can minimise the KLD to obtain the PHD filters, closed-form formulas for the resulting KLD cannot be obtained though it can be approximated using Monte Carlo integration.

VI. Acknowledgements

Appendix A

We proceed to compute the density of the target generated measurements. According to Assumption U1, Z is the union of two independent sets Z_1 and Z_2. Here, we calculate the PDF $t(\cdot)$ of Z_1, which is given by

$$t(Z_1) = \int t(Z_1|X)\nu(X)\partial X$$

where $t(Z_1|X)$ is given by (2) removing the terms corresponding to clutter. Therefore,

$$t(\{z_1, ... z_{n_z}\})$$

$$= \sum_{n=n_z}^{\infty} \frac{1}{n!} \lambda^n e^{-\lambda} n_z! \binom{n}{n_z}$$

$$\times \prod_{i=1}^{n_z} \int l(z_i|x) \breve{\nu}(x) p_D(x) dx$$

$$\times \prod_{i=n_z}^{n} \int (1 - p_D(x)) \breve{\nu}(x) dx$$

$$= e^{-\lambda} n_z! \prod_{i=1}^{n_z} \int l\left(z_i|x\right) \breve{\nu}\left(x\right) p_D\left(x\right) dx$$

$$\times \sum_{n=n_z}^{\infty} \frac{1}{n!} \lambda^n \begin{pmatrix} n \\ n_z \end{pmatrix} \left[\int \left(1 - p_D\left(x\right)\right) \breve{\nu}\left(x\right) dx \right]^{n-n_z} \tag{17}$$

In the first step of the previous derivation, we have used the fact given n targets and n_z measurements and no clutter, there are

$$n_z! \begin{pmatrix} n \\ n_z \end{pmatrix}$$

possible measurement-to-target associations. In addition, as the targets are Poisson distributed, which implies that they are IID, all the associations produce the same value of the integral so we can just take one and multiply it by the previous number. Using (18) in Appendix B, we get

$$\sum_{n=n_z}^{\infty} \begin{pmatrix} n \\ n_z \end{pmatrix} \frac{e^{-\lambda} \lambda^n}{n!} \left[\int \left(1 - p_D\left(x\right)\right) \breve{\nu}\left(x\right) dx \right]^{n-n_z}$$
$$= e^{-\lambda \int p_D(x) \breve{\nu}(x) dx} \frac{\lambda^{n_z}}{n_z!}$$

Plugging this equation into (17), we get

$$t\left(\{z_1, \ldots z_{n_z}\}\right) = \lambda^{n_z} e^{-\lambda \int p_D(x)\breve{\nu}(x) dx}$$
$$\times \prod_{i=1}^{n_z} \int l\left(z_i|x\right) \breve{\nu}\left(x\right) p_D\left(x\right) dx$$

Comparing this equation with (7), we deduct that Z_t is Poisson distributed with intensity

$$D_t\left(z\right) = \lambda \int p_D\left(x\right) l\left(z|x\right) \breve{\nu}\left(x\right) dx$$

APPENDIX B

In this appendix we prove that, given $k \in \mathbb{N}$, $a \in [0, 1]$, $\lambda > 0$, the following equality is met

$$e^{-\lambda} \sum_{n=k}^{\infty} \begin{pmatrix} n \\ k \end{pmatrix} \frac{\lambda^n}{n!} a^k \left(1 - a\right)^{n-k} = e^{-\lambda a} \frac{\left(\lambda a\right)^k}{k!} \tag{18}$$

This is done as follows

$$e^{-\lambda} \sum_{n=k}^{\infty} \begin{pmatrix} n \\ k \end{pmatrix} \frac{\lambda^n}{n!} a^k \left(1 - a\right)^{n-k}$$
$$= \frac{e^{-\lambda} a^k}{k!} \sum_{n=k}^{\infty} \frac{\lambda^n}{\left(n-k\right)!} \left(1 - a\right)^{n-k}$$
$$= \frac{e^{-\lambda} \left(\lambda a\right)^k}{k!} \sum_{n=k}^{\infty} \frac{1}{\left(n-k\right)!} \left(\lambda \left(1 - a\right)\right)^{n-k}$$
$$= \frac{e^{-\lambda} \left(\lambda a\right)^k}{k!} e^{\lambda(1-a)}$$
$$= e^{-\lambda a} \frac{\left(\lambda a\right)^k}{k!}$$

REFERENCES

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

[2] ——, *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.

[3] ——, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[4] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, Oct. 2005.

[5] J. Mullane, B.-N. Vo, M. D. Adams, and B.-T. Vo, "A random-finite-set approach to Bayesian SLAM," *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 268–282, April 2011.

[6] E. Maggio, M. Taj, and A. Cavallaro, "Efficient multitarget visual tracking using random finite sets," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1016–1027, Aug. 2008.

[7] C. Lundquist, L. Hammarstrand, and F. Gustafsson, "Road intensity based mapping using radar measurements with a probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1397–1408, April 2011.

[8] B. Ristic, B.-N. Vo, and D. Clark, "A note on the reward function for PHD filters with sensor control," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1521–1529, April 2011.

[9] H. Hoang, B.-N. Vo, B.-T. Vo, and R. Mahler, "The Cauchy-Schwarz divergence for Poisson point processes," *IEEE Transactions on Information Theory*, vol. 61, no. 8, pp. 4475–4485, Aug. 2015.

[10] R. Mahler, "PHD filters of higher order in target number," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1523–1543, October 2007.

[11] D. Clark and R. Mahler, "Generalized PHD filters via a general chain rule," in *15th International Conference on Information Fusion*, July 2012, pp. 157–164.

[12] D. E. Clark and J. Houssineau, "Faà di Bruno's formula and spatial cluster modelling," *Spatial Statistics*, vol. 6, pp. 109–117, 2013.

[13] B.-N. Vo and S. Singh, "On the Bayes filtering equations of finite set statistics," in *5th Asian Control Conference*, July 2004, pp. 1264–1269.

[14] F. Caron, P. D. Moral, A. Doucet, and M. Pace, "On the conditional distributions of spatial point processes," *Advanced Applied Probability*, vol. 43, no. 2, pp. 301–307, 2011.

[15] S. S. Singh, B.-N. Vo, A. Baddeley, and S. Zuyev, "Filters for spatial point processes," *SIAM Journal of Control and Optimization*, vol. 48, no. 4, pp. 2275–2295, 2009.

[16] O. Erdinc, P. Willett, and Y. Bar-Shalom, "The bin-occupancy filter and its connection to the PHD filters," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4232–4246, Nov. 2009.

[17] A. F. García-Fernández and B.-N. Vo, "Derivation of the PHD and CPHD filters based on direct Kullback-Leibler divergence minimisation," *accepted for publication in IEEE Transactions on Signal Processing*, 2015.

[18] P. Maybeck, *Stochastic models, estimation and control*. Academic Press, 1982, vol. 2.

[19] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013.

[20] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 2009.

[21] T. P. Minka, "Expectation propagation and approximate Bayesian inference," in *Proceedings of the 17th Conference in Uncertainty in Artifical Intelligence*, 2001, pp. 362–369.

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

[23] J. F. C. Kingman, *Poisson processes*. Oxford University Press, 1993.

[24] D. Daley and D. Vere-Jones, *An introduction to the theory of point processes: Volume I: Elementary Theory and methods*. Springer, 2003.

[25] D. Franken, M. Schmidt, and M. Ulmke, ""Spooky action at a distance" in the cardinalized probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1657–1664, Oct. 2009.

[26] M. R. Morelande and A. F. García-Fernández, "Analysis of Kalman filter approximations for nonlinear measurements," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5477–5484, Nov. 2013.