Poisson Multi-Bernoulli Mixture Filter: Direct Derivation and Implementation

ÁNGEL F. GARCÍA-FERNÁNDEZ (iD)
University of Liverpool, Liverpool, U.K.

JASON L. WILLIAMS, Senior Member, IEEE
Defence Science and Technology Group, Edinburgh, SA, Australia and also School of Electrical and Electronic Engineering at the University of Adelaide, Adelaide, SA, Australia

KARL GRANSTRÖM (iD), Member, IEEE
LENNART SVENSSON (iD)
Chalmers University of Technology, Gothenburg, Sweden

We provide a derivation of the Poisson multi-Bernoulli mixture (PMBM) filter for multitarget tracking with the standard point target measurements without using probability generating functionals or functional derivatives. We also establish the connection with the δ-generalized labeled multi-Bernoulli (δ-GLMB) filter, showing that a δ-GLMB density represents a multi-Bernoulli mixture with labeled targets so it can be seen as a special case of PMBM. In addition, we propose an implementation for linear/Gaussian dynamic and measurement models and how to efficiently obtain typical estimators in the literature from the PMBM. The PMBM filter is shown to outperform other filters in the literature in a challenging scenario.

I. INTRODUCTION

Multiple target tracking (MTT) is an important problem with many different uses, for example, in aerospace applications, surveillance, air traffic control, computer vision, and autonomous driving [1]–[6]. In MTT, a variable and unknown number of targets appear, move, and disappear from a scene of interest. At each time step, these targets are observed through noisy measurements, possibly coming from multiple sensors [7], [8], and the aim is to infer where the targets are at each time step.

The random finite set (RFS) framework is widely used to model this problem in a Bayesian way [9]. Here, the usual setup is to consider the state of the system at the current time as a set of targets. There are a variety of dynamic models [10] for this set of targets but it is usually assumed that it evolves in time according to a Markov process, which also accounts for target births/deaths. There are also different widely used measurement models, for example, standard (point target) [9], extended target [11], [12], or track-before-detect [13], [14] measurement models.

As in any Bayesian setting, the information of interest about the targets at the current time step is contained in the (multitarget) density of the current set of targets given present and past measurements. In theory, this density can be computed via the prediction and update steps of the Bayesian filtering recursion. However, in general, this computation is intractable and general, computationally expensive approximations such as particle filters should be used [15]. Nevertheless, as we explain next, there are families of multitarget densities that are conjugate prior for some models that enable easier and more efficient computation.

In Bayesian probability theory, a family of probability distributions is conjugate for a given likelihood function if the posterior distribution for any member of this family also belongs to the same family [16]. In MTT filtering, it is especially useful for computational reasons to consider conjugate priors in which the posterior distributions can be written explicitly in terms of single target Bayesian updates, which might not admit a closed-form expression [17], [18]. Additionally, in MTT, it is convenient to introduce conjugacy for the prediction step. That is, a multitarget density is conjugate with respect to a dynamic model if the same family is preserved after performing the prediction step. This conjugacy property for the prediction and update steps is quite important in the RFS context as it allows the posterior to be written in terms of single target predictions and updates, which are much easier to compute/approximate than full multitarget predictions and updates. Due to this important characteristic, in general, when we refer to MTT conjugacy, we are referring to a family of distributions which is closed under both prediction and update steps. Note that, in MTT, we are generally dealing with conjugate prior mixtures in which the number of mixture components can grow, due to the data association. This implies that the conjugate prior does not have a fixed dimensional sufficient statistic, even if the single target densities have it.

We proceed to describe the two conjugate priors in the literature for the standard (point target) measurement model, in which the set of measurements at a given time comprises of clutter and one or zero measurements per target. The first conjugate prior consists of the union of a Poisson process and a multi-Bernoulli mixture (MBM) [18]. Importantly, the MBM, which considers all the data association hypotheses, can be implemented efficiently using a track-oriented multiple hypotheses tracking (MHT) formulation [19]. The Poisson part considers all targets that have never been detected and enables an efficient management of the number of hypotheses covering potential targets [18]. The second conjugate prior was presented for labeled targets in [17]. In the usual radar tracking case, in which targets do not have a unique ID, labels are artificial variables that are added to the target states with the objective of estimating target trajectories [13], [17], [20]–[22]. With them, we can also obtain conjugate priors, as in the δ-generalized labeled multi-Bernoulli (δ-GLMB) filter [17], [21].

The Poisson multi-Bernoulli mixture (PMBM) filter in [18], which is based on the previously mentioned conjugate prior, was derived by using probability generating functionals (PGFLs) and functional derivatives [23]. These are very important tools for deriving RFS filters, such as the probability hypothesis density (PHD) or cardinalized PHD (CPHD) filters [23], [24]. However, non-PGFL derivations are also useful as they can provide insights about the structure of the filter and make the understanding of the filter accessible to more researchers, as was done in [25] for the PHD and CPHD filters.

The main aim of this paper is to make the PMBM filter accessible to a wider audience from a theoretical and practical point of view. In order to do so, we make the following contributions. In Section III, we provide a derivation of the PMBM filter for point measurements that does not rely on PGFLs or functional derivatives, improving the accessibility of these results and providing more insight into the structure of the solution. In Section IV, we show that the δ-GLMB (multitarget) density can be seen as a special case of a PMBM on a labeled state space, and discuss the benefits of the PMBM form. Section V proposes an implementation of the PMBM filter for linear/Gaussian dynamic and measurement models. In Section VI, we provide tractable methods for obtaining the estimators used in MHT and the δ-GLMB filter using the PMBM distribution form. We also provide a third estimator that improves performance for high probability of detection. Finally, Section VII demonstrates the PMBM implementation on a challenging scenario, comparing performance between the three estimators and other multitarget filters. Conclusion is given in Section VIII.

II. BAYESIAN FILTERING WITH RFSS

In Section II-A, we review the Bayesian filtering recursion with RFSs. In Section II-B, we present the likelihood function for the standard point target measurement model.

A. Filtering Recursion

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 for target tracking, we have a single target state $x \in \mathbb{R}^{n_x}$ and a multitarget state $X \in \mathcal{F}(\mathbb{R}^{n_x})$, where X is a set whose elements are single target state vectors and $\mathcal{F}(\mathbb{R}^{n_x})$ denotes the space of all finite subsets of $\mathbb{R}^{n_x}$. In the 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 (multitarget) density $f(\cdot)$ and the (multitarget) density $l(Z|X)$ of the measurement Z given the state X, the posterior multitarget density of X after observing Z is given by Bayes' rule [24]

$$q(X) = \frac{l(Z|X)f(X)}{\rho(Z)} \tag{1}$$

where the normalizing constant is

$$\rho(Z) = \int l(Z|X)f(X)\delta X \tag{2}$$

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

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

$$\omega(X') = \int \gamma(X'|X) q(X) \delta X \tag{4}$$

where $X' \in \mathcal{F}(\mathbb{R}^{n_x})$ denotes the state at the next time step and $\gamma(X'|X)$ is the transition density of the state X' given the state X. We consider the conventional dynamic assumptions for MTT used in the RFS framework [26]: at each time step, a target follows a Markovian process such that it survives with a probability $p_s(\cdot)$ and moves with a transition density $g(\cdot|\cdot)$. New born targets follow a Poisson RFS with intensity $\lambda^b(\cdot)$.

B. Standard Point Target Measurement Model

In this section, we provide the likelihood $l(Z|X)$ for the standard point target measurement model, which is described next. At different parts of this paper, we will make use of different representations of the likelihood, which require the introduction of extra notation. To aid the reader, a summary of this notation is found in Table I.

Given the set $X = \{x_1, \ldots, x_n\}$ of targets, the set Z of measurements is $Z = Z^c \uplus Z_1 \uplus \ldots \uplus Z_n$, where Z^c, Z_1,..., Z_n are independent sets, Z^c is the set of clutter measurements, and Z_i is the set of measurements produced by target i. Symbol $\uplus$ stands for disjoint union, which is used to represent that $Z = Z^c \cup Z_1 \cup \ldots \cup Z_n$ and Z^c, Z_1, $\ldots$, Z_n are mutually disjoint (and possibly empty) [9]. Set Z^c is a Pois-

1) $l(Z|X)$: Density of measurement set Z given set X of targets, defined in (5).
2) $\hat{l}(Z|x)$: Density of measurement set Z given target x, defined in (6).
3) $\bar{l}(z|Y)$: Likelihood of set Y after observing measurement z, defined in (14).
4) $l_o(Z|Y, X_1, \ldots, X_n)$: Density of measurement set Z given sets $Y, X_1, \ldots, X_n$ $|X_i| \le 1$, defined in (25).
5) $t(Z_i|X_i)$: Density of measurement Z_i without clutter given set X_i, $|X_i| \le 1$, defined in (26).

son point process with intensity/PHD $c(\cdot)$. We get $Z_i = \emptyset$ with probability $1 - p_d(x_i)$, which corresponds to the case where the target is not detected, and $Z_i = \{z\}$ where z has a density $p(z|x_i)$ with probability $p_d(x_i)$, which corresponds to the case where the target is detected.

Using the convolution formula for multiobject densities [9, eq. (4.17)], the resulting density $l(\cdot|\cdot)$ of Z given X can be written as

$$l(Z|\{x_1, \ldots, x_n\}) = e^{-\lambda_c} \sum_{Z^c \uplus Z_1 \ldots \uplus Z_n = Z} [c(\cdot)]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i) \tag{5}$$

$$\hat{l}(Z|x) = \begin{cases} p_d(x) \, p(z|x) & Z = \{z\} \\ 1 - p_d(x) & Z = \emptyset \\ 0 & |Z| > 1 \end{cases} \tag{6}$$

where $\lambda_c = \int c(z) \, dz$ and we use the multiobject exponential notation $[c(\cdot)]^Z = \prod_{z \in Z} c(Z)$, $[c(\cdot)]^{\emptyset} = 1$ [17]. The notation in (5) means that for a given Z, we perform a sum that goes through all possible sets Z^c, Z_1,..., Z_n that meet the requirement $Z^c \uplus Z_1 \uplus \ldots \uplus Z_n = Z$. In other words, each term of the sum considers a measurement-to-target association hypothesis. Note that any hypothesis that assigns more than one measurement to a target has zero likelihood, as indicated in the last row of (6). In the next example, we illustrate how the sum in (5) is interpreted as it is widely used in this paper.

EXAMPLE 1 Let us consider $Z = \{z_1, z_2\}$ and $n = 1$ so the sum in (5) goes through all possible sets Z^c and Z_1, such that $Z^c \uplus Z_1 = \{z_1, z_2\}$. These are as follows:

1) $Z^c = \emptyset$ and $Z_1 = \{z_1, z_2\}$;
2) $Z^c = \{z_1\}$ and $Z_1 = \{z_2\}$;
3) $Z^c = \{z_2\}$ and $Z_1 = \{z_1\}$; and
4) $Z^c = \{z_1, z_2\}$ and $Z_1 = \emptyset$.

Nevertheless, as pointed out before, hypotheses that assign two measurements to a target have probability zero, so case 1) can be removed.

III. PROOF OF THE CONJUGACY OF THE PMBM

In this section, we provide a non-PGFL proof of the conjugate prior in [18] for the standard point target measurement model. We first review the conjugate prior in Section III-A. Then, we proceed to derive the update for a Poisson

prior in Section III-B. Based on this preliminary derivation, we perform a Bayesian update on the conjugate prior to show its conjugacy in Section III-C. The prediction step is addressed in Section III-D. We also establish the conjugacy property for MBMs in Section III-E.

A. Conjugate Prior

It was proved in [18] using PGFLs that the union of two independent RFS, one Poisson and another an MBM, is conjugate with respect to the standard point target measurement model. Before reviewing the mathematical form of the conjugate prior, we give an overview of its key components and the underlying structure.

1) *Interpretation:* The Poisson part of the conjugate prior models the undetected targets, which represent targets that exist at the current time but have never been detected. Each measurement at each time step gives rise to a new potentially detected target. That is, there is the possibility that a new measurement is the first detection of a target, but it can also correspond to another previously detected target or clutter, in which case there is no new target. As this target may exist or not, its resulting distribution is Bernoulli and we refer to it as "potentially detected target."

In addition, for each potentially detected target, there are single target association history hypotheses (single target hypotheses), which represent possible histories of target-to-measurement (or misdetections) associations. A single target hypothesis along with the existence probability of the corresponding Bernoulli RFS incorporates information about the events: the target never existed, the target exists at the current time, the target did exist but death occurred at some point since the last detection. Finally, a global association history hypothesis (global hypothesis) contains one single target hypotheses for each potential target with the constraints that each of the measurements has to be contained in only one of the single target hypotheses.

2) *Mathematical Representation:* Due to the independence property, the considered density is [9]

$$f(X) = \sum_{Y \uplus W = X} f^p(Y) f^{\text{mbm}}(W) \tag{7}$$

where $f^p(\cdot)$ is a Poisson density and $f^{mbm}(\cdot)$ is a MBM [18]. The Poisson density is

$$f^p(X) = e^{-\int \mu(x) dx} [\mu(\cdot)]^X \tag{8}$$

where $\mu(\cdot)$ represents its intensity. The MBM has multiplicative weights such that

$$f^{\text{mbm}}(X) \propto \sum_j \sum_{X_1 \uplus \ldots \uplus X_n = X} \prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i) \tag{9}$$

where $\propto$ stands for proportionality, j is an index over all global hypotheses (components of the mixtures) [18], n is the number of potentially detected targets, and $w_{j,i}$ and $f_{j,i}(\cdot)$ are the weight and the Bernoulli density of potentially detected target i under the jth global hypothesis. The

GARCÍA-FERNÁNDEZ ET AL.: POISSON MULTI-BERNOULLI MIXTURE FILTER: DIRECT DERIVATION AND IMPLEMENTATION 1885

Bernoulli densities have the expression

$$f_{j,i}(X) = \begin{cases} 1 - r_{j,i} & X = \emptyset \\ r_{j,i}\, p_{j,i}(x) & X = \{x\} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

where $r_{j,i}$ is the probability of existence and $p_{j,i}(\cdot)$ is the state density given that it exists. Note that if there is only one mixture component in the MBM in (9), i.e., j can only take value 1, we obtain a multi-Bernoulli density

$$f^{\text{mb}}(X) = \sum_{X_1 \uplus \dots \uplus X_n = X} \prod_{i=1}^{n} f_{1,i}(X_i). \quad (11)$$

The derivation demonstrates that a new Bernoulli component should be created for each new measurement, where its existence corresponds to the event that the measurement is the first detection of a new target (which, prior to detection, was modeled by the Poisson component), and nonexistence corresponds to the event that the measurement is a false alarm, or it corresponded to a different, previously detected target. In addition, as each target can create at maximum one measurement, the number of potentially detected targets corresponds to the number of measurements up to the current time. The weight of global hypothesis j is proportional to the product of the hypothesis weights $\prod_{i=1}^{n} w_{j,i}$ for the n potentially detected targets. If potentially detected target i is not considered in global hypothesis j, which implies that its originating measurement was assigned to another target, $w_{j,i} = 1$ and the probability of existence of $f_{j,i}(\cdot)$ is zero. We do not make global hypotheses explicit in the notation as it is not necessary to prove conjugacy. A notation that explicitly states both these hypotheses and the data association history is provided in [18].

Plugging (9) into (7), we can also write (7) as

$$f(X) \propto \sum_{Y \uplus X_1 \uplus \dots \uplus X_n = X} f^p(Y) \sum_{j} \prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i). \quad (12)$$

Note that, given X, X_i can be either empty or a single element set (otherwise the density $f_{j,i}(\cdot)$ is zero) and Y can have any cardinality that meets the constraint $Y \uplus X_1 \uplus \dots \uplus X_n = X$.

B. Update of a Poisson Prior

In this section, we prove the update for a Poisson prior using the likelihood (5). This result will be used in Section III-C to update the Poisson component of the conjugate prior (12).

1) *Likelihood Representation:* For $Z = \{z_1, \dots, z_m\}$, we prove in Appendix A that we can write the likelihood (5) as

$$l(\{z_1, \dots, z_m\} | X) = e^{-\lambda_c} \sum_{U \uplus Y_1 \uplus \dots \uplus Y_m = X} [1 - p_d(\cdot)]^U$$
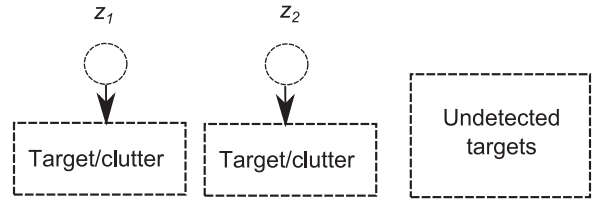$$\times \prod_{i=1}^{m} \tilde{l}(z_i | Y_i) \quad (13)$$

Fig. 1. Example of the likelihood decomposition for $\{z_1, z_2\}$. Each measurement may have been produced by a target or clutter. The likelihood also accounts for the set of undetected targets.

where

$$\tilde{l}(z|Y) = \begin{cases} p_d(y)\, p(z|y) & Y = \{y\} \\ c(z) & Y = \emptyset \\ 0 & |Y| > 1. \end{cases} \quad (14)$$

The interpretation of (13) is as follows. We decompose the set X of targets into all possible sets $U, Y_1, \dots, Y_m$ such that $X = U \uplus Y_1 \dots \uplus Y_m$. Set U represents the undetected targets and set Y_i represents the origin of the ith measurement, which can be a single-element set containing the state of the target that gave rise to the measurement, or an empty set if the measurement is clutter. This is a different but equivalent way of expressing the data association hypotheses considered in (5). An example is illustrated in Fig. 1.

2) *Update:* Given a Poisson prior $f^p(\cdot)$ and $Z = \{z_1, \dots, z_m\}$, we use Bayes' rule to compute the posterior $q^p(\cdot|Z)$ given the measurement set Z

$$q^p(X|Z) \propto l(Z|X)\, f^p(X). \quad (15)$$

Note that $q^p(X|Z)$ denotes the updated Poisson process with set Z but this density is not Poisson unless Z is empty. We show in Appendix B that substituting (8) and (13) into (15), we find that the updated posterior is a union of a Poisson process and a multi-Bernoulli RFS such that

$$q^p(X|Z)$$
$$\propto \sum_{U \uplus Y_1 \uplus \dots \uplus Y_m = X} q^p(U) \prod_{i=1}^{m} \rho^p(z_i)\, q^p(Y_i | z_i) \quad (16)$$
$$\propto \sum_{U \uplus Y_1 \uplus \dots \uplus Y_m = X} q^p(U) \prod_{i=1}^{m} q^p(Y_i | z_i) \quad (17)$$

where the Poisson component has the intensity of the prior multiplied by $(1 - p_d(\cdot))$

$$q^p(U) \propto [(1 - p_d(\cdot))\, \mu(\cdot)]^U \quad (18)$$

and the Bernoulli components are given by

$$q^p(Y_i | z_i) = \tilde{l}(z_i | Y_i)\, f^p(Y_i) \Big/ \left(e^{-\int \mu(x)dx} \rho^p(z_i) \right) \quad (19)$$
$$= \begin{cases} 1 - r^p(z_i) & Y_i = \emptyset \\ r^p(z_i)\, p^p(y | z_i) & Y_i = \{y\} \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

where

$$\rho^p (z_i) = \int \tilde{l}(z_i|Y_i) \, f^p (Y_i) \, \delta Y_i / e^{-\int \mu(x)dx}$$

$$= c(z_i) + e(z_i) \tag{21}$$

$$e(z_i) = \int p(z_i|y) \, p_d(y) \, \mu(y) \, dy \tag{22}$$

$$r^p(z_i) = e(z_i) / \rho^p(z_i) \tag{23}$$

$$p^p(y|z_i) = p_d(y) \, p(z_i|y) \, \mu(y) / e(z_i). \tag{24}$$

Note that, we define $\rho^p(z_i)$ by normalizing it by $e^{-\int \mu(x)dx}$ as (21) will be used later on and there is no need to compute this exponential in the resulting filter.

The explanation of the resulting updated density (17) is as follows. Given $Z = \{z_1, \dots, z_m\}$ and a Poisson process with intensity $\mu(\cdot)$, the updated density is the union of $m+1$ independent RFSs, represented by $U, Y_1, \dots, Y_m$. The RFS U is Poisson with intensity $(1 - p_d(\cdot)) \, \mu(\cdot)$ and represents the undetected part of the prior. The RFS Y_j is the Bernoulli RFS coming from the jth measurement. Its density is given by (19), which has a probability of existence given by (23).

C. Update of Conjugate Prior

In order to show the update of the conjugate prior, we first propose another likelihood representation in Section III-C1. Then, we show the update of one Bernoulli component in Section III-C2 and utilize this result to obtain the whole update in Section III-C3.

1) Likelihood Representation: Here we represent the likelihood in a way that is suitable to update the PMBM. For any sets $Y, X_1, \dots, X_n$, such that $|X_i| \le 1$ for $i = 1, \dots, n$, we define the function

$$l_o(Z|Y, X_1, \dots, X_n) = \sum_{Z_1 \uplus \dots \uplus Z_n \uplus Z^y = Z} l\left(Z^y|Y\right)$$

$$\times \prod_{i=1}^{n} t(Z_i|X_i). \tag{25}$$

where Z^y represents both measurements from targets in Y and clutter, and $t(Z_i|X_i)$ is the likelihood for a set with zero or one measurement elements without clutter

$$t(Z_i|X_i) = \begin{cases} p_d(x) \, l(z|x) & Z_i = \{z\}, X_i = \{x\} \\ 1 - p_d(x) & Z_i = \emptyset, X_i = \{x\} \\ 1 & Z_i = \emptyset, X_i = \emptyset \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

We show in Appendix C that for any sets $Y, X_1, \dots, X_n$, such that $|X_i| \le 1$ for $i = 1, \dots, n$, we have

$$l_o(Z|Y, X_1, \dots, X_n) = l(Z|X) \tag{27}$$

where $X = Y \uplus X_1 \uplus \dots \uplus X_n$. That is, the evaluation of function $l_o(Z|\cdot, \cdot, \dots, \cdot)$ at any sets $Y, X_1, \dots, X_n$, such that $|X_i| \le 1$ for $i = 1, \dots, n$, is equivalent to the evaluation of the likelihood $l(Z|\cdot)$ at set $X = Y \uplus X_1 \uplus \dots \uplus X_n$.

2) Update of One Bernoulli Component: As will be seen in the next section, one part of the update of the conjugate prior requires the update of the Bernoulli components.

Therefore, we proceed to derive this update in this section so that we have the result available for the next section. In the update of the conjugate prior, we will need to compute the update of Bernoulli component $f_{j,i}(\cdot)$, which is given by (10), by measurement Z_i considering the likelihood $t(Z_i|\cdot)$. We denote the corresponding updated density as

$$q_{j,i}(X_i|Z_i) = t(Z_i|X_i) \, f_{j,i}(X_i) / \rho_{j,i}(Z_i) \tag{28}$$

where the numerator is the joint density of Z_i and X_i and

$$\rho_{j,i}(Z_i) = \int t(Z_i|X) \, f_{j,i}(X) \, \delta X. \tag{29}$$

According to $t(Z_i|X)$ in (26), Z_i can only take values $Z_i = \{z\}$ or $Z_i = \emptyset$ so that the likelihood is different from zero, so we proceed to compute (28) in these two cases. For $Z_i = \{z\}$, $t(Z_i|X)$ is only different from zero if $X = \{x\}$ so, using (10), (26), and (29), we obtain

$$\rho_{j,i}(\{z\}) = r_{j,i} \int p_d(x) \, l(z|x) \, p_{j,i}(x) \, dx. \tag{30}$$

Substituting the previous equations into (28), we find that $q_{j,i}(\cdot|\{z\})$ is Bernoulli with probability of existence 1 and target state density proportional to $p_d(x) \, l(z|x) \, p_{j,i}(x)$. For $Z_i = \emptyset$, $t(Z_i|X)$ can be different from zero if $X = \{x\}$ or $X = \emptyset$. Now, using (10), (26), and (29), we have

$$\rho_{j,i}(\emptyset) = 1 - r_{j,i} + r_{j,i} \int (1 - p_d(x)) \, p_{j,i}(x) \, dx. \tag{31}$$

Then, substituting the previous equations into (28), we find that $q_{j,i}(\cdot|\emptyset)$ is Bernoulli with probability of existence

$$r_{j,i} \left[\int (1 - p_d(x)) \, p_{j,i}(x) \, dx \right] / \rho_{j,i}(\emptyset)$$

and target state density proportional to $(1 - p_d(x)) \, p_{j,i}(x)$.

3) Update of the Conjugate Prior: Substituting the prior (12) into Bayes' rule (1), we have that

$$q(X|Z)$$

$$\propto \sum_{Y \uplus X_1 \uplus \dots \uplus X_n = X} l(Z|X) \, f^p(Y) \sum_j \prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i)$$

$$= \sum_{Y \uplus X_1 \uplus \dots \uplus X_n = X} l(Z|Y \uplus X_1 \uplus \dots \uplus X_n) \, f^p(Y)$$

$$\times \sum_j \prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i).$$

As $f_{j,i}(\cdot)$ is Bernoulli, the corresponding term in the previous sum is different from zero if and only if $|X_i| \le 1$. Therefore, we can add this constraint to the sum

$$q(X|Z)$$

$$\propto \sum_{Y \uplus X_1 \uplus \dots \uplus X_n = X: |X_i| \le 1 \forall i} l(Z|Y \uplus X_1 \uplus \dots \uplus X_n) \, f^p(Y)$$

$$\times \sum_j \prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i). \tag{32}$$

Now, substitute (27) in (32), so that

$$q(X|Z)$$

$$\propto \sum_{Y \uplus X_1 \uplus \ldots \uplus X_n = X : |X_i| \leq 1 \quad \forall i} l_o(Z|Y, X_1, \ldots, X_n) f^p(Y)$$

$$\times \sum_j \prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i).$$

$$= \sum_{Y \uplus X_1 \uplus \ldots \uplus X_n = X} \sum_{Z = Z_1 \uplus \ldots \uplus Z_n \uplus Z^y} \left[l(Z^y|Y) f^p(Y) \right]$$

$$\times \sum_j \left[\prod_{i=1}^{n} w_{j,i} t(Z_i|X_i) f_{j,i}(X_i) \right]. \tag{33}$$

Factor $l(Z^y|Y) f^p(Y)$ in (33) represents the unnormalized update of a Poisson prior. In (16), we obtained the result for such an update so we can apply it in (33). Therefore, we have that

$$q(X|Z)$$

$$\propto \sum_{Y \uplus X_1 \uplus \ldots \uplus X_n = X} \sum_{Z = Z_1 \uplus \ldots \uplus Z_n \uplus Z^y} \sum_{U \uplus Y_1 \ldots \uplus Y_m = Y} q^p(U)$$

$$\times \prod_{i=1}^{m} \left[\chi_{Z^y}(z_i) \rho^p(z_i) q^p(Y_i|z_i) + (1 - \chi_{Z^y}(z_i)) \delta_\emptyset(Y_i) \right]$$

$$\times \sum_j \left[\prod_{i=1}^{n} w_{j,i} t(Z_i|X_i) f_{j,i}(X_i) \right] \tag{34}$$

where $\chi_A(\cdot)$ denotes the indicator function on set A

$$\chi_A(z) = \begin{cases} 0 & z \notin A \\ 1 & z \in A \end{cases}$$

and $\delta_\emptyset(\cdot)$ is the multitarget Dirac delta centered at $\emptyset$ [26, eq. (11.124)]

$$\delta_\emptyset(Y) = \begin{cases} 0 & Y \neq \emptyset \\ 1 & Y = \emptyset. \end{cases}$$

We should note that for the update of the Poisson RFS Y, we only consider the measurements that are hypothesized to be coming from Y, which are represented by Z^y in (34). Therefore, in the third line of (34), we use a product over measurements $z_1, \ldots, z_m$ but setting the probability of existence of the Bernoulli RFS associated to z_i to zero, if z_i is not included in Z^y, $\chi_{Z^y}(z_i) = 0$.

Simplifying (34), we have

$$q(X|Z)$$

$$\propto \sum_{U \uplus X_1 \uplus \ldots \uplus X_n \uplus Y_1 \uplus \ldots \uplus Y_m = X} q^p(U) \sum_j \sum_{Z_1 \uplus \ldots \uplus Z_n \uplus Z^y = Z}$$

$$\times \prod_{i=1}^{m} \left[\chi_{Z^y}(z_i) \rho^p(z_i) q(Y_i|z_i) + (1 - \chi_{Z^y}(z_i)) \delta_\emptyset(Y_i) \right]$$

$$\times \left[\prod_{i=1}^{n} w_{j,i} \rho_{j,i}(Z_i) q_{j,i}(X_i|Z_i) \right]. \tag{35}$$

Merging the two inner summations into one, rearranging the indices, and comparing with the prior (12), we see that the posterior is also the union of two independent processes: one Poisson and the other an MBM. This proves that this density is conjugate with respect to the standard point target measurement model.

We would also like to comment on the weights of the new potentially detected targets, which are considered in the product over m factors in (35). If a new potentially detected target i does not exist in a new global hypothesis, which implies that $\chi_{Z^y}(z_i) = 0$, then its hypothesis weight is one and its density $\delta_\emptyset(Y_i)$ can also be represented as Bernoulli with zero probability of existence. On the contrary, if a new potentially detected target i exists in a new global hypothesis, $\chi_{Z^y}(z_i) = 1$, its hypothesis weight is $\rho^p(z_i)$ and its Bernoulli density is given by $q(Y_i|z_i)$. The weight for a previous potentially detected target corresponds to the same weight $w_{j,i}$ multiplied by $\rho_{j,i}(Z_i)$, see (29). Depending on the hypothesis Z_i can be either empty or has one element, the resulting weights and Bernoulli components in these two cases are discussed after (29).

D. Prediction of the Conjugate Prior

In this section, we prove that, if the posterior is a PMBM of the form (7)–(9), then the prior at the next time step is also PMBM with the following parameters. The Poisson part of the predicted density is obtained using the PHD filter prediction equation [24], so that its intensity is

$$\mu(x) = \lambda^b(x) + \int g(x|y) p_s(y) \lambda^u(y) dy$$

where $\lambda^u(\cdot)$ denotes the intensity of the Poisson part of the posterior. In addition, if the parameters of the posterior MBM are $w_{j,i}^u$, $p_{j,i}^u(\cdot)$, $r_{j,i}^u$, the predicted parameters are given by the multitarget multi-Bernoulli filter prediction equation [23]

$$w_{j,i} = w_{j,i}^u$$

$$r_{j,i} = r_{j,i}^u \int p_{j,i}^u(y) p_s(y) dy$$

$$p_{j,i}(x) \propto \int g(x|y) p_s(y) p_{j,i}^u(y) dy.$$

In order to prove this result, we first note the equivalences between the dynamic/measurement processes [26, Ch. 13]. In the standard models, each target is detected/survives with probability $p_d(\cdot)/p_s(\cdot)$ and generates a measurement/new target state according to $l(\cdot|\cdot)/g(\cdot|\cdot)$ and there are additional independent clutter measurements/new born targets distributed according to a Poisson process with intensity $c(\cdot)/\lambda^b(\cdot)$. In other words, the density of the measurement, denoted as $\rho(\cdot)$ in (2), is equivalent to the predicted density, denoted as $\omega(\cdot)$ in (4), by making the previous equivalences [25]. As we have explained the notation for proving the update step, we will first compute the density of the measurements and then establish the equivalence with the prediction step. Before doing so, we establish the following corollary.

COROLLARY 2 Let us consider an RFS $X = X_1 \uplus ... \uplus X_n$ where $X_1, \ldots, X_n$ are independent, so the density $f(\cdot)$ of X can be written as

$$f(X) = \sum_{X_1 \uplus ... \uplus X_n = X} \prod_{i=1}^{n} f_i(X_i)$$

where $f_i(\cdot)$ is the density of X_i. For an arbitrary set-valued function $v(\cdot)$

$$\int v(X) f(X) \delta X$$

$$= \int ... \int v(X_1 \cup ... \cup X_n) \prod_{i=1}^{n} f_i(X_i) \delta X_1 ... \delta X_n.$$

The proof of the corollary is straightforward using [27, eq. (63)] $n - 1$ times. Substituting (12) into (2), we obtain

$$\rho(Z) \propto \sum_j \left[\prod_{i=1}^{n} w_{j,i} \right] \int l(Z|X)$$

$$\times \sum_{Y \uplus X_1 \uplus ... \uplus X_n = X} f^p(Y) \prod_{i=1}^{n} f_{j,i}(X_i) \delta X.$$

where $l(\cdot|X)$ is the density of the measurements (including clutter) given X. Using Corollary 2, we find

$$\rho(Z) \propto \sum_j \left[\prod_{i=1}^{n} w_{j,i} \right] \int \int ... \int l(Z|Y \cup X_1 \cup \cdots \cup X_n)$$

$$\times f^p(Y) \prod_{i=1}^{n} f_{j,i}(X_i) \delta Y \delta X_1 \cdots \delta X_n.$$

As $f_{j,i}(\cdot)$ are Bernoulli, we can apply (27) and then (25), so that

$$\rho(Z) \propto \sum_j \left[\prod_{i=1}^{n} w_{j,i} \right] \int \int ... \int l_o(Z|Y, X_1, \ldots, X_n)$$

$$\times f^p(Y) \prod_{i=1}^{n} f_{j,i}(X_i) \delta Y \delta X_1 ... \delta X_n$$

$$= \sum_j \sum_{Z_1 \uplus ... \uplus Z_n \uplus Z^y = Z} \int l\left(Z^y|Y\right) f^p(Y) \delta Y$$

$$\times \left[\prod_{i=1}^{n} w_{j,i} \int t(Z_i|X_i) f_{j,i}(X_i) \delta X_i \right]$$

$$= \sum_j \sum_{Z_1 \uplus ... \uplus Z_n \uplus Z^y = Z} \int l\left(Z^y|Y\right) f^p(Y) \delta Y$$

$$\times \left[\prod_{i=1}^{n} w_{j,i} \rho_{j,i}(Z_i) \right]$$

where we recall that $\rho_{j,i}(\cdot)$ is a Bernoulli density previously specified in (30) and (31), and $t(\cdot|X)$ is the density of the measurement generated by a set X, which can have cardinality zero or one, without clutter. From the PHD filter recursion [24], [25], we know that $\int l(Z^y|Y) f^p(Y) \delta Y$ is a Poisson density on Z^y with intensity $c(x) + \int p(x|y) p_d(y) \mu(y) dy$.

In summary, the density of the measurement is the union of a Poisson process and an MBM with the same weights as the prior and the parameters specified above. Due to the equivalence of parameters in the prediction/update steps mentioned at the beginning of this section, the proof of the conjugacy of the PMBM is finished.

E. Conjugacy for MBMs

In this section, we establish the conjugacy property of MBMs, which results in the MBM filter. This result will help us establish relations between PMBM and labeled conjugate priors, see Section IV.

COROLLARY 3 If the birth process is multi-Bernoulli or MBM, the family of MBMs is a conjugate prior for the standard point target measurement and dynamic models.

The update step can be performed as above by setting the intensity of the Poisson density to zero and the prediction step is proved in Appendix D. In the prediction step, for multi-Bernoulli birth, we incorporate additional multi-Bernoulli components to each term in the mixture. For MBM birth, a new term is created for each combination of a term in the old mixture and a term in the birth mixture, where the new term combines the Bernoulli components from each.

IV. CONNECTION BETWEEN THE PMBM FILTER AND THE δ-GLMB FILTER

In this section, we establish the connection between the PMBM filter and the δ-GLMB filter. In order to do so, we first discuss an alternative parameterization of MBMs in Section IV-A. Then, we introduce the conjugacy properties of labeled MBMs in Section IV-B. Section IV-C proves that the δ-GLMB density is in fact a labeled MBM, but with a less efficient parameterization from a storage and computational point of view. A discussion on both parameterizations and the advantages of the PMBM form is given in Section IV-D.

A. MBM 01 Parameterization

In this section, we explain the MBM_{01} parameterization, which is an alternative parameterization of an MBM in which the Bernoulli densities have existence probabilities that are either zero or one. The MBM_{01} parameterization is relevant to the connection between the PMBM filter and the δ-GLMB filter, as will be explained in the following sections. The MBM parameterization in (9) is simply referred to as the MBM parameterization.

We first explain the MBM_{01} parameterization of a single Bernoulli density. A Bernoulli density $f_{j,i}(\cdot)$, see (10), can be written as a mixture of Bernoulli densities with existence probabilities that are either zero or one as

$$f_{j,i}(X_i) = (1 - r_{j,i}) f^0_{j,i}(X_i) + r_{j,i} f^1_{j,i}(X_i) \qquad (36)$$

where

$$f_{j,i}^{\theta_i}(X_i) = \begin{cases} 1 - \theta_i & X_i = \emptyset \\ \theta_i\, p_{j,i}(x) & X_i = \{x\} \\ 0 & \text{otherwise} \end{cases} \qquad (37)$$

for $\theta_i \in \{0, 1\}$. It should be noted that if $r_{j,i} \in (0, 1)$, the mixture in (36) has two components, otherwise, it has one component. We say that $f_{j,i}^0(\cdot)$ and $f_{j,i}^1(\cdot)$ have deterministic existence, since $X_i = \emptyset$ and $|X_i| = 1$ have probability one for $f_{j,i}^0(\cdot)$ and $f_{j,i}^1(\cdot)$, respectively.

In an MBM, we can expand all Bernoulli densities in a similar way, such that existence probabilities of all Bernoulli densities are either 0 or 1. For instance, the MBM in (9) can be written in MBM$_{01}$ parameterization as

$$f^{\mathrm{mbm}}(X) \propto \sum_j \sum_{\theta \in \{0,1\}^n} \sum_{X_1 \uplus \ldots \uplus X_n = X} \prod_{i=1}^n w_{j,i} v_{j,i,\theta_i} f_{j,i}^{\theta_i}(X_i) \qquad (38)$$

where $\theta = (\theta_1, \ldots, \theta_n)$, $v_{j,i,\theta_i} = (1 - r_{j,i})^{1-\theta_i} r_{j,i}^{\theta_i}$, and $\{0, 1\}^n$ represents n Cartesian products of $\{0, 1\}$. From (38), we can directly establish the following proposition.

PROPOSITION 4 Consider an MBM with m mixture components. Let n_j denote the number of Bernoulli densities, in component j of the MBM, with existence probability in the interval $(0, 1)$. Then, the MBM$_{01}$ parameterization of the MBM requires $\sum_{j=1}^m 2^{n_j}$ mixture components.

Let us illustrate the increase in the number of mixture components (global hypotheses) with the following example.

EXAMPLE 5 Consider an MB density (MBM with one mixture component) with three targets and existence probabilities $r_{1,1} = 0.8$, $r_{1,2} = 0.2$ and $r_{1,3} = 1$. The corresponding MBM$_{01}$ parameterization contains four mixture components (global hypotheses) with weights $r_{1,1} r_{1,2}$, $(1 - r_{1,1}) r_{1,2}$, $r_{1,1}(1 - r_{1,2})$, and $(1 - r_{1,1})(1 - r_{1,2})$.

It should be noted that, according to Proposition 4, the MBM$_{01}$ parameterization can give rise to a tremendous increase in the number of components in the mixture (global hypotheses), which is an inefficient way to represent an MBM distribution. In fact, we can use the PMBM filter with an MBM$_{01}$ parameterization, but a standard brute-force implementation would yield much higher computational complexity due to the increase in the number of global hypotheses. For instance, as will be clarified in Section V, we need to solve a data-association problem for each global hypothesis so it is desirable to have as few global hypotheses as possible.

B. Conjugacy of Labeled MBMs

In this section, we prove the conjugacy for labeled MBMs. In the labeled approach, we augment the single target state space with a label, which is a variable that is unique for each new born target and fixed with time [13], [17]. A labeled MBM is therefore obtained by adding (unique) labels to an MBM, see (9), which results in a density of the form

$$f(X) \propto \sum_j \sum_{X_1 \uplus \ldots \uplus X_n = X} \prod_{i=1}^n w_{j,i} f_{j,i}^{lb}(X_i) \qquad (39)$$

where $f_{j,i}^{lb}(\cdot)$ is the labeled Bernoulli density for target i for mixture component j given by

$$f_{j,i}^{lb}(X) = \begin{cases} 1 - r_{j,i} & X = \emptyset \\ r_{j,i}\, p_{j,i}(x)\, \delta[\ell - \ell_i] & X = \{(x, \ell)\} \\ 0 & \text{otherwise.} \end{cases} \qquad (40)$$

Here, $\delta[\cdot]$ represents a Kronecker delta, ℓ_i is the deterministic label of target i, and $r_{j,i}$ and $p_{j,i}(\cdot)$ are its existence probability and state density for global hypothesis j. In addition, in (39), we have $\ell_i \neq \ell_{i'}$ for $i \neq i'$ to ensure unique labels. The main difference between (40) and its unlabeled counterpart (10) is that the state space has been expanded to incorporate a unique label that is known for each i. Note that the labeled MBM in (39) can also be written in (labeled) MBM$_{01}$ parameterization analogously to how (9) was expressed in (38).

We establish the following corollary.

COROLLARY 6 If the birth process is labeled multi-Bernoulli or labeled MBM, whose targets have unique labels, and labels are fixed with time, the family of labeled MBM is a conjugate prior for the standard point target measurement and dynamic models.

As we explain in this paragraph, Corollary 6 is a particular case of Corollary 3 by considering the specific properties of the labels: they are unique and fixed with time. Note that, in this paper, we have denoted the single target state as x, without any assumptions on it so it is flexible enough to include a label, without specifying it explicitly. In order to prove conjugacy for labeled MBM, we just need to model that one component of the target state (the label) is unique and fixed using the general birth/dynamic models. This is done by considering labeled MB or MBM birth process and a single target transition density $g(\cdot|\cdot)$ that has the constraint that the label does not change with time. Therefore, the conjugacy for labeled MBM is just a particular case of MBM conjugacy, with the previous constraints in the birth model and single transition density. As a result, the prediction and update equations for the general MBM filter are also valid for the labeled MBM filter.

C. Relation Between δ-GLMB Densities and Labeled MBMs

The most common conjugate prior for labeled RFSs is the δ-GLMB density [17], and in the following proposition, which is proved in Appendix E, we relate a δ-GLMB density to a labeled MBM density.

PROPOSITION 7 δ-GLMB and labeled MBM with MBM$_{01}$ parameterization can represent the same labeled multitarget

densities with the same number of global hypotheses, in which target existence is deterministic.

As indicated in the previous proposition, δ-GLMB and labeled MBM with MBM_{01} parameterizations have the same type of global hypotheses, in the sense that both consider global hypotheses with deterministic target existence and labeled targets. One difference, however, is that the δ-GLMB notation [17], [21] can only consider labeled targets, while the MBM_{01} notation can handle labeled and unlabeled targets. According to Proposition 7, the number of global hypotheses (mixture components) in the δ-GLMB density in relation to a (labeled) MBM parameterization is the same as in the (labeled) MBM_{01} parameterization, which is given by Proposition 4. This is illustrated in the next example.

EXAMPLE 8 Suppose distinct labels ℓ_1, ℓ_2, ℓ_3 are added to the three Bernoulli components in Example 5, such that we have a labeled MB density (labeled MBM with one mixture component). As in Example 5, its MBM_{01}/ δ-GLMB parameterizations have four mixture components (global hypotheses), with the same weights as in Example 5.

D. Discussion

We proceed to discuss some computational and implementational advantages of the MBM parameterization (either labeled or not) compared to the MBM_{01} and δ-GLMB parameterizations with multi-Bernoulli births. In the MBM filter (either labeled or not), the prediction step is straightforward, see Section III-D. This is in stark contrast with the δ-GLMB filter prediction implementation in [21], which truncates the predicted density by a K-shortest path algorithm. This approximation is introduced due to an inefficient representation of the MBM. For instance, for probability of survival lower than one, Bernoulli components that have existence probability 1 have a smaller existence probability after the prediction step, see Section III-D. Because of this, a multi-Bernoulli density that contains n Bernoulli components, all with existence probability 1, is represented after the prediction step by an MBM_{01}/δ-GLMB with 2^n global hypotheses, see Proposition 4. These MBM_{01}/δ-GLMB representations are highly inefficient as the predicted density is simply one multi-Bernoulli process with existence probabilities in (0,1).

In the update step, as can be seen in (35), we need to solve a data-association problem for each mixture component, that is, for every global hypothesis in the prior. In this case, the MBM parameterization is also advantageous due to the lower number of mixture components, compared to the MBM_{01}/δ-GLMB parameterizations. The reason for these advantages in the prediction and update steps in the MBM filter is mainly due to the inefficient MBM_{01}/δ-GLMB parameterizations. One MBM global hypothesis can efficiently represent many δ-GLMB global hypotheses and this extra degree of flexibility in the MBM filter simplifies the prediction and update steps and it is independent of whether or not we use labels.
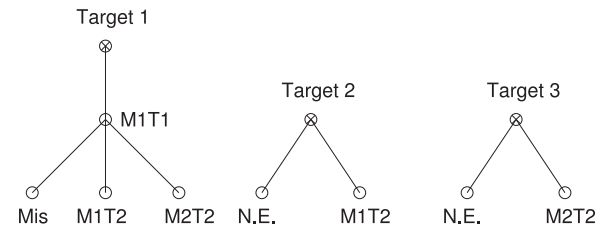

Fig. 2. Illustration of the single-target hypothesis tree. We consider there is one measurement at time 1 (M1T1) and two measurements at time 2 (M1T2 and M2T2). The hypothesis tree at time 2 considers that potentially detected target 1 is associated to M1T1 at time 1. At time 2, it can be associated with a misdetection (Mis) or with M1T2 or M2T2. Potentially detected target 2 might not exist (N.E.) or be associated to M1T2. Potentially detected target 3 might not exist or be associated to M2T2. There are three global hypotheses at time 2. All the global hypotheses associate M1T1 to potentially detected target 1. At time 2, the measurement associations to potentially detected targets 1, 2, and 3 in the global hypotheses are: (Mis, M1T2, M2T2), (M1T2, N.E, M2T2), and (M2T2, M1T2, N.E.).

In addition, if there are Poisson births, the PMBM characterizes the Poisson part by its intensity, which is an efficient way of representing a Poisson distribution. In contrast, if we were to use a labeled Poisson process to model target births, the δ-GLMB parameterization would need an infinite number of global hypotheses to represent the Poisson part, since each global hypothesis in the δ-GLMB density has a deterministic cardinality.

V. IMPLEMENTATION FOR LINEAR/GAUSSIAN DYNAMIC AND MEASUREMENT MODELS

In this section, we propose an implementation of the PMBM filter for linear Gaussian dynamic and measurement models with Poisson births. We first provide an overview of the structure of the hypotheses in Section V-A. Then, we explain the prediction and update in Sections V-B and V-C, respectively.

A. Structure of the Hypotheses

In the conjugate prior, see (12), there is an index j for the MBM. Each j corresponds to a global hypothesis, which represents possible association of measurements to potentially detected targets. As explained in [18], global hypotheses can be expressed in terms of single-target hypothesis. A single-target hypothesis corresponds to a sequence of measurements associated to a potentially detected target. Given a single-target hypothesis, this potentially detected target follows a Bernoulli distribution, as explained in Section III. Therefore, each measurement starts a new single-target hypothesis. At following time steps, new single-target hypotheses are created by associating previous single-target hypotheses with current measurements or with a misdetection. By doing this, global hypotheses are a collection of these single-target hypotheses, with the conditions that no measurement is left without being associated and a measurement can only be assigned to one single target hypothesis. This hypothesis structure resembles the one in track-oriented MHT [19] and is illustrated in Fig. 2. We proceed

to explain the prediction and update steps.

B. Prediction

We assume that, in the posterior at the previous time step, intensity of the Poisson component is a Gaussian mixture

$$\lambda^u(x) = \sum_{i=1}^{N_u} w_{u,i} \mathcal{N}\left(x; \overline{x}_{u,i}^p, P_{u,i}^p\right)$$

and the MBM parameters are $w_{j,i}^u$, $p_{j,i}^u(x) = \mathcal{N}(x; \overline{x}_{j,i}^u, P_{j,i}^u)$, $r_{j,i}^u$.

We also assume constant probability of survival p_s, linear/Gaussian dynamics $g(x|y) = \mathcal{N}(x; Fy, Q)$ and new born target intensity

$$\lambda^b(x) = \sum_{i=1}^{N_b} w_{b,i}^p \mathcal{N}\left(x; \overline{x}_{b,i}^p, P_{b,i}^p\right).$$

Then, from Section III-D and using known results from the Kalman filter prediction step [28], we find that the predicted intensity is a Gaussian mixture

$$\mu(x) = \lambda^b(x) + p_s \sum_{i=1}^{N_u} w_{u,i} \mathcal{N}\left(x; F\overline{x}_{u,i}^p, FP_{u,i}^p F^T + Q\right). \tag{41}$$

The predicted Bernoulli components have the same weights as in the previous time step with existence $r_{j,i} = r_{j,i}^u p_s$ and

$$p_{j,i}(x) = \mathcal{N}\left(x; F\overline{x}_{j,i}^u, FP_{j,i}^u F^T + Q\right).$$

Clearly, the implementation of the prediction step is straightforward, contrary to the prediction step of the δ-GLMB filter in [21], as discussed in Section IV-D.

C. Update

We assume that p_d is constant and $p(z|x) = \mathcal{N}(z; Hx, R)$. We rewrite the predicted intensity of the Poisson part (41) as

$$\mu(x) = \sum_{i=1}^{N_\mu} w_{\mu,i} \mathcal{N}\left(x; \overline{x}_{\mu,i}, P_{\mu,i}\right) \tag{42}$$

and the MBM parameters as $w_{j,i}$, $p_{j,i}(x) = \mathcal{N}(x; \overline{x}_{j,i}, P_{j,i})$, $r_{j,i}$.

From the conjugate prior update, see Section III-C3, we have that three different types of updates: update for undetected targets (Poisson component), update for potential targets detected for the first time, and update for previously potentially detected targets. The update of the Poisson component is straightforward. Using (18), the updated intensity for undetected targets is (42) multiplied by $1 - p_d$. We proceed to explain the other two updates.

1) Potential Targets Detected for the First Time: We first go through all components of the Poisson prior and perform ellipsoidal gating [19] on the measurements to lower the computational complexity. For those measurements that can create a new track according to the gating output, we perform the Bayesian update (19). For measurement z, this

gives a Bernoulli component with existence $r^p(z)$ and target state density $p^p(x|z)$ such that

$$r^p(z) = e(z)/\rho^p(z) \tag{43}$$

$$p^p(y|z) = p_d p(z|y)\mu(y)/e(z)$$

$$= \sum_{i=1}^{N_\mu} w_i(z)\mathcal{N}\left(x; \overline{x}_{\mu,i}^u(z), P_{\mu,i}^u\right) \tag{44}$$

where

$$e(z) = p_d \int p(z|y)\mu(y)\,dy$$

$$= p_d \sum_{i=1}^{N_\mu} w_{\mu,i}\mathcal{N}\left(z; H\overline{x}_{\mu,i}, S_{\mu,i}\right)$$

$$\rho^p(z) = e(z) + c(z) \tag{45}$$

$$w_i(z) \propto w_{\mu,i}\mathcal{N}\left(z; H\overline{x}_{\mu,i}, S_{\mu,i}\right)$$

$$\overline{x}_{\mu,i}^u(z) = \overline{x}_{\mu,i} + \Psi_{\mu,i} S_{\mu,i}^{-1}\left(z - H\overline{x}_{\mu,i}\right)$$

$$P_{\mu,i}^u = P_{\mu,i} - \Psi_{\mu,i} S_{\mu,i}^{-1}\Psi_{\mu,i}^T$$

$$\Psi_{\mu,i} = P_{\mu,i} H^T$$

$$S_{\mu,i} = HP_{\mu,i} H^T + R$$

and we recall that $c(\cdot)$ is the clutter intensity. Note that $\overline{x}_{\mu,i}^u(z)$, $P_{\mu,i}^u$ are the updated mean and covariance matrix of a Kalman filter with prior $\overline{x}_{\mu,i}$ and $P_{\mu,i}$ [28]. To lower the computational complexity, we approximate the Gaussian mixture in (44) as a Gaussian by performing moment matching.

We still have to determine the hypothesis weight of the newly created components of the MBM. According to (35), the hypothesis weight $w_{j,i}$ of a potential target detected for the first time with measurement z in a global hypothesis j that considers it is $\rho^p(z)$, which is given by (45). If the global hypothesis j does not consider this potentially detected target $w_{j,i} = 1$ and its existence probability is set to zero.

2) Previous Potentially Detected Targets: According to Section III-C2, we go through all potentially detected targets and their single target hypotheses in (9) and create the new single target hypotheses. In order to explain this procedure, let us consider that a single target hypothesis with indices j, i which has weight $w_{j,i}$, existence probability $r_{j,i}$, and Gaussian density for the target

$$p_{j,i}(x) = \mathcal{N}\left(x; \overline{x}_{j,i}, P_{j,i}\right). \tag{46}$$

For this single target hypothesis, we first create a new misdetection hypothesis, which has a weight $w_{j,i}$ $(1 - r_{j,i} + r_{j,i}(1 - p_d))$. The associated Bernoulli component has an existence probability $r_{j,i}(1 - p_d)/(1 - r_{j,i} + r_{j,i}(1 - p_d))$ and the density given that the target exists remains the same, $p_{j,i}(\cdot)$. We then perform ellipsoidal gating [19] using (46) to consider only the relevant measurements. For each of the chosen measurements and this Bernoulli component, we perform the update (28), which has a closed-form expression given by the update step of the Kalman filter [28]. For measurement z, we have that the

corresponding hypothesis weight is

$$w_{j,i} r_{j,i} p_d \mathcal{N}\left(z; H\overline{x}_{j,i}, S_{j,i}\right)$$

and the Bernoulli component has existence probability one and density

$$\mathcal{N}\left(x; \overline{x}_{j,i}^u\left(z\right), P_{j,i}^u\right)$$

where

$$\overline{x}_{j,i}^u\left(z\right) = \overline{x}_{j,i} + \Psi_{j,i} S_{j,i}^{-1}\left(z - H\overline{x}_{j,i}\right)$$
$$P_{j,i}^u = P_{j,i} - \Psi_{j,i} S_{j,i}^{-1} \Psi_{j,i}^T$$
$$\Psi_{j,i} = P_{j,i} H^T$$
$$S_{j,i} = H P_{j,i} H^T + R.$$

3) Selection of k-Best Global Hypotheses: At this point, we have calculated all possible new single-target hypotheses but we still have to form the global hypotheses. We can see in (35) that, for each global hypothesis j at the previous time step, we must go through all possible data association hypotheses that give rise to the updated global hypotheses. This high increase in the number the global hypotheses is the bottleneck of the computation of the conjugate prior. However, based on the literature on labeled RFSs and MHT, we approximate this update by pruning the number of hypotheses using Murty's algorithm [29]. With this algorithm, we can select the k new global hypotheses with highest weight for a given global hypothesis j without evaluating all the newly generated global hypotheses [17], [21], [30], [31]. An interesting alternative would be to use the generalized Murty's algorithm for multiple frames [32].

For global hypothesis j, all measurements (excluding those removed by gating) must be associated either to an existing track in hypothesis j or to a new track, i.e., no measurement is left unassigned. We can then construct the corresponding cost matrix using the updated weights of the conjugate prior. Let us assume there are n_o old tracks in global hypothesis j and m measurements $z_1, \ldots, z_m$ after gating. The cost matrix is

$$C = -\left[\ln\left(W_{ot}\right), \ln\left(W_{nt}\right)\right] \tag{47}$$

where

$$W_{nt} = \text{diag}\left(\rho^p\left(z_1\right), \ldots, \rho^p\left(z_m\right)\right)$$

with $\rho^p\left(z_i\right)$ given by (45). Matrix W_{nt} represents the weight matrix for new potentially detected targets and $W_{ot} \in \mathbb{R}^{m \times n_j}$ represents the weight matrix for old targets, where n_j are the number of potentially detected targets at the previous time steps in global hypothesis j. Component p, i of W_{ot} represents the weight of the pth measurement associated to ith target, which is

$$w_{j,i} \rho_{j,i}\left(\{z_p\}\right) / \rho_{j,i}\left(\emptyset\right)$$
$$= \frac{w_{j,i} r_{j,i} p_d \mathcal{N}\left(z_p; H\overline{x}_{j,i}, S_{j,i}\right)}{w_{j,i}\left(1 - r_{j,i} + r_{j,i}\left(1 - p_d\right)\right)}$$

according to Section V-C2. Note that, we normalize the previous weights by $\rho_{j,i}\left(\emptyset\right)$ so that the weight of a hypothesis that does not assign a measurement to a target is the same for

an old and a new target. This is just done so that we can obtain the k-best global hypotheses efficiently using Murty's algorithm but we do not alter the real weights, which are unnormalized. Each new global hypothesis that originates from hypothesis j can be written as an $m \times (m + n_0)$ assignment matrix S consisting of 0 or 1 entries such that each row sums to one and each column sums to zero or one. Then, we select the k best global hypotheses that minimize $\text{tr}\left(S^T C\right)$ using Murty's algorithm [29]. For global hypothesis j, whose weight is $w_j \propto \prod_{i=1}^n w_{j,i}$, we suggest choosing $k = \lceil N_h \cdot w_j \rceil$, where it is assumed that we want a maximum number N_h of global hypotheses as in [21]. This way, global hypotheses with higher weights will give rise to more global hypotheses. Note that this part of the algorithm is quite similar to the δ-GLMB filter update with just some modifications in the cost matrix [21, Sec. IV]. Finally, the pseudocode of a prediction and an update is given in Algorithm 1.

VI. ESTIMATION

In this section, we discuss how to perform target state estimation in the PMBM filter. In a multiple target system, an optimal estimator is given by minimizing a multitarget metric, for example, the optimal subpattern assignment (OSPA) metric [27], [33], [34]. Nevertheless, there are suboptimal estimators that are easy to compute and can work very well in many cases. In this section, we provide tractable methods for obtaining the (suboptimal) estimators used in MHT (Estimator 3) and the δ-GLMB filter (Estimator 2) using the PMBM distribution form. We also propose an additional estimator based on the PMBM (Estimator 1).

A. Estimator 1

In Estimator 1, we first select the global hypothesis of the MBM in (9) with highest weight, which corresponds to obtaining index

$$j^* = \arg\max_j \prod_{i=1}^n w_{j,i}.$$

Then, we report the mean of the Bernoulli components in hypothesis j^* whose existence probability is above a threshold Γ. Given the probabilities of detection and survival, this threshold determines the number of consecutive misdetections we can have from a target to report its estimate, see prediction and update for missed targets in Sections III-D and III-C2.

B. Estimator 2

Estimator 2 is the same kind of estimator as the one proposed in the δ-GLMB filter [21], which we proceed to describe. The δ-GLMB filter estimator first obtains the maximum a posteriori (MAP) estimate of the cardinality. Then, it finds the global hypothesis with this cardinality with highest weight and reports the mean of the targets in this hypothesis.

The same type of estimate can be constructed from the MBM in (7) by first calculating its cardinality distribution

Algorithm 1: Pseudocode for One Prediction and Update for PMBM Filter.

Input: Parameters of the PMBM posterior at the previous time step, see Section V-B, and measurement set Z at current time step.
Output: Parameters of the PMBM posterior at the current time step.
  - Perform prediction, see Section V-B.
                                    ▷ Update
  for $z \in Z$ **do**   ▷ Targets detected for first time
    - Perform ellipsoidal gating of z w.r.t. Gaussian components of Poisson prior (42).
    if z meets ellipsoidal gating for at least one component **then**
        - Create a new Bernoulli component, see Section V-C1.
    end if
  end for
  for $i = 1$ to n **do**   ▷ We go through all possible targets
    for $j_i = 1$ to l_i **do**   ▷ l_i is the number of single-target hypotheses for possible target i
        - Create new misdetection hypothesis, see Section V-C2.
        - Perform gating on Z and create new detection hypotheses, see Section V-C2.
    end for
  end for
  for all j **do** ▷ We go through all previous global hypotheses
    - Create cost matrix (47).
    - Run Murty's algorithm to select $k = \lceil N_h \cdot w_j \rceil$ new global hypotheses, see Section V-C3.
  end for
  - Estimate target states, see Section VI.
                                    ▷ Pruning
  - Prune the Poisson part by discarding components whose weight is below a threshold.
  - Prune global hypotheses by keeping the highest N_h global hypotheses.
  - Remove Bernoulli components whose existence probability is below a threshold or do not appear in the pruned global hypotheses.

[26, eq. (11.115)]

$$p(n) \propto \sum_j \left[\prod_i w_{j,i} \right] p_j(n) \qquad (48)$$

where $p_j(n)$ is the cardinality distribution of term j of the mixture. The cardinality distribution $p_j(n)$ can be calculated efficiently using a discrete Fourier transform as the cardinality distribution of a multi-Bernoulli RFS is the convolution of the cardinality distributions of its Bernoulli components [35]. By finding the value of n that maximises (48), we obtain the MAP cardinality n^*. We can then

obtain the highest weight global hypothesis with deterministic cardinality, implicitly represented by the MBM, from the global hypothesis

$$j^* = \arg \max_j \prod_{l=1}^{n^*} w_{j,i_l} r_{j,i_l} \prod_{l=n^*+1}^{n} w_{j,i_l}(1 - r_{j,i_l}) \qquad (49)$$

where $i_1, \ldots, i_n$ is an ordering, such that $r_{j,i_l} \geq r_{j,i_{l+1}} \, \forall \, l$. Note that given a MBM hypothesis j, the weight of the deterministic hypothesis with highest weight is given by the term inside the argmax in (49). Once we have found the global hypothesis j^*, the set estimate is formed by the means of the n^* Bernoulli components with highest existence in this hypothesis.

C. Estimator 3

Estimator 3 is the same type of estimator as the one proposed in the MHT of [36], [37], which has also been suggested for the δ-GLMB filter [21]. This estimate first obtains the global hypothesis with a deterministic cardinality with highest weight, i.e., the MAP estimate of the global hypotheses with deterministic cardinality. Note that the global hypotheses (and their weights) with deterministic cardinality (no uncertainty in the cardinality distribution) can be obtained from the MBM (9) by expanding each Bernoulli component so that, in each of the resulting mixture components, either a target exists or not. Then, the estimate is constructed by reporting the mean of the targets in this hypothesis.

We proceed to explain how to obtain this kind of estimate directly from the MBM. We obtain the MAP estimate of the global hypotheses with deterministic cardinality by finding

$$j^* = \arg \max_j \prod_{i|r_{j,i} \geq 0.5} w_{j,i} r_{j,i} \prod_{i|r_{j,i} < 0.5} w_{j,i}(1 - r_{j,i}). \qquad (50)$$

It should be noted that the term inside the argmax in (50) corresponds to the weight of the deterministic hypothesis with highest weight for the jth MBM hypothesis. The set estimate is formed by the means of the Bernoulli components for global hypothesis j^* whose existences are above 0.5, as indicated in (50). In summary, we find that both the δ-GLMB style and the MHT style estimators can be easily constructed from the MBM representation.

VII. SIMULATIONS

In this section, we show simulation results that compare the PMBM filter with the Gaussian mixture PHD, CPHD filters [38], [39], and track-oriented and measurement-oriented multi-Bernoulli/Poisson (TOMB/MOMB) filters in [18]. We also analyze the behaviors of the three estimators proposed in Section VI. We consider an area $[0, 300] \times [0, 300]$ and all the units in this section are in international system. Target states consist of two-dimensional position and velocity $[p_x, v_x, p_y, v_y]^T$ and are born according to a Poisson process of intensity 0.005 and Gaussian density with mean $[100, 0, 100, 0]^T$ and covariance
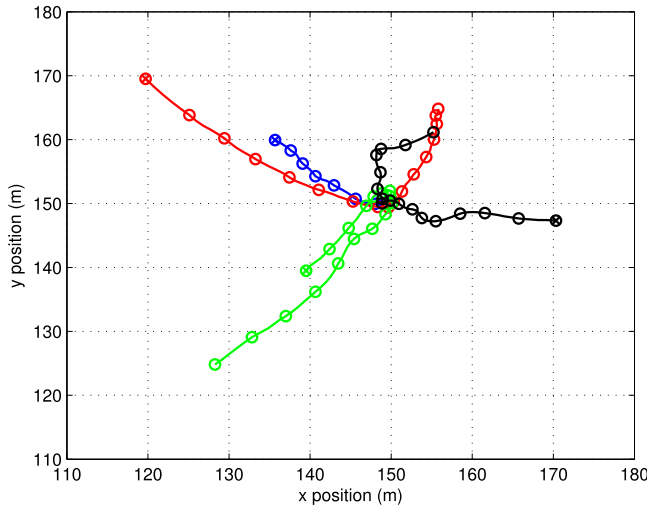
Fig. 3. Scenario of simulations. There are four targets, all born at time step 1 and alive throughout the simulation, except the blue target that dies at time step 40, when all targets are in close proximity. Initial target positions have a cross and target positions every five time steps have a circle.
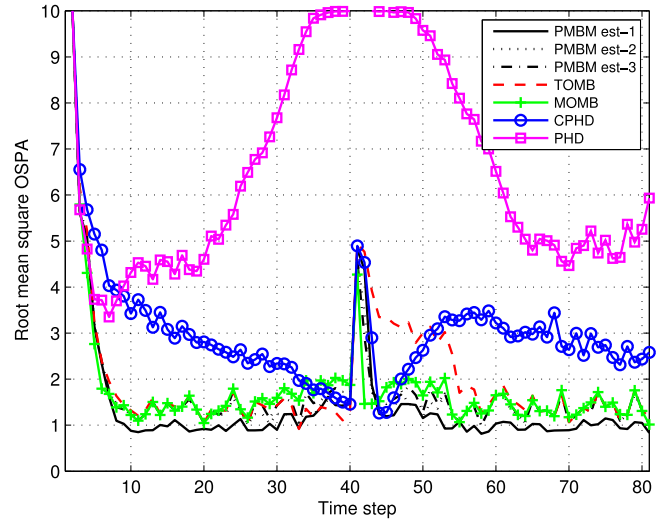

Fig. 4. Mean OSPA error for the algorithms for $p_d = 0.9$ and. The PMBM filter outperforms the rest of the algorithms. Estimator 1 of the PMBM filter provides lowest error and Estimators 2 and 3 perform similarly.

diag $\left(\left[150^2, 1, 150^2, 1 \right] \right)$, which covers the region of interest. We use the following parameters for the simulation:

$$F = I_2 \otimes \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}, Q = q I_2 \otimes \begin{pmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{pmatrix}$$

$$H = I_2 \otimes \begin{pmatrix} 1 & 0 \end{pmatrix}, R = I_2$$

where $\otimes$ is the Kronecker product, $q = 0.01$, $T = 1$, $p_s = 0.99$. We also consider Poisson clutter uniform in the region of interest with $\lambda_c = 10$, which implies 10 expected false alarms per time step, and $p_d = 0.9$. The filters consider that there are no targets at time 0.

The PMBM filter implementation uses a maximum number of global hypotheses $N_h = 200$, estimation threshold for estimator 1 is $\Gamma = 0.4$, which allows two consecutive misdetections for $p_d = 0.9$ and $p_s = 0.99$ to report an estimate, see Section VI. In the Poisson part, we use a pruning threshold of 10^{-5}. For the MB part, we remove Bernoulli components whose existence probability is lower than 10^{-5}. We also use ellipsoidal gating [19] with threshold 20. TOMB/MOMB report estimates for targets with existence probability higher than 0.7.

We consider 81 time steps and the scenario in Fig. 3. These trajectories were generated as indicated in [18, Sec. VI]. For each trajectory, we initiate the midpoint (state at time step 41) from a Gaussian with mean $[150, 0, 150, 0]^T$ and covariance matrix $0.1 I_4$ and the rest of the trajectory is generated running forward and backward dynamics. This scenario is challenging due to the broad Poisson prior that covers the region of interest, the high number of targets in close proximity, and the fact that one target dies when they are in close proximity. We perform 100 Monte Carlo runs and obtain the root mean square OSPA error ($p = 2$, $c = 10$) [33], [40] at each time step for each algorithm, as shown in Fig. 4. Estimator 1 applied to the PMBM filter provides the lowest errors followed by Estimators 2

and 3, which behave similarly. MOMB performs as accurately as Estimators 2 and 3 of the PMBM. It takes TOMB a long time to determine that one target disappears at time step 40. PHD and CPHD are rougher approximations and do not perform well in this scenario.

We also show the root mean square OSPA error averaged over all time steps of the algorithms for different values of p_d and $\lambda_c = 10$ in Table II. On the whole, the PMBM filter performs better than the rest regardless of the estimator. Estimator 1 has lower error than Estimator 2 and 3 for p_d equal or higher than 0.9. For lower values of p_d, Estimator 2 provides lowest errors. The MOMB has the second best performance followed by the TOMB algorithm. The CPHD and PHD filters perform much worse than the other filters.

VIII. CONCLUSION

In this paper, we have first provided a non-PGFL derivation of the PMBM filter in [18], showing its conjugacy property. In order to attain this, we have used a suitable representation of the prior density, which is the union of a Poisson and an MBM, as well as different representations of the likelihood function at several steps. In addition, we have also proved that this derivation can be directly extended to the labeled case by removing the Poisson component and adding unique labels to the Bernoulli components. We have also explained that the PMBM filter parameterization has important benefits compared to the δ-GLMB filter parameterization, which considers hypotheses with deterministic cardinality.

We have also provided an implementation of the PMBM filter for linear/Gaussian measurement models and Poisson births and clutter. The MBM is a more efficient parameterization of the filtering density than the δ-GLMB form and, consequently, the prediction step is greatly simplified. Based on the MTT literature on MHT and labeled RFSs, we have suggested three suboptimal estimators for the PMBM

(p_d, λ_c)	PMBM Est 1	PMBM Est 2	PMBM Est 3	TOMB	MOMB	CPHD	PHD
(0.95, 10)	**2.10**	**2.10**	**2.10**	2.32	**2.10**	2.83	6.34
(0.95, 15)	**2.15**	2.17	**2.15**	2.48	2.17	2.97	6.44
(0.95, 20)	**2.26**	2.27	**2.26**	2.61	2.27	3.00	6.51
(0.9, 10)	**2.23**	2.34	2.36	2.65	2.37	3.39	7.05
(0.9, 15)	**2.30**	2.42	2.44	2.75	2.45	3.45	7.04
(0.9, 20)	**2.37**	2.48	2.50	2.80	2.53	2.57	7.18
(0.8, 10)	2.67	**2.64**	2.66	2.95	2.78	4.19	8.22
(0.8, 15)	2.80	**2.78**	2.80	3.15	2.88	4.25	8.23
(0.8, 20)	2.93	**2.90**	2.92	3.18	3.00	4.48	8.34
(0.7, 10)	3.02	**2.99**	3.01	3.47	3.15	4.83	8.80
(0.7, 15)	3.10	**3.07**	3.09	3.57	3.24	4.99	8.86
(0.7, 20)	3.29	**3.25**	3.28	3.67	3.41	5.09	8.87
(0.6, 10)	3.42	**3.39**	3.42	3.81	3.55	5.30	9.09
(0.6, 15)	3.62	**3.60**	3.62	4.03	3.72	5.52	9.14
(0.6, 20)	3.71	**3.69**	3.71	4.09	3.82	5.61	9.18

filter and how they can be obtained efficiently. Finally, we have compared the performance of the PMBM filter with other RFS filters in a challenging scenario, in which new born targets are distributed according to a Poisson RFS with an intensity that covers the surveillance area and several targets get in close proximity. PMBM outperforms the rest of the filters in this scenario.

APPENDIX A

In this appendix, we prove (13). We denote

$$l_s(\{z_1, \ldots, z_m\}|X) = e^{-\lambda_c} \sum_{U \uplus Y_1 \ldots \uplus Y_m = X} [1 - p_d(\cdot)]^U$$
$$\times \prod_{i=1}^{m} \tilde{l}(z_i|Y_i). \tag{51}$$

We perform a proof by induction. In the rest of this appendix, we denote $Z = \{z_1, \ldots, z_m\}$ and $X = \{x_1, \ldots, x_n\}$ for notational simplicity. First, we note that

$$l(\emptyset|\emptyset) = l_s(\emptyset|\emptyset) = e^{-\lambda_c}. \tag{52}$$

The result is proved if we prove that

$$l(\{z_1, \ldots, z_j\} | \{x_1, \ldots, x_i\})$$
$$= l_s(\{z_1, \ldots, z_j\} | \{x_1, \ldots, x_i\}) \tag{53}$$

for $j \leq m$ and $i \leq n$, implies that

$$l(Z \uplus \{z_{m+1}\}|X) = l_s(Z \uplus \{z_{m+1}\}|X) \tag{54}$$

and

$$l(Z|X \uplus \{x_{n+1}\}) = l_s(Z|X \uplus \{x_{n+1}\}). \tag{55}$$

A. First Part

We proceed to prove (54). We have that

$$l_s(Z \uplus \{z_{m+1}\}|X)$$
$$= e^{-\lambda_c} \sum_{U \uplus Y_1 \ldots \uplus Y_m \uplus Y_{m+1} = X} [1 - p_d(\cdot)]^U \prod_{i=1}^{m+1} \tilde{l}(z_i|Y_i)$$
$$= e^{-\lambda_c} \sum_{Y_{m+1} \subseteq X} \tilde{l}(z_{m+1}|Y_{m+1}) \sum_{U \uplus Y_1 \uplus \ldots \uplus Y_m = X \setminus Y_{m+1}}$$
$$\times [1 - p_d(\cdot)]^U \prod_{i=1}^{m} \tilde{l}(z_i|Y_i)$$
$$= \sum_{Y_{m+1} \subseteq X} \tilde{l}(z_{m+1}|Y_{m+1}) l_s(Z|X \setminus Y_{m+1})$$
$$= \tilde{l}(z_{m+1}|\emptyset) l_s(Z|X) + \sum_{j=1}^{n} \tilde{l}(z_{m+1}|\{x_j\}) l_s(Z|X \setminus \{x_j\}). \tag{56}$$

We also have

$$l(Z \uplus \{z_{m+1}\}|X)$$
$$= e^{-\lambda_c} \sum_{Z^c \uplus Z_1 \ldots \uplus Z_n = Z \uplus \{z_{m+1}\}} [c(\cdot)]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i)$$
$$= e^{-\lambda_c} \left[\sum_{Z^c \uplus Z_1 \ldots \uplus Z_n = Z \uplus \{z_{m+1}\}: z_{m+1} \in Z^c} [c(\cdot)]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i) \right.$$
$$+ \sum_{j=1}^{n} \sum_{Z^c \uplus Z_1 \ldots \uplus Z_n = Z \uplus \{z_{m+1}\}: z_{m+1} \in Z_j} [c(\cdot)]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i) \left. \right]$$
$$= e^{-\lambda_c} \left[\tilde{l}(z_{m+1}|\emptyset) \sum_{Z^c \uplus Z_1 \ldots \uplus Z_n = Z} [c(\cdot)]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i) \right.$$
$$+ \sum_{j=1}^{n} \hat{l}(\{z_{m+1}\}|x_j) \sum_{Z^c \uplus Z_1 \ldots \uplus Z_n = Z: Z_j = \emptyset} [c(\cdot)]^{Z^c}$$

$$\times \prod_{i=1:i\neq j}^{n} \hat{l}(Z_i|x_i) \Bigg]$$

$$= \tilde{l}(z_{m+1}|\emptyset) \, l(Z|X) + \sum_{i=1}^{n} \tilde{l}(z_{m+1}|\{x_i\}) \, l(Z|X \setminus \{x_i\}).$$

$$\text{(57)}$$

Using the induction hypothesis (53), (57) equals (56), so we finish the proof of (54).

B. Second Part

We proceed to prove (55). In this part, we denote $p_d'(\cdot) = 1 - p_d(\cdot)$. We have that

$$l_s(Z|X \uplus \{x_{n+1}\})$$

$$= e^{-\lambda_c} \sum_{U \uplus Y_1 ... \uplus Y_m = X \uplus \{x_{n+1}\}} \left[p_d'(\cdot) \right]^U \prod_{i=1}^{m} \tilde{l}(z_i|Y_i)$$

$$= e^{-\lambda_c} \Bigg[\sum_{U \uplus Y_1 ... \uplus Y_m = X \uplus \{x_{n+1}\}: x_{n+1} \in U} \left[p_d'(\cdot) \right]^U \prod_{i=1}^{m} \tilde{l}(z_i|Y_i)$$

$$+ \sum_{j=1}^{m} \sum_{U \uplus Y_1 ... \uplus Y_m = X \uplus \{x_{n+1}\}: x_{n+1} \in Y_j} \left[p_d'(\cdot) \right]^U \prod_{i=1}^{m} \tilde{l}(z_i|Y_i) \Bigg]$$

$$= e^{-\lambda_c} \Bigg[p_d'(x_{n+1}) \sum_{U \uplus Y_1 ... \uplus Y_m = X} \left[p_d'(\cdot) \right]^U \prod_{i=1}^{m} \tilde{l}(z_i|Y_i)$$

$$+ \sum_{j=1}^{m} \tilde{l}(z_j|\{x_{n+1}\}) \sum_{U \uplus Y_1 ... \uplus Y_m = X: Y_j = \emptyset} \left[p_d'(\cdot) \right]^U$$

$$\times \prod_{i=1:i\neq j}^{n} \tilde{l}(z_i|Y_i) \Bigg]$$

$$= p_d'(x_{n+1}) \, l_s(Z|X) + \sum_{j=1}^{m} \tilde{l}(z_j|\{x_{n+1}\}) \, l_s(Z \setminus \{z_j\}|X).$$

$$\text{(58)}$$

We also have that

$$l(Z|X \uplus \{x_{n+1}\})$$

$$= e^{-\lambda_c} \sum_{Z^c \uplus Z_1 ... \uplus Z_{n+1} = Z} \left[c(\cdot) \right]^{Z^c} \prod_{i=1}^{n+1} \hat{l}(Z_i|x_i)$$

$$= e^{-\lambda_c} \sum_{Z_{n+1} \subseteq Z} \hat{l}(Z_{n+1}|x_{n+1})$$

$$\times \sum_{Z^c \uplus Z_1 ... \uplus Z_n = Z \setminus Z_{n+1}} \left[c(\cdot) \right]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i)$$

$$= e^{-\lambda_c} \Bigg[p_d'(x_{n+1}) \sum_{Z^c \uplus Z_1 ... \uplus Z_n = Z} \left[c(\cdot) \right]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i)$$

$$+ \sum_{j=1}^{m} \tilde{l}(z_j|\{x_{n+1}\}) \sum_{Z^c \uplus Z_1 ... \uplus Z_n = Z \setminus \{z_j\}} \left[c(\cdot) \right]^{Z^c} \prod_{i=1}^{n} \hat{l}(Z_i|x_i) \Bigg]$$

$$= p_d'(x_{n+1}) \, l(Z|X) + \sum_{j=1}^{m} \tilde{l}(z_j|\{x_{n+1}\}) \, l(Z \setminus \{z_j\}|X).$$

$$\text{(59)}$$

Given that the induction hypothesis (53) holds, (58) and (59) are identical, so we finish the proof of (55).

APPENDIX B

In this appendix, we show how to update a Poisson prior, whose result is given in (16)–(24). Substituting (13) into (15), we find

$$q^p(X|Z)$$

$$\propto f^p(X) \sum_{U \uplus Y_1 ... \uplus Y_m = X} \left[1 - p_d(\cdot) \right]^U \prod_{i=1}^{m} \tilde{l}(z_i|Y_i)$$

$$= \sum_{U \uplus Y_1 ... \uplus Y_m = X} \left[1 - p_d(\cdot) \right]^U \left[\prod_{i=1}^{m} \tilde{l}(z_i|Y_i) \right]$$

$$\times f^p(U \uplus Y_1 ... \uplus Y_m)$$

$$\propto \sum_{U \uplus Y_1 ... \uplus Y_m = X} \left[1 - p_d(\cdot) \right]^U f^p(U) \left[\prod_{i=1}^{m} \tilde{l}(z_i|Y_i) \, f^p(Y_i) \right]$$

$$\propto \sum_{U \uplus Y_1 \uplus ... \uplus Y_m = X} q^p(U) \left[\prod_{i=1}^{m} \rho^p(z_i) \, q^p(Y_i|z_i) \right].$$

In the previous derivation, we have used that $f^p(U \uplus Y_1 ... \uplus Y_m) \propto f^p(U) \prod_{i=1}^{m} f^p(Y_i)$, see (8), and (18) and (19). The specific form of $q^p(Y_i|z_i)$, which is given in (20), is obtained straightforwardly by calculating (19).

APPENDIX C

In this appendix, we prove (27). By definition, we know that (27) is met for $n = 0$ as $l_o(Z|Y) = l(Z|Y)$. By induction, (27) is proved if the equality

$$l_o(Z|Y, X_1, \ldots, X_n) = l(Z|Y \uplus X_1 \uplus ... \uplus X_n)$$

implies

$$l_o(Z|Y, X_1, \ldots, X_n, X_{n+1})$$
$$= l(Z|Y \uplus X_1 \uplus ... \uplus X_n \uplus X_{n+1}).$$

We have to prove two cases: $X_{n+1} = \emptyset$ and $X_{n+1} = \{x\}$. For $X_{n+1} = \emptyset$, we have that $Z_{n+1} = \emptyset$ so that $t(Z_{n+1}|X_{n+1}) \neq 0$. Therefore

$$l_o(Z|Y, X_1, \ldots, X_n, \emptyset)$$

$$= \sum_{Z_1 \uplus ... \uplus Z_n \uplus Z^y = Z} l(Z^y|Y) \prod_{i=1}^{n} t(Z_i|X_i)$$

$$= l_o(Z|Y, X_1, \ldots, X_n)$$
$$= l(Z|X \uplus \emptyset)$$

where $X = Y \uplus X_1 \uplus ... \uplus X_n$. This proves the first case.

For $X_{n+1} = \{x\}$, we have

$$l_o(Z|Y, X_1, \ldots, X_n, \{x\})$$

$$= \sum_{Z_1 \ldots \uplus Z_n \uplus Z_{n+1} \uplus Z^y = Z} l\left(Z^y|Y\right) t\left(Z_i|\{x\}\right) \prod_{i=1}^{n} t\left(Z_i|X_i\right)$$

$$= t\left(\emptyset|\{x\}\right) \sum_{Z_1 \ldots \uplus Z_n \uplus Z^y = Z} l\left(Z^y|Y\right) t\left(Z_i|\{x\}\right) \prod_{i=1}^{n} t\left(Z_i|X_i\right)$$

$$+ \sum_{z \in Z} t\left(\{z\}|\{x\}\right) \sum_{Z_1 \uplus \ldots \uplus Z_n \uplus Z^y = Z \setminus \{z\}} l\left(Z^y|Y\right) t\left(Z_i|\{x\}\right)$$

$$\times \prod_{i=1}^{n} t\left(Z_i|X_i\right)$$

$$= (1 - p_d(x)) l(Z|X)$$

$$+ p_d(x) \sum_{z \in Z} l(z|x) l(Z \setminus \{z\}|X)$$

$$= l(Z|X \uplus \{x\})$$

where $X = Y \uplus X_1 \uplus ... \uplus X_n$. This proves the second case.

APPENDIX D

In this appendix, we prove the prediction step of Corollary 3. We consider that the new born targets follow an MBM with parameters

$$f_b^{\text{mbm}}(Y) \propto \sum_{j_b} \sum_{Y_1 \uplus \ldots \uplus Y_{n_b} = Y} \prod_{i_b=1}^{n_b} w_{j_b, i_b}^b f_{j_b, i_b}^b(Y_i). \quad (60)$$

As indicated in Section III-D, the predicted density of the survival targets when the Poisson intensity is zero is an MBM. We denote the parameters of this MBM as in (9). Then, the output of the prediction step is the multi-target density of the union of the survival targets and the new born targets, which can be computed using the convolution formula [9, eq. (4.17)]

$$f_{\text{pred}}(W) = \sum_{X \uplus Y = W} f^{\text{mbm}}(X) f_b^{\text{mbm}}(Y)$$

$$\propto \sum_{X \uplus Y = W} \left[\sum_j \sum_{X_1 \uplus \ldots \uplus X_n = X} \prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i) \right]$$

$$\times \left[\sum_{j_b} \sum_{Y_1 \uplus \ldots \uplus Y_{n_b} = Y} \prod_{i_b=1}^{n_b} w_{j_b, i_b}^b f_{j_b, i_b}^b(Y_i) \right]$$

$$= \sum_j \sum_{j_b} \sum_{X_1 \uplus \ldots \uplus X_n \uplus Y_1 \uplus \ldots \uplus Y_{n_b} = W}$$

$$\left[\prod_{i=1}^{n} w_{j,i} f_{j,i}(X_i) \right] \left[\prod_{i_b=1}^{n_b} w_{j_b, i_b}^b f_{j_b, i_b}^b(Y_i) \right]$$

which corresponds to an MBM.

APPENDIX E

In this appendix, we prove Proposition 7. We first prove how a labeled MBM, which contains the labeled MBM$_{01}$ as a particular case, can be written as a δ-GLMB density.

We write (39) as

$$f(X) = \sum_j w_j \sum_{X_1 \uplus \ldots \uplus X_n = X} \prod_{i=1}^{n} f_{j,i}^{lb}(X_i) \quad (61)$$

where we have normalized the weights of the global hypotheses such that $\sum_j w_j = 1$ and $w_j \propto \prod_{i=1}^{n} w_{j,i}$. Let $\mathbb{L} = \{\ell_1, \ldots, \ell_n\}$ denote the set with all the possible target labels according to the density (39).

Both the δ-GLMB density and the labeled MBM are zero if; first, they are evaluated on a set that includes more than one target with the same label, or second, if they are evaluated on a set that includes a target whose label does not belong to the label space $\mathbb{L}$. Therefore, the case of interest is when we evaluate the density with a set of targets with distinct labels that belong to $\mathbb{L}$. We evaluate the labeled MBM (61) on a labeled set $\{(x_1, \ell_{a_1}), \ldots, (x_p, \ell_{a_p})\}$ where $\ell_{a_1}, \ldots, \ell_{a_p}$ are p distinct labels that belong to $\mathbb{L}$. We also denote by $\ell_{a_{p+1}}, \ldots, \ell_{a_n}$ the rest of distinct labels in $\mathbb{L}$. As labels $\ell_{a_1}, \ldots, \ell_{a_p}$ are distinct, there is only one combination in the sum over $X_1 \uplus \ldots \uplus X_n = X$ that is nonzero. This yields

$$f\left(\{(x_1, \ell_{a_1}), \ldots, (x_p, \ell_{a_p})\}\right)$$

$$= \sum_j w_j \left[\prod_{m=1}^{p} r_{j,a_m} p_{j,a_m}(x_m) \right] \prod_{i=p+1}^{n} \left(1 - r_{j,a_i}\right). \quad (62)$$

We proceed to write this density in the δ-GLMB form [21]. We denote

$$w_j\left(\{\ell_{a_1}, \ldots, \ell_{a_p}\}\right) = w_j \left[\prod_{m=1}^{p} r_{j,a_m} \right] \prod_{i=p+1}^{n} \left(1 - r_{j,a_i}\right). \quad (63)$$

In the δ-GLMB filter, this weight is written as (see sentence that contains (9) in [21])

$$w_j\left(\{\ell_{a_1}, \ldots, \ell_{a_p}\}\right) = \sum_{I \subseteq \mathbb{L}} w_j(I) \delta_I\left(\{\ell_{a_1}, \ldots, \ell_{a_p}\}\right) \quad (64)$$

where [21]

$$\delta_I(L) \triangleq \begin{cases} 1 & \text{if } I = L \\ 0 & \text{otherwise} \end{cases}$$

and it can be verified that $\sum_j \sum_{I \subseteq \mathbb{L}} w_j(I) = 1$. The previous step is direct, as there is only one summand in (64) that is different from zero, which corresponds to (63). Following [21], we also denote $p_\xi(x, \ell) = p_{\xi, i(\ell)}(x)$ where $i(\ell) = i$, such that $\ell = \ell_i$ and index j is denoted as ξ. Substituting this notation into (62), we find

$$f\left(\{(x_1, \ell_{a_1}), \ldots, (x_p, \ell_{a_p})\}\right)$$

$$= \sum_\xi \sum_{I \subseteq \mathbb{L}} w_\xi(I) \delta_I\left(\{\ell_{a_1}, \ldots, \ell_{a_p}\}\right) \prod_{m=1}^{p} p_\xi\left(x_m, \ell_{a_m}\right) \quad (65)$$

which corresponds to the δ-GLMB density [21, eq. (9)] evaluated on a set of targets with different labels.

In order to finish the proof of Proposition 7, we write a δ-GLMB density as a labeled MBM with MBM_{01} parameterization.

We consider that the label space is $\mathbb{L} = \{\ell_1, \ldots, \ell_n\}$, the δ-GLMB single target densities are $p_\xi(\cdot, \ell_i)$ for all ξ and $\ell_i \in \mathbb{L}$, and the global hypothesis weights are $w_\xi(I)$ for $I \subseteq \mathbb{L}$. In order to prove the equivalence, we evaluate a δ-GLMB density $f(\cdot)$ at $\{(x_1, \ell_{a_1}), \ldots, (x_p, \ell_{a_p})\}$, which is given by (65), with $\{\ell_{a_1}, \ldots, \ell_{a_p}\} \subseteq \mathbb{L}$. We also denote by $\ell_{a_{p+1}}, \ldots, \ell_{a_n}$ the rest of distinct labels in $\mathbb{L}$. It should be noted that the pair (ξ, I) represents a δ-GLMB global hypothesis [21] and that, in this global hypothesis, all targets whose label belongs to I exist and the rest do not exist, which is represented by $\delta_I(\{\ell_{a_1}, \ldots, \ell_{a_p}\})$ in (65). For global hypothesis (ξ, I), this factor can be written as a product of existence probabilities, which are either 0 or 1, as

$$
\delta_I\left(\{\ell_{a_1}, \ldots, \ell_{a_p}\}\right)
$$
$$
= \left[\prod_{m=1}^{p} r_{(\xi,I),a_m}\right]\left[\prod_{i=p+1}^{n}\left(1 - r_{(\xi,I),a_i}\right)\right] \quad (66)
$$

where $r_{(\xi,I),a_m} = 1$ if $\ell_{a_m} \in I$ and $r_{(\xi,I),a_m} = 0$ if $\ell_{a_m} \notin I$ for $m \in \{1, \ldots, n\}$. We can write the two sums in (65) as one sum over $j = (\xi, I)$ such that

$$
f\left(\{(x_1, \ell_{a_1}), \ldots, (x_p, \ell_{a_p})\}\right)
$$
$$
= \sum_j w_j \left[\prod_{m=1}^{p} r_{j,a_m}\right]\left[\prod_{i=p+1}^{n}\left(1 - r_{j,a_i}\right)\right]\prod_{m=1}^{p} p_{j,a_m}(x_m)
$$
$$
\quad (67)
$$

where $p_{(\xi,I),a_m}(\cdot) = p_\xi(\cdot, \ell_{a_m})$, $w_{(\xi,I)} = w_\xi(I)$. It should be noted that, in the δ-GLMB density, we have $\sum_\xi \sum_{I \subseteq \mathbb{L}} w_\xi(I) = 1$, which implies that $\sum_j w_j = 1$, as required. Also note that $r_{j,i}$ is the existence probability of Bernoulli component i, with label ℓ_i, and global hypothesis j, which is either 0 or 1. Equation (67) corresponds to the evaluation of a labeled MBM, see (62). In particular, the resulting global hypotheses (mixture components) of the δ-GLMB density are equivalent to the global hypotheses in an MBM_{01} parameterization, which have deterministic target existence.

REFERENCES

[1] S. S. Blackman
     Multiple hypothesis tracking for multiple target tracking
     IEEE Aerosp. Electron. Syst. Mag., vol. 19, no. 1, pp. 5–18, Jan. 2004.
[2] W. Koch and F. Govaers
     On accumulated state densities with applications to out-of-sequence measurement processing
     IEEE Trans. Aerosp. Electron. Syst., vol. 47, no. 4, pp. 2766–2778, Oct. 2011.
[3] J. García, A. Berlanga, and J. M. M. López
     Effective evolutionary algorithms for many-specifications attainment: Application to air traffic control tracking filters
     IEEE Trans. Evol. Comput., vol. 13, no. 1, pp. 151–168, Feb. 2009.
[4] H. Bhaskar, K. Dwivedi, D. P. Dogra, M. Al-Mualla, and L. Mihaylova
     Autonomous detection and tracking under illumination changes, occlusions and moving camera
     Signal Process., vol. 117, pp. 343–354, 2015.
[5] A. Petrovskaya and S. Thrun
     Model based vehicle detection and tracking for autonomous urban driving
     Auton. Robots, vol. 26, no. 2, pp. 123–139, 2009.
[6] F. Kunz *et al.*
     Autonomous driving at Ulm university: A modular, robust, and sensor-independent fusion approach
     In *Proc. IEEE Intell. Veh. Symp.*, Jun. 2015, pp. 666–673.
[7] C. Fantacci and F. Papi
     Scalable multisensor multitarget tracking using the marginalized δ-GLMB density
     IEEE Signal Process. Lett., vol. 23, no. 6, pp. 863–867, Jun. 2016.
[8] F. Meyer, P. Braca, P. Willett, and F. Hlawatsch
     A scalable algorithm for tracking an unknown number of targets using multiple sensors
     IEEE Trans. Signal Process., vol. 65, no. 13, pp. 3478–3493, Jul. 2017.
[9] R. P. S. Mahler
     Advances in Statistical Multisource-Multitarget Information Fusion. Norwood, MA, USA: Artech House, 2014.
[10] X. Li and V. Jilkov
     Survey of maneuvering target tracking. Part I: Dynamic models
     IEEE Trans. Aerosp. Electron. Syst., vol. 39, no. 4, pp. 1333–1364, Oct. 2003.
[11] A. Swain and D. Clark
     Extended object filtering using spatial independent cluster processes
     In *Proc. 13th Conf. Inf. Fusion*, Jul. 2010, pp. 1–8.
[12] K. Granström, C. Lundquist, and O. Orguner
     Extended target tracking using a Gaussian-mixture PHD filter
     IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 4, pp. 3268–3286, Oct. 2012.
[13] A. F. García-Fernández, J. Grajal, and M. R. Morelande
     Two-layer particle filter for multiple target detection and tracking
     IEEE Trans. Aerosp. Electron. Syst., vol. 49, no. 3, pp. 1569–1588, Jul. 2013.
[14] S. J. Davey, M. G. Rutten, and B. Cheung
     A comparison of detection performance for several track-before-detect algorithms
     EURASIP J. Adv. Signal Process., vol. 2008, 2008, Art. no. 41.
[15] B. Ristic, S. Arulampalam, and N. Gordon
     Beyond the Kalman Filter: Particle Filters for Tracking Applications. Norwood, MA, USA: Artech House, 2004.
[16] C. P. Robert
     The Bayesian Choice. New York, NY, USA: Springer, 2007.
[17] B. T. Vo and B. N. Vo
     Labeled random finite sets and multi-object conjugate priors
     IEEE Trans. Signal Process., vol. 61, no. 13, pp. 3460–3475, Jul. 2013.
[18] J. L. Williams
     Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA and association-based MeMBer
     IEEE Trans. Aerosp. Electron. Syst., vol. 51, no. 3, pp. 1664–1687, Jul. 2015.
[19] T. Kurien
     Issues in the design of practical multitarget tracking algorithms
     In *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed. Norwood, MA, USA: Artech House, 1990.
[20] A. F. García-Fernández and J. Grajal
     Multitarget tracking using the joint multitrack probability density
     In *Proc. 12th Int. Conf. Inf. Fusion*, Jul. 2009, pp. 595–602.

[21] B.-N. Vo, B.-T. Vo, and D. Phung
Labeled random finite sets and the Bayes multi-target tracking filter
IEEE Trans. Signal Process., vol. 62, no. 24, pp. 6554–6567, Dec. 2014.

[22] E. H. Aoki, P. K. Mandal, L. Svensson, Y. Boers, and A. Bagchi
Labeling uncertainty in multitarget tracking
IEEE Trans. Aerosp. Electron. Syst., vol. 52, no. 3, pp. 1006–1020, Jun. 2016.

[23] R. Mahler
PHD filters of higher order in target number
IEEE Trans. Aerosp. Electron. Syst., vol. 43, no. 4, pp. 1523–1543, Oct. 2007.

[24] R. P. S. Mahler
Multitarget Bayes filtering via first-order multitarget moments
IEEE Trans. Aerosp. Electron. Syst., vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[25] A. F. García-Fernández and B.-N. Vo
Derivation of the PHD and CPHD filters based on direct Kullback-Leibler divergence minimization
IEEE Trans. Signal Process., vol. 63, no. 21, pp. 5812–5820, Nov. 2015.

[26] R. P. S. Mahler
Statistical Multisource-Multitarget Information Fusion. Norwood, MA, USA: Artech House, 2007.

[27] J. L. Williams
An efficient, variational approximation of the best fitting multi-Bernoulli filter
IEEE Trans. Signal Process., vol. 63, no. 1, pp. 258–273, Jan. 2015.

[28] S. Särkkä
Bayesian Filtering and Smoothing. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[29] K. G. Murty
An algorithm for ranking all the assignments in order of increasing cost
Oper. Res., vol. 16, no. 3, pp. 682–687, 1968.

[30] I. J. Cox and M. L. Miller
On finding ranked assignments with application to multitarget tracking and motion correspondence
IEEE Trans. Aerosp. Electron. Syst., vol. 31, no. 1, pp. 486–489, Jan. 1995.

[31] I. J. Cox and S. L. Hingorani
An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking
IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 2, pp. 138–150, Feb. 1996.

[32] E. Fortunato, W. Kreamer, S. Mori, C.-Y. Chong, and G. Castanon
Generalized Murty's algorithm with application to multiple hypothesis tracking
In *Proc. Int. Conf. Inf. Fusion*, Jul. 2007, pp. 1–8.

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

[34] M. Guerriero, L. Svensson, D. Svensson, and P. Willett
Shooting two birds with two bullets: How to find minimum mean OSPA estimates
In *Proc. 13th Conf. Inf. Fusion*, Jul. 2010, pp. 1–8.

[35] M. Fernandez and S. Williams
Closed-form expression for the Poisson-binomial probability density function
IEEE Trans. Aerosp. Electron. Syst., vol. 46, no. 2, pp. 803–817, Apr. 2010.

[36] D. Reid
An algorithm for tracking multiple targets
IEEE Trans. Automat. Control, vol. 24, no. 6, pp. 843–854, Dec. 1979.

[37] S. Mori, C.-Y. Chong, E. Tse, and R. Wishner
Tracking and classifying multiple targets without a priori identification
IEEE Trans. Automat. Control, vol. 31, no. 5, pp. 401–409, May 1986.

[38] B.-N. Vo and W.-K. Ma
The Gaussian mixture probability hypothesis density filter
IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

[39] B.-T. Vo, B.-N. Vo, and A. Cantoni
Analytic implementations of the cardinalized probability hypothesis density filter
IEEE Trans. Signal Process., vol. 55, no. 7, pp. 3553–3567, Jul. 2007.

[40] A. S. Rahmathullah, A. F. García-Fernández, and L. Svensson
Generalized optimal sub-pattern assignment metric
In *Proc. 20th Int. Conf. Inf. Fusion*, 2017, pp. 1–8.

Ángel F. García-Fernández received the M.Sc. (with hons.) and the Ph.D. degrees in telecommunication engineering from Universidad Politécnica de Madrid, Madrid, Spain, in 2007 and 2011, respectively.

He is currently a Lecturer with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, U.K. He previously held postdoctoral positions with Universidad Politécnica de Madrid, Chalmers University of Technology, Gothenburg, Sweden, Curtin University, Perth, Australia, and Aalto University, Espoo, Finland. His main research activities and interests include the area of Bayesian nonlinear inference, with emphasis on dynamic systems. He was the recipient of the best paper award at the International Conference on Information Fusion in 2017.

Jason L. Williams (S'01–M'07–SM'16) received the B.E. degree in electronics and the B.Inf.Tech. degree from Queensland University of Technology, Brisbane, QLD, Australia, in 1999, the M.S.E.E. degree from the United States Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, USA, in 2003, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2007, respectively.

He worked for several years as an Engineering Officer with the Royal Australian Air Force, before joining Australia's Defence Science and Technology Group (formerly Defence Science and Technology Organisation), in 2007. He is also an Adjunct Associate Professor with Queensland University of Technology. His research interests include tracking, sensor resource management, Markov random fields, and convex optimization.

Karl Granström (M'08) received the M.Sc. degree in applied physics and electrical engineering and the Ph.D. degree in automatic control from Linköping University, Linköping, Sweden, in May 2008 and November 2012, respectively.

He is currently a Postdoctoral Research Fellow with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. He previously held Postdoctoral positions with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT, USA, from September 2014 to August 2015, and with the Department of Electrical Engineering, Linköping University, from December 2012 to August 2014. His research interests include Bayesian probability theory, multiple model estimation, sensor fusion and target tracking, especially for extended targets. He has organised several workshops and tutorials on the topic multiple extended target tracking.

Dr. Granstrom was the recipient of paper awards at the Fusion 2011 and Fusion 2012 conferences.

Lennart Svensson was born in Älvängen, Sweden, in 1976. He received the M.S. degree in electrical engineering in 1999 and the Ph.D. degree in 2004, both from Chalmers University of Technology, Gothenburg, Sweden.

He is currently Professor of Signal Processing, again at Chalmers University of Technology. His main research interests include machine learning and Bayesian inference in general, and nonlinear filtering, deep learning and tracking in particular.

Dr. Svensson was the recipient of the best paper award at the International Conference on Information Fusion in 2010 and 2017, and the second-best paper award in 2009.