Poisson Multi-Bernoulli Mixtures for Sets of Trajectories

Karl Granström, *Member, IEEE,* Lennart Svensson, *Senior Member, IEEE,* Yuxuan Xia, Jason Williams, *Senior Member, IEEE,* and Ángel F. García-Fernández

***Abstract*—The Poisson Multi-Bernoulli Mixture (PMBM) density is a conjugate multi-target density for the standard point target model with Poisson point process birth. This means that both the filtering and predicted densities for the set of targets are PMBM. In this paper, we first show that the PMBM density is also conjugate for sets of trajectories with the standard point target measurement model. Second, based on this theoretical foundation, we develop two trajectory PMBM filters that provide recursions to calculate the posterior density for the set of all trajectories that have ever been present in the surveillance area, and the posterior density of the set of trajectories present at the current time step in the surveillance area. These two filters therefore provide complete probabilistic information on the considered trajectories enabling optimal trajectory estimation. Third, we establish that the density of the set of trajectories in any time window, given the measurements in a possibly different time window, is also a PMBM. Finally, the trajectory PMBM filters are evaluated via simulations, and are shown to yield state-of-the-art performance compared to other multi-target tracking algorithms based on random finite sets and multiple hypothesis tracking.**

***Index Terms*—Multiple target tracking, point targets, sets of trajectories, conjugate priors, Poisson multi-Bernoulli mixtures.**

I. Introduction

Multi-target tracking (MTT) refers to the processing of sensor measurements to estimate the unknown and time-varying number of targets and their trajectories [1]. In this paper, we consider discrete-time multi-target tracking in which measurements at taken at discrete time steps, and each trajectory is characterised by a start time, and a sequence of target states. An alternative definition of a trajectory is to consider it a function of time, which can then be estimated using optimisation techniques [2]–[4]. This approach has the benefit of estimating continuous-time trajectories, but it does not provide associated uncertainties.

This paper uses the Bayesian framework of random finite sets (RFS) of trajectories to address the MTT problem [5]. Within this framework, we are given probabilistic models for multi-target birth, dynamics and death, and also models for the sensor measurements, and the goal is to compute or approximate the posterior density over the set of trajectories, as it is the density that enables us to answer all trajectory-related questions we may have, such as the estimation of the set of trajectories, and their associated uncertainties.

This paper focuses on the sequential calculation of two important densities: the density for the set of all trajectories that have ever been present in the surveillance area, and the density for the set of trajectories present in the surveillance area at the current time step. We show that, in both cases, for the standard point target measurement model and standard dynamic model with Poisson point process (PPP) birth, the conjugate density of the set of trajectories is a Poisson Multi-Bernoulli Mixture (PMBM), which was first introduced for sets of targets in [6]. The resulting algorithms are referred to as trajectory PMBM (TPMBM) filters.

Solutions to the MTT problem presuppose the ability to solve single target Bayesian filtering. Let $x_k \in \mathbb{R}^{n_x}$ and $z_k \in \mathbb{R}^{n_z}$ denote the single target state and the measurement state at time step k. In single target Bayesian estimation, linear and Gaussian models are important because they yield closed-form solutions to Bayesian filtering, provided by the Kalman filter, and smoothing, provided by the Rauch-Tung-Striebel (RTS) smoother or the two-filter smoother [7]. Specifically, the density of the state x_k conditioned on the sequence $z_{1:k'} = (z_1, ..., z_{k'})$ is Gaussian [7]

$$p(x_k|z_{1:k'}) = \mathcal{N}\left(x_k\,;\, m_{k|k'}, P_{k|k'}\right) \tag{1}$$

where $m_{k|k'}$ and $P_{k|k'}$ are the mean and covariance matrix, and $k > k'$ for prediction, $k = k'$ for filtering and $k < k'$ for smoothing.

Even more generally, if $x_{\alpha:\gamma}$ and $z_{\xi:\chi}$ denote the concatenation of target state and measurements in any (finite) time intervals $\alpha:\gamma$ and $\xi:\chi$, it holds that [7], [8]

$$p(x_{\alpha:\gamma}|z_{\xi:\chi}) = \mathcal{N}\left(x_{\alpha:\gamma}\,;\, m_{\alpha:\gamma|\xi:\chi}, P_{\alpha:\gamma|\xi:\chi}\right) \tag{2}$$

where $m_{\alpha:\gamma|\xi:\chi}$ and $P_{\alpha:\gamma|\xi:\chi}$ are the corresponding mean and covariance matrix. That is, regardless of the time intervals we consider in the state or in the measurement, the posterior density is Gaussian.

This paper is an extension of the conference paper [9]. In this paper, in addition to deriving the TPMBM filters, we also show that a result similar to (2) holds for MTT with the standard point target models. Specifically, the density of the set of trajectories in some time interval $\alpha:\gamma$ given the measurements in any time interval $\xi:\chi$ is PMBM. Consequently, for prediction, filtering and smoothing, the set

K. Granström is with Zoox, San Francisco, CA, USA. He was previously with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. E-mail: kagranstrom@gmail.com.

L. Svensson is with the Department of Electrical Engineering, Chalmers university of Technology, Gothenburg, Sweden. E-mail: lennart.svensson@chalmers.se.

Y. Xia is with the Department of Automation, Shanghai Jiaotong University, Shanghai, China. Email: yuxuan.xia@sjtu.edu.cn.

J. Williams is with Whipbird Signals, Australia. E-mail: jlw@ieee.org.

A. F. García-Fernández is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, UK. E-mail: angel.f.garcia.fernandez@gmail.com.

of trajectories in a given time interval is PMBM distributed. A special case of this result is obtained if we consider the density of the current set of targets and the sequence of measurements up to the current time step, which is a PMBM and can be calculated by the PMBM filtering recursion [6].

To summarise the contributions in this paper, for the standard point target model, we

1) extend PMBM conjugacy to sets of trajectories and derive the TPMBM filters for the set of all trajectories, and the set of alive trajectories,
2) present computationally efficient algorithms for computing the PMBM posterior,
3) show the time marginalisation theorem for PMBM densities. This theorem states that the posterior density for the set of trajectories in any time window whose trajectories are alive at some point in a time sub-interval is a PMBM. In addition, the associated PMBM parameters can be computed from the posterior PMBM on the set of all trajectories.
4) show that the posterior density for the set of trajectories in any time window, given the measurements in a possibly different time window, is a PMBM.
5) derive relevant properties of the probability mass functions (pmfs) of the start time step and of the end time step for each Bernoulli trajectory.

The rest of the paper is organised as follows. Related work is discussed in the next section. In Section III, we present the two considered problem formulations and the background theory on sets of trajectories. The prediction and update steps of the TPMBM filter are presented in Section IV. In Section V, we show the time marginalisation theorem and that the distribution of the set of trajectories over any time window is PMBM. Linear and Gaussian implementations of the TPMBM filters are presented in Section VI, and Section VII analyses the performance of TPMBM filters via simulations. The paper is concluded in Section VIII.

II. Related work

Approaches to the MTT problem include the joint probabilistic data association filter (JPDA) [1], multiple hypothesis tracking (MHT) [10], and filters based on RFS [11]. Links between RFS, JPDA, and MHT have been established in [6], [12], [13].

MHT refers to a collection of algorithms that address the data association problem in MTT by forming hypotheses across multiple scans [14]. In the original MHT [10], each measurement could potentially be the first detection of a new target, and the number of newly detected targets was given a Poisson distribution. MHT can be implemented by forming tracks for each target and assigning them a score, which is then used for track confirmation and deletion, and to determine the score of a data association hypothesis [15, Chap. 16] [16]. In [17], the MHT model was made rigorous using a multi-target state-space which allows for an unknown number of targets, under the assumption that the number of targets is unknown but constant, with a Poisson prior. Target state sequences were formed under each global hypothesis, and the Poisson distribution of targets remaining to be detected provided a Bayes prior for events involving newly detected targets. Hypotheses were constructed as being data-to-data, since no a priori data was assumed on target identity. This formulation was revisited in [18], and compared to random finite sets, and finite point processes. MHT was extended in [19], to include a time-varying number of targets by using a target birth distribution in the measurement space and assuming stationary object dynamics.

The Bayesian multi-target filter, using a PPP birth model and the standard point target models, was derived in [6] resulting in a PMBM posterior. This PMBM provides all information on the current set of targets, but not their trajectories. The PMBM posterior represents the union of two independent RFS processes: a PPP that has information on targets that have not been detected yet, and a multi-Bernoulli mixture (MBM) that has information on the potential targets that have been detected at some time up to now. Information on potentially undetected targets is important, for example, in search-and-track sensor management applications so that sensor resources can be directed to the regions of the surveillance area that may contain undetected targets [20], [21]. The MBM contains information on all data-to-data global hypotheses, their weights, and a multi-Bernoulli (MB) density for each global hypothesis with information on the detected targets that may be present. Therefore, the data-to-data global hypotheses structure of PMBM resembles the MHT approach. Nevertheless, to our knowledge, no MHT algorithm considers a MB density for each global hypothesis, which results in computational benefits. Therefore, the PMBM filter [6], [22] can be regarded as a state-of-the-art Bayesian MHT-type multi-target filter. Explicit links between PMBM and MHT are provided in [6], [12].

Three important advantages of the PMBM filter are its efficient representation of the global hypotheses [22], that it maintains information on undetected targets, and that it is the result of the Bayesian filtering recursion. The PMBM filter has been applied to solve MTT problems with different types of data such as radar [23], lidar [23]–[25], audio [26] and images [23], [27]. The PMBM filter has also been successfully applied to mapping of stationary objects [28], simultaneous localisation and mapping [29], [30], and joint tracking and sensor localisation [31].

If the birth model is an MB rather than a PPP, the posterior of the current set of targets becomes an MBM with the same recursion as the PMBM filter with the difference that the birth Bernoulli components are added to the MBM in the prediction step [22]. That is, tracks in the MBM filter are initiated using a priori birth information, while in the PMBM filter, track initiation is measurement-driven, which improves performance [32]. After each MBM filter prediction step, we can also represent the MBM with global hypotheses that consider targets that either exist or not (deterministic target existence). This approach has the drawback of requiring an exponential increase in the number of global hypotheses, giving rise to the MBM_{01} filter [22]. Deterministic target existence for each global hypothesis is also the approach in standard MHT algorithms [15].

In MTT, apart from having information on the set of

targets, it is important to represent information on the target trajectories [8]. One approach to estimate target trajectories sequentially is to augment the target states with labels, which are unique for each potential target and not allowed to change in time. Labels have been used in non-Bayesian MTT algorithms in computer vision to design an optimisation problem to estimate trajectories [2]. In a Bayesian setting with labelled targets, we can calculate the corresponding posterior, estimate the labelled targets at each time step, and estimate trajectories by connecting target state estimates with the same label [33]–[36]. This approach gives rise to MTT algorithms based on labelled RFS if the multi-target state is a set [33], [34]. There are also related approaches for sequential track formation such as the use of PMBM filter metadata (auxiliary variables) [37] or the methods in [38], [39]. Labelled versions of the MBM and MBM_{01} filters can be obtained by using an MB birth model where each Bernoulli of the birth process gives rise to a unique and fixed label [34]. The posterior can then be calculated using the equations of the MBM or the MBM_{01} filters [32]. The δ-generalised labelled MB (δ-GLMB) recursion [34] is equivalent to the MBM_{01} recursion, which means that the δ-GLMB filter makes use of an less efficient representation of the global hypotheses than MBM and PMBM filters [22, Sec. IV].

The δ-GLMB filter often estimates trajectories well. However, δ-GLMB filter trajectory estimation has some weaknesses. For example, the δ-GLMB filter sequential trajectory estimator can produce trajectory estimates with holes even though the existence of holes in the trajectories is not a possible event according to the standard dynamic model, in which targets may die but are not allowed to resurrect. In addition, the δ-GLMB filter has problems estimating trajectories of multiple targets born at the same time if the birth process is an independent and identically distributed cluster process covering a large area [5, Example 2]. A practical solution to solve this problem that has been used in large-scale problems [40] is to use an adaptive birth model [41], in which each measurement gives rise to a birth Bernoulli, resembling how the PMBM filter initiates Bernoulli components. However, using adaptive birth implies that the δ-GLMB filter is no longer a fully Bayesian filter. For example, target birth is no longer independent of past measurements, as required in δ-GLMB modelling assumptions, and an extra user-defined parameter is needed [42].

One way to avoid the above issues with trajectory estimation is to estimate trajectories directly from the posterior on the sequence of sets of labelled targets [43] or the posterior on the set of trajectories [5]. The benefits of using sets of trajectories instead of sequences of sets of labelled targets are thoroughly discussed in [5]. The paper [5] introduced the MBM_{01} filter for the set of all trajectories: the trajectory MBM_{01} (TMBM_{01}) filter. In the TMBM_{01} filter, the set of trajectories can also be labelled, without adding relevant information or changes to the TMBM_{01} recursion [5, Sec. IV.A], and the marginal distribution on the current set of targets corresponds to the posterior δ-GLMB density [5, Sec. IV.C]. In addition, the TMBM_{01} filter has also been derived using the sequence of labelled sets of targets and δ-GLMB notation in [44], and was referred to as multi-scan δ-GLMB. However, the TMBM_{01} (and the multi-scan δ-GLMB) filter still suffers from an inefficient global hypothesis representation, since hypotheses have deterministic target existence. The trajectory MBM (TMBM) filter [45] improves the global hypothesis representation of the TMBM_{01} filter by enabling probabilistic target existence.

In this paper, we show that the benefits of the global hypotheses representation of the PMBM filter can be combined with sets of trajectories to derive trajectory PMBM (TPMBM) filters. These filters provide full information on the set of trajectories, including undetected ones, and make use of a more efficient representation of the global hypotheses than TMBM_{01} and TMBM filters. The proposed TPMBM filters are fully Bayesian methods that can be considered state-of-the-art MHT-type algorithms that provide full trajectory information on all targets.

Several subsequent papers[1] have been proposed based on the derivations presented in this paper, through the preprint [46] or the preliminary conference version [9]. For instance, the trajectory Poisson MB (TPMB) filter in [37] approximates the TPMBM posterior at each update step as a TPMB, which only has one mixture component, via a Kullback-Leibler divergence minimisation, see [37, Figure 1]. The extended target TPMB filter in [47] obtains the TPMB approximation at each update step using set-type belief propagation. TPMB filters are computationally faster than TPMBM filters but they not provide a closed-form solution to the posterior. Reference [45] proposed a multi-scan implementation of the TPMBM filter, as well as the TMBM filtering recursion, which results from an MB birth model instead of a PPP birth model. When target spawning is considered, rather than sets of trajectories, it is more appropriate to consider sets of tree trajectories to keep full genealogy information. The resulting TPMBM filter for sets of tree trajectories is proposed in [48]. A continuous-discrete TPMBM filter that optimally processes out-of-sequence measurements is proposed in [49].

Differently from the above-mentioned works, this paper provides the theoretical foundation for the TPMBM filters, presents three possible single-trajectory filtering implementations, includes a thorough comparison with other state-of-the-art RFS filters and also with an MHT implementation, derives pertinent properties of the pmfs of the start time step and of the end time step for each Bernoulli trajectory, derives the time marginalisation theorem for PMBM densities on sets of trajectories, and proves the theorem on TPMBM densities for arbitrary time intervals, which shows that the posterior of the set of trajectories on any time window given measurements in a possibly different time window is a TPMBM. Paper [50] is based this theorem, and shows that if we add spatio-temporal constraints to a TPMBM density, the density remains TPMBM.

III. Problem formulations and background

In this section, we first present the problem formulations. Then, we review the trajectory state representation, and present the densities for sets of trajectories. All details on how to define densities and integration for sets of trajectories are given in [5] using finite set statistics [11], and in [45] using measure theory.

[1] MATLAB code of these algorithms is provided at https://github.com/yuhsuansia and https://github.com/Agarciafernandez.

A. Problem formulations

We use notation $x_k \in \mathcal{X}$ to denote a single target state at time step k, where $\mathcal{X}$ is the single target state space. Similarly, we use $z_k \in \mathcal{Z}$ to denote a single measurement at time step k, where $\mathcal{Z}$ is the single measurement state space.

Assumption 1. *The multi-target dynamic model is [11]:*

- *Targets are born at each time step following an independent* PPP *with intensity* $\lambda^{\mathrm{b}}(x_k)$.
- *A target with state* x_k *survives to the next time step with probability of survival* $P^{\mathrm{S}}(x_k)$ *and single-target transition density* $\pi^{x}(x_k|x_{k-1})$*, independently of the rest of the targets.*

Assumption 2. *The measurement model is [11]:*

- *A target* x_k *is detected with a probability* $P^{\mathrm{D}}(x_k)$*, independently of the rest of the targets, and its associated single measurement has a density* $\varphi^{z}(z_k|x_k)$.
- *Clutter measurements are modelled by an independent* PPP *with intensity* $\lambda^{\mathrm{FA}}(z_k)$.

Under these assumptions, we consider the following two problem formulations (PF) of MTT:

Problem Formulation 1. The set of all trajectories: *the aim is to compute the density of the set of trajectories of all targets that have been present in the surveillance area at some point up to the current time step.*

Problem Formulation 2. The set of current trajectories: *the aim is to compute the density of the set of trajectories of the targets that are currently present in the surveillance area.*

B. Single trajectories

We first present the notation for time intervals. The (ordered) sequence of consecutive time steps between time steps α and γ is $\alpha : \gamma = (\alpha, \alpha+1, \ldots, \gamma-1, \gamma)$. The corresponding (unordered) set of time steps is $\mathbb{N}_{\alpha}^{\gamma} = \{\alpha, \alpha+1, \ldots, \gamma-1, \gamma\}$.

The single trajectory state is $X = (\beta, \varepsilon, x_{\beta:\varepsilon})$ where β is the time step when the trajectory starts, ε is the time step when the trajectory ends, $x_{\beta:\varepsilon} = (x_{\beta}, x_{\beta+1}, \ldots, x_{\varepsilon-1}, x_{\varepsilon})$, where $x_k \in \mathcal{X},\ \forall k \in \mathbb{N}_{\beta}^{\varepsilon}$. The length in time steps of a trajectory X is $\ell = \varepsilon - \beta + 1$. Note that it is also possible to define the trajectory state X without including ε, as ε is implicitly given in the length of $x_{\beta:\varepsilon}$ [5].

The single trajectory space in the time interval $\alpha : \gamma$ is [5]

$$\mathcal{T}_{\alpha:\gamma} = \uplus_{(\beta,\varepsilon)\in I_{\alpha:\gamma}} \{\beta\} \times \{\varepsilon\} \times \mathcal{X}^{\varepsilon-\beta+1}, \tag{3}$$

where $\uplus$ is the union of disjoint sets, $I_{\alpha:\gamma} = \{(\beta, \varepsilon) : \alpha \leq \beta \leq \varepsilon \leq \gamma\}$ and $\mathcal{X}^{\ell}$ represents ℓ Cartesian products of $\mathcal{X}$. This is a slight generalisation of the definition in [5], where $\mathcal{T}_k$ is used to denote the trajectory state space for trajectories in $0 : k$. The possible lengths of trajectories in $\mathcal{T}_{\alpha:\gamma}$ are constrained to $1 \leq \ell \leq \gamma - \alpha + 1$. It should be noted that state spaces that are the union of different, bounded size state spaces, as the trajectory space, have been used in interacting multiple models [51] and unknown clutter models [52].

A density on the single trajectory space $\mathcal{T}_{\alpha:\gamma}$ factorises as

$$p(X) = p(x_{\beta:\varepsilon}|\beta, \varepsilon)P(\beta, \varepsilon), \tag{4}$$

where the domain of $P(\beta, \varepsilon)$ is $I_{\alpha:\gamma}$. Integration is defined as [5]

$$\int_{\mathcal{T}_{\alpha:\gamma}} p(X)\mathrm{d}X = \sum_{(\beta,\varepsilon)\in I_{\alpha:\gamma}} \left[\int_{\mathcal{X}^{\ell}} p(x_{\beta:\varepsilon}|\beta, \varepsilon)\mathrm{d}x_{\beta:\varepsilon}\right] P(\beta, \varepsilon). \tag{5}$$

C. Sets of trajectories

The set of trajectories in the time interval $\alpha : \gamma$ is $\mathbf{X}_{\alpha:\gamma}$, such that $\mathbf{X}_{\alpha:\gamma} \in \mathcal{F}(\mathcal{T}_{\alpha:\gamma})$, the set of all finite subsets of $\mathcal{T}_{\alpha:\gamma}$. The subset of the trajectories in $\mathbf{X}_{\alpha:\gamma}$ that were alive at some time in the time interval $\eta : \zeta$, where $\alpha \leq \eta \leq \zeta \leq \gamma$, is

$$\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta} = \left\{(\beta, \varepsilon, x_{\beta:\varepsilon}) \in \mathbf{X}_{\alpha:\gamma} \ : \ \mathbb{N}_{\beta}^{\varepsilon} \cap \mathbb{N}_{\eta}^{\zeta} \neq \emptyset\right\}. \tag{6}$$

For example, in PF 2, we consider (6) with $\eta = \zeta = \gamma$. As in RFS theory for sets of targets [11], we assume that two different trajectories can never have the same state at the same time step [5, Sec. II.A]. This is actually an event of probability zero in most settings of interest, that is, when at least one variable of the single-target space has a continuous distribution, e,g., one variable takes values in $\mathbb{R}$.

Let $g\left(\mathbf{X}_{\alpha:\gamma}\right)$ be a real-valued function on a set of trajectories $\mathbf{X}_{\alpha:\gamma}$. Its corresponding set integral is [5], [11]

$$\int g\left(\mathbf{X}_{\alpha:\gamma}\right) \delta\mathbf{X}_{\alpha:\gamma} \triangleq \sum_{n=0}^{\infty} \frac{1}{n!} \int g(\{X^1, \ldots, X^n\})\mathrm{d}X^{1:n} \tag{7}$$

where $X^{1:n} = (X^1, \ldots, X^n)$.

A (multi-trajectory) density $f\left(\mathbf{X}_{\alpha:\gamma}\right)$, which represents the information on an RFS of trajectories, is non-negative and integrates to one using (7). For instance, a trajectory PPP has density

$$f\left(\mathbf{X}_{\alpha:\gamma}\right) = e^{-\int \lambda(X)\mathrm{d}X} \lambda^{\mathbf{X}_{\alpha:\gamma}}, \tag{8}$$

with intensity $\lambda(X)$, $\lambda^{\mathbf{X}_{\alpha:\gamma}} = \prod_{X\in\mathbf{X}_{\alpha:\gamma}} \lambda(X)$ if $\mathbf{X}_{\alpha:\gamma} \neq \emptyset$, and $\lambda^{\mathbf{X}_{\alpha:\gamma}} = 1$ if $\mathbf{X}_{\alpha:\gamma} = \emptyset$.

A trajectory Bernoulli process has a density

$$f\left(\mathbf{X}_{\alpha:\gamma}\right) = \begin{cases} 1 - r, & \mathbf{X}_{\alpha:\gamma} = \emptyset, \\ rf(X), & \mathbf{X}_{\alpha:\gamma} = \{X\}, \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

where $f(X)$ is a single trajectory density (4), and r is the probability of existence. In the same way, the densities of trajectory MBM and PMBM RFS are natural extensions of their definitions for RFS of targets.

IV. TPMBM FILTERS

In this section we present the prediction and update steps of the TPMBM filters for the two problem formulations. For the set of all trajectories, we seek the density for the RFS $\mathbf{X}_{0:k}$, in other words, all trajectories existing at some point in time between the initial time step 0 to the current time step k. For the set of current trajectories, we seek the density for the RFS $\mathbf{X}_{0:k}^{k} \triangleq \mathbf{X}_{0:k}^{k:k}$, see (6).

A. PMBM *density*

As in [6], we hypothesise that the set of trajectories density is a multi-target conjugate prior of the PMBM form, and we show that the PMBM form is preserved through prediction and update. The PPP represents trajectories that are hypothesised to exist at some time in the considered time interval, but have not been detected. The MBM contains information on the trajectories that have been detected at some point up to the current time step. Each MB in the MBM represents the information on the detected trajectories given a global data association hypothesis.

Using the shorthand notation $\mathbf{X}_k \triangleq \mathbf{X}_{0:k}$, the PMBM density is[2]

$$f_{k|k'}(\mathbf{X}_k) = \sum_{\mathbf{X}_k^{\mathrm{u}} \uplus \mathbf{X}_k^{\mathrm{d}} = \mathbf{X}_k} f_{k|k'}^{\mathrm{u}}(\mathbf{X}_k^{\mathrm{u}}) f_{k|k'}^{\mathrm{d}}(\mathbf{X}_k^{\mathrm{d}}) \tag{10a}$$

$$f_{k|k'}^{\mathrm{d}}(\mathbf{X}_k^{\mathrm{d}}) = \sum_{a_{k|k'} \in \mathcal{A}_{k|k'}} w_{k|k'}^{a} \sum_{\uplus_{i \in \mathbb{T}_{k|k'}} \mathbf{X}_k^i = \mathbf{X}_k^{\mathrm{d}}} \prod_{i \in \mathbb{T}_{k|k'}} f_{k|k'}^{i,a^i}(\mathbf{X}_k^i) \tag{10b}$$

where $\uplus$ is the union of disjoint sets, $f_{k|k'}^{\mathrm{u}}(\mathbf{X}_k^{\mathrm{u}})$ is a PPP density (8) with intensity $\lambda_{k|k'}^{\mathrm{u}}(X)$, $f_{k|k'}^{i,a^i}(\mathbf{X}_k^i)$ are Bernoulli densities (9) with probabilities of existence $r_{k|k'}^{i,a^i}$ and trajectory densities $f_{k|k'}^{i,a^i}(X)$. In the MBM density (10b), $\mathbb{T}_{k|k'}$ is a track table with $n_{k|k'}$ tracks indexed 1 to $n_{k|k'}$, and $a \in \mathcal{A}_{k|k'}$ is a global data association hypothesis. For each global hypothesis a, there is a track local hypothesis a^i for each track $i \in \mathbb{T}_{k|k'}$. The number of local hypotheses for track i is denoted as $h_{k|k'}^i$. The weight of the global hypothesis a is proportional to the product of the local hypothesis weights $w_{k|k'}^{i,a^i}$ for each track such that $w_{k|k'}^{a} \propto \prod_{i \in \mathbb{T}_{k|k'}} w_{k|k'}^{i,a^i}$.

Following [6], the set of all measurement indices up to time k is $\mathcal{M}^k$, and $\mathcal{M}^k(i, a^i)$ is the subset of indices that are associated to track i with local hypothesis a^i. The set of global data association hypotheses $\mathcal{A}_{k|k'}$ can be written in terms of $\mathcal{M}^k$, and $\mathcal{M}^k(i, a^i)$, see [6, Eq. 35].

A PMBM density (10) is then characterised by

$$\lambda_{k|k'}^{\mathrm{u}}(\cdot), \left\{ \left(w_{k|k'}^{i,a^i}, r_{k|k'}^{i,a^i}, f_{k|k'}^{i,a^i}(\cdot) \right) \right\}_{a \in \mathcal{A}_{k|k'},\ i \in \mathbb{T}_{k|k'}}, \tag{11}$$

The following subsections show the prediction and update steps of the trajectory PMBM density for the two considered problem formulations.

[2]For the sake for brevity, we use sub-scripts $_k$ and $_{k|k'}$ instead of sub-scripts $_{0:k\ 0:k|0:k'}$ to denote a density for trajectories in $\mathcal{T}_{0:k}$, conditioned on measurements in the interval $0:k'$.

B. Prediction step

In this section we describe the time evolution of the set of trajectories under Assumption 1. The birth PPP in terms of trajectories is

$$\lambda_k^{\mathrm{B}}(X) = \begin{cases} \lambda_k^{\mathrm{b}}(x_k) & (\beta, \varepsilon) = (k, k), \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

The trajectory state dependent probability of survival at time step k is $P_k^{\mathrm{S}}(X) = P^{\mathrm{S}}(x_\varepsilon)\Delta_k(\varepsilon)$, where $\Delta_k(\cdot)$ denotes Kronecker's delta function located at k.

1) Dynamic model for the set of all trajectories: Given a set of trajectories $\mathbf{X}_{0:k-1}$ with at maximum one trajectory, the Bernoulli transition density without birth is

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

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

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

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

where $\delta(\cdot)$ denotes the Dirac delta function. In this problem formulation, every trajectory that has ever existed remains in the set of all trajectories. The probability of survival $P_{k-1}^{\mathrm{S}}(\cdot)$ is used to determine whether the trajectory ends, or if it is extended by one more time step with a new state.

The prediction step is presented in the theorem below, where $\langle f; g \rangle = \int f(X)g(X)\mathrm{d}X$.

Theorem 1 (TPMBM prediction for all trajectories)**.** *If the posterior density of the set of all trajectories at the previous time step is a* PMBM *of the form* (10)*, the predicted density of the set of all trajectories is a* PMBM *of the form* (10) *with:*

$$\lambda_{k|k-1}^{\mathrm{u}}(X) = \lambda_k^{\mathrm{B}}(X) + \left\langle \lambda_{k-1|k-1}^{\mathrm{u}}; \pi^{\mathrm{a}} \right\rangle, \tag{14a}$$

$$n_{k|k-1} = n_{k-1|k-1}, \tag{14b}$$

$$h_{k|k-1}^i = h_{k-1|k-1}^i \ \forall\, i, \tag{14c}$$

$$w_{k|k-1}^{i,a^i} = w_{k-1|k-1}^{i,a^i} \ \forall\, i, a^i, \tag{14d}$$

$$r_{k|k-1}^{i,a^i} = r_{k-1|k-1}^{i,a^i}, \ \forall\, i, a^i, \tag{14e}$$

$$f_{k|k-1}^{i,a^i} = \left\langle f_{k-1|k-1}^{i,a^i}; \pi^{\mathrm{a}} \right\rangle, \ \forall\, i, a^i. \tag{14f}$$

Proof. Analogous to the PMBM prediction proof [6, Thm. 1] [22], using the transition density (13) and the PPP birth intensity (12). □

2) Dynamic model for the set of current trajectories: Given a set of trajectories $\mathbf{X}_{0:k-1}^{k-1}$ with at maximum one trajectory that is present at time step $k-1$, the Bernoulli transition density without birth is

$$f_{k|k-1}^{\mathrm{c}}(\mathbf{X}_{0:k}^{k}|\mathbf{X}_{0:k-1}^{k-1}) = \tag{15a}$$
$$\begin{cases} 1, & \mathbf{X}_{0:k-1}^{k-1}=\emptyset, \mathbf{X}_{0:k}^{k}=\emptyset, \\ 1-P_{k-1}^{\mathrm{S}}(X'), & \mathbf{X}_{0:k-1}^{k-1}=\{X'\}, \mathbf{X}_{0:k}^{k}=\emptyset, \\ P_{k-1}^{\mathrm{S}}(X')\pi^{\mathrm{c}}(X|X'), & \mathbf{X}_{0:k-1}^{k-1}=\{X'\}, \mathbf{X}_{0:k}^{k}=\{X\}, \\ 0, & \text{otherwise}, \end{cases}$$
$$\pi^{\mathrm{c}}(X|X') = \pi^{\mathrm{c},x}(x_{\beta:\varepsilon}|\beta,\varepsilon,X')\Delta_{\varepsilon'+1}(\varepsilon)\Delta_{\beta'}(\beta), \tag{15b}$$
$$\pi^{\mathrm{c},x}(x_{\beta:\varepsilon}|\beta,\varepsilon,X') = \pi^{x}(x_{\varepsilon}|x'_{\varepsilon'})\delta_{x'_{\beta':\varepsilon'}}(x_{\beta:\varepsilon-1}). \tag{15c}$$

In this problem formulation, if a trajectory dies, it will no longer be a element of the set of current trajectories. If the trajectory survives, then it is kept in the set of current trajectories and is extended by one time step. For the set of current trajectories, the probability of survival $P_{k-1}^{\mathrm{S}}(\cdot)$ is therefore used as in tracking a set of targets [11].

The prediction step is given in the following theorem.

Theorem 2 (TPMBM prediction for current trajectories)**.** *If the posterior density of the set of current trajectories at the previous time step is a* PMBM *of the form* (10)*, the predicted density of the set of current trajectories is a* PMBM *of the form* (10) *with:*

$$\lambda_{k|k-1}^{\mathrm{u}}(X) = \lambda_{k}^{\mathrm{B}}(X) + \left\langle \lambda_{k-1|k-1}^{\mathrm{u}}; \pi^{\mathrm{c}}P_{k-1}^{\mathrm{S}} \right\rangle, \tag{16a}$$
$$n_{k|k-1} = n_{k-1|k-1}, \tag{16b}$$
$$h_{k|k-1}^{i} = h_{k-1|k-1}^{i} \ \forall\, i, \tag{16c}$$
$$w_{k|k-1}^{i,a^i} = w_{k-1|k-1}^{i,a^i} \ \forall\, i, a^i, \tag{16d}$$
$$r_{k|k-1}^{i,a^i} = r_{k-1|k-1}^{i,a^i} \left\langle f_{k-1|k-1}^{i,a^i}; P_{k-1}^{\mathrm{S}} \right\rangle, \ \forall\, i, a^i, \tag{16e}$$
$$f_{k|k-1}^{i,a^i} = \frac{\left\langle f_{k-1|k-1}^{i,a^i}; \pi^{\mathrm{c}}P_{k-1}^{\mathrm{S}} \right\rangle}{\left\langle f_{k-1|k-1}^{i,a^i}; P_{k-1}^{\mathrm{S}} \right\rangle}, \ \forall\, i, a^i. \tag{16f}$$

Proof. Analogous to the PMBM prediction proof [6, Thm. 1] [22], using the transition density (15) and the PPP birth intensity (12). □

C. Update step

The measurement model is the same regardless of the problem formulation. Given a set of trajectories $\mathbf{X}$ with at most one trajectory, the Bernoulli measurement density for detections of Assumption 2 can be written in terms of $\mathbf{X}$ as:

$$\varphi_k(\mathbf{w}_k|\mathbf{X}) = \tag{17a}$$
$$\begin{cases} 1, & \mathbf{X}=\emptyset, \mathbf{w}_k=\emptyset, \\ 1-P_k^{\mathrm{D}}(X), & \mathbf{X}=\{X\}, \mathbf{w}_k=\emptyset, \\ P_k^{\mathrm{D}}(X)\varphi(z_k|X), & \mathbf{X}=\{X\}, \mathbf{w}_k=\{z_k\}, \\ 0, & \text{otherwise}, \end{cases}$$
$$P_k^{\mathrm{D}}(X) = P^{\mathrm{D}}(x_{\varepsilon})\Delta_k(\varepsilon), \tag{17b}$$
$$\varphi(z|X) = \varphi^{z}(z|x_{\varepsilon}). \tag{17c}$$

The measurement model for sets of trajectories is of the same form as for sets of targets, so the update is analogous to the target measurement update in [6], [22].

Theorem 3 (TPMBM update)**.** *Assume that the predicted density of the set of trajectories is a* PMBM *of the form* (10) *and the set of measurements is* $\{z_k^1,\dots,z_k^{m_k}\}$*. Then, the posterior density of the set of trajectories is a* PMBM *of the form* (10) *with* $n_{k|k} = n_{k|k-1} + m_k$*, and*

$$\lambda_{k|k}^{\mathrm{u}}(X) = \left(1-P_k^{\mathrm{D}}(X)\right)\lambda_{k|k-1}^{\mathrm{u}}(X), \tag{18}$$
$$\mathcal{M}^k = \mathcal{M}^{k-1} \cup \left\{(k,j)|j\in\{1,\dots,m_k\}\right\}. \tag{19}$$

For tracks continuing from previous time steps ($i \in \{1,\dots,n_{k|k-1}\}$)*, a local hypothesis is created for each combination of a previous local hypothesis and either a missed detection or an update with one of the* m_k *measurements. This implies that the number of local hypotheses is* $h_{k|k}^{i} = h_{k|k-1}^{i}(1+m_k)$*. For missed detection hypotheses* ($i \in \{1,\dots,n_{k|k-1}\}$, $a^i \in \{1,\dots,h_{k|k-1}^{i}\}$)*, the parameters are*

$$\mathcal{M}^k(i,a^i) = \mathcal{M}^{k-1}(i,a^i), \tag{20a}$$
$$w_{k|k}^{i,a^i} = w_{k|k-1}^{i,a^i}\left(1-r_{k|k-1}^{i,a^i}\left\langle f_{k|k-1}^{i,a^i}; P_k^{\mathrm{D}}\right\rangle\right), \tag{20b}$$
$$r_{k|k}^{i,a^i} = \frac{r_{k|k-1}^{i,a^i}\left\langle f_{k|k-1}^{i,a^i}; 1-P_k^{\mathrm{D}}\right\rangle}{1-r_{k|k-1}^{i,a^i}\left\langle f_{k|k-1}^{i,a^i}; P_k^{\mathrm{D}}\right\rangle}, \tag{20c}$$
$$f_{k|k}^{i,a^i}(X) = \frac{\left(1-P_k^{\mathrm{D}}(X)\right)f_{k|k-1}^{i,a^i}(X)}{\left\langle f_{k|k-1}^{i,a^i}; 1-P_k^{\mathrm{D}}\right\rangle}. \tag{20d}$$

For a previous local hypothesis $\tilde{a}^i \in \{1,\dots,h_{k|k-1}^{i}\}$ *of an existing track* ($i \in \{1,\dots,n_{k|k-1}\}$)*, the local hypothesis generated by the association with measurement* z_k^j *has* $a^i = \tilde{a}^i + h_{k|k-1}^{i}j$ *and*

$$\mathcal{M}^k(i,a^i) = \mathcal{M}^{k-1}(i,\tilde{a}^i) \cup \{(k,j)\}, \tag{21a}$$
$$w_{k|k}^{i,a^i} = w_{k|k-1}^{i,\tilde{a}^i} r_{k|k-1}^{i,\tilde{a}^i}\left\langle f_{k|k-1}^{i,\tilde{a}^i}; \varphi(z_k^j|\cdot)P_k^{\mathrm{D}}\right\rangle, \tag{21b}$$
$$r_{k|k}^{i,a^i} = 1, \tag{21c}$$
$$f_{k|k}^{i,a^i}(X) = \frac{\varphi(z_k^j|X)P_k^{\mathrm{D}}(X)f_{k|k-1}^{i,\tilde{a}^i}(X)}{\left\langle f_{k|k-1}^{i,\tilde{a}^i}; \varphi(z_k^j|\cdot)P_k^{\mathrm{D}}\right\rangle}. \tag{21d}$$

Finally, the new track initiated by measurement z_k^j *has* $i = n_{k|k-1}+j$*,* $h_{k|k}^{i} = 2$*, and*

$$\mathcal{M}^k(i,1) = \emptyset, \quad w_{k|k}^{i,1} = 1, \quad r_{k|k}^{i,1} = 0, \tag{22a}$$
$$\mathcal{M}^k(i,2) = \{(t,j)\} \tag{22b}$$
$$w_{k|k}^{i,2} = \lambda^{\mathrm{FA}}(z_k^j) + \left\langle \lambda_{k|k-1}^{\mathrm{u}}; \varphi(z_k^j|\cdot)P_k^{\mathrm{D}}\right\rangle, \tag{22c}$$
$$r_{k|k}^{i,2} = \frac{\left\langle \lambda_{k|k-1}^{\mathrm{u}}; \varphi(z_k^j|\cdot)P_k^{\mathrm{D}}\right\rangle}{\lambda^{\mathrm{FA}}(z_k^j) + \left\langle \lambda_{k|k-1}^{\mathrm{u}}; \varphi(z_k^j|\cdot)P_k^{\mathrm{D}}\right\rangle}, \tag{22d}$$
$$f_{k|k}^{i,2}(X) = \frac{\varphi(z_k^j|X)P_k^{\mathrm{D}}(X)\lambda_{k|k-1}^{\mathrm{u}}(X)}{\left\langle \lambda_{k|k-1}^{\mathrm{u}}; \varphi(z_k^j|\cdot)P_k^{\mathrm{D}}\right\rangle}. \tag{22e}$$

Proof. Analogous to the PMBM update proof [6, Thm. 2] [22] considering the measurement model (17). □

D. Properties of the resulting TPMBM filters

Two TPMBM filters are obtained from the theorems:

1) A TPMBM filter for the set of all trajectories (PF 1). Its filtering recursion is provided by the prediction in Theorem 1 and the update in Theorem 3.
2) A TPMBM filter for the set of current trajectories (PF 2). Its filtering recursion is provided by the prediction in Theorem 2 and the update in Theorem 3.

Both TPMBM filters are track oriented. For each measurement, a potential new track is initiated, see (22). In the update, additional hypotheses are created, as indicated in (20) and (21). In the prediction, the number of tracks and hypotheses remains constant, see (16b) and (16c).

We proceed to discuss the representation of the PPP intensity and the Bernoulli densities in the TPMBM filters. We also present relevant properties of the pmf of the start time step and the pmf of the end time step for each Bernoulli trajectory.

1) Density/intensity representation: Consider a trajectory mixture density of the form

$$f(X) = \sum_{j\in\mathbb{J}} \nu^j f^j(X;\theta^j), \tag{23a}$$

$$f^j(X;\theta^j) = \begin{cases} p^j(x_{b^j:e^j}) & (\beta,\varepsilon) = (b^j, e^j), \\ 0 & \text{otherwise}, \end{cases} \tag{23b}$$

$$\theta^j = \left(b^j, e^j, p^j(\cdot)\right), \tag{23c}$$

where $\mathbb{J}$ is an index set, and each mixture component is characterised by a weight ν^j and a parameter θ^j. The parameter consists of a birth time b^j, a most recent time e^j, where $b^j \leq e^j$, and a density $p^j(x_{b^j:e^j})$. For the density (23a), the pmf of (β,ε) is

$$P(\beta,\varepsilon) = \int f(X)\mathrm{d}x_{\beta:\varepsilon} = \sum_{j\in\mathbb{J}} \nu^j \Delta_{b^j}(\beta)\Delta_{e^j}(\varepsilon). \tag{23d}$$

For the mixture weights, it holds that $\sum_j \nu^j = 1$ if $f(X)$ is a density, and $\sum_j \nu^j \geq 0$ if $f(X)$ is a PPP intensity. Note that there is no restriction in (23) that b^j and e^j must be unique, i.e., we may have $b^j = b^{j'}$ and/or $e^j = e^{j'}$ for $j, j' \in \mathbb{J}$, $j \neq j'$. Densities/intensities of the form (23) facilitate representations for the state sequence $x_{\beta:\varepsilon}$, conditioned on β and ε.

The target birth PPP intensity $\lambda_k^{\mathrm{b}}(x_k)$ is often modelled as an unnormalised distribution mixture, often a Gaussian mixture. It then follows that the trajectory birth PPP intensity $\lambda_k^{\mathrm{B}}(X)$, cf. (12), is of the form (23), with the special structure that $b^j = e^j = k$. From this, it follows that the PPP intensity $\lambda_{k|k'}^{\mathrm{u}}(X)$, and all Bernoulli densities $f_{k|k'}^{i,a^i}(X)$ will be of the form (23).

In other words, the parameters of the TPMBM posterior are

$$\lambda^{\mathrm{u}}(\cdot), \left\{\left(w^{i,a^i}, r^{i,a^i}, f^{i,a^i}(\cdot)\right)\right\}_{a\in\mathcal{A},\ i\in\mathbb{T}}, \tag{24a}$$

with intensity and state densities of the form (23),

$$\lambda^{\mathrm{u}}(X) = \sum_{j\in\mathbb{J}^{\mathrm{u}}} \nu^{\mathrm{u},j} f^{\mathrm{u},j}(X;\theta^{\mathrm{u},j}), \tag{24b}$$

$$f^{i,a^i}(X) = \sum_{j\in\mathbb{J}^{i,a^i}} \nu^{i,a^i,j} f^{i,a^i,j}(X;\theta^{i,a^i,j}), \tag{24c}$$

where $\mathbb{J}^{\mathrm{u}}$ and $\mathbb{J}^{i,a^i}$ are index sets for the mixture density components, and $\nu^{\mathrm{u},j}$ and $\nu^{i,a^i,j}$ are weights such that $\sum_{j\in\mathbb{J}^{\mathrm{u}}} \nu^{\mathrm{u},j} \geq 0$ and $\sum_{j\in\mathbb{J}^{i,a^i}} \nu^{i,a^i,j} = 1$.

2) Properties of the pmf of the time of birth: Consider a data association in which a measurement z_k at time step k is associated to a potential new target. Conditioned on the association, for the trajectory end time ε, we have that $\Pr(\varepsilon = k) = 1$. We proceed to focus on the time of birth pmf. The new Bernoulli track density is of the form (22e),

$$f_{k|k}(X) = \frac{\varphi(z_k|X)P_k^{\mathrm{D}}(X)\lambda_{k|k-1}^{\mathrm{u}}(X)}{\left\langle \lambda_{k|k-1}^{\mathrm{u}}; \varphi(z_k|\cdot)P_k^{\mathrm{D}}\right\rangle}. \tag{25}$$

With a PPP intensity of the mixture form (23),

$$\lambda_{k|k-1}^{\mathrm{u}}(X) = \sum_{j\in\mathbb{J}_{k|k-1}^{\mathrm{u}}} \nu_{k|k-1}^{\mathrm{u},j} f_{k|k-1}^{\mathrm{u},j}(X;\theta_{k|k-1}^{\mathrm{u},j}), \tag{26}$$

we get a multi-modal posterior trajectory density $f_{k|k}(X)$, which can be pruned if necessary. The pmf for β is

$$P_{k|k}(\beta) = \begin{cases} \frac{\sum_{j\in\mathbb{J}_{k|k-1}^{\mathrm{u},\varepsilon=k}} \nu_{k|k-1}^{\mathrm{u},j} q_k^{\mathrm{u},j}(z_k)\Delta_{b_{k|k-1}^{\mathrm{u},j}}(\beta)}{\sum_{j\in\mathbb{J}_{k|k-1}^{\mathrm{u},\varepsilon=k}} \nu_{k|k-1}^{\mathrm{u},j} q_k^{\mathrm{u},j}(z_k)} & \beta \leq k, \\ 0 & \beta > k, \end{cases} \tag{27}$$

where $\mathbb{J}_{k|k-1}^{\mathrm{u},\varepsilon=k} = \left\{j \in \mathbb{J}_{k|k-1}^{\mathrm{u}} : e_{k|k-1}^{\mathrm{u},j} = k\right\}$, and

$$q_k^{\mathrm{u},j}(z_k) = \int \varphi^z(z_k|x_k)P_k^{\mathrm{D}}(x_k)p_{k|k-1}^{\mathrm{u},j}(x_k)\,\mathrm{d}x_k. \tag{28}$$

This implies that the time of birth of a newly detected trajectory is not deterministic, but has a certain pmf. In addition, as more measurements are associated, the trajectory density changes, which implies that the pmf of the time of birth can also change.

3) Properties of the pmf of the trajectory end time: Consider a posterior Bernoulli density $f_{k|k}(X)$ of the form (23), to which a measurement was associated at time k. For the sake of brevity, assume that it has a single mixture component with parameter $\theta = \left(b, k, p_{k|k}(\cdot)\right)$. In addition, the probability of survival is a constant $P^{\mathrm{S}}(x_k) = P^{\mathrm{S}}$, and $Q^{\mathrm{S}} = 1 - P^{\mathrm{S}}$. Then, the predicted density at time $k+1$ is a mixture

$$f_{k+1|k}(X_{k+1}) = Q^{\mathrm{S}} f^0(X_{k+1};\theta^0) + P^{\mathrm{S}} f^1(X_{k+1};\theta^1), \tag{29a}$$

$$\theta^0 = \left(b, k, p_{k|k}(x_{b:k})\right), \tag{29b}$$

$$\theta^1 = \left(b, k+1, p_{k|k}(x_{b:k})\pi^x(x_{k+1}|x_k)\right). \tag{29c}$$

Note that the state sequence density for $\varepsilon = k$, (29b), is given by marginalising x_{k+1} from the state sequence density for $\varepsilon = k+1$, (29c). This has important implications for

the implementation of the TPMBM filter for the set of all trajectories: during the prediction step, the hypothesis space for the trajectory density increases, due to the fact that we do not know if the trajectory ended at time k, (29b), or continued to time $k+1$, (29c). However, it is not necessary to explicitly represent both state sequence densities, instead it is sufficient to explicitly represent the state sequence density that continued to time $k+1$, as well as a pmf $P_{k+1|k}(\varepsilon)$. This is especially important when there are several consecutive misdetections associated: for N consecutive misdetections, with a single pmf $P_{k|k'}(\varepsilon)$ and a single density $p_{k|k'}(x_{b:k})$ we can compactly represent $N+1$ different hypotheses for ε and $x_{b:\varepsilon}$. Similar observations were made for MHT in [19, Sec. IV].

We now also assume that $P^{\mathrm{D}}(x_k) = P^{\mathrm{D}}$, with $Q^D = 1 - P^{\mathrm{D}}$, as this facilitates the closed-form expression for the pmf of the time of the latest state $P_{k|k}(\varepsilon)$. Consider a Bernoulli density where τ is the last time step that a measurement was associated. Given this association, we have that $\Pr(\varepsilon \geq \tau) = 1$. Then, at time $k > \tau$, if $\phi = Q^{\mathrm{D}} P^{\mathrm{S}} < 1$,[3] the posterior pmf of ε is

$$P_{k|k}(\varepsilon) = \begin{cases} 0 & \varepsilon < \tau \text{ or } \varepsilon > k, \\ \frac{1}{C} Q^{\mathrm{S}} \phi^i & \varepsilon = \tau + i, i \in \mathbb{N}_0^{k-\tau-1}, \\ \frac{1}{C} \phi^{k-\tau} & \varepsilon = k, \end{cases} \tag{30}$$

where $C = Q^{\mathrm{S}} \left(1 - \phi^{k-\tau}\right) (1-\phi)^{-1} + \phi^{k-\tau}$. When only misdetections are associated, in the limit, the pmf of ε is

$$\lim_{k\to\infty} P_{k|k}(\varepsilon) = \begin{cases} 0 & \varepsilon < \tau, \\ \phi^i (1-\phi) & \varepsilon = \tau + i, i \geq 0, \end{cases} \tag{31}$$

which is the pmf of a geometric distribution with success probability $1 - \phi$. From this it follows that, eventually, for the given association (only misdetections after time step τ), the maximum a posteriori (MAP) estimate of ε is τ and the expected value of ε is $\tau + \frac{\phi}{1-\phi}$, which can be rounded to the nearest integer. Note that $\Pr(\varepsilon > \tau) = \phi$, which may be significant for high P^{S} and/or low P^{D}. The pmf $P_{k|k}(\varepsilon)$ can also be computed recursively.

E. Estimator

We proceed to explain how to estimate the set of trajectories.

1) Single trajectory estimator: Given a trajectory density $f(X) = p(x_{\beta:\varepsilon}|\beta,\varepsilon)P(\beta,\varepsilon)$, in this work, we obtain a trajectory estimate $\hat{X} = (\hat{\beta},\ \hat{\varepsilon},\ \hat{x}_{\hat{\beta}:\hat{\varepsilon}})$ as follows. The MAP estimates of the trajectory start and end times are

$$\left(\hat{\beta}, \hat{\varepsilon}\right) = \arg\max_{(\beta,\varepsilon)} P(\beta,\varepsilon). \tag{32}$$

For trajectory densities of the form (23), this corresponds to finding the component with maximum weight. The trajectory estimate from the estimated birth time to the estimated most recent state time is computed as the conditional expectation

$$\hat{x}_{\hat{\beta}:\hat{\varepsilon}} = \mathbb{E}\left[x_{\hat{\beta}:\hat{\varepsilon}} \middle| \hat{\beta}, \hat{\varepsilon}\right]. \tag{33}$$

[3] $Q^{\mathrm{D}} P^{\mathrm{S}} = 1 \Rightarrow P^{\mathrm{S}} = 1,\ P^{\mathrm{D}} = 0$. Note that MTT without detections ($P^{\mathrm{D}} = 0$) is a problem of little interest.

2) Multiple trajectory estimator: This paper uses the following estimator of the set of trajectories given the TPMBM posterior density. We first select the global hypothesis with the highest weight

$$a_\star = \arg\max_a w^a_{k|k}. \tag{34a}$$

Then, the estimated set of trajectories is

$$\hat{\mathbf{X}}_{k|k} = \left\{ \hat{X}^{i,a^i_\star}_{k|k} \right\}_{i \in \mathbb{T}_{k|k} \,:\, r^{i,a^i_\star}_{k|k} > r^{\mathrm{e}}}, \tag{34b}$$

where the detection threshold r^{e} on the probability of existence is a parameter, and the trajectory estimate $\hat{X}^{i,a^i_\star}_{k|k}$ is computed as in Section IV-E1. The estimator (34) corresponds to PMBM Estimator 1 [22, Sec. VI.A].

V. Properties of PMBM densities on sets of trajectories

This section provides some properties on PMBM densities on sets of trajectories. Section V-A states the time marginalisation theorem. Section V-B provides a result analogous to (2) for PMBM densities. Finally, Section V-C establishes relations between the TPMBM filters and the PMBM filter.

A. TPMBM time marginalisation

Given a density on a set of trajectories defined on a time interval, time marginalisation is the process of marginalising states for particular times from the trajectories, such that only states related to a different time interval remain.

We use $\int \mathrm{d}x_{\beta:\varepsilon \setminus b:e}$ to denote the integral of the states in the sequence $x_{\beta:\varepsilon}$ that are not in the interval $b:e$, with $\beta \leq b \leq e \leq \varepsilon$. The following theorem explains the time marginalisation operation for PMBM densities on sets of trajectories.

Theorem 4. *Given a posterior* PMBM *density, with parameters* (24)*, for the set of all trajectories at time k, $\mathbf{X}_{0:k}$, the posterior density for the set of trajectories in the interval $\alpha:\gamma$ whose trajectories were present in the surveillance area at least in one time step in the interval $\eta:\zeta$, $\mathbf{X}^{\eta:\zeta}_{\alpha:\gamma}$ with $\mathbb{N}^\zeta_\eta \subseteq \mathbb{N}^\gamma_\alpha \subseteq \mathbb{N}^k_0$, is a* PMBM *characterised by*

$$\tilde{\lambda}^{\mathrm{u}}(\cdot), \left\{ \left(w^{i,a^i}, \tilde{r}^{i,a^i}, \tilde{f}^{i,a^i}(\cdot) \right) \right\}_{a \in \mathcal{A},\ i \in \mathbb{T}:\tilde{r}^{i,a^i} \neq 0} \tag{35a}$$

with PPP *intensity*

$$\tilde{\lambda}^{\mathrm{u}}(X) = \sum_{j \in \mathbb{J}^{\mathrm{u},\eta:\zeta}} \nu^{\mathrm{u},j} \tilde{f}^{\mathrm{u},j}(X; \tilde{\theta}^{\mathrm{u},j}), \tag{35b}$$

$$\mathbb{J}^{\mathrm{u},\eta:\zeta} = \left\{ j \in \mathbb{J}^{\mathrm{u}} : \mathbb{N}^{e^{\mathrm{u},j}}_{b^{\mathrm{u},j}} \cap \mathbb{N}^\zeta_\eta \neq \emptyset \right\}, \tag{35c}$$

$$\tilde{\theta}^{\mathrm{u},j} = \left(\tilde{b}^{\mathrm{u},j}, \tilde{e}^{\mathrm{u},j}, \tilde{p}^{\mathrm{u},j}(\cdot) \right), \tag{35d}$$

$$\tilde{b}^{\mathrm{u},j} = \max\left(b^{\mathrm{u},j}, \alpha\right), \quad \tilde{e}^{\mathrm{u},j} = \min\left(e^{\mathrm{u},j}, \gamma\right), \tag{35e}$$

$$\tilde{p}^{\mathrm{u},j}(x_{\tilde{b}^{\mathrm{u},j}:\tilde{e}^{\mathrm{u},j}}) = \int p^{\mathrm{u},j}(x_{b^{\mathrm{u},j}:e^{\mathrm{u},j}}) \mathrm{d}x_{b^{\mathrm{u},j}:e^{\mathrm{u},j} \setminus \tilde{b}^{\mathrm{u},j}:\tilde{e}^{\mathrm{u},j}}, \tag{35f}$$

and Bernoulli parameters

$$\tilde{r}^{i,a^i} = r^{i,a^i} \sum_{j \in \mathbb{J}^{i,a^i,\eta:\zeta}} \nu^{i,a^i,j}, \tag{35g}$$

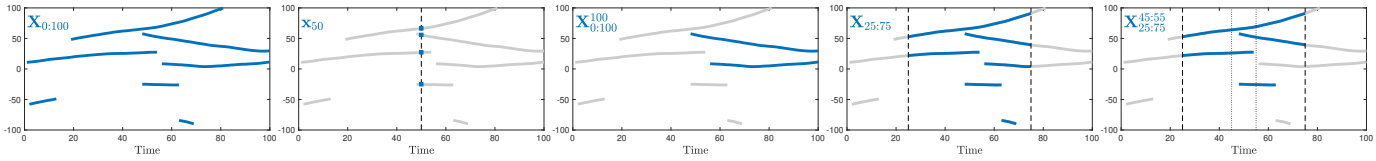


Fig. 1. Illustration of the trajectories considered in Theorem 4. Left to right: set of all (1-D) trajectories up to the current time step 100, $\mathbf{X}_{0:100}$; corresponding target set at time step 50, $\mathbf{x}_{50}$; corresponding set of current trajectories $\mathbf{X}^{100}_{0:100}$; corresponding set of trajectories in the time interval 25 to 75, $\mathbf{X}_{25:75}$; corresponding set of trajectories in the time interval 25 to 75 that alive at some time in the interval 45 to 55, $\mathbf{X}^{45:55}_{25:75}$. If $\mathbf{X}_{0:100}$ is PMBM distributed, so are the rest of the sets of trajectories.

$$\tilde{f}^{i,a^i}(X) = \frac{\sum_{j\in\mathbb{J}^{i,a^i,\eta:\zeta}} \nu^{i,a^i,j} \tilde{f}^{i,a^i,j}(X;\tilde{\theta}^{i,a^i,j})}{\sum_{j\in\mathbb{J}^{i,a^i,\eta:\zeta}} \nu^{i,a^i,j}}, \tag{35h}$$

$$\mathbb{J}^{i,a^i,\eta:\zeta} = \left\{ j \in \mathbb{J}^{i,a^i} : \mathbb{N}^{e^{i,a^i,j}}_{b^{i,a^i,j}} \cap \mathbb{N}^{\zeta}_{\eta} \neq \emptyset \right\}, \tag{35i}$$

$$\tilde{\theta}^{i,a^i,j} = \left(\tilde{b}^{i,a^i,j}, \tilde{e}^{i,a^i,j}, \tilde{p}^{i,a^i,j}(\cdot) \right), \tag{35j}$$

$$\tilde{b}^{i,a^i,j} = \max\left(b^{i,a^i,j}, \alpha\right), \tilde{e}^{i,a^i,j} = \min\left(e^{i,a^i,j}, \gamma\right), \tag{35k}$$

$$\tilde{p}^{i,a^i,j}(x_{\tilde{b}^{i,a^i,j}:\tilde{e}^{i,a^i,j}}) = \int p^{i,a^i,j}(x_{b^{i,a^i,j}:e^{i,a^i,j}}) \mathrm{d}x_{b^{i,a^i,j}:e^{i,a^i,j}\setminus\tilde{b}^{i,a^i,j}:\tilde{e}^{i,a^i,j}}. \tag{35l}$$

Proof. See Appendix A. □

Theorem 4 is understood as follows: from (35c)/(35i), we obtain the densities of the trajectories that are alive in the time interval $\eta:\zeta$. From (35e)/(35k), we get the time of birth and time of most recent state parameters, and from (35f)/(35l), we get the state sequence densities corresponding to $\alpha:\gamma$. Some examples of time marginalisation are illustrated in Fig. 1.

B. TPMBM density for arbitrary time intervals

It follows from Theorem 1 and Theorem 3 that the density

$$f_{k|k'}(\mathbf{X}_{0:k}) \triangleq f_{0:k|0:k'}(\mathbf{X}_{0:k}) \tag{36}$$

is PMBM, regardless if we are doing prediction ($k > k'$) or filtering ($k = k'$). In other words, in terms of prediction and filtering, we have a result that is analogous to (1) for the Gaussian distribution in linear Gaussian models. The following theorem presents more general results.

Theorem 5. *Given the measurements in any time interval $\xi:\chi$, the density of the set of trajectories in any time interval $\alpha:\gamma$ that are alive at any point in the time interval $\eta:\zeta$, such that $\mathbb{N}^{\zeta}_{\eta} \subseteq \mathbb{N}^{\gamma}_{\alpha}$, is* PMBM. *This density is denoted by $f^{\eta:\zeta}_{\alpha:\gamma|\xi:\chi}\left(\mathbf{X}^{\eta:\zeta}_{\alpha:\gamma}\right)$.*

The following corollary to Theorem 5 is analogous to (2).

Corollary 1. *If $\eta = \alpha$ and $\zeta = \gamma$ in Theorem 5, then*

$$f_{\alpha:\gamma|\xi:\chi}\left(\mathbf{X}_{\alpha:\gamma}\right) \tag{37}$$

is PMBM *for any $\alpha:\gamma$ and $\xi:\chi$.*

We prove Theorem 5 in Appendix A, making use of the time marginalisation theorem for TPMBM densities (Theorem 4). To illustrate Theorem 5, let us consider for example the sets $\mathbf{X}_{0:100}$, $\mathbf{x}_{50}$, and $\mathbf{X}^{45:55}_{25:75}$ in Figure 1. Theorem 5 indicates that these sets are PMBM distributed given the measurements in any time interval.

C. Relation between the TPMBM filters and the PMBM filter

Theorems 4 and 5 provide a way to compute the density $f^{\eta:\zeta}_{\alpha:\gamma|\xi:\chi}\left(\mathbf{X}^{\eta:\zeta}_{\alpha:\gamma}\right)$: first use the TPMBM filter for the set of all trajectories, and then use time marginalisation to obtain the desired density. However, we note that the application of the TPMBM filter followed by time marginalisation is possibly not the only way to compute the density $f^{\eta:\zeta}_{\alpha:\gamma|\xi:\chi}\left(\mathbf{X}^{\eta:\zeta}_{\alpha:\gamma}\right)$. Compare to the Gaussian result in (2): it is possible to compute the Gaussian density $p(x_k|z_{0:K})$ by computing the joint density $p(x_{0:K}|z_{0:K})$ and then marginalizing, however, one can also use forward-backward RTS-smoothing, see, e.g., [7].

Using Theorem 4 for $\alpha = \gamma = \eta = \zeta = k$, the relations between the two TPMBM filters, and between the TPMBM filters and the PMBM filter [6], become clear. This is illustrated in the following diagram indicating how several densities can be obtained by applying different theorems:

$$\begin{array}{ccccc} f_{k|k}(\mathbf{X}_{0:k}) & \overset{\text{Thm. 1}}{\rightarrow} & f_{k+1|k}(\mathbf{X}_{0:k+1}) & \overset{\text{Thm. 3}}{\rightarrow} & f_{k+1|k+1}(\mathbf{X}_{0:k+1}) \\ \downarrow \text{Thm. 4} & & \downarrow \text{Thm. 4} & & \downarrow \text{Thm. 4} \\ f_{k|k}(\mathbf{x}_k) & \overset{\text{[6, Thm. 1]}}{\rightarrow} & f_{k+1|k}(\mathbf{x}_{k+1}) & \overset{\text{[6, Thm. 2]}}{\rightarrow} & f_{k+1|k+1}(\mathbf{x}_{k+1}) \end{array}$$

It can be seen that marginalising out the past states in the TPMBM filter posterior, we obtain the PMBM filter posterior [6]. Predicting and updating the PMBM posterior yields the same result as predicting and updating the TPMBM filter, and then marginalising out past states. Thus, the hypotheses in the PMBM filter are implicitly operating on the latest time step of the trajectory hypotheses in the TPMBM filter. Then, the target estimates of the PMBM filter corresponding to the same Bernoulli component (created by a specific measurement) are implicitly estimating the current state of the associated trajectory in the TPMBM filter. This indicates that we can sequentially connect target state estimates at different time steps to estimate trajectories using PMBM filter metadata. This sequential track building procedure for the PMBM filter is analogous to linking target state estimates with the same auxiliary variable, which can be used to mark each track (Bernoulli component) [37]. It should anyway be noted that this sequential procedure to estimate the set of trajectories is sub-optimal. The optimal estimation of the set of trajectories must be done directly from the TPMBM posterior density.

VI. Linear Gaussian implementation

Let $\mathcal{X} = \mathbb{R}^{n_x}$, $\mathcal{Z} = \mathbb{R}^{n_z}$, and let the transition density and measurement model be linear and Gaussian,

$$\pi^x(x|x') = \mathcal{N}\left(x\,;\, Fx', Q\right), \tag{38a}$$

$$\varphi^z(z|x) = \mathcal{N}(z\,;\, Hx, R)\,, \tag{38b}$$

where $F \in \mathbb{R}^{n_x \times n_x}$ is a state transition matrix, $H \in \mathbb{R}^{n_z \times n_x}$ is a measurement matrix, $Q \in \mathbb{S}_{n_x}^{++}$ and $R \in \mathbb{S}_{n_z}^{++}$ are the covariance matrices of the process noise and measurement noise, respectively, and $\mathbb{S}_d^{++}$ denotes the space of positive definite matrices of size $d \times d$.

For linear and Gaussian models (38) we present three representations of the state sequence density. Let $m_{k|k}^{[a]}$ and $P_{k|k}^{[a]}$ be the parts of the mean and the covariance matrix for time step a. Let $P_{k|k}^{[a:b,c:d]}$ denote the part of the covariance matrix with rows for time steps a to b and columns for time steps c to d. The predictions and updates can be generalised to non-linear models by linearization [7], [53].

A. Gaussian moment form

The state sequence density can be expressed on Gaussian moment form, with mean $m_{k|k'}$ and covariance $P_{k|k'}$,

$$p_{k|k'}(x_{\beta:k}) = \mathcal{N}\left(x_{\beta:k}\ ;\ m_{k|k'}, P_{k|k'}\right), \tag{39}$$

1) Prediction: Given the motion model (38) and posterior parameters $m_{k|k}$ and $P_{k|k}$, the predicted parameters are

$$m_{k+1|k} = \begin{bmatrix} m_{k|k} \\ F m_{k|k}^{[k]} \end{bmatrix}, \tag{40a}$$

$$P_{k+1|k} = \begin{bmatrix} P_{k|k} & P_{k|k}^{[\beta:k,k]} F^{\mathrm{T}} \\ F P_{k|k}^{[k,\beta:k]} & F P_{k|k}^{[k]} F^{\mathrm{T}} + Q \end{bmatrix}. \tag{40b}$$

2) Update: Given the measurement model (38), predicted parameters $m_{k+1|k}$ and $P_{k+1|k}$, and an associated detection z, the posterior parameters are

$$m_{k+1|k+1} = m_{k+1|k} + K(z - H m_{k+1|k}^{[k]}), \tag{41a}$$

$$P_{k+1|k+1} = P_{k+1|k} - K H P_{k+1|k}^{[k,\beta:k]}, \tag{41b}$$

$$K = P_{k+1|k}^{[\beta:k,k]} H^{\mathrm{T}} (H P_{k+1|k}^{[k]} H^{\mathrm{T}} + R)^{-1}. \tag{41c}$$

B. Gaussian information form

The density (39) can be expressed in information form with information vector $y_{k|k'}$ and information matrix $Y_{k|k'}$, such that $y_{k|k'} = P_{k|k'}^{-1} m_{k|k'}$ and $Y_{k|k'} = P_{k|k'}^{-1}$ [53].

1) Prediction: Given the motion model (38) and posterior parameters $y_{k|k}$ and $Y_{k|k}$, the predicted parameters are [53]

$$y_{k+1|k} = \begin{bmatrix} y_{k|k} \\ 0_{n_x \times 1} \end{bmatrix}, \tag{42a}$$

$$Y_{k+1|k} = \begin{bmatrix} Y_{k|k}^{[\beta:k-1]} & Y_{k|k}^{[\beta:k-1,k]} & 0_{(\ell-1)n_x \times n_x} \\ Y_{k|k}^{[k,\beta:k-1]} & Y_{k|k}^{[k]} + F^T Q^{-1} F & -F^T Q^{-1} \\ 0_{n_x \times (\ell-1)n_x} & -Q^{-1} F & Q^{-1} \end{bmatrix}, \tag{42b}$$

where $0_{m \times n}$ is an m by n all-zero matrix.

2) Update: Given the measurement model (38), predicted parameters $y_{k+1|k}$ and $Y_{k+1|k}$, and an associated detection z, the posterior parameters are [53]–[55]

$$y_{k+1|k+1} = y_{k+1|k} + \begin{bmatrix} 0_{(\ell-1)n_x \times 1} \\ H^T R^{-1} z \end{bmatrix}, \tag{43a}$$

$$Y_{k+1|k+1} = Y_{k+1|k} + \begin{bmatrix} 0_{(\ell-1)n_x \times (\ell-1)n_x} & 0_{(\ell-1)n_x \times n_x} \\ 0_{n_x \times (\ell-1)n_x} & H^T R^{-1} H \end{bmatrix}. \tag{43b}$$

C. Gaussian L-scan approximation

In the L-scan approximation, the covariance matrix of (39) is approximated as [56, Sec. VI.C]

$$P_{k|k'} = \operatorname{diag}\left(P_{k|k'}^{[\beta]},\ \ldots, P_{k|k'}^{[\varepsilon-L]},\ P_{k|k'}^{[\varepsilon-L+1:\varepsilon]}\right). \tag{44}$$

In other words, states before the last L time steps are assumed independent of the most recent L states.

1) Prediction: Given the motion model (38) and posterior parameters $m_{k|k}$ and $P_{k|k}$, the predicted parameters are

$$m_{k+1|k} = \begin{bmatrix} m_{k|k} \\ F m_{k|k}^{[k]} \end{bmatrix}, \tag{45a}$$

$$P_{k+1|k} = \operatorname{diag}\left(P_{k|k}^{[\beta]},\ \ldots,\ P_{k|k}^{[k-L]},\ P_{k|k}^{[k-L+1]},\ \Psi\right), \tag{45b}$$

$$\Psi = \begin{bmatrix} P_{k|k}^{[k-L+2:k]} & P_{k|k}^{[k-L+2:k,k]} F^{\mathrm{T}} \\ F P_{k|k}^{[k,k-L+2:k]} & F P_{k|k}^{[k]} F^{\mathrm{T}} + Q \end{bmatrix}. \tag{45c}$$

2) Update: Given the measurement model (38), predicted parameters $m_{k+1|k}$ and $P_{k+1|k}$ and an associated detection z, the posterior parameters are

$$m_{k+1|k+1} = \begin{bmatrix} m_{k+1|k}^{[\beta:k-L]} \\ m_{k+1|k}^{[k-L+1:k]} + K(z - H m_{k+1|k}^{[k]}) \end{bmatrix}, \tag{46a}$$

$$P_{k+1|k+1}^{[\beta:k-L]} = P_{k+1|k}^{[\beta:k-L]}, \tag{46b}$$

$$P_{k+1|k+1}^{[k-L+1:k]} = P_{k+1|k}^{[k-L+1:k]} - K H P_{k+1|k}^{[k,k-L+1:k]}, \tag{46c}$$

$$K = P_{k+1|k}^{[k-L+1:k,k]} H^{\mathrm{T}} (H P_{k+1|k}^{[k,k]} H^{\mathrm{T}} + R)^{-1}. \tag{46d}$$

D. Discussion

In all three state sequence predictions, (40), (42) and (45), the mean/covariance, or information vector/matrix, are augmented to account for the additional time step that, following the prediction, is represented by the trajectory. The moment form update (41) and the L-scan update (46) affect all states in the state sequence, and the L latest states, respectively. In comparison, the information form update (43) only affects the parts of the information vector and information matrix corresponding to the single most recent state.

For the information form, a key result is that the bottom left and top right corners of the predicted information matrix (42b) are exactly zero, and that the update (43b) only affects the part of the information matrix related to the current state. This means that the information matrix is sparse [8].

In theory, when using the Gaussian information form, computing the weights (21b) and (22c) as well as an estimate $\hat{x}_{\hat{\beta}:\hat{\varepsilon}}$ involves the inverse of the information matrix. However, it is not necessary to compute the inverse in practice. Instead, multiplications with the inverse information matrix are solved efficiently as sparse, symmetric, positive-definite, linear system of equations. Utilising the sparseness makes the computations significantly faster but the computational cost increases with the trajectory length. An alternative is to compute, in parallel to the information vector and matrix, the mean and covariance for the current state, making them available for computation of the weights (21b) and (22c).

For the L-scan approximation, the appropriate choice of L depends on the linear and Gaussian models (38), and how well the cross-covariances $P_{k|k}^{[k,k']}$ can be approximated by all-zero matrices $0_{n_x \times n_x}$ for time steps $k' < k - L$. The larger L is, the more accurate the resulting tracking filter is, at the price of an increased computational cost. One implementation alternative that is computationally cheap and has low memory requirements is to perform a fixed-lag RTS-smoothing for selected Bernoulli densities $f_{k|k}^{i,a^i}(X)$ (keeping multiple hypotheses between scans) when it is deemed necessary, e.g., to perform trajectory estimation [45]. In addition, for TPMBM implementation for the set of all trajectories, we do not update single trajectory hypotheses with a small probability $P(\epsilon = k)$ of being present at the current time step k.

Lastly, as with MHT, δ-GLMB and PMBM filters, the number of global hypotheses in the TPMBM filters grows fast in time, so we use pruning methods for computational tractability. The PMBM posterior density (10) remains a PMBM, keeping the symmetry of the density over the elements of the set, if we 1) prune MBs with small global hypothesis weights (and then re-normalise remaining weights), 2) prune Bernoulli components with a small probability of existence, and 3) prune PPP intensity components with small weights, see details in [57], [58]. That is, if pruning keeps the set of global hypotheses $\widetilde{\mathcal{A}}_{k|k'} \subseteq \mathcal{A}_{k|k'}$ the PMBM is of the form (10a) with an MBM density

$$f_{k|k'}^{\mathrm{d}}(\mathbf{X}_k^{\mathrm{d}}) \propto \sum_{a_{k|k'} \in \widetilde{\mathcal{A}}_{k|k'}} w_{k|k'}^{a} \sum_{\uplus_{i \in \mathbb{T}_{k|k'}} \mathbf{X}_k^i = \mathbf{X}_k^{\mathrm{d}}} \prod_{i \in \mathbb{T}_{k|k'}} f_{k|k'}^{i,a^i}(\mathbf{X}_k^i). \tag{47}$$

In PMBM pruning, keeping the MBs with largest weights minimises an upper bound of the mean absolute error [57, Sec. V.D]. Standard methods can then be used to select local and global hypotheses with non-negligible weights. For example, we can use gating [59], single-scan data association solvers such as the Hungarian method and Jonker-Volgenant-Castanon algorithm [60], M-best single-scan data association methods, such as Murty's algorithm [61], single-scan sampling (e.g. Gibbs sampling) [34], [62], [63], multi-scan data associations via Lagrangian relaxation [45], [64], [65], multi-scan generalised Murty's algorithm [66] and multi-scan sampling (e.g. Markov chain Monte Carlo or Gibbs sampling) [44], [67], [68].

The above pruning algorithms have different pros and cons related to accuracy and computational complexity, and which one to use can be application dependent. Importantly, all the pruning strategies can be used in the different filtering frameworks to address the data association problem. For example, Murty's algorithm is ensured to obtain the k-best global hypotheses arising from a previous global hypothesis with a quartic complexity in the number of measurements $O(m_k^4)$. Gibbs sampling can achieve a linear complexity in the number of measurements $O(m_k)$, but it is not ensured to obtain the k-best global hypotheses [34].

VII. Simulation study

We present results from three scenarios that include interesting tracking challenges: Scenario 1) has many targets and several births and deaths; Scenario 2) has long trajectories; and Scenario 3) includes complicated target coalescence. We compare the following MTT algorithms:

1) TPMBM filters for the set of all/alive trajectories, implemented using *Murty's algorithm* [61] and the Gaussian information form, abbreviated TPMBM-IF/TPMBM-A-IF.
2) TPMBM filters for the set of all/alive trajectories, implemented using *Murty's algorithm* and $L = 1$ scan Gaussian approximation, abbreviated TPMBM-L1 and TPMBM-A-L1.
3) Multi-scan TPMBM filters for the set of all/alive trajectories, implemented using *dual decomposition* [45], with Gaussian information form, abbreviated M-TPMBM-IF and M-TPMBM-A-IF.
4) Multi-scan TPMBM filters for the set of all/alive trajectories, implemented using *dual decomposition*, with $L = 1$ scan Gaussian approximation, abbreviated M-TPMBM-L1 and M-TPMBM-A-L1.
5) δ-GLMB with joint prediction and update steps [34], implemented using *Murty's algorithm*.
6) δ-GLMB with joint prediction and update steps as well as *adaptive birth* [41], abbreviated δ-GLMB-AB, implemented using *Murty's algorithm*.
7) Trajectory MBM_{01} [5], implemented using *multi-scan Gibbs sampling* [44], with $L = 1$ scan Gaussian approximation, abbreviated M-TMBM_{01}-L1. Adding labels to the trajectories in the M-TMBM_{01}-L1 filter does not change the estimated set of trajectories [5, Sec. IV]. This filter is also referred to as multi-scan δ-GLMB in [44].
8) Track-oriented MHT implementation [16] from MathWorks' Sensor Fusion and Tracking Toolbox, abbreviated TOMHT.

Note that M-TPMBM-IF, M-TPMBM-A-IF, M-TPMBM-L1, M-TPMBM-A-L1 and M-TMBM_{01}-L1 solve a multi-scan data association problem, whereas the rest solve a single-scan data association problem. In addition, TPMBM-IF, TPMBM-A-IF, M-TPMBM-IF and M-TPMBM-A-IF perform smoothing-while-filtering and therefore provide smoothed trajectory estimates.

A. Simulation setup

In all scenarios, a 2D constant velocity motion model [69, Sec. III] with Gaussian acceleration standard deviation σ_v was used. Target measurements were simulated using linear position measurements with Gaussian noise with $R = \mathrm{diag}\left([\sigma_r^2,\ \sigma_r^2]\right)$. Clutter measurements were simulated uniformly distributed in the surveillance area, with average

number of false alarms per time step $\mu^{\mathrm{FA}} = \int \lambda^{\mathrm{FA}}(z)\mathrm{d}z$. The simulation parameters for all three scenarios are listed in Table I, where K denotes the number of total time steps, and $|\mathbf{X}_{0:K}|$ is the cardinality of the set of all trajectories at the final time step. The ground truth set of trajectories in each scenario is shown in Figure 2. The target cardinalities over time for the three scenarios are shown in Appendix B. The birth models are described in the following.

1) Scenario with a large number of trajectories: The true trajectories were generated by simulating the models in the surveillance area $[-5\times10^3, 5\times10^3]\times[-5\times10^3, 5\times10^3]$. The trajectories vary in length from 2 to 98 time steps, with mean/median length $38.9/38$. The birth density had nine components, with mean positions given by the columns in the following matrix,

$$\begin{bmatrix} -\sqrt{2} & \sqrt{2} & -\sqrt{2} & \sqrt{2} & 1 & -1 & 0 & 0 & 0 \\ -\sqrt{2} & -\sqrt{2} & \sqrt{2} & \sqrt{2} & 0 & 0 & 1 & -1 & 0 \end{bmatrix} \cdot \frac{10^4}{4},$$

zero velocities, and covariances $\mathrm{diag}\left([500^2, 500^2, 10^2, 10^2]\right)$. The expected number of births per time step for each component was set to $1/9$ for all filters. That is, the MB birth has 9 components, each with a Gaussian single-target density and probability of existence $1/9$. The PPP birth intensity is a Gaussian mixture, each one with a weight $1/9$. This ensures that the Kullback-Leibler divergence between the MB birth, used by the δ-GLMB filters, and the PPP used by the PMBM filters [11], [37], is minimised.

2) Scenario with long trajectories: The trajectory birth and death time steps were set deterministically to $10,\ 20, \ldots, 100$ and $990,\ 980, \ldots, 900$, respectively. The trajectories vary in length from 801 to 981 time steps. The true trajectories were generated by sampling the initial state from the birth process, and then simulating the motion model in the surveillance area $[-4\times10^4, 4\times10^4]\times[-4\times10^4, 4\times10^4]$. The birth density had a single component with mean $[0,0,0,0]^{\mathrm{T}}$, and covariance $\mathrm{diag}\left([4\times10^6, 4\times10^6, 1, 1]\right)$. The expected number of births per time step was set to 0.01 for the filters.

3) Scenario with target coalescence: As noted in several previous publications [6], [22], [47], many spatially separated trajectories, or very long trajectories, is not necessarily the most difficult scenario. Therefore, we also simulated a scenario in the surveillance area $[-5\times10^2, 5\times10^2]\times[-5\times10^2, 5\times10^2]$ with several targets that start separated, move in proximity to each other at the mid-point of the scenario, and then separate again. The birth density had a single component with mean $[0,0,0,0]^{\mathrm{T}}$, and covariance $\mathrm{diag}\left([10^3, 10^3, 10^2, 10^2]\right)$. The expected number of births per time step was set to 0.04 for the filters.

4) Performance evaluation: We evaluate the trajectory estimation performance using the generalised optimal subpattern assignment (GOSPA) metric for sets of trajectories in [70], which we refer to as trajectory metric (TM) in this paper. Specifically, we use its linear programming implementation [70], which is reviewed in Appendix C. It should be noted that the filters based on sets of trajectories directly estimate a set of trajectories, which can be compared with the ground truth set of trajectories with the trajectory metric. In contrast, δ-GLMB and TOMHT require some extra processing. At each time step, δ-GLMB and TOMHT estimate a set of targets with an associated label and track ID, respectively. By linking target state estimates with the same labels or track IDs over time, one can derive an estimated set of trajectories. The labels or track IDs are then discarded after linking.

TABLE I
SIMULATION PARAMETERS

Scenario	K	$\lvert\mathbf{X}_{0:K}\rvert$	σ_v	σ_r	P^{S}	P^{D}	μ^{FA}
1	100	100	1	1	0.99	0.9	30
2	1000	10	1	1	0.99	0.7	60
3	100	4	0.5	1	0.98	0.98	10

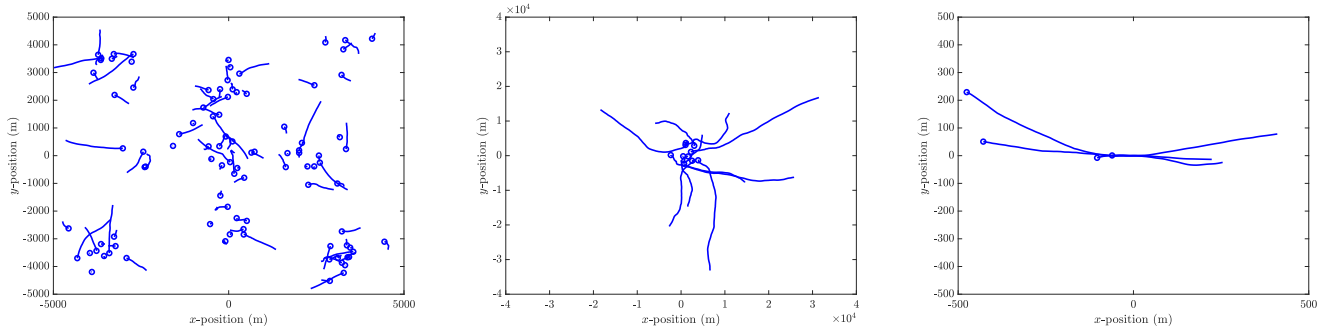
The parameters of the TM are: 1-norm base metric, location error cut-off $c = 20$, order $p = 1$, and switch cost $\gamma = 2$. The TM penalises four types of error: location error (LE), missed target error (ME), false target error (FE), and switch error (SE), see [70] for metric definitions and further details. The TM error is computed at each time step, and the result is normalised by the time step. This enables a comparison of how the trajectory estimation error changes over time.

B. Parameters of the filters

All the implementations used gating with gate probability 0.999. In all the TPMBM filters, we prune Bernoulli components with probability of existence smaller than 10^{-5} and Gaussian components in the PPP intensity with weight smaller than 10^{-5}. In addition, we do not update single trajectory hypotheses whose probability of being alive at current time step is smaller than 10^{-5}. In TPMBM-IF, TPMBM-L1 and the three δ-GLMB filters, we prune global hypotheses with weight smaller than 10^{-5}, and their number is capped to 10^3. In TPMBM-IF, TPMBM-L1, δ-GLMB and δ-GLMB-AB, for global hypothesis a in the prediction step, the $M = \lceil 10^3 w^a \rceil$ best global hypotheses are found using Murty's algorithm.

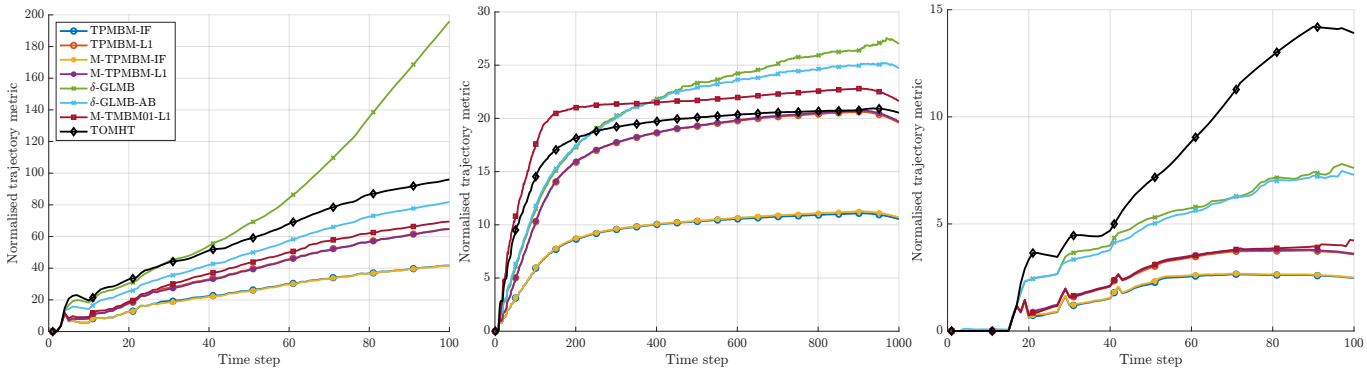
In M-TPMBM-IF and M-TPMBM-L1, we consider the data association problem over 3 scans of measurements ($N = 2$ scan pruning [71]), and the resulting 4D assignment problem is solved using dual decomposition with convergence threshold 0.01. Moreover, we limit the number of single trajectory hypotheses from the same Bernoulli component to 30. In M-TMBM$_{01}$-L1, we also consider the 3-scan data association problem, and the number of iterations in multi-scan Gibbs sampling is set to $\lceil 3\times10^4 w^a \rceil$ for global hypothesis a in the prediction step. In δ-GLMB-AB, parameter $r_{B,max}$ of adaptive birth [41] is set to the expected number of births per time step in each scenario. For the δ-GLMB-AB user-defined single target density, the mean state is the position indicated by the measurement with zero velocity, and the covariance matrix is $\mathrm{diag}\left([10^2, 10^2, 10^2, 10^2]\right)$. For extracting the estimates, TPMBM filters use the estimator described in Section IV-E2 with $r^e = 0.5$, and δ-GLMB filters use the estimator in [34].

As for TOMHT, we set the maximum number of hypotheses to 100 and the maximum number of track branches to 50. In addition, we use N-scan pruning with maximum 3 scans and the minimum probability required to keep a branch is 10^{-4}. The performance of TOMHT is sensitive w.r.t. the choice of track confirmation and deletion thresholds which are tailored



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 2. Ground truth set of trajectories in each of the three scenarios in the simulations. The start time of each trajectory is marked with a circle.



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 3. Trajectory estimation performance for the set of all trajectories in terms of the normalised TM over time.

TABLE II
TRAJECTORY ESTIMATION PERFORMANCE FOR THE SET OF ALL TRAJECTORIES IN TERMS OF THE NORMALISED TM (SUMMED OVER ALL THE TIME STEPS) AND ITS DECOMPOSITION

	Scenario 1					Scenario 2					Scenario 3				
Filter	TM	LE	ME	FE	SE	TM	LE	ME	FE	SE	TM	LE	ME	FE	SE
TPMBM-IF	2514.6	2159.4	339.6	15.1	0.5	9351.0	9118.5	231.1	0.7	0.7	176.2	160.7	10.3	0	5.2
TPMBM-L1	3796.0	3362.6	378.7	54.2	0.5	17323.7	16834.3	359.5	129.1	0.7	241.9	225.5	10.3	0	6.1
M-TPMBM-IF	2490.4	2162.4	312.5	15.2	0.2	9443.8	9111.6	330.9	0.6	0.7	179.8	160.7	13.5	0	5.6
M-TPMBM-L1	3717.5	3373.1	287.1	56.8	0.5	17387.0	16795.9	460.4	130.1	0.7	246.8	226.7	13.6	0.2	6.4
δ-GLMB	8164.4	2841.0	3462.8	1900.7	52.0	21046.1	15383.4	5739.5	42.4	2.3	453.0	209.1	248.3	2.1	12.3
δ-GLMB-AB	4870.8	3158.1	1414.7	310.0	7.8	20474.4	15650.7	4871.5	136.0	9.3	438.9	212.5	223.9	7.4	13.0
M-TMBM$_{01}$-L1	4141.8	3306.3	655.9	178.7	0.9	20545.0	16644.3	1200.0	2696.9	3.9	250.2	225.9	15.3	2.8	6.2
TOMHT	5861.1	2970.7	2432.3	457.2	0.9	18608.0	16913.6	1548.4	145.9	0.1	731.8	283.8	370.9	66.0	11.0

to each scenario to obtain good performance. Specifically, the confirmation and deletion thresholds are $(10, -8.2)$ for Scenario 1, $(10, -11.7)$ for Scenario 2, and $(20, -7)$ for Scenario 3.

C. Results for estimating the set of all trajectories

We analyse the performance of the algorithms to track the set of all trajectories using Monte Carlo simulation with 100 runs. The trajectory estimation performance is evaluated in terms of the normalised TM error over time, shown in Fig. 3. In addition, the TM error and its decomposition into location error (LE), missed target error (ME), false target error (FE), and switch error (SE) summed over all time steps and normalised over the time window is given in Table II. The average runtime per time step with a MATLAB implementation on an Apple M2

TABLE III
AVERAGE RUNTIME PER TIME STEP (IN SECONDS)

Filter	Scenario 1	Scenario 2	Scenario 3
TPMBM-IF	0.25	0.22	0.15
TPMBM-L1	0.24	0.21	0.15
M-TPMBM-IF	0.51	0.23	0.03
M-TPMBM-L1	0.50	0.22	0.02
TPMBM-A-IF	0.14	0.06	0.09
TPMBM-A-L1	0.14	0.05	0.08
M-TPMBM-A-IF	0.43	0.17	0.02
M-TPMBM-A-L1	0.42	0.15	0.02
δ-GLMB	1.91	0.11	0.09
δ-GLMB-AB	6.35	2.13	0.98
M-TMBM$_{01}$-L1	11.89	2.13	1.44
TOMHT	0.63	0.29	0.48

Pro is provided in Table III, which also includes the runtimes of the filters that track the set of alive trajectories, whose results will be analysed in the next section. In addition, the number of targets at each time step and the TM evaluated at the final time step are given in Appendix B.

The results show that for all three scenarios, the TPMBM filters outperform the two δ-GLMB filters, M-TMBM$_{01}$-L1 and TOMHT. In particular, the two TPMBM filters with Gaussian information form outperform their counterparts with Gaussian $L = 1$ scan approximation by providing smoothed trajectory estimates, with slightly increased runtime. TPMBM filters are generally faster than δ-GLMB filter implementations, mainly due to a more efficient structure of the global hypotheses [22]. TPMBM filters are also faster than the TOMHT implementation.

TOMHT uses a track management algorithm that does not fully take into account the target birth and death model information. In Scenarios 1 and 2, TOMHT performs better than δ-GLMB, but worse than the TPMBM filters. In Scenario 3, which deals with a challenging data association problem, TOMHT performs worse than the rest of the filters, especially when targets get in close proximity.

Compared to TPMBM filters using Murty's algorithm, multi-scan TPMBM filters using dual decomposition present similar estimation performance in Scenario 1 and 3, worse performance in Scenario 2, and much lower runtime in Scenario 3. By using an adaptive birth model, the performance of δ-GLMB is improved in all three scenarios, in particular Scenario 1, at the price of significantly increased runtime. M-TMBM$_{01}$-L1 outperforms δ-GLMB-AB by 1) keeping the full trajectory information and 2) solving a more complex data association problem using Gibbs sampling, but it also has the highest runtime. Moreover, we can observe that M-TMBM$_{01}$-L1 has difficulty initiating new trajectories in Scenario 2, where the expected number of births per time step is low.

In general, Scenario 1 and 2 illustrate that the TPMBM filters are not limited to handling a high number of relatively short trajectories, but are capable of handling long trajectories. Scenario 3 is challenging due to the complex data association problem when the targets are in proximity for several consecutive time steps, and the TPMBM filters produce reasonable trajectory estimates. However, the δ-GLMB filters have a higher ME and FE, and are susceptible to unrealistic track switching given the motion model, see [5, Fig. 3] and [9, Fig. 2].

D. Results for estimating the set of alive trajectories

In this section, we use the same scenarios and parameters as in the previous section, but we analyse the performance to estimate the set of alive trajectories at each time step. In this section, we therefore also include the considered TPMBM filters that directly approximate the posterior of the set of alive trajectories, namely, TPMBM-A-IF, TPMBM-A-L1, M-TPMBM-A-IF and M-TPMBM-A-L1. In addition, to estimate the set of alive trajectories we have also tested the previously considered filters, including the TPMBM filters that approximate the posterior over the set of all trajectories. For these filters, the estimated set of alive trajectories is obtained by extracting the alive trajectories from the estimated set of all trajectories. The average runtime per time step was shown in Table III, where we can see that the TPMBM filters for alive trajectories are faster than the TPMBM filters for all trajectories.

The normalised TM error over time for the set of alive trajectories is shown in Fig. 4. Also, the TM error and its decomposition across all time steps is provided in Table IV. We can draw the similar main conclusion as before. TPMBM filters generally work better than the other algorithms. Specifically, the M-TPMBM-IF filter is the one that provides the most accurate trajectories in the three scenarios. The TPMBM filters that directly estimate alive trajectories have a very similar performance to the corresponding TPMBM filters that track all trajectories but only report the alive trajectories.

VIII. Conclusions

In this paper we have shown that the PMBM density for the set of trajectories is conjugate for MTT with standard point target models. We have presented two TPMBM filtering recursions for the set of trajectories. The benefits of the proposed MTT filters compared to relevant MTT algorithms is shown via simulations. Further, we have showed that, regardless of what time window we consider, the multi-trajectory posterior is PMBM, a result analogous to the Gaussian distribution in linear/Gaussian systems.

An interesting line of future work is the development of algorithms that make use of the TPMBM time marginalisation theorem (Theorem 4) and the TPMBM for arbitrary time interval theorem (Theorem 5) to answer certain types of trajectory-related questions in MTT.

References

[1] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishing, 2011.

[2] A. Milan, K. Schindler, and S. Roth, "Multi-target tracking by discrete-continuous energy minimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 10, pp. 2054–2068, 2016.

[3] T. Li, H. Chen, S. Sun, and J. M. Corchado, "Joint smoothing and tracking based on continuous-time target trajectory function fitting," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 3, pp. 1476–1483, 2019.

[4] J. Zhou, T. Li, X. Wang, and L. Zheng, "Target tracking with equality/inequality constraints based on trajectory function of time," *IEEE Signal Processing Letters*, vol. 28, pp. 1330–1334, 2021.

[5] A. F. García-Fernández, L. Svensson, and M. R. Morelande, "Multiple target tracking based on sets of trajectories," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 3, pp. 1685–1707, 2020. [Online]. Available: https://arxiv.org/abs/1605.08163

[6] 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, July 2015.

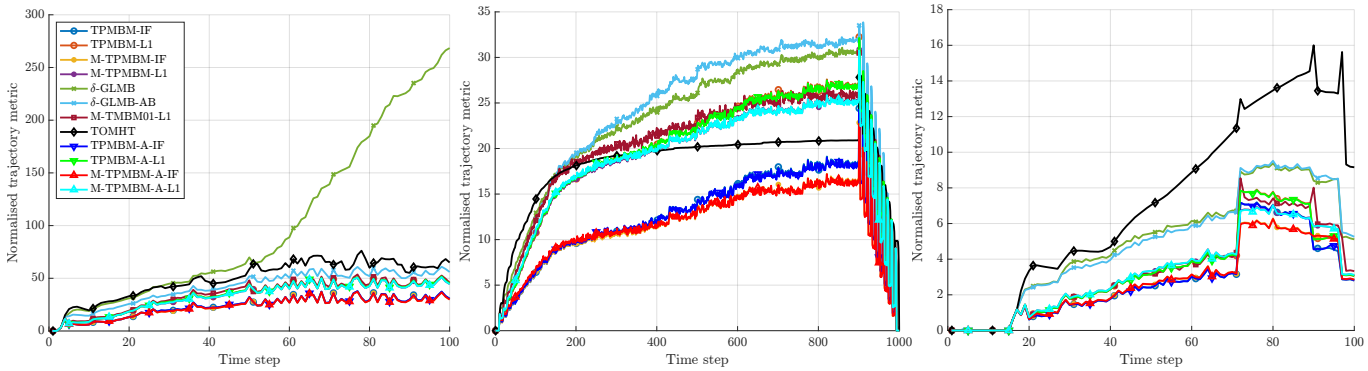
[7] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[8] W. Koch and F. Govaers, "On accumulated state densities with applications to out-of-sequence measurement processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 4, pp. 2766–2778, 2011.

[9] K. Granström, L. Svensson, Y. Xia, J. Williams, and A. F. García-Fernández, "Poisson multi-Bernoulli mixture trackers: Continuity through random finite sets of trajectories," in *Proc. Int. Conf. Inform. Fusion*, Cambridge, UK, Jul. 2018.

[10] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Automatic Control*, vol. AC-24, no. 6, pp. 843–854, December 1979.

[11] R. Mahler, *Advances in Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2014.



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 4. Trajectory estimation performance for the set of alive trajectories in terms of the normalised TM over time

TABLE IV
TRAJECTORY ESTIMATION PERFORMANCE FOR THE SET OF ALIVE TRAJECTORIES IN TERMS OF THE NORMALISED TM (SUMMED OVER ALL THE TIME STEPS) AND ITS DECOMPOSITION

	Scenario 1					Scenario 2					Scenario 3				
Filter	TM	LE	ME	FE	SE	TM	LE	ME	FE	SE	TM	LE	ME	FE	SE
TPMBM-IF	2334.7	1695.1	453.2	186.3	0	12434.5	8098.0	4111.9	224.5	0	280.7	144.1	57.7	74.4	3.8
TPMBM-L1	3329.8	2630.7	483.0	216.1	0	19472.5	14969.9	4195.0	307.6	0	331.2	194.5	57.7	74.5	4.5
M-TPMBM-IF	2297.0	1700.3	409.7	186.9	0	11569.5	8219.0	3117.8	232.6	0	272.4	145.0	52.3	70.9	4.2
M-TPMBM-L1	3294.1	2634.6	441.2	218.4	0	18705.8	15176.1	3207.5	322.3	0	325.1	197.1	52.3	70.9	4.7
δ-GLMB	10122.3	2188.1	2960.1	4972.2	13.5	22381.4	13764.7	8473.8	222.1	0	499.2	190.4	244.7	69.6	8.7
δ-GLMB-AB	4149.9	2442.5	1427.3	284.5	0.8	23495.4	13473.4	9889.5	218.5	0.1	491.8	196.1	225.4	75.9	9.0
M-TMBM$_{01}$-L1	3483.4	2600.7	599.2	283.5	0	19873.2	14900.5	4544.9	427.8	0	336.9	198.3	53.3	80.7	4.6
TOMHT	4963.7	2326.9	2055.3	580.8	0.6	18353.9	15846.5	1610.2	897.2	0	728.1	252.6	305.7	170.0	8.8
TPMBM-A-IF	2334.7	1695.2	453.2	186.3	0	12408.3	8099.3	4084.2	224.8	0	280.1	143.9	58.0	74.4	3.8
TPMBM-A-L1	3329.5	2630.6	482.9	216.0	0	19449.8	14974.8	4167.2	307.8	0	331.2	194.3	58.0	74.4	4.5
M-TPMBM-A-IF	2299.3	1700.0	412.5	186.8	0	11651.7	8199.0	3222.0	230.7	0	272.7	144.8	53.3	70.4	4.2
M-TPMBM-A-L1	3295.6	2634.2	443.5	217.9	0	18761.4	15135.3	3308.7	317.4	0	325.1	196.6	53.3	70.4	4.7

[12] E. Brekke and M. Chitre, "Relationship between finite set statistics and the multiple hypothesis tracker," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1902–1917, Aug. 2018.

[13] E. F. Brekke, A. G. Hem, and L.-C. N. Tokle, "Multitarget tracking with multiple models and visibility: Derivation and verification on maritime radar data," *IEEE Journal of Oceanic Engineering*, vol. 46, no. 4, pp. 1272–1287, 2021.

[14] C.-Y. Chong, S. Mori, and S. Coraluppi, "Forty years of multiple hypothesis tracking," *Journal of Advances in Information Fusion. Special Issue on Multiple Hypothesis Tracking.*, vol. 14, no. 2, pp. 131–151, Dec. 2019.

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

[16] J. R. Werthmann, "Step-by-step description of a computationally efficient version of multiple hypothesis tracking," in *Signal and Data Processing of Small Targets 1992*, vol. 1698. SPIE, 1992, pp. 288–300.

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

[18] S. Mori, C.-Y. Chong, and K. C. Chang, "Three formalisms of multiple hypothesis tracking," in *Proc. 19th International Conference on Information Fusion*, July 2016.

[19] 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, July 2014.

[20] P. Boström-Rost, D. Axehill, and G. Hendeby, "Sensor management for search and track using the Poisson multi-Bernoulli mixture filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 2771–2783, 2021.

[21] P. Boström-Rost, D. Axehill, and G. Hendeby, "PMBM filter with partially grid-based birth model with applications in sensor management," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 1, pp. 530–540, 2022.

[22] A. F. García-Fernández, J. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: direct derivation and implementation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, Aug. 2018.

[23] L. Cament, M. Adams, and J. Correa, "A multi-sensor, Gibbs sampled, implementation of the multi-Bernoulli Poisson filter," in *Proc. Int. Conf. Inform. Fusion*, Cambridge, UK, Jul. 2018, pp. 2580–2587.

[24] K. Granström, L. Svensson, S. Reuter, Y. Xia, and M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods," *IEEE Trans. Intel. Veh.*, vol. 3, no. 1, Mar. 2018.

[25] S. Pang and H. Radha, "Multi-object tracking using Poisson multi-Bernoulli mixture filtering for autonomous vehicles," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 7963–7967.

[26] J. Zhao, P. Wu, X. Liu, Y. Xu, L. Mihaylova, S. Godsill, and W. Wang, "Audio-visual tracking of multiple speakers via a PMBM filter," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 5068–5072.

[27] S. Scheidegger, J. Benjaminsson, E. Rosenberg, A. Krishnan, and K. Granström, "Mono-camera 3D multi-object tracking using deep learning detections and PMBM filtering," in *Proc. IEEE Intel. Veh. Symp.*, Changshu, Suzhou, China, Jun. 2018.

[28] M. Fatemi, K. Granström, L. Svensson, F. Ruiz, and L. Hammarstrand, "Poisson multi-Bernoulli mapping using Gibbs sampling," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2814–2827, Jun. 2017.

[29] Y. Ge, O. Kaltiokallio, H. Kim, F. Jiang, J. Talvitie, M. Valkama, L. Svensson, S. Kim, and H. Wymeersch, "A computationally efficient EK-PMBM filter for bistatic mmWave radio SLAM," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 7, pp. 2179–2192, 2022.

[30] H. Kim, K. Granstrom, L. Svensson, S. Kim, and H. Wymeersch, "PMBM-based SLAM filters in 5G mmwave vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 8646–8661, 2022.

[31] M. Fröhle, C. Lindberg, K. Granström, and H. Wymeersch, "Multisensor Poisson multi-Bernoulli filter for joint target-sensor state tracking," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 4, pp. 609–621, 2019.

[32] A. F. García-Fernández, Y. Xia, K. Granström, L. Svensson, and J. L.

Williams, "Gaussian implementation of the multi-Bernoulli mixture filter," in *Proc. Int. Conf. Inform. Fusion*, Ottawa, Canada, Jul. 2019.
[33] 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, Apr. 2013.
[34] 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, Apr. 2017.
[35] A. F. Garcia-Fernandez, J. Grajal, and M. R. Morelande, "Two-layer particle filter for multiple target detection and tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 3, pp. 1569–1588, Jul. 2013.
[36] E. H. Aoki, P. K. Mandal, L. Svensson, Y. Boers, and A. Bagchi, "Labeling uncertainty in multitarget tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 3, pp. 1006–1020, Jun. 2016.
[37] A. F. García-Fernández, L. Svensson, J. L. Williams, Y. Xia, and K. Granström, "Trajectory Poisson multi-Bernoulli filters," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4933–4945, 2020.
[38] K. Panta, D. Clark, and B.-N. Vo, "Data association and track management for the Gaussian mixture probability hypothesis density filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 3, pp. 1003–1016, Jul. 2009.
[39] J. Houssineau and D. E. Clark, "Multitarget filtering with linearized complexity," *IEEE Trans. Signal Process.*, vol. 66, no. 18, pp. 4957–4970, Sep. 2018.
[40] M. Beard, B. T. Vo, and B. Vo, "A solution for large-scale multi-object tracking," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2754–2769, 2020.
[41] 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, Jul. 2014.
[42] A. F. García-Fernández, Y. Xia, and L. Svensson, "A comparison between PMBM Bayesian track initiation and labelled RFS adaptive birth," in *25th International Conference on Information Fusion*, 2022.
[43] T. Vu, B.-N. Vo, and R. Evans, "A particle marginal Metropolis-Hastings multi-target tracker," *IEEE Transactions on Signal Processing*, vol. 62, no. 15, pp. 3953–3964, Aug. 2014.
[44] B. N. Vo and B. T. Vo, "A multi-scan labeled random finite set model for multi-object state estimation," *IEEE Trans. Signal Process.*, vol. 67, no. 19, pp. 4948–4963, Oct. 2019.
[45] Y. Xia, K. Granström, L. Svensson, A. F. García-Fernández, and J. L. Williams, "Multi-scan implementation of the trajectory Poisson multi-Bernoulli mixture filter," *Jour. Adv. Inform. Fusion*, vol. 14, no. 2, pp. 213–235, Dec. 2019.
[46] K. Granström, L. Svensson, Y. Xia, J. Williams, and A. F. García-Fernández, "Poisson multi-Bernoulli mixtures for sets of trajectories," 2019. [Online]. Available: https://arxiv.org/abs/1912.08718
[47] Y. Xia, Á. F. García-Fernández, F. Meyer, J. L. Williams, K. Granström, and L. Svensson, "Trajectory PMB filters for extended object tracking using belief propagation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 6, pp. 9312–9331, 2023.
[48] A. F. García-Fernández and L. Svensson, "Tracking multiple spawning targets using Poisson multi-Bernoulli mixtures on sets of tree trajectories," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1987–1999, 2022.
[49] A. F. García-Fernández and W. Yi, "Continuous-discrete multiple target tracking with out-of-sequence measurements," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4699–4709, 2021.
[50] K. Granström, L. Svensson, Y. Xia, A. F. García-Fernández, and J. L. Williams, "Spatiotemporal constraints for sets of trajectories with applications to PMBM densities," in *23rd International Conference on Information Fusion*, 2020, pp. 1–8.
[51] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: a survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 103–123, Jan. 1998.
[52] R. Mahler, B.-T. Vo, and B.-N. Vo, "CPHD filtering with unknown clutter rate and detection profile," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3497–3513, Aug. 2011.
[53] 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.
[54] M. R. Walter, R. M. Eustice, and J. J. Leonard, "Exactly sparse extended information filters for feature-based SLAM," *International Journal of Robotics Research*, vol. 26, no. 4, pp. 335–359, Apr. 2007.
[55] I. Mahon, S. B. Williams, O. Pizarro, and M. Johnson-Roberson, "Efficient view-based SLAM using visual loop closures," *IEEE Trans. Robot.*, vol. 24, no. 5, pp. 1002–1014, Oct. 2008.
[56] A. F. García-Fernández and L. Svensson, "Trajectory PHD and CPHD filters," *IEEE Trans. Signal Process.*, vol. 67, no. 22, pp. 5702–5714, Nov. 2019.
[57] K. Granström, M. Fatemi, and L. Svensson, "Poisson multi-Bernoulli conjugate prior for multiple extended object filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 208–225, 2020.
[58] A. F. García-Fernández and S. Maskell, "Continuous-discrete multiple target filtering: PMBM, PHD and CPHD filter implementations," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1300–1314, 2020.
[59] S. Challa, M. R. Morelande, D. Musicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge University Press, 2011.
[60] D. F. Crouse, "On implementing 2D rectangular assignment algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 4, pp. 1679–1696, August 2016.
[61] K. Murty, "An algorithm for ranking all the assignments in order of increasing cost," *Operations Research*, vol. 16, no. 3, pp. 682–687, 1968.
[62] M. R. Morelande, "Joint data association using importance sampling," in *12th International Conference on Information Fusion*, 2009, pp. 292–299.
[63] C. Hue, J. Le Cadre, and P. Perez, "Tracking multiple objects with particle filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 791–812, Jul 2002.
[64] A. B. Poore and A. J. Robertson III, "A new Lagrangian relaxation based algorithm for a class of multidimensional assignment problems," *Computational Optimization and Applications*, vol. 8, pp. 129–150, 1997.
[65] S. Deb, M. Yeddanapudi, K. Pattipati, and Y. Bar-Shalom, "A generalized S-D assignment algorithm for multisensor-multitarget state estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 2, pp. 523–538, April 1997.
[66] E. Fortunato, W. Kreamer, S. Mori, C.-Y. Chong, and G. Castanon, "Generalized Murty's algorithm with application to multiple hypothesis tracking," in *International Conference on Information Fusion*, July 2007.
[67] S. Oh, S. Russell, and S. Sastry, "Markov chain Monte Carlo data association for multi-target tracking," *IEEE Trans. Autom. Control*, vol. 54, no. 3, pp. 481–497, Mar. 2009.
[68] S. He, H.-S. Shin, and A. Tsourdos, "Track-oriented multiple hypothesis tracking based on tabu search and Gibbs sampling," *IEEE Sensors Journal*, vol. 18, no. 1, pp. 328–339, 2018.
[69] X. Rong-Li and V. Jilkov, "Survey of maneuvering target tracking: Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.
[70] Á. F. García-Fernández, A. S. Rahmathullah, and L. Svensson, "A metric on the space of finite sets of trajectories for evaluation of multi-target tracking algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3917–3928, 2020.
[71] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, 2004.
[72] J. Williams, "An efficient, variational approximation of the best fitting multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 258–273, Jan. 2015.

SUPPLEMENTAL MATERIAL

APPENDIX A
PROOFS OF THEOREM 4 AND THEOREM 5

In this appendix, we first prove Theorem 4, and then we prove Theorem 5.

A. Preliminaries

For trajectories $X \in \mathcal{T}_{0:k}$ we define the function

$$\tau_{\alpha:\gamma}^{\eta:\zeta}(X) = \begin{cases} \{(b, e, x_{b:e})\} & \mathbb{N}_\beta^\varepsilon \cap \mathbb{N}_\eta^\zeta \neq \emptyset, \\ \emptyset & \text{otherwise}, \end{cases} \tag{48a}$$

$$b = \max(\beta, \alpha), \quad e = \min(\varepsilon, \gamma). \tag{48b}$$

where $0 \leq \alpha \leq \eta \leq \zeta \leq \gamma \leq k$. Note that a function corresponding to $\tau_{k:k}^{k:k}(X)$ was defined in [5, Sec. II.A]. For set inputs, like in [5, Sec. II.A], we have

$$\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}) = \begin{cases} \bigcup_{X \in \mathbf{X}} \tau_{\alpha:\gamma}^{\eta:\zeta}(X) & \mathbf{X} \neq \emptyset \\ \emptyset & \mathbf{X} = \emptyset \end{cases} \tag{49}$$

For $\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X})$ it holds that $\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}^1 \cup \mathbf{X}^2) = \tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}^1) \cup \tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}^2)$.

The multi-target Dirac delta is defined as [11]

$$\delta_{\mathbf{X}'}(\mathbf{X}) \triangleq \begin{cases} 0 & |\mathbf{X}| \neq |\mathbf{X}'| \\ 1 & \mathbf{X} = \mathbf{X}' = \emptyset \\ \sum_{\sigma \in \Gamma_n} \prod_{i=1}^{n} \delta_{X'_{\sigma_i}}(X_i) & \begin{cases} \mathbf{X} = \{X_i\}_{i=1}^n \\ \mathbf{X}' = \{X'_i\}_{i=1}^n \end{cases} \end{cases} \tag{50}$$

From [5, Thm. 11] we know that using the multi-target Dirac delta function and $\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X})$ we can formulate a transition density from the set $\mathbf{X}_{0:k}$ to the set $\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta} = \tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}_{0:k})$ as

$$\delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}_{0:k})}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right). \tag{51}$$

Lemma 1 shows how to find the density of $\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X})$.

Lemma 1. *Suppose $f(\mathbf{X})$ is a* PMBM *density parameterised as in Section IV. Then, the density of $\mathbf{Y} = \tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X})$ is*

$$g(\mathbf{Y}) = \int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X})}(\mathbf{Y}) f(\mathbf{X}) \delta\mathbf{X} \tag{52a}$$

$$= \sum_{\left(\uplus_{i \in \mathbb{T}} \mathbf{Y}^i\right) \uplus \mathbf{Y}^{\mathrm{u}} = \mathbf{Y}} \int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}^{\mathrm{u}})}(\mathbf{Y}^{\mathrm{u}}) f^{\mathrm{u}}(\mathbf{X}^{\mathrm{u}}) \delta\mathbf{X}^{\mathrm{u}} \times \sum_{a \in \mathcal{A}} w_a \prod_{i \in \mathbb{T}} \int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}^i)}(\mathbf{Y}^i) f^{i,a^i}(\mathbf{X}^i) \delta\mathbf{X}^i \tag{52b}$$

Proof. Follows from [72, Lem. 2] and [5, Thm. 11]. □

Note that in Lemma 1 we have integrals with the transition density and PPP densities and Bernoulli densities. The following Lemmas provide the solutions to these integrals.

Lemma 2. *Let $f(\mathbf{X}_{0:k})$ be a Bernoulli density, with parameters r and $f(X)$. The trajectory* RFS *density*

$$f(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}) = \int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}_{0:k})}(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}) f(\mathbf{X}_{0:k}) \delta\mathbf{X}_{0:k} \tag{53}$$

is a trajectory Bernoulli density with parameters

$$\tilde{r} = r \Pr\left(\mathbb{N}_\beta^\varepsilon \cap \mathbb{N}_\eta^\zeta \neq \emptyset\right) \tag{54a}$$

$$\tilde{f}(X) = \frac{\sum_{\substack{\beta,\varepsilon: \\ \mathbb{N}_\beta^\varepsilon \cap \mathbb{N}_\eta^\zeta \neq \emptyset}} \int p(x_{\beta:\varepsilon}|\beta,\varepsilon) P(\beta,\varepsilon) \mathrm{d}x_{\beta:\varepsilon \setminus b:e}}{\Pr\left(\mathbb{N}_\beta^\varepsilon \cap \mathbb{N}_\eta^\zeta \neq \emptyset\right)} \tag{54b}$$

where

$$\Pr\left(\mathbb{N}_\beta^\varepsilon \cap \mathbb{N}_\eta^\zeta \neq \emptyset\right) = \sum_{\substack{\beta,\varepsilon: \\ \mathbb{N}_\beta^\varepsilon \cap \mathbb{N}_\eta^\zeta \neq \emptyset}} P(\beta,\varepsilon). \tag{55}$$

Specifically, for

$$f(X) = \sum_{j \in \mathbb{J}} \nu^j f^j(X; \theta^j), \tag{56}$$

we get

$$\tilde{r} = r \sum_{j \in \mathbb{J}^{\eta:\zeta}} \nu^j \tag{57a}$$

$$\tilde{f}(X) = \frac{\sum_{j \in \mathbb{J}^{\eta:\zeta}} \nu^j \tilde{f}^j(X; \tilde{\theta}^j)}{\sum_{j \in \mathbb{J}^{\eta:\zeta}} \nu^j} \tag{57b}$$

$$\tilde{\theta}^j = \left(\tilde{b}^j, \tilde{e}^j, \tilde{p}^j(x_{\tilde{b}^j:\tilde{e}^j})\right) \tag{57c}$$

$$\tilde{b}^j = \max(b^j, \alpha), \quad \tilde{e}^j = \min(e^j, \gamma) \tag{57d}$$

$$\tilde{p}^j(x_{\tilde{b}^j:\tilde{e}^j}) = \int p^j(x_{b^j:e^j}) \mathrm{d}x_{b^j:e^j \setminus \tilde{b}^j:\tilde{e}^j} \tag{57e}$$

where $\mathbb{J}^{\eta:\zeta} = \left\{j \ : \ \mathbb{N}_{b^j}^{e^j} \cap \mathbb{N}_\eta^\zeta \neq \emptyset\right\}$.

Proof. See (58). □

Lemma 3. *Let $f(\mathbf{X}_{0:k})$ be a trajectory* PPP *density, with intensity $\lambda(X) = \mu p(x_{\beta:\varepsilon}|\beta,\varepsilon) P(\beta,\varepsilon)$. The trajectory* RFS *density*

$$f(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}) = \int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}_{0:k})}(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}) f(\mathbf{X}_{0:k}) \delta\mathbf{X}_{0:k} \tag{59}$$

is a trajectory PPP *density with intensity*

$$\tilde{\lambda}(X) = \sum_{\substack{\beta,\varepsilon: \\ \mathbb{N}_\beta^\varepsilon \cap \mathbb{N}_\eta^\zeta \neq \emptyset}} \int \mu p(x_{\beta:\varepsilon}|\beta,\varepsilon) P(\beta,\varepsilon) \mathrm{d}x_{\beta:\varepsilon \setminus b:e} \tag{60}$$

Specifically, for

$$\lambda(X) = \sum_{j \in \mathbb{J}} \nu^j f^j(X; \theta^j), \tag{61}$$

we get

$$\tilde{\lambda}(X) = \sum_{j \in \mathbb{J}^{\eta:\zeta}} \nu^j \tilde{f}^j(X; \tilde{\theta}^j) \tag{62a}$$

$$\tilde{\theta}^j = \left(\tilde{b}^j, \tilde{e}^j, \tilde{p}^j(x_{\tilde{b}^j:\tilde{e}^j})\right) \tag{62b}$$

$$\tilde{b}^j = \max(b^j, \alpha), \quad \tilde{e}^j = \min(e^j, \gamma) \tag{62c}$$

$$\tilde{p}^j(x_{\tilde{b}^j:\tilde{e}^j}) = \int p^j(x_{b^j:e^j}) \mathrm{d}x_{b^j:e^j \setminus \tilde{b}^j:\tilde{e}^j} \tag{62d}$$

where $\mathbb{J}^{\eta:\zeta} = \left\{j \ : \ \mathbb{N}_{b^j}^{e^j} \cap \mathbb{N}_\eta^\zeta \neq \emptyset\right\}$.

$$f\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right)=\int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}_{0:k})}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right) f(\mathbf{X}_{0:k})\delta\mathbf{X}_{0:k}=\delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\emptyset)}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right)(1-r)+\int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\{X\})}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right) rf(X)\mathrm{d}X \tag{58a}$$

$$=\delta_{\emptyset}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right)(1-r)+\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}=\emptyset}}\int \delta_{\emptyset}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right) rf((x_{\beta:\varepsilon},\beta,\varepsilon))\mathrm{d}x_{\beta:\varepsilon}+\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}\int \delta_{\{(b,e,x_{b:e})\}}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right) rf((x_{\beta:\varepsilon},\beta,\varepsilon))\mathrm{d}x_{\beta:\varepsilon} \tag{58b}$$

$$=\delta_{\emptyset}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right)\left(1-r\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}P(\beta,\varepsilon)\right)+r\left(\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}P(\beta,\varepsilon)\right)\frac{\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}\int \delta_{\{(b,e,x_{b:e})\}}\left(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}\right)p(x_{\beta:\varepsilon}|\beta,\varepsilon)P(\beta,\varepsilon)\mathrm{d}x_{\beta:\varepsilon}}{\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}P(\beta,\varepsilon)} \tag{58c}$$

$$=\begin{cases}1-r\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}P(\beta,\varepsilon) & \text{if } \mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}=\emptyset\\ r\left(\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}P(\beta,\varepsilon)\right)\frac{\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}\int p(x_{\beta:\varepsilon}|\beta,\varepsilon)\mathrm{d}x_{\beta:\varepsilon\setminus b:e}P(\beta,\varepsilon)}{\sum_{\substack{\beta,\varepsilon:\\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset}}P(\beta,\varepsilon)} & \text{if } \mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}=\{X\}\end{cases} \tag{58d}$$

Proof. From [11, p. 99, Rmk. 12] we know that the PPP $\mathbf{X}_{0:k}$ with intensity $\lambda(X)$ can be divided into two disjoint and independent PPP subsets

$$\mathbf{X}_{0:k}=\mathbf{X}_{0:k}^{\eta:\zeta}\cup\mathbf{X}_{0:k}^{\not\eta:\not\zeta},\quad \mathbf{X}_{0:k}^{\eta:\zeta}\cap\mathbf{X}_{0:k}^{\not\eta:\not\zeta}=\emptyset \tag{63a}$$

$$\mathbf{X}_{0:k}^{\eta:\zeta}=\left\{X\in\mathbf{X}_{0:k}\ :\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset\right\} \tag{63b}$$

$$\mathbf{X}_{0:k}^{\not\eta:\not\zeta}=\left\{X\in\mathbf{X}_{0:k}\ :\ \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}=\emptyset\right\} \tag{63c}$$

with PPP intensities

$$\lambda^{\eta:\zeta}(X)=\mathbf{1}_{\mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}}\lambda(X), \tag{64a}$$

$$\lambda^{\not\eta:\not\zeta}(X)=\left(1-\mathbf{1}_{\mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}}\right)\lambda(X), \tag{64b}$$

where

$$\mathbf{1}_{\mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}}=\begin{cases}1 & \mathbb{N}_{\beta}^{\varepsilon}\cap\mathbb{N}_{\eta}^{\zeta}\neq\emptyset\\ 0 & \text{otherwise.}\end{cases} \tag{65}$$

It is straightforward to verify that $\lambda^{\eta:\zeta}(X)+\lambda^{\not\eta:\not\zeta}(X)=\lambda(X)$. The rest of the proof follows in (66). □

B. Proof of Theorem 4

Given a PMBM density $f(\mathbf{X}_{0:k})$ for the set of all trajectories from 0 to k, the density for $\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta}$, where $0\leq\alpha\leq\eta\leq\zeta\leq\gamma\leq k$, is [5, Thm. 11]

$$f(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta})=\int \delta_{\tau_{\alpha:\gamma}^{\eta:\zeta}(\mathbf{X}_{0:k})}(\mathbf{X}_{\alpha:\gamma}^{\eta:\zeta})f(\mathbf{X}_{0:k})\delta\mathbf{X}_{0:k} \tag{67}$$

Theorem 4 then follows from Lemmas 1, 2 and 3.

C. Proof of Theorem 5

It follows from Theorem 1 and Theorem 3 that

$$f_{\min(\alpha,\xi):\max(\gamma,\chi)|\xi:\chi}\left(\mathbf{X}_{\min(\alpha,\xi):\max(\gamma,\chi)}\right), \tag{68}$$

is a PMBM density, where no update (only prediction) is performed for time steps outside the interval $\xi:\chi$. Applying Theorem 4 to (68) concludes the proof.

APPENDIX B
ADDITIONAL DETAILS ON EXPERIMENTAL RESULTS

In this appendix, we provide further details on the experimental results. First, we show the number of targets at each time step in each of the scenarios in Figure 5. Second, we show the TM for the estimated trajectories at the final time step in Figure 6 and Table V. In this case, the M-TPMBM-IF is also the best performing algorithm in Scenario 1, and TPMBM-IF in Scenarios 2 and 3.

APPENDIX C
GOSPA METRIC FOR SETS OF TRAJECTORIES

This appendix reviews the GOSPA metric for sets of trajectories based on linear programming [70, Prop. 2], which we abbreviate as TM in this paper. We consider two sets of trajectories $\mathbf{X}$ and $\mathbf{Y}$ from time step 0 to T. The number of trajectories is these sets is $n_{\mathbf{X}}$ and $n_{\mathbf{Y}}$, respectively. For $1\leq p<\infty$, $c>0$, $\gamma>0$ and a base metric $d_b(\cdot,\cdot)$ for single targets, the TM is

$$d\left(\mathbf{X},\mathbf{Y}\right)=\min_{\substack{W^k\in\overline{\mathcal{W}}_{\mathbf{X},\mathbf{Y}}\\ k=1,\ldots,T}}\left(\sum_{k=1}^{T}\mathrm{tr}\left[\left(D_{\mathbf{X},\mathbf{Y}}^{k}\right)^{\dagger}W^k\right]+\frac{\gamma^p}{2}\sum_{k=1}^{T-1}\sum_{i=1}^{n_{\mathbf{X}}}\sum_{j=1}^{n_{\mathbf{Y}}}\left|W^k(i,j)-W^{k+1}(i,j)\right|\right)^{\frac{1}{p}}, \tag{69}$$

where $D_{\mathbf{X},\mathbf{Y}}^{k}$ is a $(n_{\mathbf{X}}+1)\times(n_{\mathbf{Y}}+1)$ matrix whose (i,j) element is

$$D_{\mathbf{X},\mathbf{Y}}^{k}(i,j)=d_G\left(\mathbf{x}_i^k,\mathbf{y}_j^k\right)^p. \tag{70}$$

$$f^{\gamma}(\mathbf{Y}) = \int \delta_{\tau^{\eta:\zeta}_{\alpha:\gamma}(\mathbf{X}_{0:k})}(\mathbf{Y}) f(\mathbf{X}_{0:k}) \delta\mathbf{X}_{0:k} \tag{66a}$$

$$= \sum_{\substack{\mathbf{X}^{\eta:\zeta}_{0:k}, \mathbf{X}^{\eta':\zeta'}_{0:k} \\ \mathbf{X}^{\eta:\zeta}_{0:k} \uplus \mathbf{X}^{\eta':\zeta'}_{0:k} = \mathbf{Y}}} \int \delta_{\tau^{\eta:\zeta}_{\alpha:\gamma}(\mathbf{X}^{\gamma}_{0:k})}(\mathbf{X}^{\eta:\zeta}_{0:k}) f(\mathbf{X}^{\gamma}_{0:k}) \delta\mathbf{X}^{\gamma}_{0:k} \int \delta_{\tau^{\eta:\zeta}_{\alpha:\gamma}(\mathbf{X}^{\gamma'}_{0:k})}(\mathbf{X}^{\eta':\zeta'}_{0:k}) f(\mathbf{X}^{\gamma'}_{0:k}) \delta\mathbf{X}^{\gamma'}_{0:k} \tag{66b}$$

$$= \int \delta_{\tau^{\eta:\zeta}_{\alpha:\gamma}(\mathbf{X}^{\gamma}_{0:k})}(\mathbf{Y}) f(\mathbf{X}^{\gamma}_{0:k}) \delta\mathbf{X}^{\gamma}_{0:k} \tag{66c}$$

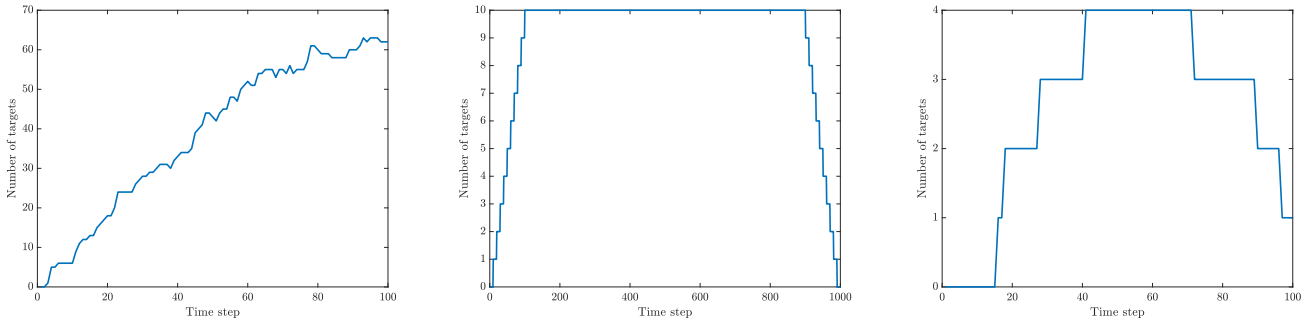
$$= \delta_{\emptyset}(\mathbf{Y}) e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle} + \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int \delta_{\tau^{\eta:\zeta}_{\alpha:\gamma}(\{X^1, \ldots, X^n\})}(\mathbf{Y}) e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle} \prod_{i=1}^{n} \lambda^{\eta:\zeta}(X^i) \mathrm{d}X^1 \cdots \mathrm{d}X^n \tag{66d}$$

$$= \delta_{\emptyset}(\mathbf{Y}) e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle} + \sum_{n=1}^{\infty} \frac{e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle}}{n!} \sum_{\substack{\mathbf{Y}^1, \ldots, \mathbf{Y}^n: \\ \uplus_{i=1}^{n} \mathbf{Y}^i = \mathbf{Y}}} \prod_{i=1}^{n} \int \delta_{\tau^{\eta:\zeta}_{\alpha:\gamma}(\{X^i\})}(\mathbf{Y}^i) \lambda^{\eta:\zeta}(X^i) \mathrm{d}X^i \tag{66e}$$

$$= \delta_{\emptyset}(\mathbf{Y}) e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle} + \sum_{n=1}^{\infty} \frac{e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle}}{n!} \sum_{\substack{\mathbf{Y}^1, \ldots, \mathbf{Y}^n: \\ \uplus_{i=1}^{n} \mathbf{Y}^i = \mathbf{Y}}} \prod_{i=1}^{n} \sum_{\substack{\beta, \varepsilon: \\ \mathbb{N}^{\varepsilon}_{\beta} \cap \mathbb{N}^{\zeta}_{\eta}}} \int \delta_{\{X^i\}}(\mathbf{Y}^i) \lambda(X^i) \mathrm{d}x^i_{\beta:\varepsilon \setminus b:e} \tag{66f}$$

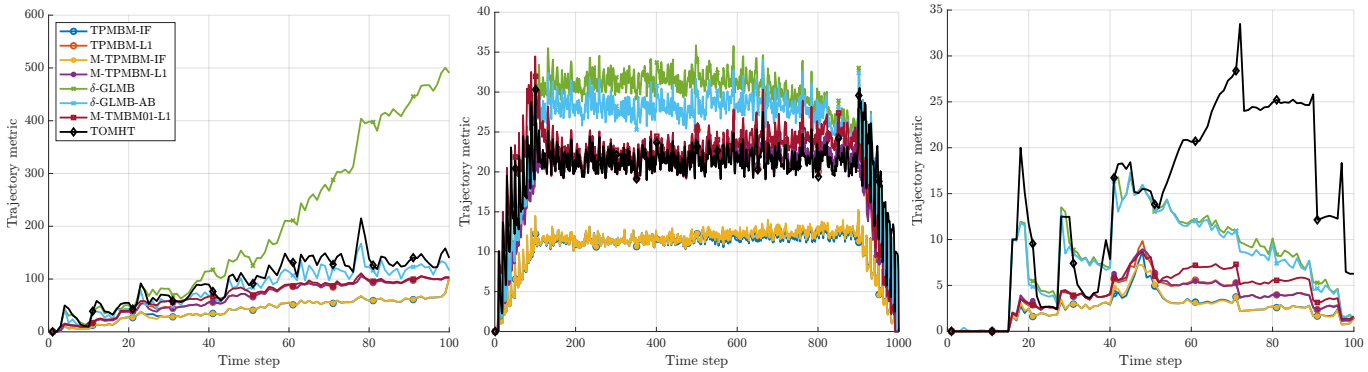
$$= e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle} \delta_{\emptyset}(\mathbf{Y}) + e^{-\langle \lambda^{\eta:\zeta}; 1 \rangle} \sum_{n=1}^{\infty} \delta_{\{Y^1, \ldots, Y^n\}}(\mathbf{Y}) \prod_{i=1}^{n} \sum_{\substack{\beta, \varepsilon: \\ \mathbb{N}^{\varepsilon}_{\beta} \cap \mathbb{N}^{\zeta}_{\eta}}} \int \delta_{\{X^i\}}(\{Y^i\}) \lambda(X^i) \mathrm{d}x^i_{\beta:\varepsilon \setminus b:e} \tag{66g}$$

$$= e^{-\mu \sum_{\beta, \varepsilon: \mathbb{N}^{\varepsilon}_{\beta} \cap \mathbb{N}^{\zeta}_{\eta}} P(\beta, \varepsilon)} \prod_{Y \in \mathbf{Y}} \mu p(Y) \sum_{\beta, \varepsilon: \mathbb{N}^{\varepsilon}_{\beta} \cap \mathbb{N}^{\zeta}_{\eta}} P(\beta, \varepsilon) \tag{66h}$$



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 5. Target cardinalities against time in each of the three scenarios in the simulations.



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 6. Trajectory estimation performance (evaluated at the final time step) in terms of the normalised TM over time

TABLE V
TRAJECTORY ESTIMATION PERFORMANCE IN TERMS OF THE NORMALISED TM (EVALUATED AT THE FINAL TIME STEP) AND ITS DECOMPOSITION

	Scenario 1					Scenario 2					Scenario 3				
Filter	TM	LE	ME	FE	SE	TM	LE	ME	FE	SE	TM	LE	ME	FE	SE
TPMBM-IF	4170.1	3884.8	273.4	10.5	1.5	10532.8	10182.7	349.7	0	0.4	247.3	235.5	4.3	0	7.5
TPMBM-L1	6487.2	6116.4	316.1	53.2	1.5	19584.2	18918.5	507.5	157.8	0.4	357.6	344.5	4.3	0	8.8
M-TPMBM-IF	4158.4	3885.7	262.7	9.3	0.7	10692.7	10165.3	527.0	0	0.4	250.1	235.5	6.5	0	8.0
M-TPMBM-L1	6444.5	6112.8	272.0	58.0	1.7	19698.8	18859.0	683.2	156.2	0.4	361.6	345.9	6.6	0.1	9.1
δ-GLMB	19590.7	5157.7	6433.3	8134.9	78.3	26998.5	16496.8	10356.0	282.5	1.1	760.8	325.1	438.4	19.5	20.2
δ-GLMB-AB	8175.5	5855.7	1641.7	718.7	13.3	24696.2	17269.2	7337.5	364.0	8.7	729.0	325.2	393.8	24.8	21.7
M-TMBM$_{01}$-L1	6939.8	6033.3	706.7	197.7	2.1	21628.8	18608.0	1934.2	1085.0	1.6	422.6	344.8	38.9	30.2	8.7
TOMHT	9606.2	5650.2	2730.8	1223.1	2.1	20518.0	18967.3	587.4	963.1	0.1	1390.1	507.8	542.2	325.4	15.4

with $\mathbf{x}_i^k$ being the target set of the i-th trajectory in $\mathbf{X}$ at time step k and $\mathbf{y}_j^k$ being the target set of the j-th trajectory in $\mathbf{Y}$ at time step k, and

$$d_G(\mathbf{x},\mathbf{y}) \triangleq \begin{cases} \min(c, d_b(x,y)) & \mathbf{x}=\{x\}, \mathbf{y}=\{y\} \\ 0 & \mathbf{x}=\mathbf{y}=\emptyset \\ \frac{c}{2^{1/p}} & \text{otherwise.} \end{cases} \tag{71}$$

A matrix $W^k \in \overline{\mathcal{W}}_{\mathbf{X},\mathbf{Y}}$ whose (i,j) element is $W^k(i,j)$ meets

$$\sum_{i=1}^{n_{\mathbf{X}}+1} W^k(i,j) = 1, \ j = 1,\ldots,n_{\mathbf{Y}} \tag{72}$$

$$\sum_{j=1}^{n_{\mathbf{Y}}+1} W^k(i,j) = 1, \ i = 1,\ldots,n_{\mathbf{X}} \tag{73}$$

$$W^k(n_{\mathbf{X}}+1, n_{\mathbf{Y}}+1) = 0, \tag{74}$$

$$W^k(i,j) \geq 0, \forall\ i,j. \tag{75}$$

Detailed explanations of these terms are provided in [70]. The TM normalised by the time window is then $d(\mathbf{X},\mathbf{Y})/T$.