Multiscan Implementation of the Trajectory Poisson Multi-Bernoulli Mixture Filter

YUXUAN XIA
KARL GRANSTRÖM
LENNART SVENSSON
ÁNGEL F. GARCÍA-FERNÁNDEZ
JASON L. WILLIAMS

The Poisson multi-Bernoulli mixture (PMBM) and the multi-Bernoulli mixture (MBM) are two multitarget distributions for which closed-form filtering recursions exist. The PMBM has a Poisson birth process, whereas the MBM has a multi-Bernoulli birth process. This paper considers a recently developed formulation of the multitarget tracking problem using a random finite set of trajectories, through which the track continuity is explicitly established. A multiscan trajectory PMBM filter and a multiscan trajectory MBM filter, with the ability to correct past data association decisions to improve current decisions, are presented. In addition, a multiscan trajectory MBM_{01} filter, in which the existence probabilities of all Bernoulli components are either 0 or 1, is presented. This paper proposes an efficient implementation that performs track-oriented *N*-scan pruning to limit computational complexity, and uses dual decomposition to solve the involved multiframe assignment problem. The performance of the presented multitarget trackers, applied with an efficient fixed-lag smoothing method, is evaluated in a simulation study.

Manuscript received January 2, 2019; revised May 6, 2019, September 3, 2019; released for publication November 14, 2019.

Refereeing of this contribution was handled by Chee-Yee Chong.

Y. Xia, K. Granström, and L. Svensson are with the Department of Electrical Engineering, Chalmers University of Technology, 41296, Göteborg, Sweden.
Á. F. García-Fernández is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool J69 3GJ, U.K.
J. L. Williams is with the Commonwealth Scientific and Industrial Research Organisation, Brisbane QLD 4006 and Queensland University of Technology, Brisbane QLD 4000, Australia.

1557-6418/19/$17.00

I. INTRODUCTION

Multitarget tracking (MTT) refers to the problem of jointly estimating the number of targets and their trajectories from noisy sensor measurements [1]. The number of targets and their trajectories can be time-varying due to targets appearing and disappearing. In a general MTT system, a multitarget tracker needs to tackle the modeling of births and deaths of targets, as well as the partitioning of noisy sensor measurements into potential tracks and false alarms; the latter is also referred to as data association. The major approaches to MTT include the joint probabilistic data association (JPDA) filter [2], the multiple hypothesis tracker (MHT) [3]–[5], and random finite set (RFS) [6] based multitarget filters [7, Ch. 6].

The JPDA filter [2] seeks to calculate the marginal distribution of each track. To accommodate for an unknown and time-varying number of targets, the joint integrated probabilistic data association [8] extends the basic JPDA [2] by incorporating target existence as an additional random variable to be estimated. It has recently been shown that the marginal data association probabilities can be efficiently approximated using message passing algorithms [9], [10].

MHT is described in a number of books; e.g., see [3, Ch. 16] and [4, Chs. 6 and 7]. The model was made rigorous in [11] through random finite sequences, under the assumption that the number of targets present is constant but unknown, with an a priori distribution that is Poisson. In MHT, multiple data association hypotheses are formed to explain the source of the measurements. Each data association hypothesis assigns measurements to previously detected targets, newly detected targets, or false alarms. Data association uncertainty is captured by the data hypothesis weight, and the target state uncertainty is captured by the target state density distribution conditioned on each hypothesis.

There are two types of MHT algorithms: the hypothesis-oriented MHT (HOMHT) [12] and the track-oriented MHT (TOMHT) [13], [14]. In HOMHT, multiple global hypotheses are formed and evaluated between consecutive time scans; the complete algorithmic approach was first developed by Reid [12]. The TOMHT operates by maintaining a number of single target hypothesis trees, each of which contains a number of single target hypotheses explaining the measurement association history of a potential target.

A TOMHT algorithm usually uses a deferred decision logic to consider the data associations of measurements from more than one scan, in the sense that the hypotheses are propagated into the future in anticipation that subsequent data will resolve the uncertainty [5]. Intuitively, measurements in more than one scan may provide more accurate data association than those in a single scan. The number of single target hypotheses can be limited by performing *N*-scan pruning [5], and the involved multiframe assignment problem is typi-

cally solved using Lagrangian relaxation-based methods [15]–[17]. Track management (target initiation and termination) is usually performed using some external procedures; see, e.g., [18].

RFSs and finite set statistics (FISST) were developed to provide a systematic methodology for dealing with MTT problems involving a time-varying number of targets [6]. The relationship between RFS-based approaches to MTT and MHT has been discussed in [19] and [20]. In the RFS formulation of MTT, the multitarget filtering density contains the information of the target states at the current time step. Exact closed-form solutions of RFS-based multitarget Bayes filter are given by multitarget conjugate priors. The concept of multitarget conjugate prior was defined in [21] as "If we start with the proposed conjugate initial prior, then all subsequent predicted and posterior distributions have the same form as the initial prior."

Two well-established MTT conjugate priors for the standard point target measurement model are the Poisson multi-Bernoulli mixture (PMBM) [22] based on unlabeled RFSs and the generalized labeled multi-Bernoulli (GLMB) [21] based on labeled RFSs. The PMBM consists of a Poisson distribution representing targets that are hypothesized to exist but have not been detected and a multi-Bernoulli mixture (MBM) representing targets that have been detected at some stage. The resulting PMBM filter [23] is a computationally tractable filter for the standard point target dynamic model, where the birth model is a Poisson RFS. If the birth process is a multi-Bernoulli RFS, the multitarget conjugate prior is of the form MBM [23], [24]. A discussion regarding the differences between the use of a Poisson birth model and the use of a multi-Bernoulli birth model can be found in [24].

A. Track Continuity in MTT

In this section, we discuss how track continuity can be maintained in different MTT methodologies. Vector-type MTT methods, e.g., the JPDA filter and the MHT, describe the multitarget states and measurements by random vectors. They are able to explicitly maintain track continuity; i.e., they associate a state estimate with a previous state estimate or declare the appearance of a new target [10]. For multitarget filters based on unlabeled RFS, time sequences of tracks cannot be constructed easily due to the set representation of the multitarget states that are order independent. The PMBM filter (as well as the MBM filter) seemingly does not provide explicit track continuity between time steps,[1] although a hypothesis structure in analogy to MHT was observed in [22] and [23].

[1]The PMBM filter and the MBM filter are able to maintain track continuity implicitly, in a practical setting, based on information provided by metadata.

One approach to addressing the lack of track continuity is to add unique labels to the target states and estimate target states from the multitarget filtering density [21], [25], [26]. This procedure can work well in some cases, but it becomes problematic in challenging situations, for example, when target birth is independent and identically distributed, and when targets get in close proximity and then separate [27]. The δ-GLMB filter [28] (and its approximation the labeled multi-Bernoulli (LMB) filter [29]) is an example of the resulting labeled filter when the birth model is an LMB (mixture) RFS. The δ-GLMB density is similar in structure to labeled MBM using MBM_{01} parameterization [23], in which Bernoulli components are uniquely labeled, and their existence probability is restricted to either 0 or 1. It was shown in [23] that the MBM parameterization has computational and implementational advantages over the MBM_{01} parameterization.

B. Trajectory PMBM Filter and Its Relation to MHT

In this section, we give a brief introduction to the trajectory PMBM filter and discuss its relation to MHT. More details of the trajectory PMBM filter will be given in Section III.

Compared to augmenting target states with unique labels, a more appealing approach to ensuring track continuity for RFS-based multitarget filters is to generalize the concept of RFSs of targets to RFSs of trajectories. The theoretical background to perform MTT using RFS of trajectories was provided in [27] and [30]. Within the set of trajectories framework, the goal of MTT is to recursively compute the posterior density over the set of trajectories, which contains full information about the target trajectories, and can be used to estimate the best set of trajectories at each time step.

Closed-form PMBM filtering recursions based on the set of trajectories framework have been derived in [31], which enables us to leverage on the benefits of the PMBM filter recursion based on sets of targets, while also obtaining track continuity. Assuming standard point target dynamic [32, Sec. 13.2.4] and measurement models (defined in Section II-A), two different trajectory PMBM filters were proposed in [31]: one in which the set of current (i.e., alive) trajectories is tracked, and one in which the set of all trajectories (dead and alive) up to the current time step is tracked. In both cases, finite trajectories, i.e., trajectories of finite length in time, are considered.

The implementation of the trajectory PMBM filter in [31] considers the single-scan data association problem, and the best global hypotheses are found using Murthy's algorithm [33]. As a complement to [31], an approximation to the exact trajectory PMBM filter that considers multiscan data association was developed in [34]. It operates by performing track-oriented N-scan pruning [5] to limit computational complexity, and using dual

decomposition [17] to solve the involved multiframe assignment problem. The proposed algorithm therefore shares some of the key properties of certain TOMHT algorithms [5], [17], but is derived using RFSs of trajectories and birth/death models. As a comparison, TOMHT algorithms typically use heuristics to take into account the appearance and disappearance of targets [4, Ch. 7].

Numerical results in [34] show that the proposed multiscan trajectory PMBM filter has better tracking performance than the fast implementation of the δ-GLMB filter using Gibbs sampling [35] in terms of estimation error and computational time. These two filters use different birth models, Poisson RFS and multi-Bernoulli RFS, respectively. A multi-Bernoulli birth can be suitable if one is certain that a known maximum of targets will enter the area of interest and the targets appear around some known locations. With multi-Bernoulli birth, the PMBM conjugate prior becomes an MBM conjugate prior [23]. An implementation of the MBM filter for sets of targets was proposed in [24]. The case in which the probability distribution of the number of targets is not necessarily Poisson was discussed in [36] for the batch-processing formulation used for TOMHT; however, a practical implementation was not provided in [36].

The data association is explicitly represented in both the trajectory PMBM filter and the trajectory MBM filter, in a data structure analogous to TOMHT. Compared to conventional MHT formalism, as described in [5] and [14], one important difference is that the presented trajectory PMBM filters include a Poisson RFS that models undetected trajectories. The modeling of undetected targets allows for newly discovered targets to have been born at earlier time steps [20]. Therefore, the trajectory PMBM filters give a higher effective birth rate than general TOMHT. The modeling of undetected targets was incorporated into TOMHT in [37]. In comparison, in the trajectory PMBM filters the hypotheses are purely data-to-data assignments and they are more efficiently represented using Bernoulli RFSs with probabilistic target existence. More importantly, in the PMBM trajectory filters the estimates of the set of trajectories can be directly extracted from the multitarget densities in addition to the target current states.

C. Contributions and Organization

This paper is an extension of [34]. In this paper, we present the trajectory PMBM and the trajectory MBM filter with multiscan data association. The main novelties of the proposed algorithms, compared to previous work based on sets of trajectories [27], [31], [38], [39], are that they consider the multiscan data association problem. The main novelties of the proposed algorithms, compared to TOMHT, are that they produce full trajectory estimates, i.e., smoothed estimates, upon receipt of each new set of measurements, and that the filters based on sets of trajectories model the targets that remain to be detected and the target death subsequent to the final detection.

The contributions can be summarized as follows:

1) We present the filtering recursions for the trajectory MBM filter and the trajectory MBM_{01} filter using a multi-Bernoulli birth model. Two variants are considered for each filter: the set of current trajectories and the set of all trajectories.
2) We show that the ideas from the efficient TOMHT in [17] can be utilized in trajectory filters based on PMBM, MBM, and MBM_{01} conjugate priors, resulting in so-called multiscan trajectory filters.
3) We explain how to efficiently perform fixed-lag smoothing to extract smoothed trajectory estimates for the presented algorithms.
4) We evaluate the performance of the presented algorithms in a simulation study, in terms of target state/trajectory estimation error and computational time.

The paper is organized as follows. In Section II, we introduce the modeling assumption and background on sets of trajectories. In Section III, we review the PMBM conjugate prior on the set of trajectories. In Section IV, we present the filtering recursion for trajectory MBM filter. In Section V, we present implementations of the multiscan trajectory filters. In Section VI, we present how to efficiently perform fixed-lag smoothing when extracting trajectory estimates. Simulation results are presented in Section VII, and conclusions are drawn in Section VIII.

II. MODELING

In this section, we first outline the modeling assumptions utilized in this work. Next, we give a brief introduction to RFSs of trajectories. Then, we introduce the generalized transition and measurement models in the framework of set of trajectories; the precise mathematical definitions can be found in [27]. The modeling is probabilistic, and the interested reader can find the necessary details about FISST, measure theory, probability generating functionals (PGFLs), and functional derivatives for sets of trajectories in Appendices A and D.

A. Modeling Assumptions

We assume that for each discrete time k (a nonnegative integer), a continuous time t_k is assigned, such that $t_k > t_{k'}$ for $k > k'$. In the traditional formulation for RFSs of targets, target states and measurements are represented in the form of finite sets [6]. A random single target state x_k is a random element of the state (Euclidean) space $\mathcal{X} = \mathbb{R}^n$, and a random measurement z_k is a random element of the measurement space $\mathcal{Z} = \mathbb{R}^m$, all at discrete time k. The random set of measurements obtained by a single sensor, including clutter and target measurements with unknown origin, at time

step k is denoted as $\mathbf{z}_k \in \mathcal{F}(\mathcal{Z})$, where $\mathcal{F}(\mathcal{Z})$ denotes the set of all the finite subsets of $\mathcal{Z}$.

We proceed by introducing two families of RFSs that will have prominent roles throughout the paper: the Poisson RFS [6, Sec. 4.3.1] and the Bernoulli RFS [6, Sec. 4.3.3]. A Poisson RFS $\boldsymbol{\Psi}$ has multi-object density distribution

$$f^{\text{ppp}}(\boldsymbol{\Psi}) = e^{-\int \lambda(\Psi) d\Psi} \prod_{\Psi \in \boldsymbol{\Psi}} \lambda(\Psi), \tag{1}$$

where $\lambda(\cdot)$ is the intensity function and the number of objects is Poisson distributed. An RFS $\boldsymbol{\Psi}$ is a Bernoulli RFS if $|\boldsymbol{\Psi}| \leq 1$, and a Bernoulli RFS has multi-object density distribution

$$f^{\text{ber}}(\boldsymbol{\Psi}) = \begin{cases} 1-r, & \boldsymbol{\Psi} = \emptyset, \\ r f(\Psi), & \boldsymbol{\Psi} = \{\Psi\}, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $f(\cdot)$ is a single object probability density and r is the probability of existence. A multi-Bernoulli RFS is the union of a finite number of independent Bernoulli RFSs.

In previous works [27], [31], [38], [39], two different birth models have been used. In this paper, we present multiscan trajectory filter implementations for both birth models: the Poisson birth model defined in Assumption 1 and the multi-Bernoulli birth model defined in Assumption 2. The standard point target measurement model is defined in Assumption 3.

Assumption 1. *The multitarget state evolves according to the following standard dynamic process with a Poisson birth model:*

1) *New targets appear in the surveillance area independently of any existing targets. Targets arrive at each time step according to a Poisson RFS with birth intensity $\lambda_k^{\text{b}}(x_k)$ defined on the target state space $\mathcal{X}$.*
2) *Given a target with state x_k, the target survives with a probability $P^{\text{S}}(x_k)$ and moves with a Markov state transition density $\pi(x_{k+1}|x_k)$ defined on the target state space $\mathcal{X}$. The state transition density is the density of the target state at time step $k+1$, given that the target had state x_k at time step k.*

Assumption 2. *The multitarget state evolves according to the following modified dynamic process with a multi-Bernoulli birth model:*

1) *New targets appear in the surveillance area independently of any existing targets. Targets arrive at time step k according to a multi-Bernoulli RFS, which has n_k^{b} Bernoulli components. The lth Bernoulli component has existence probability $r_k^{\text{b},l}$ and state density $f_k^{\text{b},l}(x_k)$ defined on the target state space $\mathcal{X}$.*
2) *Same as Assumption 1, point 2.*

Assumption 3. *The multitarget measurement process is as follows:*

1) *Each target may give rise to at most one measurement, and each measurement is the result of at most one target. The probability of detection of a target with state x_k is $P^{\text{D}}(x_k)$, and the single measurement density is $f(z_k|x_k)$ from the target space $\mathcal{X}$ to the measurement space $\mathcal{Z}$, which is the probability density of the measurement z_k, given that there is a target with state x_k in the scene.*
2) *Clutter measurements arrive according to a Poisson RFS with intensity $\lambda^{\text{FA}}(z_k)$ defined on the measurement space $\mathcal{Z}$, independently of targets and target-oriented measurements.*

B. Random Finite Sets of Trajectories

In this section, we first explain how the single trajectory state and its density are defined. Then, we briefly introduce some basic types of RFSs of trajectories.

1) Trajectory State: We use the trajectory state model presented in [27] and [30], in which the trajectory state is a tuple

$$X = (\beta, \varepsilon, x_{\beta:\varepsilon}), \tag{3}$$

where β is the discrete time of the trajectory birth, i.e., the time the trajectory begins; ε is the discrete time of the trajectory's end time. If k is the current time, $\varepsilon = k$ means that the trajectory is alive; $x_{\beta:\varepsilon}$ is, given β and ε, the (finite) sequence of states

$$x_{\beta:\varepsilon} = (x_\beta, x_{\beta+1}, \ldots, x_{\varepsilon-1}, x_\varepsilon), \tag{4}$$

where $x_\kappa \in \mathcal{X}$ for all $\kappa \in \{\beta, \ldots, \varepsilon\}$. This gives a trajectory of length $l = \varepsilon - \beta + 1$ time steps.

The single trajectory state can be considered a hybrid state consisting of discrete states β and ε representing the start and end time indices, and a continuous state $x_{\beta:\varepsilon}$ that evolves according to a stochastic model dependent on the discrete states.[2] The trajectory state space at time step k is [27]

$$\mathcal{T}_k = \uplus_{(\beta,\varepsilon)\in I_k} \{\beta\} \times \{\varepsilon\} \times \mathcal{X}^{\varepsilon-\beta+1}, \tag{5}$$

where $\uplus$ denotes the union of (possibly empty) sets that are mutually disjoint, $I_k = \{(\beta, \varepsilon) : 0 \leq \beta \leq \varepsilon \leq k\}$ is the set of all possible start and end times of trajectories up to time step k, and $\mathcal{X}^l$ denotes l Cartesian products of $\mathcal{X}$, i.e., the Cartesian products of spaces of different sizes. A trajectory state density $p(\cdot)$ of X factorizes as follows:

$$p(X) = p(x_{\beta:\varepsilon}|\beta, \varepsilon) P(\beta, \varepsilon), \tag{6}$$

where, if $\varepsilon < \beta$, then $P(\beta, \varepsilon)$ is zero. Integration for single trajectory densities is performed as follows [27]:

[2] We remark that the use of such a hybrid state, i.e., a combination of one (or more) discrete state and one (or more) continuous state, is not uncommon in MTT: a typical example is the interacting multiple model [40], in which the identification of multiple models, which can be of different dimensionality [41], is governed by a discrete stochastic process.

$$\int p(X)dX = \sum_{(\beta,\varepsilon)\in I_k} \left[\int \cdots \int p(x_{\beta:\varepsilon}|\beta,\varepsilon)dx_\beta \cdots dx_\varepsilon \right] P(\beta,\varepsilon). \tag{7}$$

2) Sets of Trajectories: A set of trajectories is denoted as $\mathbf{X}_k \in \mathcal{F}(\mathcal{T}_k)$, where $\mathcal{F}(\mathcal{T}_k)$ is the set of all the finite subsets of $\mathcal{T}_k$. Let $g(\mathbf{X}_k)$ be a real-valued function on a set of trajectories, then the set integral is

$$\int g(\mathbf{X}_k)\delta\mathbf{X}_k \triangleq g(\emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int g\left(\left\{X_k^1, \ldots, X_k^n\right\}\right) dX_k^1 \cdots dX_k^n. \tag{8}$$

A trajectory Poisson RFS has (multitrajectory) density of the form (1), where the trajectory Poisson RFS intensity $\lambda(\cdot)$ is defined on the trajectory state space $\mathcal{T}_k$; i.e., realizations of the Poisson RFS are trajectories with a birth time, a time of the most recent state, and a state sequence [38]. A trajectory Bernoulli RFS has density of the form (2), where $f(\cdot)$ is a single trajectory density (6). Trajectory multi-Bernoulli RFS and trajectory MBM RFS are both defined analogously to target multi-Bernoulli RFS and target MBM RFS [27]: a trajectory multi-Bernoulli is the disjoint union of a multiple trajectory Bernoulli RFS; trajectory MBM RFS is an RFS whose density is a mixture of trajectory multi-Bernoulli densities.

C. Transition Models for Sets of Trajectories

In the standard multitarget dynamic model with Poisson birth (see Assumption 1), target birth at time step k is modeled by a Poisson RFS, with intensity

$$\lambda_k^{\mathrm{B}}(X) = \lambda_k^{\mathrm{B},x}(x_{\beta:\varepsilon}|\beta,\varepsilon)\Delta_k(\varepsilon)\Delta_k(\beta), \tag{9a}$$

$$\lambda_k^{\mathrm{B},x}(x_{k:k}|k,k) = \lambda_k^{\mathrm{b}}(x_k), \tag{9b}$$

where $\Delta(\cdot)$ denotes the Kronecker delta function. In the modified multitarget dynamic model with multi-Bernoulli birth (see Assumption 2), target birth at time step k is modeled by a multi-Bernoulli RFS, with the trajectory state density in the lth Bernoulli component

$$f_k^{\mathrm{B},l}(X) = f_k^{\mathrm{B},l,x}(x_{\beta:\varepsilon}|\beta,\varepsilon)\Delta_k(\varepsilon)\Delta_k(\beta), \tag{10a}$$

$$f_k^{\mathrm{B},l,x}(x_{k:k}|k,k) = f_k^{\mathrm{b},l}(x_k), \tag{10b}$$

and the existence probability $r_k^{\mathrm{b},l}$.

We focus on two different MTT problem formulations: the set of current trajectories, where the objective is to estimate the trajectories of targets that are still present in the surveillance area at the current time, and the set of all trajectories, where the objective is to estimate the trajectories of both the targets that are still present in the surveillance area at the current time and the targets that once were in (but have since left) the surveillance area at some previous time. The probability of survival as a function on trajectories at time step k is defined as

$$P_k^{\mathrm{S}}(X) = P^{\mathrm{S}}(x_\varepsilon)\Delta_k(\varepsilon). \tag{11}$$

The transition density for the trajectories depends on the problem formulation.

1) Transition Model for the Set of Current Trajectories: The Bernoulli RFS transition density for a single potential target without birth is

$$f_{k|k-1}^{\mathrm{c}}(\mathbf{X}|\mathbf{X}') = \begin{cases} 1, & \mathbf{X}' = \emptyset, \mathbf{X} = \emptyset, \\ 1 - P_{k-1}^{\mathrm{S}}(X'), & \mathbf{X}' = \{X'\}, \mathbf{X} = \emptyset, \\ P_{k-1}^{\mathrm{S}}(X')\pi^{\mathrm{c}}(X|X'), & \mathbf{X}' = \{X'\}, \mathbf{X} = \{X\}, \\ 0, & \text{otherwise}, \end{cases} \tag{12a}$$

$$\pi^{\mathrm{c}}(X|X') = \pi^{\mathrm{c},x}(x_{\beta:\varepsilon}|\beta,\varepsilon,X')\Delta_{\varepsilon'+1}(\varepsilon)\Delta_{\beta'}(\beta), \tag{12b}$$

$$\pi^{\mathrm{c},x}(x_{\beta:\varepsilon}|\beta,\varepsilon,X') = \pi^{x}(x_\varepsilon|x'_{\varepsilon'})\delta_{x'_{\beta':\varepsilon'}}(x_{\beta:\varepsilon-1}), \tag{12c}$$

where $\delta(\cdot)$ denotes Dirac delta function and X' denotes the single trajectory state at time step $k-1$. In this model, $P^{\mathrm{S}}(\cdot)$ is used as follows. If the target disappears, or "dies," then the entire trajectory will no longer be a member of the set of current trajectories. If the trajectory survives, then the trajectory is extended by one time step.

2) Transition Model for the Set of All Trajectories: The Bernoulli RFS transition density for a single potential target without birth is

$$f_{k|k-1}^{\mathrm{a}}(\mathbf{X}|\mathbf{X}') = \begin{cases} 1, & \mathbf{X}' = \emptyset, \mathbf{X} = \emptyset, \\ \pi^{\mathrm{a}}(X|X'), & \mathbf{X}' = \{X'\}, \mathbf{X} = \{X\}, \\ 0, & \text{otherwise}, \end{cases} \tag{13a}$$

$$\pi^{\mathrm{a}}(X|X') = \pi^{\mathrm{a},x}(x_{\beta:\varepsilon}|\beta,\varepsilon,X')\pi^{\varepsilon}(\varepsilon|\beta,X')\Delta_{\beta'}(\beta), \tag{13b}$$

$$\pi^{\varepsilon}(\varepsilon|\beta,X') = \begin{cases} 1, & \varepsilon = \varepsilon' < k-1, \\ 1 - P_{k-1}^{\mathrm{S}}(X'), & \varepsilon = \varepsilon' = k-1, \\ P_{k-1}^{\mathrm{S}}(X'), & \varepsilon = \varepsilon' + 1 = k, \\ 0, & \text{otherwise}, \end{cases} \tag{13c}$$

$$\pi^{\mathrm{a},x}(x_{\beta:\varepsilon}|\beta,\varepsilon,X') = \begin{cases} \delta_{x'_{\beta':\varepsilon'}}(x_{\beta:\varepsilon}), & \varepsilon = \varepsilon', \\ \pi^{x}(x_\varepsilon|x'_{\varepsilon'})\delta_{x'_{\beta':\varepsilon'}}(x_{\beta:\varepsilon-1}), & \varepsilon = \varepsilon' + 1. \end{cases} \tag{13d}$$

In this model, the interpretation of the probability of survival is that it governs whether the trajectory ends or it is extended by one more time step. However, importantly, regardless of whether or not the trajectory ends,

the trajectory remains in the set of all trajectories with probability one.

The complete transition model for sets of trajectories is analogous to the complete transition model for sets of targets, by using sets of trajectories and the corresponding Bernoulli transition density for each problem formulation. Given the set $\mathbf{X}_{k-1} = \{X_{k-1}^1, \ldots, X_{k-1}^n\}$ of trajectories at time step $k-1$, the set $\mathbf{X}_k$ of trajectories at time step k is $\mathbf{X}_k = \mathbf{X}_k^{\mathrm{b}} \uplus \mathbf{X}_k^1 \uplus \cdots \uplus \mathbf{X}_k^n$, where $\mathbf{X}_k^{\mathrm{b}}, \mathbf{X}_k^1, ..., \mathbf{X}_k^n$ are independent sets, $\mathbf{X}_k^{\mathrm{b}}$ is the set of newborn trajectories, and $\mathbf{X}_k^i$ is the set of trajectories resulting from X_{k-1}^i. Using the convolution formula for multi-object densities [6, eq. (4.17)], the resulting multitrajectory density $f(\cdot|\cdot)$ of $\mathbf{X}_k$ given $\mathbf{X}_{k-1}$ can be written as

$$f(\mathbf{X}_k|\mathbf{X}_{k-1}) = \sum_{\mathbf{X}_k^{\mathrm{b}} \uplus \mathbf{X}_k^1 \uplus \cdots \uplus \mathbf{X}_k^n = \mathbf{X}_k} f_k^{\mathrm{birth}}(\mathbf{X}_k^{\mathrm{b}}) \times \prod_{i=1}^{n} f_{k|k-1}^{\mathrm{persist}}(\mathbf{X}_k^i|\{X_{k-1}^i\}), \quad (14)$$

where $f_k^{\mathrm{birth}}(\cdot)$ is either a trajectory Poisson RFS or a trajectory multi-Bernoulli RFS, and $f_{k|k-1}^{\mathrm{persist}}(\cdot|\cdot)$ is a Bernoulli transition density for a single potential target without birth, with the form $f_{k|k-1}^{\mathrm{a}}(\cdot|\cdot)$ or $f_{k|k-1}^{\mathrm{c}}(\cdot|\cdot)$.

D. Single Trajectory Measurement Model

According to the point target measurement model in Assumption 3, the multi-object density of a target-generated measurement at time step k given a set of trajectories with 0 or 1 element is Bernoulli, with the form

$$\varphi_k(\mathbf{w}_k|\mathbf{X}) = \begin{cases} 1, & \mathbf{X} = \emptyset, \mathbf{w}_k = \emptyset, \\ 1 - P_k^{\mathrm{D}}(X), & \mathbf{X} = \{X\}, \mathbf{w}_k = \emptyset, \\ P_k^{\mathrm{D}}(X)\varphi(z_k|X), & \mathbf{X} = \{X\}, \mathbf{w}_k = \{z_k\}, \\ 0, & \text{otherwise}, \end{cases} \quad (15a)$$

$$P_k^{\mathrm{D}}(X) = P^{\mathrm{D}}(x_\varepsilon)\Delta_k(\varepsilon), \quad (15b)$$

$$\varphi(z|X) = f(z|x_\varepsilon). \quad (15c)$$

Note that trajectories that do not exist at the current time cannot be detected. The complete measurement model for sets of trajectories is similar to the measurement model for sets of targets by using the proper probability of detection and single measurement density for trajectories [27]. Given the set $\mathbf{X}_k = \{X_k^1, \ldots, X_k^n\}$ of trajectories at time step k, the set $\mathbf{z}_k$ of measurements at time step k is $\mathbf{z}_k = \mathbf{w}_k^{\mathrm{c}} \uplus \mathbf{w}_k^1 \uplus \cdots \uplus \mathbf{w}_k^n$, where $\mathbf{w}_k^{\mathrm{c}}$, $\mathbf{w}_k^1$, ..., $\mathbf{w}_k^n$ are independent sets, $\mathbf{w}_k^{\mathrm{c}}$ is the set of clutter measurements, and $\mathbf{w}_k^i$ is the set of measurements produced by trajectory i. The resulting measurement set density $f(\cdot|\cdot)$ of $\mathbf{z}_k$ given $\mathbf{X}_k$ can be written as

$$f(\mathbf{z}_k|\mathbf{X}_k) = \sum_{\mathbf{w}_k^{\mathrm{c}} \uplus \mathbf{w}_k^1 \uplus \cdots \uplus \mathbf{w}_k^n = \mathbf{z}_k} f_k^{\mathrm{ppp}}(\mathbf{w}_k^{\mathrm{c}}) \prod_i^n \varphi_k(\mathbf{w}_k^i|\{X_k^i\}). \quad (16)$$

III. TRAJECTORY PMBM FILTER

The PMBM conjugate prior was developed for point targets in [22] and for extended targets in [42], and it was further generalized to trajectories in [31] and [43]. Given the sequence of measurements up to time step k' and Assumptions 1 and 3, the density of the set of trajectories at time step $k \in \{k', k'+1\}$ is given by the PMBM density of the form

$$f_{k|k'}(\mathbf{X}_k) = \sum_{\mathbf{X}_k^{\mathrm{u}} \uplus \mathbf{X}_k^{\mathrm{d}} = \mathbf{X}} f_{k|k'}^{\mathrm{ppp}}(\mathbf{X}_k^{\mathrm{u}}) \sum_{a \in \mathcal{A}_{k|k'}} w_{k|k'}^a f_{k|k'}^a(\mathbf{X}_k^{\mathrm{d}}), \quad (17a)$$

$$f_{k|k'}^{\mathrm{ppp}}(\mathbf{X}_k^{\mathrm{u}}) = e^{-\int \lambda_{k|k'}^{\mathrm{u}}(X)dX} \prod_{X \in \mathbf{X}_k^{\mathrm{u}}} \lambda_{k|k'}^{\mathrm{u}}(X), \quad (17b)$$

$$f_{k|k'}^a(\mathbf{X}_k^{\mathrm{d}}) = \sum_{\uplus_{i \in \mathbb{T}_{k|k'}} \mathbf{X}_k^i = \mathbf{X}_k^{\mathrm{d}}} \prod_{i \in \mathbb{T}_{k|k'}} f_{k|k'}^{i,a^i}(\mathbf{X}_k^i), \quad (17c)$$

where the RFS of trajectories $\mathbf{X}_k$ is an independent union of a Poisson RFS $\mathbf{X}_k^{\mathrm{u}}$ with intensity $\lambda_{k|k'}^{\mathrm{u}}$ and an MBM RFS $\mathbf{X}_k^{\mathrm{d}}$ with Bernoulli parameters $r_{k|k'}^{i,a^i}$ and $f_{k|k'}^{i,a^i}(\cdot)$, cf. (2), and $\mathcal{A}_{k|k'}$ is the set of all global hypotheses, which will be explained in the next section. A trajectory PMBM RFS can be defined by the parameters of the density,

$$\lambda_{k|k'}^{\mathrm{u}}, \mathcal{A}_{k|k'}, \left\{\Theta_{k|k'}^a\right\}_{a \in \mathcal{A}_{k|k'}}, \quad (18a)$$

$$\Theta_{k|k'}^a = \left\{\left(w_{k|k'}^{i,a^i}, r_{k|k'}^{i,a^i}, f_{k|k'}^{i,a^i}\right)\right\}_{i \in \mathbb{T}}. \quad (18b)$$

A. Structure of the Trajectory PMBM Filter

The structure of the trajectory PMBM (17) is in analogy to the structure of the target PMBM [22]. The Poisson RFS represents trajectories that are hypothesized to exist, but have never been detected; i.e., no measurement has been associated with them. In the track-oriented trajectory PMBM filter, a new track is initiated for each measurement received. In the MBM in (17), $\mathbb{T}_{k|k'} = \{1, \ldots, n_{k|k'}\}$ is a track table with $n_{k|k'}$ tracks, $a = (a^1, \ldots, a^{n_{k|k'}}) \in \mathcal{A}_{k|k'}$ is a possible global data association hypothesis, and for each global hypothesis a and for each track $i \in \mathbb{T}_{k|k'}$, a^i indicates which track hypothesis is used in the global hypothesis. For each track, there are $h_{k|k'}^i$ single trajectory hypotheses.[3] The weight of global hypothesis a is $w_{k|k'}^a \propto \prod_{i \in \mathbb{T}_{k|k'}} w_{k|k'}^{i,a^i}$, where $w_{k|k'}^{i,a^i}$ is the weight of single trajectory hypothesis a^i from track i.

Let m_k be the number of measurements at time step $k \in \{1, \ldots, \tau\}$ and $j \in \mathbb{M}_k = \{1, \ldots, m_k\}$ be an index to each measurement. Let $\mathcal{M}_k$ denote the set of all measurement indices up to and including time step k; the elements of $\mathcal{M}_k$, if not empty, are of the form (τ, j), where

[3] The "track" defined here is different from the convention used in MHT algorithms, where "track" is referred to as single trajectory hypothesis.

$j \in \{1, \ldots, m_\tau\}$ is an index of a measurement at time step $\tau \leq k$. Further, let $\mathcal{M}^k(i, a^i)$ denote the history of measurements that are hypothesized to belong to hypothesis a^i from track i at time step k. Under the standard point target measurement model assumption (see Assumption 3), there can be at maximum one measurement corresponding to the same time step in $\mathcal{M}^k(i, a^i)$.

For a global hypothesis to be correct, we have the following constraints. Each global hypothesis should explain the association of each measurement received so far. In addition, every measurement should be associated with one and only one track in each global hypothesis. In other words, the single trajectory hypotheses included in a given global hypothesis cannot have any shared measurement. Under these constraints, the set of global hypotheses at time step k can be expressed as

$$\mathcal{A}_{k|k'} = \left\{ a = (a^1, \ldots, a^{n_{k|k'}}) \middle| \bigcup_{i \in \mathbb{T}_{k|k'}} \mathcal{M}^k(i, a^i) = \mathcal{M}_k, \right. \\ \left. \mathcal{M}^k(i, a^i) \cap \mathcal{M}^k(j, a^j) = \emptyset \; \forall i \neq j, i, j \in \mathbb{T}_{k|k'} \right\}. \tag{19}$$

B. PMBM Filtering Recursion

The form of the PMBM conjugate prior on the sets of trajectories is preserved through prediction and update. The two different trajectory PMBM filters based on the two different transition models for sets of trajectories are both track-oriented. For each track, there is a hypothesis tree, where each hypothesis corresponds to different data association sequences for the track. The prediction step preserves the number of tracks and the number of hypotheses. By using a Poisson RFS birth model, the density of newborn trajectories $\lambda_k^{\mathrm{B}}(X_k)$ can be easily incorporated into the predicted density of Poisson distributed trajectories $\lambda_{k|k-1}^{\mathrm{u}}(X_k)$ that have never been detected. The two different trajectory PMBM filters have different prediction steps; the difference is that whether dead trajectories are still maintained in the set of trajectories. In the update step, a potential new track is initiated for each measurement, and additional hypotheses are created due to data association. The two different trajectory PMBM filters have the same update step. Explicit expressions for how the PMBM parameters (18) are predicted and updated, using the two different problem formulations, can be found in [31]; they are omitted here.

IV. TRAJECTORY MBM FILTER

It is shown in [23] that the MBM RFS of targets is a multitarget conjugate prior if the birth model is a multi-Bernoulli RFS, as in Assumption 2. In this section, we extend this result to RFS of trajectories. Given the sequence of measurements up to time step k' and Assumptions 2 and 3, the density of the set of trajectories at time step $k \in \{k', k'+1\}$ is given by the MBM of the form

$$f_{k|k'}(\mathbf{X}_k) = \sum_{a \in \mathcal{A}_{k|k'}} w_{k|k'}^a \sum_{\uplus_{i \in \mathbb{T}_{k|k'}} \mathbf{X}_k^i = \mathbf{X}_k} \prod_{i \in \mathbb{T}_{k|k'}} f_{k|k'}^{i,a^i}(\mathbf{X}_k^i), \tag{20}$$

where the MBM RFS $\mathbf{X}_k$ has Bernoulli parameters $r_{k|k'}^{i,a^i}$ and $f_{k|k'}^{i,a^i}(\cdot)$, cf. (2). A trajectory MBM RFS can be defined by the parameters of the density

$$\mathcal{A}_{k|k'}, \left\{ \Theta_{k|k'}^a \right\}_{a \in \mathcal{A}_{k|k'}}, \tag{21a}$$

$$\Theta_{k|k'}^a = \left\{ \left(w_{k|k'}^{i,a^i}, r_{k|k'}^{i,a^i}, f_{k|k'}^{i,a^i} \right) \right\}_{i \in \mathbb{T}}. \tag{21b}$$

A. Structure of the Trajectory MBM Filter

The structure of the trajectory MBM is similar to the MBM maintained in the trajectory PMBM. The difference lies in how tracks (i.e., Bernoulli components) are initiated. In the trajectory PMBM filter, a new track is initiated for each measurement, whereas in the trajectory MBM filter, a new track is initiated for each Bernoulli component in the multi-Bernoulli birth model; i.e., MBM hypotheses explicitly enumerate potential targets that remain to be detected. Both the trajectory PMBM filter and the trajectory MBM filter can explicitly represent trajectories that remain to be detected. In the PMBM representation, these trajectories are efficiently represented through the trajectory Poisson intensity $\lambda_{k|k'}^{\mathrm{u}}(\cdot)$, whereas in the MBM representation, they are split across many single trajectory hypotheses (trajectory Bernoulli RFSs) with empty measurement association history, i.e., $\mathcal{M}^k(i, a^i) = \emptyset$.

In each global hypothesis $a \in \mathcal{A}_{k|k}$, each measurement, at each time step, is associated with at most one track, and each track is associated with at most one measurement. Measurements that are not associated with any tracks in a global hypothesis are considered to be clutter under this global hypothesis. Tracks that are not associated with any measurements in a global hypothesis are considered to be misdetected under this global hypothesis. Under these constraints, the set of global hypotheses at time step k can be expressed as

$$\mathcal{A}_{k|k'} = \left\{ a = (a^1, \ldots, a^{n_{k|k'}}) \middle| \bigcup_{i \in \mathbb{T}_{k|k'}} \mathcal{M}^k(i, a^i) \subseteq \mathcal{M}_k, \right. \\ \left. \mathcal{M}^k(i, a^i) \cap \mathcal{M}^k(j, a^j) = \emptyset \; \forall i \neq j, i, j \in \mathbb{T}_{k|k'} \right\}. \tag{22}$$

Compared to (19), here $\mathcal{M}_k \backslash \bigcup_{i \in \mathbb{T}_{k|k'}} \mathcal{M}^k(i, a^i)$ consists of indices of measurements received so far that are clutter under global hypothesis $a \in \mathcal{A}_{k|k'}$. This is an important difference from the trajectory PMBM filter, in which the question whether a measurement corresponds to clutter, or to the initialization of a new target trajectory, is

captured by the existence probability of the created trajectory Bernoulli RFS.

In the rest of the section, we present the prediction and update steps for recursively computing (20) for the MBM parameterization. Similar to the trajectory PMBM filter, the two different trajectory MBM filters, based on the set of current trajectories formulation and the set of all trajectories formulation, have the same update step. For compactness, we denote the inner product of two functions $h(\cdot)$ and $g(\cdot)$ as $\langle h; g\rangle = \int h(x)g(x)dx$.

B. MBM Filtering Recursion

We first present the prediction steps, respectively, for the two different problem formulations, and then we present the update step.

1) Prediction Step for the Set of Current Trajectories: The prediction step is given in the following theorem.

Theorem 1. *Assume that the distribution from the previous time step $f_{k-1|k-1}(\mathbf{X}_{k-1})$ is given by (20), that the transition model is (12), and that the birth model is a trajectory multi-Bernoulli RFS with n_k^{b} Bernoulli components, each of which has density of the form (10). Then, the predicted distribution for the next step $f_{k|k-1}(\mathbf{X}_k)$ is given by (20), with $n_{k|k-1} = n_{k-1|k-1} + n_k^{\mathrm{b}}$. For tracks continuing from previous time ($i \in \{1, \dots, n_{k-1|k-1}\}$), the parameters of the MBM are*

$$h_{k|k-1}^{i} = h_{k-1|k-1}^{i}, \tag{23a}$$

$$w_{k|k-1}^{i,a^i} = w_{k-1|k-1}^{i,a^i} \ \forall a^i, \tag{23b}$$

$$r_{k|k-1}^{i,a^i} = r_{k-1|k-1}^{i,a^i} \left\langle f_{k-1|k-1}^{i,a^i}; P_{k-1}^{\mathrm{S}} \right\rangle \forall a^i, \tag{23c}$$

$$f_{k|k-1}^{i,a^i}(X) = \frac{\left\langle f_{k-1|k-1}^{i,a^i}; \pi^{\mathrm{c}} P_{k-1}^{\mathrm{S}} \right\rangle}{\left\langle f_{k-1|k-1}^{i,a^i}; P_{k-1}^{\mathrm{S}} \right\rangle} f_{k-1|k-1}^{i,a^i}(X) \ \forall a^i. \tag{23d}$$

For new tracks ($i \in \{n_{k-1|k-1} + l\}$, $l \in \{1, \dots, n_k^{\mathrm{b}}\}$), the parameters of the MBM are

$$h_{k|k-1}^{i} = 1, \tag{24a}$$

$$\mathcal{M}^{k-1}(i, 1) = \emptyset, \tag{24b}$$

$$w_{k|k-1}^{i,1} = 1, \tag{24c}$$

$$r_{k|k-1}^{i,1} = r_k^{\mathrm{b},l}, \tag{24d}$$

$$f_{k|k-1}^{i,1}(X) = f_k^{\mathrm{B},l}(X). \tag{24e}$$

2) Prediction Step for the Set of All Trajectories: The prediction step is given in the following theorem.

Theorem 2. *Assume that the distribution from the previous time step $f_{k-1|k-1}(\mathbf{X}_{k-1})$ is given by (20), that the transition model is (13), and that the birth model is a trajectory multi-Bernoulli RFS with n_k^{b} Bernoulli components, each of which has density given by (10). Then, the predicted distribution for the next step $f_{k|k-1}(\mathbf{X}_k)$ is given by (20), with $n_{k|k-1} = n_{k-1|k-1} + n_k^{\mathrm{b}}$. For tracks continuing from previous time ($i \in \{1, \dots, n_{k-1|k-1}\}$), the parameters of the MBM are*

$$h_{k|k-1}^{i} = h_{k-1|k-1}^{i}, \tag{25a}$$

$$w_{k|k-1}^{i,a^i} = w_{k-1|k-1}^{i,a^i} \ \forall a^i, \tag{25b}$$

$$r_{k|k-1}^{i,a^i} = r_{k-1|k-1}^{i,a^i} \ \forall a^i, \tag{25c}$$

$$f_{k|k-1}^{i,a^i}(X) = \left\langle f_{k-1|k-1}^{i,a^i}; \pi^{a} \right\rangle \forall a^i. \tag{25d}$$

For new tracks ($i \in \{n_{k-1|k-1} + l\}$, $l \in \{1, \dots, n_k^{\mathrm{b}}\}$), the parameters of the MBM are the same as (24).

3) Update Step: The update step is given in the following theorem.

Theorem 3. *Assume that the predicted distribution $f_{k|k-1}(\mathbf{X}_k)$ is given by (20), that the measurement model is (15), and that the measurement set at time step k is $\mathbf{z}_k = \{z_k^1, \dots, z_k^{m_k}\}$. Then, the updated distribution $f_{k|k}(\mathbf{X}_k)$ is given by (20), with $n_{k|k} = n_{k|k-1}$. For each track ($i \in \{1, \dots, n_{k|k}\}$), a hypothesis is included for each combination of a hypothesis from a previous time and either a misdetection or an update using one of the m_k new measurements, such that the number of hypotheses becomes $h_{k|k}^{i} = h_{k|k-1}^{i}(1+m_k)$. For misdetection hypotheses ($i \in \{1, \dots, n_{k|k}\}$, $a^i \in \{1, \dots, h_{k|k-1}\}$), the parameters of the MBM are*

$$\mathcal{M}^{k}(i, a^i) = \mathcal{M}^{k-1}(i, a^i), \tag{26a}$$

$$w_{k|k}^{i,a^i} = w_{k|k-1}^{i,a^i}\left(1 - r_{k|k-1}^{i,a^i}\left\langle f_{k|k-1}^{i,a^i}; P^{\mathrm{D}}\right\rangle\right), \tag{26b}$$

$$r_{k|k}^{i,a^i} = \frac{r_{k|k-1}^{i,a^i}\left\langle f_{k|k-1}^{i,a^i}; 1 - P^{\mathrm{D}}\right\rangle}{1 - r_{k|k-1}^{i,a^i}\left\langle f_{k|k-1}^{i,a^i}; P^{\mathrm{D}}\right\rangle}, \tag{26c}$$

$$f_{k|k}^{i,a^i}(X) = \frac{(1 - P_k^{\mathrm{D}}(X)) f_{k|k-1}^{i,a^i}(X)}{\left\langle f_{k|k-1}^{i,a^i}; 1 - P^{\mathrm{D}}\right\rangle}. \tag{26d}$$

For hypotheses updating tracks ($i \in \{1, \dots, n_{k|k}\}$, $a^i = \tilde{a}^i + h_{k|k-1}^{i} j$, $\tilde{a}^i \in \{1, \dots, h_{k|k-1}^{i}\}$, $j \in \{1, \dots, m_k\}$, i.e., the previous hypothesis $\tilde{a}^i$, updated with measurement z_k^j), the parameters are

$$\mathcal{M}^{k}(i, a^i) = \mathcal{M}^{k-1}(i, \tilde{a}^i) \cup \{(k, j)\}, \tag{27a}$$

$$w_{k|k}^{i,a^i} = \frac{w_{k|k-1}^{i,a^i} r_{k|k-1}^{i,\tilde{a}^i} \left\langle f_{k|k-1}^{i,\tilde{a}^i}; \varphi\left(z_k^j|\,\cdot\right) P^{\mathrm{D}}\right\rangle}{\lambda^{\mathrm{FA}}\left(z_k^j\right)}, \tag{27b}$$

$$r_{k|k}^{i,a^i} = 1, \tag{27c}$$

$$f_{k|k}^{i,a^i}(X) = \frac{\varphi\left(z_k^j|X\right) P_k^{\mathrm{D}}(X) f_{k|k-1}^{i,\tilde{a}^i}(X)}{\left\langle f_{k|k-1}^{i,\tilde{a}^i}; \varphi\left(z_k^j|\,\cdot\right) P_k^{\mathrm{D}}\right\rangle}. \tag{27d}$$

The derivation here incorporates hypotheses updating every prior hypothesis with every measurement; however, in practical implementations, gating can be used to reduce the computational burden by excluding hypotheses with negligible weights.

C. MBM_{01} Filtering Recursion

The trajectory MBM_{01} filter can be considered as a variant of the trajectory MBM filter, in which existence probabilities of Bernoulli components are either 0 or 1. The MBM_{01} filtering recursion can be obtained from the MBM filtering recursions by expanding the MBM into its MBM_{01} equivalent [23]. The filtering recursions for the trajectory MBM_{01} filter are given in Appendix C.

D. Discussion

All the trajectory filters presented above are track-oriented. For each Bernoulli component in the multi-Bernoulli birth density, a new track is initiated. Compared to the trajectory PMBM filter with Poisson RFS birth, tracks are created in the prediction step but not the update step of trajectory MBM/MBM_{01} filter. In the trajectory MBM/MBM_{01} filter for the set of all trajectories, the predictions (25d) and (66c) result in additional mixture component in Bernoulli densities $f^{i,a^i}_{k|k'}(X_k)$, which are of the form

$$p(X) = \sum_j w^j p^j(x_{\beta:\varepsilon}|\beta,\varepsilon)\Delta_{e^j}(\varepsilon)\Delta_{b^j}(\beta), \tag{28}$$

where each mixture component is characterized by a weight w^j, a distinct birth time b^j, a distinct most recent time e^j where $b^j \leq e^j$ for all j,[4] and a state sequence density $p^j(\cdot)$. This type of state density facilitates simple representations for the state sequence $x_{\beta:\varepsilon}$ (either the state of a trajectory that is still present or the state of a dead trajectory), conditioned on β and ε.

The prediction steps, given by Theorems 5 and 6, in the trajectory MBM_{01} filter, create more single trajectory hypotheses than the prediction steps, given by Theorems 1 and 2, in the trajectory MBM filter; this is a direct result of restricting the existence probability of Bernoulli components to either 0 or 1. The existence probability of trajectory Bernoulli RFS r has different meanings in the four different trajectory filters: in the trajectory MBM filter for the set of current trajectories, r is the probability that the trajectory exists at the current time and has not ended yet; in the trajectory MBM filter for the set of all trajectories, r represents the probability that the trajectory existed at any time before including the current time; in the trajectory MBM_{01} filter for the set of current trajectories, r indicates whether the trajectory exists at the current time and has not ended yet; in the trajectory MBM_{01} filter for the set of all trajectories, r indicates whether the trajectory existed at any time before and including the current time.

We remark that the labeled trajectory MBM and MBM_{01} filters, which are defined over the set of labeled trajectories, can be obtained by augmenting label to single target state x [27, Sec. IV-A]. This does not affect the filtering recursion or the information in the computed posterior, compared to MBM and MBM_{01}. Therefore, the corresponding multiscan implementations in Section V are analogous.

[4]Neither the birth time β nor the most recent time ε is deterministic.

V. IMPLEMENTATION OF MULTISCAN TRAJECTORY FILTERS

In this section, we present efficient multiscan implementations of the above trajectory filters.

A. Hypothesis Reduction

The hypothesis reduction techniques for the trajectory PMBM, MBM, and MBM_{01} are quite similar, so we first explain the general formulation and then highlight the differences. As a first step, we identify the most probable global hypothesis, from which estimates of trajectories are also typically extracted. Conditioning on the most likely global hypothesis, we make use of track-oriented N-scan pruning [5], a conventional hypothesis reduction technique used in TOMHT, to prune global hypotheses with negligible weights.

We note that hypothesis reduction is not complicated by the fact that we are working with symmetric (unlabeled) distributions. Specifically, in (20), the quantities stored are the weight of hypothesis a, i.e., $w^a_{k|k'}$, and the hypothesis-conditioned trajectory distributions $f^{i,a^i}_{k|k'}(\mathbf{X}^i_k)$ for each target. Symmetry is ensured by the sum over $\uplus_{i\in\mathbb{T}_{k|k'}}\mathbf{X}^i_k = \mathbf{X}_k$; this sum is implicit, and terms never need to be explicitly represented. Therefore, hypothesis reduction achieved by either setting $w^a_{k|k'} = 0$ for some subset of hypotheses (and renormalizing the weights of remaining hypotheses to sum to 1) or removing a subset of multi-Bernoulli components $f^{i,a^i}_{k|k'}(\mathbf{X}^i_k)$ for some hypotheses always results in valid symmetric distributions. Likewise, if the existence probability of a Bernoulli component is close to zero in all the considered global hypotheses, pruning is equivalent to setting this existence probability equal to zero, which does not affect the symmetry of the posterior.

Given the most likely global hypothesis a^* at current time step k, we trace the single trajectory hypotheses included in a^* back to their local hypotheses at time step $k-N$. The assumption behind the N-scan pruning method is that the data association ambiguity is resolved before scan $k-N$ [5]. In other words, global hypotheses that do not coincide with a^* up until and including time step $k-N+1$ are assumed to have negligible weights; these global hypotheses can then be pruned. In addition, tracks (local hypothesis trees) that, after pruning, have a single nonexistence local hypothesis, i.e., $r = 0$, can be pruned. In what follows, we show that the most likely global hypothesis a^* can be obtained as the solution of a multiframe assignment problem.

B. Data Association Modeling and Problem Formulation

As indicated in the previous section, the posterior global hypothesis probability $w^a_{k|k}$ is proportional to the product of the weights of different single trajectory hypotheses $w^{i,a^i}_{k|k}$, one from each track:

$$w^a_{k|k} \propto \prod_{i\in\mathbb{T}_{k|k}} w^{i,a^i}_{k|k}, \tag{29}$$

where the proportionality denotes that normalization is required to ensure that $\sum_{a\in\mathcal{A}_{k|k}} w^a_{k|k} = 1$. Omitting time indices and introducing the notation $c^a = -\log(w^a)$ and $c^{i,a^i} = -\log(w^{i,a^i})$ yields

$$c^a = \sum_{i\in\mathbb{T}} c^{i,a^i} + C, \tag{30}$$

where C is the logarithm of the normalization constant in (29). The most likely global hypothesis is the collection of single trajectory hypotheses that minimizes the total cost, i.e.,

$$a^* = \underset{(a^i)\in\mathcal{A}}{\arg\min} \sum_{i\in\mathbb{T}} c^{i,a^i}. \tag{31}$$

Let $\mathcal{H}^i$ denote the set of single trajectory hypotheses for the ith track, and let $\mathbb{M}_\tau$ denote the set of measurement indices at time step τ. Further, let $\rho^{i,a^i} \in \{0,1\}$ be a binary indicator variable, indicating whether single trajectory hypothesis a^i in the ith track is included in a global hypothesis or not, and let

$$\boldsymbol{\rho} = \{\rho^{i,a^i} \in \{0,1\} | a^i \in \mathcal{H}^i \ \forall i \in \mathbb{T}\} \tag{32}$$

be the set of all binary indicator variables. The minimization problem (31) can be further posed as a multiframe assignment problem by decomposing the constraint $(a^i) \in \mathcal{A}$ into a set of smaller constraints [17, Sec. III], in the form of

$$\underset{\boldsymbol{\rho}\in\bigcap_{\tau=0}^{k}\mathcal{P}^\tau}{\arg\min} \sum_{i\in\mathbb{T}} \sum_{a^i\in\mathcal{H}^i} c^{i,a^i}\rho^{i,a^i}, \tag{33}$$

with the constraint sets denoted as

$$\mathcal{P}^0 = \left\{ \boldsymbol{\rho} \,\middle|\, \sum_{a^i\in\mathcal{H}^i} \rho^{i,a^i} = 1, \ \forall i \in \mathbb{T} \right\}, \tag{34a}$$

$$\mathcal{P}^\tau = \left\{ \boldsymbol{\rho} \,\middle|\, \sum_{i\in\mathbb{T}} \sum_{\substack{a^i\in\mathcal{H}^i:\\ (\tau,j)\in\mathcal{M}(i,a^i)}} \rho^{i,a^i} \le 1, \ \forall j \in \mathbb{M}_\tau \right\}, \tag{34b}$$

where k is the current time step and $\tau = 1, \ldots, k$. The first constraint (34a) enforces that each global hypothesis should include one and only one single trajectory hypothesis from each track. The set of k constraints (34b) differs in the trajectory PMBM filter and the trajectory MBM/MBM$_{01}$ filter. In the trajectory PMBM filter, each measurement from each time should be associated with exactly one track, i.e., the $\le$ sign becomes an $=$ sign in (34b), whereas in the trajectory MBM/MBM$_{01}$ filter, each measurement from each time should be associated with at most one track, which explains the $\le$ sign.

C. Multiframe Assignment via Dual Decomposition

The multidimensional assignment problem (33) is NP-hard for two or more scans of measurements. An effective approach to solving this problem is Lagrangian relaxation; this technique has been widely used to solve the multiscan data association problem in TOMHT algorithms; see, e.g., [15] and [16]. In this work, we focus on the dual decomposition formulation [44], i.e., a special case of Lagrangian relaxation, whose competitive performance, compared to traditional approaches [15], [16], in solving the multiframe assignment problem has been demonstrated in [17].

1) Decomposition of the Lagrangian Dual: We follow similar implementation steps as in [17]. The original (primal) problem (33) is separated into k subproblems, one for each time step, and for each subproblem a binary variable is used. The subproblem solutions

$$\boldsymbol{\rho}_\tau = \{\rho^{i,a^i}_\tau \in \{0,1\} | a^i \in \mathcal{H}^i \ \forall i \in \mathbb{T}\} \tag{35}$$

must be equal for all τ; this is enforced through Lagrange multipliers that are incorporated into the subproblems acting as penalty weights. The τth subproblem can be written as [17]

$$\underset{\boldsymbol{\rho}_\tau\in\mathcal{P}^0\cap\mathcal{P}^\tau}{\arg\min} \sum_{i\in\mathbb{T}} \sum_{a^i\in\mathcal{H}^i} \left(\frac{c^{i,a^i}}{k} + \delta^{i,a^i}_\tau \right) \rho^{i,a^i}_\tau \triangleq \underset{\boldsymbol{\rho}_\tau\in\mathcal{P}^0\cap\mathcal{P}^\tau}{\arg\min} \mathcal{S}(\boldsymbol{\rho}_\tau, \boldsymbol{\delta}_\tau), \tag{36}$$

where the Lagrange multipliers used for the τth subproblem are denoted by

$$\boldsymbol{\delta}_\tau = \left\{ \delta^{i,a^i}_\tau \middle| a^i \in \mathcal{H}^i \ \forall i \in \mathbb{T} \right\}, \tag{37}$$

and the division by k in (36) comes from the fact that the summation of the objectives that each subproblem tries to minimize should be equal to the objective of the original problem. The Lagrange multipliers $\delta^{i,a^i}_\tau \in \mathbb{R}$ have the constraint that, for each single trajectory hypothesis, they must add up to zero over different subproblems [44]. Thus, the set of Lagrange multipliers has the form

$$\Lambda = \left\{ \boldsymbol{\delta}_\tau \,\middle|\, \sum_{\tau=1}^{k} \delta^{i,a^i}_\tau = 0, \ \forall a^i \in \mathcal{H}^i \ \forall i \in \mathbb{T} \right\}. \tag{38}$$

2) Subproblem Solving: After eliminating all the constraint sets except two, i.e., $\mathcal{P}^0$ and $\mathcal{P}^\tau$, we obtain a 2D assignment problem (36). The objective of the τth assignment problem (36) is to associate each measurement received at time step $\tau \le k$, i.e., $j \in \mathbb{M}_\tau$, with either an existing track or a new track[5] at the current time

[5] In the trajectory MBM/MBM$_{01}$, "dummy" tracks are created to represent clutter.

step k, i.e., $i \in \mathbb{T}_k$, such that the total assignment cost is minimized.

For a track that is created after time step τ, no measurement from time step τ should be assigned to it; therefore, the measurement-to-track assignment cost is infinity. For a track that existed before and up to time step τ, i.e., $i \in \mathbb{T}_\tau$, if measurement z_τ^j was not associated with this track, let the measurement-to-track assignment cost be infinity; if otherwise, let the cost first be the minimum cost of the single trajectory hypothesis in this track that was updated by z_τ^j [45, Ch. VII, eq. (7.24)], i.e.,

$$\min \sum_{\substack{a^i \in \mathcal{H}^i: \\ (\tau, j) \in \mathcal{M}(i, a^i)}} \left(\frac{c^{i,a^i}}{k} + \delta_\tau^{i,a^i} \right). \tag{39}$$

In order to keep the cost of a hypothesis that does not assign a measurement to a track the same for an existing track and a new track (trajectory PMBM filter) or clutter (trajectory MBM filter), the cost (39) should then have subtracted from it the minimum cost of hypotheses that this track is not updated by any of the measurements at time step τ, i.e.,

$$\min \sum_{\substack{a^i \in \mathcal{H}^i: \\ (\tau, j) \notin \mathcal{M}(i, a^i),\ \forall j \in \mathbb{M}_\tau}} \left(\frac{c^{i,a^i}}{k} + \delta_\tau^{i,a^i} \right). \tag{40}$$

Note that, in the context of Lagrangian relaxation, the costs of single trajectory hypotheses refer to the costs that are penalized by the Lagrange multipliers.

After solving the 2D assignment problem, we can obtain the associations for each measurement at time step τ. For tracks not being associated with any measurements at time step τ, if the track is created before and up to time step τ, i.e., $i \in \mathbb{T}_\tau$, the single trajectory hypothesis

$$\arg\min_{a^i} \sum_{\substack{a^i \in \mathcal{H}^i: \\ (\tau, j) \notin \mathcal{M}(i, a^i),\ \forall j \in \mathbb{M}_\tau}} \left(\frac{c^{i,a^i}}{k} + \delta_\tau^{i,a^i} \right) \tag{41}$$

is included in the most likely global hypothesis; if otherwise, i.e., $i \in \mathbb{T}_k \setminus \mathbb{T}_\tau$, we can choose the single trajectory hypothesis

$$\arg\min_{a^i} \sum_{a^i \in \mathcal{H}^i} \left(\frac{c^{i,a^i}}{k} + \delta_\tau^{i,a^i} \right) \tag{42}$$

to be included in the most likely global hypothesis.

3) Subgradient Updates: The objective of Lagrangian relaxation is to find the tightest lower bound of the summation of the cost of each subproblem (36). The dual problem can be expressed as [17]

$$\arg\max_{\{\delta_\tau\} \in \Lambda} \left(\sum_{\tau=1}^{k} \min_{\rho_\tau \in \mathcal{P}^0 \cap \mathcal{P}^\tau} \mathcal{S}(\boldsymbol{\rho}_\tau, \boldsymbol{\delta}_\tau) \right), \tag{43}$$

where the maximum can be found using subgradient methods [46]. The Lagrange multipliers $\{\boldsymbol{\delta}_\tau\}$ are updated using

$$\delta_\tau^{i,a^i} = \delta_\tau^{i,a^i} + \alpha_t \cdot g_\tau^{i,a^i}, \tag{44}$$

where g_τ^{i,a^i} is the projected subgradient that can be calculated as

$$g_\tau^{i,a^i} = \rho_\tau^{i,a^i} - \frac{1}{k} \sum_{\tau'=1}^{k} \rho_{\tau'}^{i,a^i}, \tag{45}$$

and α_t is the step size at iteration t. There are many rules to set the step size; see [44]. In this work, we choose to use the same setting as in [17], which has the form

$$\alpha_t = \frac{C_t^{\mathrm{BP}} - C_t^{\mathrm{D}}}{\|\{g_\tau\}\|^2}, \tag{46}$$

where C_t^{BP} is the best (minimum) feasible primal cost so far obtained, C_t^{D} is the dual cost calculated at iteration t from (43), and $\{g_\tau\}$ denotes the concatenation of all the projected subgradients g_τ^{i,a^i}. The optimal solution is assumed to be attained when the relative gap between the primal cost and the dual cost $(C_t^{\mathrm{BP}} - C_t^{\mathrm{D}})/C_t^{\mathrm{BP}}$ is less than a specified threshold, e.g., 0.01 [44].

Each subproblem solution will, in general, be infeasible with respect to the primal problem (33); nevertheless, subproblem solutions will usually be nearly feasible since large constraint violations were penalized [44]. Hence, feasible solutions $\boldsymbol{\rho}$ can be obtained by correcting the minor conflicting binary elements on which subproblem solutions $\boldsymbol{\rho}_\tau$ disagree. For tracks for which we have not yet selected which single trajectory hypothesis to be included in the most likely global hypothesis, we use the branch and bound technique [47] to reconstruct the best feasible solution at each iteration of the Lagrangian relaxation. Note that there are many other ways to recover a feasible primal solution from subproblem solutions; see [44].

D. Discussion

The objective of solving the multiframe assignment problem is to know which Bernoulli components are included in the multi-Bernoulli with the highest weight. Because the data association ambiguity is assumed to be resolved before time step $k - N$, obtaining the most likely global hypothesis at time step k, which explains the origin of each measurement from time step $k - N$ to current time step k, requires the solution of an $(N + 2)$-dimensional assignment problem [5].

The computational complexity of filters can be further reduced by limiting the number of single target/trajectory hypotheses; see [23] and [31]. As for the multiscan trajectory PMBM and MBM filters, pruning single trajectory hypotheses with small existence probabilities besides N-scan pruning might sometimes harm the solvability of the multiframe assignment problem,

since the problem is formulated using the measurement assignment information contained in single trajectory hypotheses. Instead, we can choose single trajectory hypotheses $a^i \in \mathcal{H}^i, \ \forall i \in \mathbb{T}$, with small Bernoulli existence probability r at current time step to be updated only by misdetection at next time step. Then, single trajectory hypotheses with several consecutive misdetections can be pruned using N-scan pruning. Also, to limit the number of mixture components in the trajectory Poisson RFS, components with negligible weights can be pruned.

VI. EFFICIENT FIXED-LAG SMOOTHING

Multitarget filters based on sets of trajectories are able to estimate the full state sequence instead of appending the sequence of estimates at each time step. This is possible since the posterior density contains full trajectory information. The posterior density over the set of trajectories can be computed either off-line by applying fixed-interval smoothing or recursively as new measurements arrive by performing smoothing while filtering. Examples of the latter case include the Gaussian mixture trajectory (cardinalized) probability hypothesis density filter proposed in [38] and [39] and the trajectory MBM_{01} filter proposed in [27] that use an accumulated state density representation [48], and the trajectory PMBM filter proposed in [31] that uses an information form [49], to represent the joint state density.

As time progresses, the lengths of the trajectories increase. Eventually, the length may be such that it is computationally beneficial to perform approximate smoothing while filtering. An L-scan implementation is proposed in [27] and [38] that propagates the joint density of the states of the last L time steps and independent densities for the previous states for each trajectory. Still, from the perspective of N-scan pruning, a lot of unnecessary calculations might be spent on obtaining the smoothed posterior density for each single trajectory hypothesis. More specifically, when the data association ambiguity is high (e.g., targets move in proximity), we might have hundreds or even thousands of single trajectory hypotheses, and at each time instance we only need to compute the posterior trajectory mean for those that are included in the most likely global hypothesis. However, note that the prediction and update of the hypotheses weights are the same as in the implementation using smoothing while filtering, e.g., [31].

We propose an efficient fixed-lag smoothing implementation of multiscan trajectory filters that solves the above-mentioned problem by combining the L-scan trajectory density approximation with N-scan pruning. After N-scan pruning, single trajectory hypotheses in the same track share the same measurement association history at all times up to time step $k-N$. Then, we can apply $(N+L)$-scan density approximation, such that all single trajectory hypotheses in the same track share the same posterior trajectory density up until time step $k-N-L$. It is therefore sufficient to perform fixed-lag smoothing for $N+L$ steps for the most likely global hypothesis, and then store the parameters of the smoothed target state densities at time step $k-N-L+1$ before proceeding. Following this approach, the extracted posterior trajectory mean from the most likely global hypothesis at time step $k+1$ consists of the newly computed smoothed estimates for the last $N+L$ steps and the prestored smoothed estimates at all times up to $k-N-L+1$.

VII. SIMULATIONS

In this section, we show simulation results that compare five different filters[6]:

1) multiscan trajectory PMBM filter[7];
2) multiscan trajectory MBM filter (footnote 7);
3) multiscan trajectory MBM_{01} filter (footnote 7);
4) fast implementation of the δ-GLMB filter using Gibbs sampling[8] [35];
5) fast implementation of the LMB filter using Gibbs sampling (footnote 8) [50].

For all the trajectory filters, we consider the set of all trajectories problem formulation.

A. Parameter Setup

A 2D Cartesian coordinate system is used to define measurement and target kinematic parameters. The kinematic target state is a vector of position and velocity $x_k = [p_{x,k}, v_{x,k}, p_{y,k}, v_{y,k}]^T$. A single measurement is a vector of position $z_k = [z_{x,k}, z_{y,k}]^T$. Targets follow a linear Gaussian constant velocity model $\pi_{k|k-1}(x_k|x_{k-1}) = \mathcal{N}(x_k; F_k x_{k-1}, Q_k)$, with parameters

$$F_k = I_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad Q_k = 0.01 I_2 \otimes \begin{bmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{bmatrix},$$

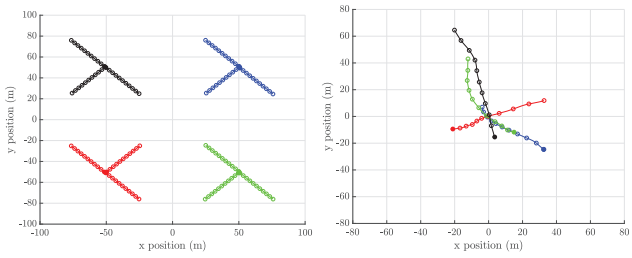
where $\otimes$ is the Kronecker product, I_m is an identity matrix of size $m \times m$, and $T = 1$. The linear Gaussian measurement likelihood model has density $f(z_k|x_k) = \mathcal{N}(z_k; H_k x_k, R_k)$, with parameters $H_k = I_2 \otimes [1, 0]$ and $R_k = I_2$.

The filters consider that there are no targets at time step 0. For multiscan trajectory filters, we use N-scan pruning ($N = 3$) to remove unlikely global

[6]The TOMHT implementation developed in [17] can be considered as a special case of the multiscan trajectory PMBM filter for sets of current trajectories where the trajectory estimates consist of target state estimates that are extracted from the marginal densities over the current set of targets. Therefore, we choose not to include the TOMHT implementation in [17] in the simulation results.

[7]MATLAB code of the multiscan trajectory PMBM, MBM, and MBM_{01} filters is available at https://github.com/yuhsuansia/Multi-scan-trajectory-PMBM-filter.

[8]We use the code that Prof. Ba-Ngu Vo and Prof. Ba-Tuong Vo share online: http://ba-tuong.vo-au.com/codes.html. The authors thank them for providing the code.



(a) Scenario 1 (b) Scenario 2

Fig. 1. True target trajectories for 81 time steps. In both scenarios, targets are born at times $\{1, 11, 21, 31\}$ and are dead at times $\{51, 61, 71, 81\}$. Targets' positions every six time steps are marked with a circle, and their initial positions with a filled circle. In Scenario 1, there are 12 targets born at four different locations. In Scenario 2, targets move in close proximity around the midpoint.

hypotheses. In addition to filtering, we also perform fixed-lag smoothing for the latest four steps. Both filtering and smoothing performances are analyzed. For the trajectory PMBM filter and the trajectory MBM filter, Bernoulli components with existence probability smaller than 10^{-3} are not updated by measurements (see Section V-D). For the trajectory PMBM filter, we remove mixture components in the trajectory Poisson RFS with weights smaller than 10^{-3}. For the δ-GLMB filter, the cap on the number of components $H^{\max} = 2000$. Ellipsoidal gating is used in all the compared filters; the gating size in probability is 0.999.

We consider two different scenarios with true trajectories shown in Fig. 1. In Scenario 1, targets are well spaced, and there is at most one target born at the same location per scan. In Scenario 2, for each trajectory, we initiate the midpoint from a Gaussian with mean $[0, 0, 0, 0]^T$ and covariance matrix I_4, and the rest of the trajectory is generated by running forward and backward dynamics. This scenario is challenging due to the fact that all the four targets move in close proximity around the midpoint. In the simulation, we consider constant target survival probability $P^{\mathrm{S}} = 0.99$, constant target detection probability $P^{\mathrm{D}} = 0.9$, and Poisson clutter uniform in the region of interest with rate $\lambda^{\mathrm{FA}} = 10$.

For the trajectory PMBM filter, the Poisson birth intensity has the form $\lambda_k^{\mathrm{b}}(x_k) = \sum_l 0.05\mathcal{N}(x; \bar{x}_k^{\mathrm{b},l}, P_k^{\mathrm{b},l})$. For the trajectory MBM filter, the trajectory MBM_{01} filter, the δ-GLMB filter, and the LMB filter, the lth Bernoulli component in the multi-Bernoulli birth has existence probability $r_k^{\mathrm{b},l} = 0.05$ and single target state density $\mathcal{N}(x; \bar{x}_k^{\mathrm{b},l}, P_k^{\mathrm{b},l})$. In Scenario 1, we set $\bar{x}_k^{\mathrm{b},1} = [50, 0, 50, 0]^T$, $\bar{x}_k^{\mathrm{b},2} = [50, 0, -50, 0]^T$, $\bar{x}_k^{\mathrm{b},3} = [-50, 0, 50, 0]^T$, $\bar{x}_k^{\mathrm{b},4} = [-50, 0, -50, 0]^T$, and $P_k^{\mathrm{b},l} = \mathrm{diag}([4, 1, 4, 1])$. In Scenario 2, we set $\bar{x}_k^{\mathrm{b},1} = [0, 0, 0, 0]^T$ and $P_k^{\mathrm{b},1} = \mathrm{diag}([100^2, 1, 100^2, 1])$, which covers the region of interest. It should be noted that the multi-Bernoulli and Poisson birth models have the same intensity (probability hypothesis density) [6, eq. (4.129)]. This implies that birth models are as close as possible in the sense of Kullback–Leibler divergence.

B. Performance Evaluation

For all the three multiscan trajectory filters, we estimate the full trajectories directly from the most likely global hypothesis. For the trajectory filters, we choose the most likely cardinality estimate $n^\star$ from the multi-Bernoulli of the most likely global hypothesis. We then report trajectory estimates from the $n^\star$ Bernoulli components with the highest existence probabilities. Given a Bernoulli state density (28), an estimate of the trajectory is obtained by selecting the most probable mixture component $j^* = \arg\max_j w_{k|k'}^j$ and reporting its mean value [31]. For the δ-GLMB filter and the LMB filter, we first obtain the maximum a posteriori estimate of the cardinality. We then find the global hypothesis with this cardinality with highest weight and report the mean of the targets in this hypothesis [28]. Trajectories are formed by connecting target estimates with the same label.

To evaluate the filtering performance, we used the generalized optimal subpattern assignment (GOSPA) metric [51], which can be decomposed into localization cost, missed target cost, and false target cost. The GOSPA metric is applied to the set of current target states at each time step. To evaluate the tracking performance, the trajectory metric in [52] based on linear programming (LP) was used, which can be decomposed into localization cost, missed target cost, false target cost, and track switch cost.

C. Results

We perform 100 Monte Carlo runs and obtain the average root-mean-square (RMS) GOSPA error (order $p = 2$, location error cutoff $c = 10$, and $\alpha = 2$), the average RMS trajectory estimation error (order $p = 2$, location error cutoff $c = 10$, switch cost $\gamma = 2$), and the average running time, summed over 81 time steps. We apply the trajectory metric [52] at each time step k, and normalize it by $\sqrt{k}$. This normalization allows a comparison of how the RMS metric evolves over time in the scenario, as opposed to only computing the metric at the final time step.

The comparison of different filters by the RMS GOSPA error and by the average running time[9] is shown in Table I for Scenario 1 and in Fig. 2 for Scenario 2. We can see that the trajectory PMBM filter arguably has the best performance in terms of target state estimation error and computational complexity, especially in Scenario 2 with coalescence. By comparing the execution time of trajectory filters with and without fixed-lag smoothing

[9]MATLAB implementations on a desktop with 3.0 GHz Intel Core i5 processor.

TABLE I
Simulation Results for Scenario 1: RMS GOSPA/LP Trajectory Metric Errors and Average Running Time (s)

Algorithm Fixed-lag smoothing	Trajectory Without	PMBM With	Trajectory Without	MBM With	Trajectory Without	MBM_{01} With	δ-GLMB (Gibbs) Without	LMB (Gibbs) Without
GOSPA	150.02	150.02	148.98	148.98	149.33	149.33	151.94	155.21
GOSPA—localization	120.73	120.73	120.76	120.76	120.74	120.74	120.82	120.93
GOSPA—missed	68.10	68.10	66.24	66.24	67.54	67.54	63.71	57.72
GOSPA—false	65.65	65.65	64.04	64.04	63.70	63.70	68.40	77.19
LP trajectory metric	141.91	128.25	141.02	127.15	141.04	127.16	167.50	168.85
LP—localization	123.23	101.72	123.40	101.87	123.35	101.72	123.01	123.01
LP—missed	98.10	98.10	93.81	93.81	93.89	93.89	131.80	128.21
LP—false	56.38	56.38	62.56	62.56	63.19	63.19	107.46	114.76
LP—track switch	9.68	9.68	7.73	7.73	6.00	6.00	22.73	30.79
Average running time (s)	4.41	4.61	8.57	8.90	10.29	10.50	12.87	2.27

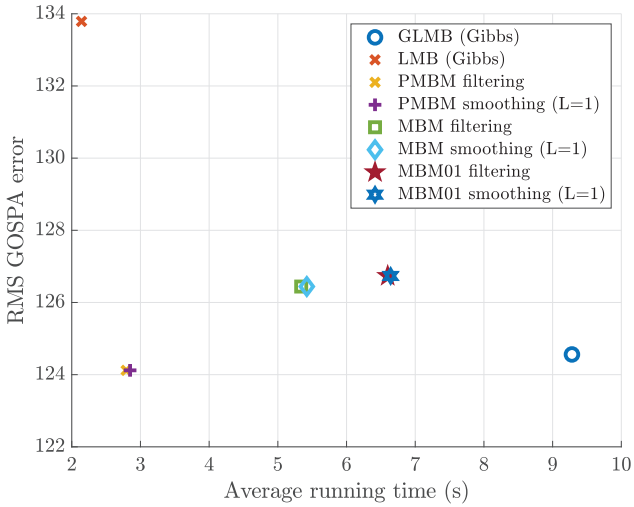


Fig. 2. Performance comparison among the δ-GLMB (Gibbs) filter, the LMB (Gibbs) filter, the trajectory PMBM filter, the trajectory MBM filter, and the trajectory MBM_{01} filter in Scenario 2: RMS GOSPA error versus average running time.

(for the latest four target states), we can find that the running time of the implemented filters is dominated by their filtering recursions.

For Scenario 1, the numerical values of the average RMS GOSPA and the trajectory estimation errors are given in Table I. For Scenario 2, the average RMS GOSPA error and its decomposed values over time are illustrated in Fig. 3, and the average RMS trajectory estimation error and its decomposed values over time are illustrated in Fig. 4. Comparing the results of the two scenarios, we can find that when the birth process is less informative, i.e., a broad birth prior density, the trajectory PMBM filter exhibits lower estimation error than the trajectory MBM and MBM_{01} filters.

While the differences in target state estimation error among different filters are not distinct in both scenarios, it is noticeable that trajectory filters yield much less trajectory estimation error than labeled RFS filters. The worse trajectory estimation performance of labeled RFS filters is a result of worse track continuity. There are two main drawbacks in forming trajectories by connecting target states with the same label: first, misdetections can lead to gaps in the trajectory formed by labeled estimates; second, physically unrealistic track switching; see [31, Fig. 2] for an example.

In addition, we can see that performing fixed-lag smoothing does not change the error due to missed/false detections and track switching; it mainly improves the localization error. This is expected since the choice of $N + L$ has a direct effect on the estimation of past states of the trajectories. From the results of the simulation study, we can conclude that the trajectory PMBM filter has the best tracking performance, and that the trajectory MBM filter is more efficient than the trajectory MBM_{01} filter.

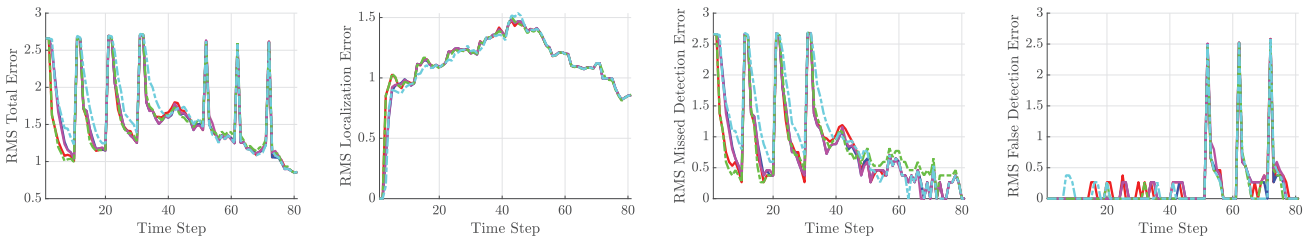


Fig. 3. Average target state estimation error in Scenario 2 evaluated using the GOSPA metric. The lines show the RMS error averaged over 100 Monte Carlo runs. Legend: trajectory PMBM filter (without smoothing) (red), trajectory MBM filter (without smoothing) (blue), trajectory MBM_{01} filter (without smoothing) (magenta), δ-GLMB (Gibbs) filter (green), and LMB (Gibbs) filter (cyan).

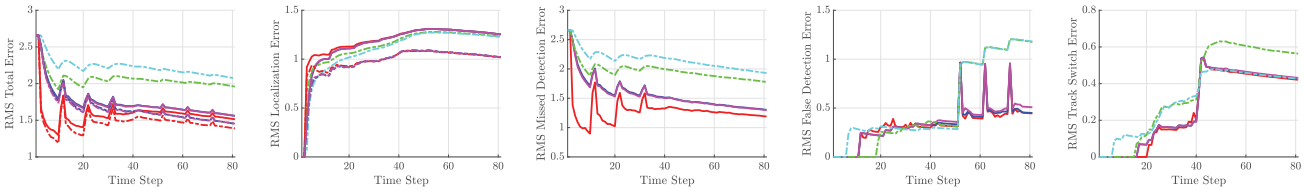


Fig. 4. Average trajectory state estimation error in Scenario 2 evaluated using the trajectory metric [52]. The lines show the RMS error averaged over 100 Monte Carlo runs. Legend: trajectory PMBM filter (without smoothing) (red solid line), trajectory PMBM filter (with smoothing) (red dash-dotted line), trajectory MBM filter (without smoothing) (blue solid line), trajectory MBM filter (with smoothing) (blue dash-dotted line), trajectory MBM_{01} filter (without smoothing) (magenta solid line), trajectory MBM_{01} filter (with smoothing) (magenta dash-dotted line), δ-GLMB (Gibbs) filter (green), and LMB (Gibbs) filter (cyan).

VIII. CONCLUSION

In this paper, we have presented the trajectory MBM filter. We have also presented an efficient implementation of multiscan trajectory PMBM, MBM, and MBM_{01} filters using N-scan pruning and dual decomposition. The performance of the presented multitarget trackers, applied with an efficient fixed-lag smoothing method, is evaluated in a simulation study. The simulation results show that the multiscan trajectory PMBM filter has improved tracking performance over the trajectory MBM filter in terms of state/trajectory estimation error and computational time.

APPENDIX A

In this appendix, we first review why FISST can be used for sets of trajectories. Then, we show how to define reference measures and measure theoretic integrals for sets of trajectories.

A1. Use of FISST for Sets of Trajectories

In this section, we review why FISST can be used for sets of trajectories. The single trajectory space is locally compact, Hausdorff, and second countable (LCHS) [27, Appendix A], where second countable is also referred to as completely separable [53]. LCHS spaces are often used in random set theory [54], and LCHS is also the type of single object space required by Mahler's FISST [6, Sec. 2.2.2].

In particular, single object/measurement spaces that are the disjoint union of spaces of different dimensionalities, similarly to the single trajectory space, have previously been used in Mahler's FISST and RFS framework in [6, Sec. 2.2.2] and [6, Sec. 11.6] for variable state space cardinalized probability hypothesis density filters, and in [6, Ch. 18], [55], and [56] for RFS filters for unknown clutter. In addition, [6, Sec. 3.5.3] explicitly explains how the set integral is constructed for this type of space. Therefore, Mahler's FISST and RFS framework on its own enables us to perform inference on sets of trajectories. For completeness, we proceed to provide also the required measure theory to define probability densities.

A2. Measure Theoretic Integrals

We begin by introducing some basic concepts in measure theory; for more details, see, e.g., [57] and [58, Appendix A]. Consider a nonempty set $\mathcal{Y}$, the pair $(\mathcal{Y}, \sigma(\mathcal{Y}))$, in which $\sigma(\mathcal{Y})$ denotes a σ-algebra of subsets of $\mathcal{Y}$, is called a measurable space. Given a topology space $\mathcal{Y}$, the Borel σ-algebra is the smallest σ-algebra of the subsets of $\mathcal{Y}$ containing the open sets of $\mathcal{Y}$ (or equivalently, by the closed sets of $\mathcal{Y}$). A set $\mathcal{B}$ is said to be measurable if $\mathcal{B} \in \sigma(\mathcal{Y})$. A function $f : \mathcal{Y} \to \mathbb{R}$ is said to be measurable if the inverse images of $\mathbb{R}$ under f are measurable. The triple $(\mathcal{Y}, \sigma(\mathcal{Y}), \mu)$ in which μ is a measure on $\sigma(\mathcal{Y})$ is called a measure space.

The integral of a measurable function $f : \mathcal{Y} \to \mathbb{R}$, $\int f(y)\mu(dy)$, is defined as a limit of integrals of simple functions. The integral of f over any measurable $\mathcal{B} \subset \mathcal{Y}$ is defined as

$$\int_{\mathcal{B}} f(y)\mu(dy) = \int \mathbf{1}_{\mathcal{B}}(y) f(y)\mu(dy), \tag{A.1}$$

where $\mathbf{1}_{\mathcal{B}}$ denotes the indicator function $\mathbf{1}_{\mathcal{B}}(y) = 1$ if $y \in \mathcal{B}$ and $\mathbf{1}_{\mathcal{B}}(y) = 0$ otherwise.

A3. Measure Theoretic Integrals for Single Object LCHS Spaces

In this section, we explain how to define measure theoretic integrals for RFSs whose single objects belong to LCHS spaces, following the steps in [58, Appendix B].

We denote an LCHS space as E. For instance, E could denote the single object space $\mathcal{X}$ or the single trajectory space $\mathcal{T}_k$. We also let $\mathcal{F}(E)$ denote the collection of finite subsets of E.[10]

A common class of RFSs is the Poisson point processes. A Poisson point process Υ is an RFS that is characterized by the property that for any k disjoint Borel subsets $S_1, \ldots, S_k$ of E, the random variables $|\Upsilon \cap S_1|, \ldots, |\Upsilon \cap S_k|$ are independent and have a Poisson distribution. The mean of the Poisson random variables $|\Upsilon \cap S_i|$ is denoted as $v_\Upsilon(S_i)$. The function $v_\Upsilon(\cdot)$ is a (unitless) measure on the Borel subsets of E and is referred

[10] We would like to clarify that the topology on $\mathcal{F}(E)$ is the myopic of Mathéron topology [59], for which we require an LCHS space. To be precise, second countability, not only separability as indicated in [58, Appendix B], is required in the Mathéron topology [59, Sec. 1.1], as it makes use of a countable base [59, p. 1].

to as the intensity measure of Υ. If the mapping from vectors to finite sets is denoted as $\chi : \uplus_{n=0}^{\infty} E^n \to \mathcal{F}(E)$, we have that $\chi((x_1, \ldots, x_n)) = \{x_1, \ldots, x_n\}$. Then, the probability distribution of Υ is [58, Appendix B]

$$P_\Upsilon(\mathcal{B}) = e^{-v_\Upsilon(E)} \sum_{n=0}^{\infty} \frac{v_\Upsilon^n(\chi^{-1}(\mathcal{B}) \cap E^n)}{n!}, \tag{A.2}$$

where $\mathcal{B}$ is a Borel subset of $\mathcal{F}(E)$, χ^{-1} is the inverse mapping of χ, and $v_\Upsilon^n(\cdot)$ is the nth product (unitless) Lebesgue measure of $v_\Upsilon(\cdot)$.

We define the measure $\mu(\cdot)$, on the Borel subsets of $\mathcal{F}(E)$, as

$$\mu(\mathcal{B}) = \sum_{n=0}^{\infty} \frac{v_\Upsilon^n(\chi^{-1}(\mathcal{B}) \cap E^n)}{n!}, \tag{A.3}$$

which is proportional to the probability distribution $P_\Upsilon(\cdot)$. The integral of a measurable function $f : \mathcal{F}(E) \to \mathbb{R}$ with respect to the measure $\mu(\cdot)$ is then [58, Appendix B]

$$\int_{\mathcal{B}} f(\mathbf{X})\mu(d\mathbf{X}) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\chi^{-1}(\mathcal{B}) \cap E^n} f(\{x_1, \ldots, x_n\}) v_\Upsilon^n(dx_1 \cdots dx_n). \tag{A.4}$$

A4. Reference Measure for Sets of Trajectories

In the previous section, we explained how to define a measure theoretic integral with respect to a measure $\mu(\cdot)$ on the Borel subsets of $\mathcal{F}(E)$ in terms of a measure $v_\Upsilon(\cdot)$ on the Borel subsets of E. We proceed to choose a specific measure $v_\Upsilon(\cdot)$ when E is the single trajectory space $\mathcal{T}_k = \uplus_{(\beta,\varepsilon)\in I_k} \{\beta\} \times \{\varepsilon\} \times \mathcal{X}^{\varepsilon-\beta+1}$ and $\mathcal{X} = \mathbb{R}^n$. This will allow us to write the measure theoretic integrals for sets of trajectories in terms of standard Lebesgue integrals and establish the correspondence with Mahler's set integral (8).

We first denote the units of the hypervolume in the single target space $\mathcal{X}$ as K. For example, if the single target state is $[p_x, v_x]$ with p_x being measured in meters (m) and v_x being measured in meters per second (m/s), then $K = \mathrm{m}^2/\mathrm{s}$.

Given a Borel subset S of $\mathcal{T}_k$, which can be written as $S = \uplus_{(\beta,\varepsilon)\in I_k} \{\beta\} \times \{\varepsilon\} \times S_{\varepsilon-\beta+1}$, $S_{\varepsilon-\beta+1} \subset \mathcal{X}^{\varepsilon-\beta+1}$, we choose the measure $v_\Upsilon(\cdot)$ in the single trajectory space as

$$v_\Upsilon(S) = \sum_{(\beta,\varepsilon)\in I_k} \frac{\lambda_{K^{\varepsilon-\beta+1}}(S_{\varepsilon-\beta+1})}{K^{\varepsilon-\beta+1}}, \tag{A.5}$$

where $\lambda_{K^{\varepsilon-\beta+1}}(\cdot)$ represents the Lebesgue measure of $S_{\varepsilon-\beta+1}$ (with units $K^{\varepsilon-\beta+1}$). Therefore, $\lambda_{K^{\varepsilon-\beta+1}}(\cdot)/K^{\varepsilon-\beta+1}$ represents the unitless Lebesgue measure on $\mathcal{X}^{\varepsilon-\beta+1}$. The normalization of each term in (A.5) by $K^{\varepsilon-\beta+1}$ is needed so that we can perform the sum; otherwise, the sum would consider terms with different units, which is erroneous. It is straightforward to check that (A.5) is a measure on the Borel subsets of $\mathcal{T}_k$. That is, $v_\Upsilon(\cdot)$ meets the following three properties that define measures [60]:

1) For any S, $v_\Upsilon(S) \geq 0$.
2) $v_\Upsilon(\emptyset) = 0$.
3) If $S^1, S^2, \ldots$ is a disjoint sequence, then $v_\Upsilon(\sum_{j=1}^{\infty} S^j) = \sum_{j=1}^{\infty} v_\Upsilon(S^j)$.

It is straightforward that the first two properties hold. For the third one, we have

$$\begin{aligned} v_\Upsilon\left(\sum_{j=1}^{\infty} S^j\right) &= \sum_{(\beta,\varepsilon)\in I_k} \frac{\lambda_{K^{\varepsilon-\beta+1}}\left(\sum_{j=1}^{\infty} S^j_{\varepsilon-\beta+1}\right)}{K^{\varepsilon-\beta+1}} \\ &= \sum_{j=1}^{\infty} \sum_{(\beta,\varepsilon)\in I_k} \frac{\lambda_{K^{\varepsilon-\beta+1}}\left(S^j_{\varepsilon-\beta+1}\right)}{K^{\varepsilon-\beta+1}} \\ &= \sum_{j=1}^{\infty} v_\Upsilon(S^j), \end{aligned} \tag{A.6}$$

where we have applied that $\lambda_{K^{\varepsilon-\beta+1}}(\cdot)$ is a measure.

We substitute (A.5) into (A.4) and integrate over the whole space, which implies that $\mathcal{B}$ satisfies that $\chi^{-1}(\mathcal{B}) \cap \mathcal{T}_k^n = \mathcal{T}_k^n$. We have that

$$\begin{aligned} &\int f(\mathbf{X})\mu(d\mathbf{X}) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{T}_k^n} f(\{X_1, \ldots, X_n\}) v_\Upsilon^n(dX_1 \cdots dX_n) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathcal{T}_k} \cdots \int_{\mathcal{T}_k^n} f(\{X_1, \ldots, X_n\}) v_\Upsilon(dX_1) \\ &\quad \cdots v_\Upsilon(dX_n) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\beta_1,\varepsilon_1)\in I_k} \cdots \sum_{(\beta_n,\varepsilon_n)\in I_k} \int_{\mathcal{X}^{\varepsilon_1-\beta_1+1} \times \cdots \times \mathcal{X}^{\varepsilon_n-\beta_n+1}} \\ &\quad f\left(\left\{\left(\beta_1, \varepsilon_1, x_1^{1:\varepsilon_1-\beta_1+1}\right), \ldots, \left(\beta_n, \varepsilon_n, x_n^{1:\varepsilon_n-\beta_n+1}\right)\right\}\right) \\ &\quad \frac{\lambda_{K^{\varepsilon_1-\beta_1+1}}\left(dx_1^{1:\varepsilon_1-\beta_1+1}\right)}{K^{\varepsilon_1-\beta_1+1}} \cdots \frac{\lambda_{K^{\varepsilon_n-\beta_n+1}}\left(dx_n^{1:\varepsilon_n-\beta_n+1}\right)}{K^{\varepsilon_n-\beta_n+1}}. \end{aligned} \tag{A.7}$$

If we further rewrite $\lambda_{K^{\varepsilon_i-\beta_i+1}}(dx_i^{1:\varepsilon_i-\beta_i+1})$ as $dx_i^{1:\varepsilon_i-\beta_i+1}$ and abbreviate $\int_{\mathcal{X}^{\varepsilon_1-\beta_1+1} \times \cdots \times \mathcal{X}^{\varepsilon_n-\beta_n+1}}$ as $\int$, then we have that

$$\begin{aligned} \int f(\mathbf{X})\mu(d\mathbf{X}) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\beta_1,\varepsilon_1)\in I_k} \cdots \sum_{(\beta_n,\varepsilon_n)\in I_k} \int \cdots \int \\ f\left(\left\{\left(\beta_1, \varepsilon_1, x_1^{1:\varepsilon_1-\beta_1+1}\right), \ldots, \left(\beta_n, \varepsilon_n, x_n^{1:\varepsilon_n-\beta_n+1}\right)\right\}\right) \\ \frac{dx_1^{1:\varepsilon_1-\beta_1+1}}{K^{\varepsilon_1-\beta_1+1}} \cdots \frac{dx_n^{1:\varepsilon_n-\beta_n+1}}{K^{\varepsilon_n-\beta_n+1}}. \end{aligned} \tag{A.8}$$

Therefore, for the reference measure $\mu(\cdot)$ in (A.3) and $v_\Upsilon(\cdot)$ in (A.5), the measure theoretic integral corresponds to Mahler's set integral over sets of trajectories (8) but normalized by the units of the differential $dx_1^{1:\varepsilon_1-\beta_1+1}, \ldots, dx_n^{1:\varepsilon_n-\beta_n+1}$, which are $K^{\varepsilon_1-\beta_1+1}, \ldots, K^{\varepsilon_n-\beta_n+1}$. The relation between set integrals and measure theoretic integrals is similar in the single target case [58]. Therefore, if probability densities on sets of trajectories are defined with respect to the reference measure $\mu(\cdot)$, with $v_\Upsilon(\cdot)$ given by (A.5), Mahler's multitrajectory densities are equivalent to measure theoretic densities, except for the normalizing units. Note that if the state space has no units, the measure theoretic integral and Mahler's set integral are alike.

APPENDIX B

In this appendix, we proceed to explain how to use PGFLs, functional derivatives, and the fundamental theorem of multi-object calculus for RFSs of trajectories. These results are important as PGFLs are useful tools to derive filters. First, the prediction and update steps can be performed in the PGFL domain. Second, the fundamental theorem of multi-object calculus indicates how to recover the corresponding multi-object density from a PGFL, which requires functional derivatives. We explain PGFLs in Appendix E and functional derivatives in Appendix F. In Appendix G, we provide and prove the fundamental theorem of multi-object calculus for RFSs of trajectories.

B1. Probability Generating Functionals

PGFLs for sets in LCHS spaces, such as the trajectory space, are defined in [6, Secs. 4.2.4 and 4.2.5]. Let $h : \mathcal{T}_k \mapsto [0, 1]$ be a test function defined on the trajectory state space $\mathcal{T}_k = \uplus_{(\beta,\varepsilon)\in I_k} \{\beta\} \times \{\varepsilon\} \times \mathcal{X}^{\varepsilon-\beta+1}$. Let $\mathbf{X}$ be an RFS of trajectories with multitrajectory density $f(\cdot)$, then its PGFL is

$$G_{\mathbf{X}}[h] = \mathrm{E}\left[h^{\mathbf{X}}\right] = \int h^{\mathbf{X}} f(\mathbf{X})\, \delta\mathbf{X}, \tag{B.1}$$

where

$$h^{\mathbf{X}} = \begin{cases} \prod_{X\in\mathbf{X}} h(X), & \mathbf{X} \neq \emptyset, \\ 1, & \mathbf{X} = \emptyset. \end{cases}$$

Note that both $h(X)$ and the PGFL are unitless functions, i.e., functions whose output has no units.

B2. Functional Derivatives

In this section, we explain (Volterra) functional derivatives for RFS of trajectories using FISST tools. We consider a unitless functional $F[h]$ defined on unitless real-valued functions $h(X)$ with $X \in \mathcal{T}_k$, e.g., a PGFL. Then, using FISST, the functional derivative of $F[h]$ with respect to a finite subset $\mathbf{Y} \in \mathcal{F}(\mathcal{T}_k)$ is defined to be [32, Sec. 11.4]

$$\frac{\delta F}{\delta \mathbf{Y}}[h] = \begin{cases} F[h], & \mathbf{Y} = \emptyset, \\ \lim_{\varepsilon\to 0} \frac{F[h+\varepsilon\delta_Y]-F[h]}{\varepsilon}, & \mathbf{Y} = \{Y\}, \\ \frac{\delta^n F}{\delta Y_1 \cdots \delta Y_n}[h], & \mathbf{Y} = \{Y_1, \ldots, Y_n\}, \end{cases} \tag{B.2}$$

where the Dirac delta on the single trajectory space is

$$\delta_{(\beta',\varepsilon',y_{\beta':\varepsilon'})}(\beta, \varepsilon, x_{\beta:\varepsilon}) = \begin{cases} \delta\left(x_{\beta:\varepsilon} - y_{\beta':\varepsilon'}\right), & \beta = \beta', \varepsilon = \varepsilon', \\ 0, & \beta \neq \beta', \varepsilon \neq \varepsilon', \end{cases}$$

and we use the notational convention

$$\frac{\delta F}{\delta\{Y\}}[h] = \frac{\delta F}{\delta Y}[h].$$

Also, note that the Dirac delta on the single trajectory space meets the following identity:

$$\int \delta_Y(X)\, f(X)\, dX = f(Y).$$

We remark that the use of δ_Y as the input of the functional is a tool of FISST that is not completely rigorous [6, p. 66], but admitted from a practical point of view. Set derivatives can be defined in terms of functional derivatives [6, p. 67].

B3. Fundamental Theorem of Multi-Object Calculus

The fundamental theorem of multi-object calculus enables the recovery of a multi-object density from its PGFL [6, Sec. 3.5.1]. This result also applies to RFS of trajectories, and we provide a proof for completeness.

Theorem 4. *Given the PGFL $G_{\mathbf{X}}[h]$ of an RFS $\mathbf{X}$ of trajectories, we can recover its multitrajectory density $f(\cdot)$ evaluated at $\mathbf{Y}$ as*

$$f(\mathbf{Y}) = \left[\frac{\delta G_{\mathbf{X}}}{\delta \mathbf{Y}}[h]\right]_{h=0}. \tag{B.3}$$

The proof of this theorem is direct for $\mathbf{Y} = \emptyset$ by substituting (B.2) into (B.1). For $\mathbf{Y} \neq \emptyset$, the theorem is a direct consequence of the following lemma.

Lemma 1. *The functional derivative of the PGFL $G_{\mathbf{X}}[h]$ of an RFS $\mathbf{X}$ of trajectories with respect to $\mathbf{Y} = \{Y_1, \ldots, Y_n\}$ is*

$$\frac{\delta^n G_{\mathbf{X}}}{\delta Y_1 \cdots \delta Y_n}[h] = \int h^{\mathbf{X}} f(\{Y_1, \ldots, Y_n\} \cup \mathbf{X})\, \delta\mathbf{X}, \tag{B.4}$$

where $f(\cdot)$ is its multitrajectory density.

The proof of Lemma 1 is given in Appendix G1. Then, by substituting $h = 0$, we directly obtain (B.3) for $\mathbf{Y} \neq \emptyset$. We also have

$$\left[\frac{\delta G_{\mathbf{X}}}{\delta Y}[h]\right]_{h=1} = \int f(\{Y\} \cup \mathbf{X})\, \delta\mathbf{X},$$

which represents the first-order moment, also called intensity and probability hypothesis density, as required.

B3.1. Proof of Lemma 1

In this section, we prove (B.4) by using induction. In part I of the proof, we prove (B.4) for $\mathbf{Y} = \{Y\}$. Then, in part II, we prove the general case $\mathbf{Y} = \{Y_1, \ldots, Y_n\}$.

B3.1.1. Part I of the proof

For $\mathbf{Y} = \{Y\}$, we proceed to prove that

$$\frac{\delta G_{\mathbf{X}}}{\delta Y}[h] = \int h^{\mathbf{X}} f(\{Y\} \cup \mathbf{X})\, \delta\mathbf{X}.$$

For $\mathbf{Y} = \{Y\}$, we have

$$\begin{aligned}
&\frac{\delta G_{\mathbf{X}}}{\delta Y}[h] \\
&= \lim_{\epsilon \to 0} \frac{G_{\mathbf{X}}[h + \epsilon \delta_Y] - G_{\mathbf{X}}[h]}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{\int [h + \epsilon \delta_Y]^{\mathbf{X}} f(\mathbf{X})\, \delta\mathbf{X} - \int [h]^{\mathbf{X}} f(\mathbf{X})\, \delta\mathbf{X}}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{\sum_{n=1}^{\infty} (1/n!) \int f(\{X_1, \ldots, X_n\}) \times \cdots}{\epsilon} \\
&\quad \times \frac{[\prod_{j=1}^{n} [h(X_j) + \epsilon \delta_Y(X_j)] - \prod_{j=1}^{n} h(X_j)] dX_{1:n}}{\epsilon},
\end{aligned}$$

where $X_{1:n} = (X_1, \ldots, X_n)$. The limit can be computed by applying L'Hôpital's rule and taking derivatives with respect to ϵ. This results in

$$\begin{aligned}
&\frac{\delta G_{\mathbf{X}}}{\delta Y}[h] \\
&= \lim_{\epsilon \to 0} \sum_{n=1}^{\infty} \frac{1}{n!} \int \sum_{j=1}^{n} \left[\delta_Y(X_j) \prod_{i=1: i \neq j}^{n} h(X_i + \epsilon \delta_Y(X_i)) \right] \\
&\quad \times f(\{X_1, \ldots, X_n\})\, dX_{1:n} \\
&= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{j=1}^{n} \int \left[\delta_Y(X_j) \prod_{i=1: i \neq j}^{n} h(X_i) \right] \\
&\quad \times f(\{X_1, \ldots, X_n\})\, dX_{1:n}.
\end{aligned}$$

The inner integral is the same for every j, so we can write

$$\begin{aligned}
&\frac{\delta G_{\mathbf{X}}}{\delta \{Y\}}[h] \\
&= \sum_{n=1}^{\infty} \frac{1}{n!} n \int \left[\delta_Y(X_1) \prod_{i=2}^{n} h(X_i) \right] \\
&\quad \times f(\{X_1, \ldots, X_n\})\, dX_{1:n} \\
&= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int \left[\prod_{i=2}^{n} h(X_i) \right] f(\{Y, X_2, \ldots, X_n\})\, dX_{2:n}.
\end{aligned}$$

We further make the change of variables $m = n - 1$ and $X^*_{1:m} = X_{2:n}$ in the previous equation, which yields

$$\begin{aligned}
&\frac{\delta G_{\mathbf{X}}}{\delta \{Y\}}[h] \\
&= \sum_{m=0}^{\infty} \frac{1}{m!} \int \left[\prod_{i=1}^{m} h(X^*_i) \right] f\left(\{Y\} \cup \{X^*_1, \ldots, X^*_m\}\right) dX^*_{1:m} \\
&= \int h^{\mathbf{X}} f(\{Y\} \cup \mathbf{X})\, \delta\mathbf{X}. \qquad \text{(B.5)}
\end{aligned}$$

B3.1.2. Part II of the proof

We proceed to prove (B.4) by induction. We assume that

$$\frac{\delta^{n-1} G_{\mathbf{X}}}{\delta Y_1 \cdots \delta Y_{n-1}}[h] = \int h^{\mathbf{X}} f(\{Y_1, \ldots, Y_{n-1}\} \cup \mathbf{X})\, \delta\mathbf{X} \qquad \text{(B.6)}$$

holds and proceed to prove (B.4). Note that the relation holds for $n = 1$, as proved in the previous section. We denote

$$L[h] = \int h^{\mathbf{X}} l(\mathbf{X})\, \delta\mathbf{X},$$

where

$$l(\mathbf{X}) = f(\{Y_1, \ldots, Y_{n-1}\} \cup \mathbf{X}).$$

Then, by making use of (B.5), we obtain

$$\begin{aligned}
\frac{\delta^n G_{\mathbf{X}}}{\delta Y_1 \cdots \delta Y_n}[h] &= \frac{\delta}{\delta Y_n} L[h] \\
&= \int h^{\mathbf{X}} l(\{Y_n\} \cup \mathbf{X})\, \delta\mathbf{X} \\
&= \int h^{\mathbf{X}} f(\{Y_1, \ldots, Y_n\} \cup \mathbf{X})\, \delta\mathbf{X}.
\end{aligned}$$

This result completes the proof of Lemma 1.

APPENDIX C

In this appendix, we present the MBM_{01} filtering recursions for both the set of current trajectories and the set of all trajectories. The MBM_{01} filtering recursions for the set of all trajectories were first given in [27]; they are presented here for completeness.

C1. Prediction Step for the Set of Current Trajectories

The prediction step is given in the following theorem.

Theorem 5. *Assume that the distribution from the previous time step $f_{k-1|k-1}(\mathbf{X}_{k-1})$ is given by (20) with $r^{i,a^i}_{k-1|k-1} \in \{0, 1\}$, that the transition model is (12), and that the birth model is a trajectory multi-Bernoulli RFS with n^{b}_k Bernoulli components, each of which has density given by (10). Then, the predicted distribution for the next step $f_{k|k-1}(\mathbf{X}_k)$ is given by (20) with $r^{i,a^i}_{k|k-1} \in \{0, 1\}$ and $n_{k|k-1} = n_{k-1|k-1} + n^{\text{b}}_k$. For tracks continuing from previous time ($i \in \{1, \ldots, n_{k-1|k-1}\}$), a hypothesis is included for each combination of a hypothesis from a previous time and either a survival or a death. For new tracks*

($i \in \{n_{k-1|k-1}+l\}$, $l \in \{1,\dots,n_k^{\mathrm{b}}\}$), a hypothesis is included for each combination of a Bernoulli component in the multi-Bernoulli birth density and either born or not born. The number of hypotheses therefore becomes $h_{k|k}^i = 2(h_{k|k-1}^i + n_k^{\mathrm{b}})$.[11] For survival hypotheses ($i \in \{1,\dots,n_{k-1|k-1}\}$, $a^i \in \{1,\dots,h_{k-1|k-1}\}$), if $r_{k-1|k-1}^{i,a^i} = 1$, the parameters are

$$w_{k|k-1}^{i,a^i} = w_{k-1|k-1}^{i,a^i}\langle f_{k-1|k-1}^{i,a^i}; P_{k-1}^{\mathrm{S}}\rangle, \tag{C.1a}$$

$$r_{k|k-1}^{i,a^i} = 1, \tag{C.1b}$$

$$f_{k|k-1}^{i,a^i}(X) = \langle f_{k-1|k-1}^{i,a^i}; \pi^{\mathrm{c}}\rangle. \tag{C.1c}$$

If $r_{k-1|k-1}^{i,a^i} = 0$, the parameters are

$$r_{k|k-1}^{i,a^i} = 0, \tag{C.2a}$$

$$w_{k|k-1}^{i,a^i} = 0. \tag{C.2b}$$

For death hypotheses ($i \in \{1,\dots,n_{k-1|k-1}\}$, $a^i = \tilde{a}^i + h_{k-1|k-1}^i$, $\tilde{a}^i \in \{1,\dots,h_{k-1|k-1}^i\}$), the parameters are

$$w_{k|k-1}^{i,a^i} = w_{k-1|k-1}^{i,a^i}\langle f_{k-1|k-1}^{i,a^i}; 1 - P_{k-1}^{\mathrm{S}}\rangle, \tag{C.3a}$$

$$r_{k|k-1}^{i,a^i} = 0. \tag{C.3b}$$

For birth hypotheses ($i \in \{n_{k-1|k-1}+l\}$, $l \in \{1,\dots,n_k^{\mathrm{b}}\}$), the parameters are

$$\mathcal{M}^{k-1}(i,1) = \emptyset, \tag{C.4a}$$

$$w_{k|k-1}^{i,1} = r_k^{\mathrm{b},l}, \tag{C.4b}$$

$$r_{k|k-1}^{i,1} = 1, \tag{C.4c}$$

$$f_{k|k-1}^{i,1}(X) = f_k^{\mathrm{B},l}(X). \tag{C.4d}$$

For nonbirth hypotheses ($i \in \{n_{k-1|k-1}+l\}$, $l \in \{1,\dots,n_k^{\mathrm{b}}\}$), the parameters are

$$\mathcal{M}^{k-1}(i,2) = \emptyset, \tag{C.5a}$$

$$w_{k|k-1}^{i,2} = 1 - r_k^{\mathrm{b},l}, \tag{C.5b}$$

$$r_{k|k-1}^{i,2} = 0. \tag{C.5c}$$

Compared to the corresponding prediction steps (23) and (24) in the trajectory MBM filter, the MBM_{01} parameterization entails an exponential increase in the number of global hypotheses.

C2. Prediction Step for the Set of All Trajectories

The prediction step is given in the following theorem.

Theorem 6. *Assume that the distribution from the previous time step $f_{k-1|k-1}(\mathbf{X}_{k-1})$ is given by (20) with $r_{k-1|k-1}^{i,a^i} \in \{0,1\}$, that the transition model is (13), and that the birth model is a trajectory multi-Bernoulli RFS with n_k^{b} Bernoulli components, each of which has density given by (10). Then, the predicted distribution for the next step $f_{k|k-1}(\mathbf{X}_k)$ is given by (20), with $r_{k|k-1}^{i,a^i} \in \{0,1\}$ and $n_{k|k-1} = n_{k-1|k-1} + n_k^{\mathrm{b}}$. For tracks continuing from previous time ($i \in \{1,\dots,n_{k-1|k-1}\}$), the number of hypotheses remains the same. For new tracks ($i \in \{n_{k-1|k-1}+l\}$, $l \in \{1,\dots,n_k^{\mathrm{b}}\}$), a hypothesis is included for each combination of a Bernoulli component in the multi-Bernoulli birth density and either born or not born. The number of hypotheses therefore becomes $h_{k|k}^i = h_{k|k-1}^i + 2n_k^{\mathrm{b}}$.*

For hypotheses in tracks continuing from previous time ($i \in \{1,\dots,n_{k-1|k-1}\}$, $a^i \in \{1,\dots,h_{k-1|k-1}\}$), the parameters are

$$w_{k|k-1}^{i,a^i} = w_{k-1|k-1}^{i,a^i} \ \forall a^i, \tag{C.6a}$$

$$r_{k|k-1}^{i,a^i} = 1, \tag{C.6b}$$

$$f_{k|k-1}^{i,a^i}(X) = \langle f_{k-1|k-1}^{i,a^i}; \pi^a\rangle \ \forall a^i. \tag{C.6c}$$

For new tracks ($i \in \{n_{k-1|k-1}+l\}$, $l \in \{1,\dots,n_k^{\mathrm{b}}\}$), the parameters of MBM_{01} parameterization are the same as (64) and (65).

C3. Update Step

The update step is given in the following theorem.

Theorem 7. *Assume that the predicted distribution $f_{k|k-1}(\mathbf{X}_k)$ is given by (20) with $r_{k|k-1}^{i,a^i} \in \{0,1\}$, that the measurement model is (15), and that the measurement set at time step k is $\mathbf{z}_k = \{z_k^1,\dots,z_k^{m_k}\}$. Then, the updated distribution $f_{k|k}(\mathbf{X}_k)$ is given by (20), with $r_{k|k}^{i,a^i} \in \{0,1\}$ and $n_{k|k} = n_{k|k-1}$. For each track ($i \in \{1,\dots,n_{k|k}\}$), a hypothesis is included for each combination of a hypothesis from a previous time with $r_{k|k-1}^{i,a^i} = 1$ and either a misdetection or an update using one of the m_k new measurements, such that the number of hypotheses becomes $h_{k|k}^i = h_{k|k-1}^i(1+m_k)$.[12] For misdetection hypotheses ($i \in \{1,\dots,n_{k|k}\}$, $a^i \in \{1,\dots,h_{k|k-1}\}$) with $r_{k|k-1}^{i,a^i} = 1$, the parameters are*

$$\mathcal{M}^k(i,a^i) = \mathcal{M}^{k-1}(i,a^i), \tag{C.7a}$$

$$w_{k|k}^{i,a^i} = w_{k|k-1}^{i,a^i}\big(1 - \langle f_{k|k-1}^{i,a^i}; P^{\mathrm{D}}\rangle\big), \tag{C.7b}$$

$$r_{k|k}^{i,a^i} = 1, \tag{C.7c}$$

$$f_{k|k}^{i,a^i}(X) = \frac{(1 - P_k^{\mathrm{D}}(X))f_{k|k-1}^{i,a^i}(X)}{\langle f_{k|k-1}^{i,a^i}; 1 - P^{\mathrm{D}}\rangle}. \tag{C.7d}$$

For hypotheses updating tracks ($i \in \{1,\dots,n_{k|k}\}$, $a^i = \tilde{a}^i + h_{k|k-1}^i j$, $\tilde{a}^i \in \{1,\dots,h_{k|k-1}^i\}$, $j \in \{1,\dots,m_k\}$, i.e.,

[11] A hypothesis at the previous time with $r_{k-1|k-1}^{i,a^i} = 0$ would be removed by setting its hypothesis weight to zero. For simplicity, the hypothesis numbering does not account for this exclusion.

[12] A hypothesis at the previous time with $r_{k|k-1}^{i,a^i} = 0$ must not be updated. For simplicity, the hypothesis numbering does not account for this exclusion.

the previous hypothesis $\tilde{a}^i$, updated with measurement z_k^j) with $r_{k|k-1}^{i,a^i} = 1$, the parameters are

$$\mathcal{M}^k(i, a^i) = \mathcal{M}^{k-1}(i, \tilde{a}^i) \cup \{(k, j)\}, \tag{C.8a}$$

$$w_{k|k}^{i,a^i} = \frac{w_{k|k-1}^{i,a^i} \left\langle f_{k|k-1}^{i,\tilde{a}^i}; \varphi\left(z_k^j | \cdot\right) P^{\mathrm{D}} \right\rangle}{\lambda^{\mathrm{FA}}(z_k^j)}, \tag{C.8b}$$

$$r_{k|k}^{i,a^i} = 1, \tag{C.8c}$$

$$f_{k|k}^{i,a^i}(X) = \frac{\varphi\left(z_k^j | X\right) P_k^{\mathrm{D}}(X) f_{k|k-1}^{i,\tilde{a}^i}(X)}{\left\langle f_{k|k-1}^{i,\tilde{a}^i}; \varphi\left(z_k^j | \cdot\right) P_k^{\mathrm{D}} \right\rangle}. \tag{C.8d}$$

REFERENCES

[1] B.-N. Vo, M. Mallick, Y. Bar-Shalom, S. Coraluppi, R. Osborne, R. Mahler, and B.-T. Vo
"Multitarget tracking," in *Wiley Encyclopedia of Electrical and Electronics Engineering*. Hoboken, NJ: Wiley, 2015.

[2] T. Fortmann, Y. Bar-Shalom, and M. Scheffe
"Sonar tracking of multiple targets using joint probabilistic data association,"
IEEE J. Ocean. Eng., vol. 8, no. 3, pp. 173–184, 1983.

[3] S. Blackman and R. Popoli
Design and Analysis of Modern Tracking Systems. Norwood, MA: Artech House, 1999.

[4] Y. Bar-Shalom, P. K. Willett, and X. Tian
Tracking and Data Fusion. Storrs, CT: YBS Publishing, 2011.

[5] S. S. Blackman
"Multiple hypothesis tracking for multiple target tracking,"
IEEE Aerosp. Electron. Syst. Mag., vol. 19, no. 1, pp. 5–18, 2004.

[6] R. P. Mahler
Advances in Statistical Multisource–Multitarget Information Fusion. Norwood, MA: Artech House, 2014.

[7] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans,
Fundamentals of Object Tracking. Cambridge, U.K.: Cambridge University Press, 2011.

[8] D. Musicki and R. Evans
"Joint integrated probabilistic data association: JIPDA,"
IEEE Trans. Aerosp. Electron. Syst., vol. 40, no. 3, pp. 1093–1099, 2004.

[9] J. Williams and R. Lau
"Approximate evaluation of marginal association probabilities with belief propagation,"
IEEE Trans. Aerosp. Electron. Syst., vol. 50, no. 4, pp. 2942–2959, 2014.

[10] F. Meyer, T. Kropfreiter, J. L. Williams, R. Lau, F. Hlawatsch, P. Braca, and M. Z. Win
"Message passing algorithms for scalable multitarget tracking,"
Proc. IEEE, vol. 106, no. 2, pp. 221–259, 2018.

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

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

[13] C. Morefield
"Application of 0–1 integer programming to multitarget tracking problems,"
IEEE Trans. Autom. Control, vol. 22, no. 3, pp. 302–312, 1977.

[14] T. Kurien, "Issues in the design of practical multitarget tracking algorithms," in *Multitarget–Multisensor Tracking: Advanced Applications*. Norwood, MA: Artech House, 1990, pp. 43–87.

[15] A. B. Poore and A. J. Robertson III
"A new Lagrangian relaxation based algorithm for a class of multidimensional assignment problems,"
Comput. Optim. Appl., vol. 8, no. 2, pp. 129–150, 1997.

[16] S. Deb, M. Yeddanapudi, K. Pattipati, and Y. Bar-Shalom
"A generalized SD assignment algorithm for multisensor–multitarget state estimation,"
IEEE Trans. Aerosp. Electron. Syst., vol. 33, no. 2, pp. 523–538, 1997.

[17] R. A. Lau and J. L. Williams, "Multidimensional assignment by dual decomposition," in *Proc. 7th Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process.*, 2011, pp. 437–442.

[18] S. Coraluppi and C. Carthel, "Modified scoring in multiple-hypothesis tracking." *J. Adv. Inf. Fusion*, vol. 7, no. 2, pp. 153–164, 2012.

[19] S. Mori, C.-Y. Chong, and K.-C. Chang, "Three formalisms of multiple hypothesis tracking," in *Proc. 19th Int. Conf. Inf. Fusion*, 2016, pp. 727–734.

[20] E. Brekke and M. Chitre
"Relationship between finite set statistics and the multiple hypothesis tracker,"
IEEE Trans. Aerosp. Electron. Syst., vol. 54, no. 4, pp. 1902–1907, 2018.

[21] 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, 2013.

[22] 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, 2015.

[23] Á. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson
"Poisson multi-Bernoulli mixture filter: Direct derivation and implementation,"
IEEE Trans. Aerosp. Electron. Syst., vol. 54, no. 4, pp. 1883–1901, 2018.

[24] Á. F. García-Fernández, Y. Xia, K. Granström, L. Svensson, and J. Williams, "Gaussian implementation of the multi-Bernoulli mixture filter," in *Proc. 22nd Int. Conf. Inf. Fusion*, 2019.

[25] Á. 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, 2013.

[26] 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, 2016.

[27] Á. F. García-Fernández, L. Svensson, and M. R. Morelande,
"Multiple target tracking based on sets of trajectories,"
IEEE Trans. Aerosp. Electron. Syst., 2019. Available: https://ieeexplore.ieee.org/document/8731733.

[28] 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, 2014.

[29] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer
"The labeled multi-Bernoulli filter,"

IEEE Trans. Signal Process., vol. 62, no. 12, pp. 3246–3260, 2014.

[30] L. Svensson and M. Morelande, "Target tracking based on estimation of sets of trajectories," in *Proc. 17th Int. Conf. Inf. Fusion*, 2014.

[31] K. Granström, L. Svensson, Y. Xia, J. Williams, and Á. F. García-Femández, "Poisson multi-Bernoulli mixture trackers: Continuity through random finite sets of trajectories," in *Proc. 21st Int. Conf. Inf. Fusion*, 2018.

[32] R. P. Mahler
Statistical Multisource–Multitarget Information Fusion. Norwood, MA: Artech House, 2007.

[33] K. Murthy
"An algorithm for ranking all the assignments in order of increasing costs,"
Oper. Res., vol. 16, no. 3, pp. 682–687, 1968.

[34] Y. Xia, K. Granström, L. Svensson, and Á. F. García-Fernández, "An implementation of the Poisson multi-Bernoulli mixture trajectory filter via dual decomposition," in *Proc. 21st Int. Conf. Inf. Fusion*, 2018.

[35] B.-N. Vo, B.-T. Vo, and H. G. Hoang
"An efficient implementation of the generalized labeled multi-Bernoulli filter,"
IEEE Trans. Signal Process., vol. 65, no. 8, pp. 1975–1987, 2017.

[36] S. Mori and C.-Y. Chong, "Evaluation of data association hypotheses: Non-Poisson iid cases," in *Proc. 7th Int. Conf. Inf. Fusion*, 2004, pp. 1133–1140.

[37] S. Coraluppi and C. A. Carthel
"If a tree falls in the woods, it does make a sound: Multiple-hypothesis tracking with undetected target births,"
IEEE Trans. Aerosp. Electron. Syst., vol. 50, no. 3, pp. 2379–2388, 2014.

[38] Á. F. García-Fernández and L. Svensson, "Trajectory probability hypothesis density filter," in *Proc. 21st Int. Conf. Inf. Fusion*, IEEE, 2018, pp. 1430–1437.

[39] Á. F. García-Fernández and L. Svensson
"Trajectory PHD and CPHD filters,"
IEEE Trans. Signal Process., vol. 67, no. 22, pp. 5702–5714, 2019.

[40] H. A. Blom and Y. Bar-Shalom
"The interacting multiple model algorithm for systems with Markovian switching coefficients,"
IEEE Trans. Autom. Control, vol. 33, no. 8, pp. 780–783, 1988.

[41] K. Granström, P. Willett, and Y. Bar-Shalom
"Systematic approach to IMM mixing for unequal dimension states,"
IEEE Trans. Aerosp. Electron. Syst., vol. 51, no. 4, pp. 2975–2986, 2015.

[42] K. Granstrom, M. Fatemi, and L. Svensson
"Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering,"
IEEE Trans. Aerosp. Electron. Syst., 2019. Available: https://ieeexplore.ieee.org/document/8730493

[43] Y. Xia, K. Granström, L. Svensson, Á. F. García-Fernández, and J. Williams, "Extended target Poisson multi-Bernoulli mixture trackers based on sets of trajectories," in *Proc. 22nd Int. Conf. Inf. Fusion*, 2019.

[44] N. Komodakis, N. Paragios, and G. Tziritas
"MRF energy minimization and beyond via dual decomposition,"
IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 3, pp. 531–552, 2011.

[45] Y. Bar-Shalom
Multitarget–Multisensor Tracking: Advanced Applications. Norwood, MA, Artech House, 1990.

[46] B. Polyak, "Subgradient methods: A survey of soviet research," in *Proc. IIASA Workshop Nonsmooth Optim.*, 1978, pp. 5–30.

[47] E. L. Lawler and D. E. Wood
"Branch-and-bound methods: A survey,"
Oper. Res., vol. 14, no. 4, pp. 699–719, 1966.

[48] 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, 2011.

[49] R. M. Eustice, H. Singh, and J. J. Leonard
"Exactly sparse delayed-state filters for view-based slam,"
IEEE Trans. Robot., vol. 22, no. 6, pp. 1100–1114, 2006.

[50] S. Reuter, A. Danzer, M. Stübler, A. Scheel, and K. Granström "A fast implementation of the labeled multi-Bernoulli filter using Gibbs sampling," in *Proc. Symp. Intell. Veh.*, IEEE, 2017, pp. 765–772.

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

[52] A. S. Rahmathullah, Á. F. García-Fernández, and L. Svensson, "A metric on the space of finite sets of trajectories for evaluation of multi-target tracking algorithms," arXiv:1605.01177, 2016.

[53] G. F. Simmons, *Topology and Modern Analysis*. New York, NY: McGraw-Hill, 1963.

[54] I. S. Molchanov, *Theory of Random Sets*. New York, NY: Springer, 2005.

[55] R. Mahler and A. El-Fallah
"CPHD and PHD filters for unknown backgrounds, part III: tractable multitarget filtering in dynamic clutter," in *Signal and Data Processing of Small Targets 2010*, vol. 7698, O. E. Drummond, Ed., Bellingham, WA: SPIE, 2010, pp. 177–188.

[56] R. P. Mahler, B.-T. Vo, and B.-N. Vo
"CPHD filtering with unknown clutter rate and detection profile,"
IEEE Trans. Signal Process., vol. 59, no. 8, pp. 3497–3513, 2011.

[57] G. Grimmett and D. Stirzaker
Probability and Random Processes. Oxford, U.K.: Oxford University Press, 2001.

[58] B.-N. Vo, S. Singh, and A. Doucet
"Sequential Monte Carlo methods for multitarget filtering with random finite sets,"
IEEE Trans. Aerosp. Electron. Syst., vol. 41, no. 4, pp. 1224–1245, 2005.

[59] G. Matheron, *Random Sets and Integral Geometry*. Hoboken, NJ: Wiley, 1975.

[60] P. Billingsley, *Probability and Measure*. Hoboken, NJ: Wiley, 2008.

Yuxuan Xia was born in Wuhan, China, in 1993. He received the B.Sc. degree in engineering of Internet of Things from the Jiangnan University, Wuxi, China, in 2015, and the M.Sc. degree in communication engineering from Chalmers University of Technology, Gothenburg, Sweden, in 2017. He is currently working toward the Ph.D. degree at the Department of Electrical Engineering, Chalmers University of Technology. His research interests include Bayesian inference in general, and target tracking in particular.

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 2008 and 2012, respectively. He is currently a Postdoctoral Research Fellow with the Department of Signals and Systems, 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 estimation theory, multiple model estimation, sensor fusion, and target tracking, especially for extended targets. He was the recipient of paper awards at the Fusion 2011 and Fusion 2012 conferences. In 2018, the International Society of Information Fusion (ISIF) awarded him the ISIF Young Investigator Award for his contributions to extended object tracking and his service to the research community.

Lennart Svensson was born in Älvängen, Sweden, in 1976. He received the M.S. degree in electrical engineering and the Ph.D. degree from Chalmers University of Technology, Gothenburg, Sweden, in 1999 and 2004, respectively. He is currently a Professor of Signal Processing with 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. He has organized a massive open online course on multiple object tracking, available on edX and YouTube, and was the recipient of paper awards at the International Conference on Information Fusion in 2009, 2010, 2017, and 2019.

Ángel F. García-Fernández received the M.Sc. degree in telecommunication engineering and the Ph.D. degree from the 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 the Universidad Politécnica de Madrid, Chalmers University of Technology, Gothenburg, Sweden, Curtin University, Perth, WA, Australia, and Aalto University, Espoo, Finland. His main research activities and interests are in the area of Bayesian nonlinear inference, with emphasis on dynamic systems and multiple target tracking. He was the recipient of paper awards at the International Conference on Information Fusion in 2017 and 2019.

Jason L. Williams (S'01–M'07–SM'16) received BE (electronics)/BInfTech degree from Queensland University of Technology, Brisbane, Australia in 1999, the MSEE degree from the United States Air Force Institute of Technology, Wright-Patterson AFB, OH, USA in 2003, and the Ph.D. degree from Massachusetts Institute of Technology, Cambridge, MA, USA in 2007. He is currently a Senior Research Scientist in Robotic Perception with the Robotics and Autonomous Systems Group of Commonwealth Scientific and Industrial Research Organisation, Brisbane, Australia. His research interests include SLAM, computer vision, multiple object tracking, and motion planning. He is also an Adjunct Associate Professor with Queensland University of Technology. He previously worked in sensor fusion and resource management as a Senior Research Scientist with the Defence Science and Technology Group, Australia, and in electronic warfare as an Engineering Officer with the Royal Australian Air Force.