Multi-Sensor Multi-Object Trajectory Estimation with Lineage Inference

Tran Thien Dat Nguyen, Hoa Van Nguyen, and Changbeom Shim

***Abstract*—In this paper, we propose an algorithm to estimate object trajectories and lineage information in the presence of spawning events using multiple sensors. For efficient computation, we approximate the predicted multi-object density with object spawning by a generalized labeled multi-Bernoulli (GLMB). To handle multi-sensor observation, we apply the multi-sensor GLMB filter update using Gibbs sampling with an approximate proposal to efficiently sample for significant GLMB components. The experimental results with a 3D tracking scenario demonstrate accurate estimation performance both in terms of object states and lineage information.**

***Index Terms*—multi-sensor multi-object tracking, spawning objects, 3D trajectory estimation.**

I. Introduction

Lineage information is important in situational awareness applications [1] where the origin of objects might bear critical information. Conversely, in cell biology, object ancestry might reveal migration patterns of cells, making it crucial in understanding their behaviors [2]. A tractable and arguably the most natural approach for lineage estimation is to assign distinct labels with ancestry information to objects and jointly estimate them with the object states [1]. Further, given each object has a distinct label, one can also construct the object trajectories from a sequence of multi-object states (sets of individual objects) by matching the labels over time, allowing the implementation of tractable estimation algorithms [3].

Labeled random finite set (LRFS) emerges as a framework to effectively handle multi-object estimation. In this approach, each individual object is assigned a unique label which could include the ancestry information [1]. Advances in LRFS provide a suite of models and mathematical apparatus to perform multi-object estimation [3], including object lineage inference [1]. Popular LRFS algorithms such as the labeled multi-Bernoulli (LMB) [4], generalized LMB (GLMB) filters/smoothers [5], [6] and their variations [7]–[10] have been developed to solve the multi-object estimation problem. Recent breakthroughs in LRFS approach include an algorithm that can track over one million objects [11], and an implementation that reduces the complexity to linear in the numbers of measurement and object [12], see the recent overview [3] for more details .

In principle, using multiple sensors could reduce the ambiguity and improve the tracking performance, especially, in 3D tracking problems [13], [14] where the uncertainty is generally higher than the 2D counterparts. Within the LRFS framework, the multi-sensor GLMB filter [10] has been proposed to efficiently address the multi-sensor tracking problem. In particular, the Gibbs sampling technique with minimally Markovian proposal achieves a complexity that is linear in the total number of measurement across all sensors. This complexity is significantly lower than solving the multi-sensor data association problem with classical multi-dimensional assignment techniques.

Leverage the advancement in LRFS, we propose an efficient algorithm that estimates trajectories with spawning objects using multiple sensors. In our technique, we approximate the predicted multi-object density induced by the spawning model with a GLMB that matches the first moment and cardinality distribution [15]. Subsequently, a Gibbs sampler with a minimally Markovian proposal is applied to sample for significant components of the updated density. In the experiment, we construct a 3D tracking scenario with four sensors where the proposed algorithm is tested over 100 Monte Carlo runs to verify its performance.

The paper is organized as follows. Section II presents backgrounds on LRFS and the multi-object filtering recursions. In Section III, we propose a closed-form multi-sensor multi-object filtering algorithm with object spawning, and present our implementation in Section IV. Section V shows the experimental results with a 3D tracking scenario. In Section VI, we provide conclusions and discuss potential future works.

II. Backgrounds

Multi-object trajectories can be constructed from the multi-object states given each object has a distinct label. In particular, a trajectory is a sequence of states of the same label over time. Early works on random finite set do not concern object labels due to the computational expense. Nevertheless, labels are necessary for the development of efficient algorithms, see the arguments in [3]. An LRFS on a state space $\mathbb{X}$ and a

T.T.D. Nguyen, H.V. Nguyen and C. Shim are with the School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Australia (email: {t.nguyen1, hoa.v.nguyen, changbeom.shim}@curtin.edu.au).
This work was supported by the Australian Research Council under Grant LP200301507. Corresponding author: Changbeom Shim.

(discrete) label space $\mathbb{L}$ is a finite-set-valued random variable on the space $\mathbb{X} \times \mathbb{L}$ such that each realization has distinct labels [5]. All information of the multi-object state is captured in the multi-object posterior density which can be propagated forward in time via the Bayes posterior recursion, see e.g. [6]. Nevertheless, propagating the multi-object posterior density is expensive since the complexity grows with the number of scans. A more efficient option is to propagate the multi-object filtering density via the Bayes filtering recursion, which is usually carried out in two steps. Given a multi-object prior $\boldsymbol{\pi}(\boldsymbol{X})$ at the current time step and the multi-object transition density $\boldsymbol{f}(\boldsymbol{X}_+|\boldsymbol{X})$–characterizes the multi-object dynamic model–where the subscript '+' is used to denote next time step, the predicted multi-object density is first computed via the Chapman-Kolmogorov equation as

$$\boldsymbol{\pi}_+(\boldsymbol{X}_+) = \int \boldsymbol{\pi}(\boldsymbol{X})\boldsymbol{f}(\boldsymbol{X}_+|\boldsymbol{X})\delta\boldsymbol{X}. \tag{1}$$

Subsequently, the updated multi-object density is computed as

$$\boldsymbol{\pi}_+(\boldsymbol{X}_+|Z_+) \propto \boldsymbol{\pi}_+(\boldsymbol{X}_+)\boldsymbol{g}(Z_+|\boldsymbol{X}_+), \tag{2}$$

where $\boldsymbol{g}(Z_+|\boldsymbol{X}_+)$ is the likelihood that $\boldsymbol{X}_+$ generates the observation Z_+, which characterizes the observation model.

A GLMB density can be written in its δ-form as [5]

$$\boldsymbol{\pi}(\boldsymbol{X}) = \Delta(\boldsymbol{X}) \sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi} w^{(I,\xi)}\delta_I[\mathcal{L}(\boldsymbol{X})] \left[p^{(\xi)}\right]^{\boldsymbol{X}}, \tag{3}$$

where $\Delta(\boldsymbol{X})$ is 1 if all labels of $\boldsymbol{X}$ are distinct and 0 otherwise, Ξ is an arbitrary discrete space, $\mathcal{F}(\mathbb{L})$ is the set of all finite subsets of $\mathbb{L}$, $\delta_X[Y]$ is a Kronecker delta function which takes 1 if $X = Y$ and 0 otherwise, $\mathcal{L}(\boldsymbol{X})$ is the set of labels of $\boldsymbol{X}$, $p^{(\xi)}$ is a probability density function on $\mathbb{X}$. Each component (I, ξ) of the sum is referred to as a "hypothesis". The GLMB filtering recursion is closed under the standard multi-object dynamic and observation models [5].

As the filtering density is propagated forward in time, the number of hypotheses grows exponentially. For tractability, ranked assignment algorithms are needed to select only significant hypotheses for propagation. Early works on GLMB filter use classical ranked assignment algorithms such as k-shortest paths and Murty [16] with relatively high complexity. Later on, Gibbs sampling [7] and tempering technique [12] reduce the complexity to linear in numbers of object and measurement.

Early techniques to process multi-sensor observation include sequentially updating the predicted multi-object density using observation from one sensor to the next. Thus, the performance depends on the order of the sensor update. Further, in the GLMB filter, the updated density is not accurate since good GLMB components might be discarded in one sensor update iterate and not included in the next. Conversely, the techniques proposed in [10], [17] directly sample the joint multi-sensor data association, alleviating the issues associating with the sequential update scheme. Especially, the Gibbs sampling technique with a sub-optimal proposal achieves state-of-the-art performance with a relatively low complexity.

Apart from the standard models, the GLMB filter and its single-term approximation–the LMB filter–can also handle different models including track-before-detect observation [18], merged/extended measurements [19], [20], spawning [1] and interacting objects [21]. They have been applied to track pedestrians [14], [22], vehicles [23], space objects [24], biological cells [25], [26], to perform simultaneous localization and mapping [27], and to control and schedule sensors [28]–[30]. The LRFS framework also provides the foundation for different multi-sensor fusion techniques, e.g., [31]–[36].

III. The Multi-Sensor GLMB Filter with Object Spawning

In this section, we propose an algorithm to perform multi-sensor multi-object tracking with spawning objects.

A. The Dynamic Model with Object Spawning

There are three types of objects that might exist at each time step: new births; survivals; and spawns. The statistics of new births are described by an LMB $\boldsymbol{f}_{B,+}(\cdot)$ with parameters $\{P_{B,+}(\ell), p_{B,+}(\cdot,\ell)\}_{\ell\in\mathbb{B}_+}$, where $P_{B,+}(\ell)$ is the birth probability of an object with birth label $\ell \in \mathbb{B}_+$ ($\mathbb{B}_+ \cap \mathbb{L} = \emptyset$) and $p_{B,+}(\cdot,\ell)$ is the state probability density function if it exists. Conversely, an object $\boldsymbol{x} = (x, \ell)$ from the current time step can survive to the next one with probability $P_S(\boldsymbol{x})$ or disappear with probability $1-P_S(\boldsymbol{x})$. If it survives, its statistics is given by the transition density $f_{S,+}(\cdot|x,\ell)$. The multi-object transition density of the surviving objects is given by [5]

$$\boldsymbol{f}_{S,+}(\boldsymbol{X}_{S,+}|\boldsymbol{X}) = \Delta(\boldsymbol{X}_{S,+})\mathbf{1}_{\mathcal{L}(\boldsymbol{X})}^{\mathcal{L}(\boldsymbol{X}_{S,+})} \left[\Phi_S\left(\boldsymbol{X}_{S,+}|\cdot\right)\right]^{\boldsymbol{X}}, \tag{4}$$

where $\mathbf{1}_S^X = \prod_{x\in X}\mathbf{1}_S(x)$ with $\mathbf{1}_S(x)$ takes 1 if $x \in S$ and 0 otherwise, and

$$\begin{aligned}\Phi_S(\boldsymbol{V}|x,\ell) = &\sum_{(x_+,\ell_+)\in\boldsymbol{V}} \delta_\ell[\ell_+]P_S(x,\ell)f_{S,+}(x_+|x,\ell) \\ &+ (1-\mathbf{1}_{\mathcal{L}(\boldsymbol{V})}(\ell))(1-P_S(x,\ell)).\end{aligned} \tag{5}$$

An object at the current time step can spawn up to N_T objects in the next time step. The transition density for spawned objects is given by [1]

$$\boldsymbol{f}_{T,+}(\boldsymbol{X}_{T,+}|\boldsymbol{X}) = \Delta(\boldsymbol{X}_{T,+})\mathbf{1}_{\mathcal{T}_+(\mathcal{L}(\boldsymbol{X}))}^{\mathcal{L}(\boldsymbol{X}_{T,+})} \left[\Phi_{T,+}(\boldsymbol{X}_{T,+}|\cdot)\right]^{\boldsymbol{X}}, \tag{6}$$

where $\mathcal{T}_+(L) = \biguplus_{\ell\in L}\mathcal{T}_+(\ell)$, $\mathcal{T}_+(\ell)$ is the set of labels that could spawn from ℓ at the next time step, and

$$\Phi_{T,+}(\boldsymbol{U}|x,\ell) = \left[\boldsymbol{\phi}_{T,+}(\boldsymbol{U}\cap(\mathbb{X}\times\mathcal{T}_+(\ell))|x,\ell;\cdot)\right]^{\mathcal{T}_+(\ell)}, \tag{7}$$

$$\begin{aligned}\boldsymbol{\phi}_{T,+}(\boldsymbol{W}|x,\ell;\varsigma) = &\sum_{(x_+,\ell_+)\in\boldsymbol{W}} \delta_\varsigma[\ell_+]P_T(x,\ell;\varsigma)f_{T,+}(x_+|x,\ell;\varsigma) \\ &+ (1-\mathbf{1}_{\mathcal{L}(\boldsymbol{W})}(\varsigma))(1-P_T(x,\ell;\varsigma)).\end{aligned} \tag{8}$$

By convention, a spawn label is the parent label concatenates with the time of spawn and the spawn index. Thus, $\mathcal{T}_+(\ell) = \{\zeta_{1,+}^{(\ell)}, ..., \zeta_{N_T,+}^{(\ell)}\} = \{(\ell, k+1, 1), ..., (\ell, k+1, N_T)\}$, where $k+1$ is the next time step. We denote the space of all spawn labels at the next time step as $\mathbb{T}_+ = \mathcal{T}_+(\mathbb{L})$, hence $\mathbb{L}_+ = \mathbb{B}_+ \uplus \mathbb{L} \uplus \mathbb{T}_+$. Since $\mathbb{B}_+$, $\mathbb{L}$, and $\mathbb{T}_+$ are disjoint, the overall multi-object dynamic model is given by

$$\boldsymbol{f}(\boldsymbol{X}_+|\boldsymbol{X}) = \boldsymbol{f}_{B,+}(\boldsymbol{X}_{B,+})\boldsymbol{f}_{S,+}(\boldsymbol{X}_{S,+}|\boldsymbol{X})\boldsymbol{f}_{T,+}(\boldsymbol{X}_{T,+}|\boldsymbol{X}). \tag{9}$$

B. The Standard Multi-Sensor Observation Model

At each time step, we receive a multi-sensor observation $Z = (Z^{(1)}, ..., Z^{(V)})$ from V sensors. An object $\boldsymbol{x} = (x, \ell)$ could be detected on the v^{th} sensor ($v \in \{1:V\}$) with probability $P_D^{(v)}(\boldsymbol{x})$ or miss-detected with probability $1 - P_D^{(v)}(\boldsymbol{x})$. If the object is detected on sensor v, it generates a measurement z according to the likelihood function $g^{(v)}(z|\boldsymbol{x})$. The multi-sensor data association map is defined as $\theta = (\theta^{(1)}, ..., \theta^{(V)})$, where $\theta^{(v)} : \mathbb{L} \to \{0 : |Z^{(v)}|\}$ [10]. If $\theta^{(v)}(\ell) = 0$, object with label ℓ is miss-detected on sensor v, otherwise, the measurement $z_{\theta^{(v)}(\ell)}$ is assigned to the object. The observation from each sensor might also include clutter measurements which is modeled by a Poisson RFS with intensity function $\kappa^{(v)}$. The multi-sensor likelihood function is given by [10]

$$\boldsymbol{g}(Z|\boldsymbol{X}) = \sum_{\theta \in \Theta(\mathcal{L}(\boldsymbol{X}))} \left[\Psi_Z^{(\theta \circ \mathcal{L}(\cdot))}(\cdot)\right]^{\boldsymbol{X}}, \tag{10}$$

where $\Theta(L) \subseteq \Theta$ is the set of data association maps with domain L, $\Psi_Z^{(j^{(1)},...,j^{(v)})}(\boldsymbol{x}) = \prod_{v=1}^{V} \psi_{Z^{(v)}}^{(v,j^{(v)})}(\boldsymbol{x})$, and

$$\psi_{\{z_{1:|Z^{(v)}|}\}}^{(v,j)}(\boldsymbol{x}) = \begin{cases} 1 - P_D^{(v)}(\boldsymbol{x}), & j = 0 \\ P_D^{(v)}(\boldsymbol{x}) \frac{g^{(v)}(z_j|\boldsymbol{x})}{\kappa^{(v)}(z_j)}, & j > 0 \end{cases}. \tag{11}$$

C. Filtering Recursion with GLMB Approximation

Our algorithm is based on an approximate prediction step and the multi-sensor GLMB update. First, applying the Chapman-Kolmogorov in (1) using the prior GLMB in (3) and the transition density in (9) yields the predicted density [1]

$$\boldsymbol{\pi}_+(\boldsymbol{X}_+) = \Delta(\boldsymbol{X}_+) \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} w_+^{(I,\xi)}(\mathcal{L}(\boldsymbol{X}_+)) p^{(I,\xi)}(\boldsymbol{X}_+), \tag{12}$$

where

$$w_+^{(I,\xi)}(L) = w^{(I,\xi)} \mathbf{1}_I^{L \cap \mathbb{L}} \mathbf{1}_{\mathcal{T}_+(I)}^{L \cap \mathbb{T}_+} [1 - P_B]^{\mathbb{B}_+ - L} [P_B]^{\mathbb{B}_+ \cap L}, \tag{13}$$

$$p^{(I,\xi)}(\boldsymbol{X}_+) = [p_B]^{\boldsymbol{X}_+ \cap \mathbb{X} \times \mathbb{B}_+} \prod_{\ell \in I} \int \Phi_{S,+}(\boldsymbol{X}_+ \cap \mathbb{X} \times \mathbb{L}|x, \ell) \times \Phi_{T,+}(\boldsymbol{X}_+ \cap \mathbb{X} \times \mathbb{T}_+|x, \ell) p^{(\xi)}(x, \ell) dx. \tag{14}$$

The predicted density is approximated by a GLMB that matches its first moment and cardinality distribution [15], i.e.,

$$\hat{\boldsymbol{\pi}}_+(\boldsymbol{X}_+) = \Delta(\boldsymbol{X}_+) \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} \sum_{I_+ \in \mathcal{F}(\mathbb{L}_+)} w_+^{(I,\xi)}(I_+) \times \delta_{I_+}[\mathcal{L}(\boldsymbol{X}_+)][\hat{p}^{(I,\xi,I_+)}]^{\boldsymbol{X}_+}, \tag{15}$$

where $\langle p \rangle(\{\ell_1, ..., \ell_n\}) = \int p(\{(x_1, \ell_1), ..., (x_n, \ell_n)\}) dx_{1:n}$,

$$\hat{p}^{(I,\xi,I_+)}(x, \ell) = \begin{cases} p_B(x, \ell), & \ell \in \mathbb{B}_+ \\ \mathbf{1}_{I_+}(\ell) \langle p^{(I,\xi)}((x, \ell) \uplus \cdot) \rangle (I_+ - \{\ell\}), & \ell \notin \mathbb{B}_+ \end{cases}. \tag{16}$$

Subsequently, given an observation Z_+, applying the multi-sensor GLMB update on (15) yields

$$\boldsymbol{\pi}_+(\boldsymbol{X}_+|Z_+) \propto \Delta(\boldsymbol{X}_+) \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} \sum_{(I_+,\theta_+) \in \mathcal{F}(\mathbb{L}_+) \times \Theta_+} w_{Z_+}^{(I,\xi,I_+,\theta_+)} \delta_{I_+}[\mathcal{L}(\boldsymbol{X}_+)][p_{Z_+}^{(I,\xi,I_+,\theta_+)}]^{\boldsymbol{X}_+}, \tag{17}$$

where

$$w_{Z_+}^{(I,\xi,I_+,\theta_+)} = w_+^{(I,\xi)}(I_+) \mathbf{1}_{\Theta_+(I_+)}^{\{\theta_+\}} [\langle \hat{p}^{(I,\xi,I_+)}, \Psi_{Z_+}^{(\theta_+ \circ \mathcal{L}(\cdot))} \rangle(\cdot)]^{I_+}, \tag{18}$$

$$p_{Z_+}^{(I,\xi,I_+,\theta_+)}(x, \ell) \propto \hat{p}^{(I,\xi,I_+)}(x, \ell) \Psi_{Z_+}^{(\theta_+ \circ \mathcal{L}(\ell))}(x, \ell). \tag{19}$$

IV. Linear Gaussian Implementation

In this section, we present our assumptions and techniques to implement the proposed algorithm.

A. GLMB Approximation

We assume the the surviving and spawning probabilities only depend on the object labels, i.e., $P_S(x, \ell) = \bar{P}_S(\ell)$ and $P_T(x, \ell; \varsigma) = \bar{P}_T(\varsigma)$. The birth probability density function is a Gaussian, i.e., $p_{B,+}(\cdot, \ell) = \mathcal{N}(\cdot, \mu_B^{(\ell)}, Q_B^{(\ell)})$ (a Gaussian with mean $\mu_B^{(\ell)}$ and covariance $Q_B^{(\ell)}$). Further, we also assume the survival and spawning transition densities are linear Gaussian, i.e., $f_{S,+}(\cdot|x, \ell) = \mathcal{N}(\cdot, F_S x, Q_S)$ and $f_{T,+}^{(i)}(\cdot|x, \ell; \zeta_i^{(\ell)}) = \mathcal{N}(\cdot, F_T^{(i)} x, Q_T^{(i)})$ (for the i^{th} spawn from ℓ). The term $p^{(I,\xi)}(\boldsymbol{X}_+)$ in (12) can be written as:

$$p^{(I,\xi)}(\boldsymbol{X}_+) = \bar{w}^{(I,\xi)}(\mathcal{L}(\boldsymbol{X}_+)) \bar{p}^{(I,\xi)}(\boldsymbol{X}_+), \tag{20}$$

where $\bar{w}^{(I,\xi)}$ is given in (21) and $\bar{p}^{(I,\xi)}$ is given in (22). The projection $\mathcal{A}: \mathbb{X} \times \mathbb{L} \mapsto \mathbb{X}$ in (22) is defined as $\mathcal{A}((x, \ell)) = x$.

Thus, the approximate prediction density in (15) becomes

$$\hat{\boldsymbol{\pi}}_+(\boldsymbol{X}_+) = \Delta(\boldsymbol{X}_+) \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} \sum_{I_+ \in \mathcal{F}(\mathbb{L}_+)} \tilde{w}_+^{(I,\xi,I_+)} \times \delta_{I_+}[\mathcal{L}(\boldsymbol{X}_+)][\tilde{p}^{(I,\xi,I_+)}]^{\boldsymbol{X}_+}, \tag{23}$$

where

$$\tilde{w}_+^{(I,\xi,I_+)} = w_+^{(I,\xi)}(I_+)\, \bar{w}^{(I,\xi)}(I_+), \tag{24}$$

$$\tilde{p}^{(I,\xi,I_+)}(x, \ell) = \mathbf{1}_{I_+}(\ell) \langle \bar{p}^{(I,\xi)}((x, \ell) \uplus \cdot) \rangle (I_+ - \{\ell\}). \tag{25}$$

$$\bar{w}^{(I,\xi)}(L) = [1-\bar{P}_S]^{I-L}[\bar{P}_S]^{I\cap L}[1-\bar{P}_T]^{\mathcal{T}(I)-L}[\bar{P}_T]^{L\cap\mathbb{T}_+}, \tag{21}$$

$$\bar{p}^{(I,\xi)}(\boldsymbol{X}_+) = [p_B]^{\boldsymbol{X}_+\cap(\mathbb{X}\times\mathbb{B}_+)}\prod_{\ell\in I}\int(1-\mathbf{1}_{\mathcal{L}(\boldsymbol{X}_+)}(\ell)+\mathbf{1}_{\mathcal{L}(\boldsymbol{X}_+)}(\ell)f_{S,+}(\mathcal{A}(\boldsymbol{X}_+\cap\mathbb{X}\times\{\ell\})|x,\ell))\times \tag{22}$$
$$\prod_{i=1}^{N_T}(1-\mathbf{1}_{\mathcal{L}(\boldsymbol{X}_+)}(\zeta_{i,+}^{(\ell)})+\mathbf{1}_{\mathcal{L}(\boldsymbol{X}_+)}(\zeta_{i,+}^{(\ell)})f_{T,+}^{(i)}(\mathcal{A}(\boldsymbol{X}_+\cap\mathbb{X}\times\{\zeta_{i,+}^{(\ell)}\})|x,\ell;\zeta_{i,+}^{(\ell)}))p^{(I,\xi)}(x,\ell)dx.$$

We also note the following Gaussian identity. If $f(x_i|x)=\mathcal{N}(x_i,F_ix,Q_i)$ and $p(x)=\mathcal{N}(x,\mu,P)$, then

$$p(x_1,...,x_n)=\int f(x_1|x)f(x_2|x)...f(x_n|x)p(x)dx$$
$$=\mathcal{N}\left(\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\,|\,\begin{bmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{bmatrix},\begin{bmatrix}Q_{11}&Q_{12}&\cdots&Q_{1n}\\Q_{21}&Q_{22}&\cdots&Q_{2n}\\\vdots&\vdots&\cdots&\vdots\\Q_{n1}&Q_{n2}&\cdots&Q_{nn}\end{bmatrix}\right),$$

where $\mu_i=F_i\mu$, $Q_{ii}=F_iPF_i^T+Q_i$, and $Q_{ij}=F_iPF_j^T$. Further, we also realize that the approximation

$$p(x_1,...,x_n)\approx\hat{p}(x_1)\hat{p}(x_2)...\hat{p}(x_n),$$

where $\hat{p}(x_i)=\int p(x_i,x_1,...,x_{n-1})dx_{1:n-1}=\mathcal{N}(\cdot,F_i\mu,Q_{ii})$, is the same as the approximation of the joint density of the unlabeled states given in (25). Thus, $\tilde{p}^{(I,\xi,I_+)}(\cdot,\ell_+)$ can be computed independently of I_+ as

$$\tilde{p}^{(I,\xi,I_+)}(\cdot,\ell_+)=\begin{cases}\mathcal{N}(\cdot,\mu_B^{(\ell_+)},Q_B^{(\ell_+)}), & \ell_+\in\mathbb{B}_+\\ \mathcal{N}(\cdot,F_S\mu,F_SPF_S^T+Q_S), & \ell_+=\ell\\ \mathcal{N}(\cdot,F_T^{(i)}\mu,F_T^{(i)}P[F_T^{(i)}]^T+Q_T^{(i)}), & \ell_+=\zeta_{i,+}^{(\ell)}\end{cases},$$

for $p^{(\xi)}(\cdot,\ell)=\mathcal{N}(\cdot,\mu,P)$ is the prior single-object density.

B. Sampling for Multi-Sensor Data Association

Given the approximate predicted density and the multi-sensor observation $Z_+=(Z_+^{(1)},..,,Z_+^{(V)})$, following [10], we define the extended association map as

$$\gamma:\mathbb{B}_+\uplus I\uplus\mathcal{T}_+(I)\to\{-1\}^V\uplus\Lambda^{(1:V)},$$

where $\Lambda^{(1:V)}=\Lambda^{(1)}\times...\times\Lambda^{(V)}$ and $\Lambda^{(v)}=\{0:|Z_+^{(v)}|\}$.

We apply the same Gibbs sampling technique with the minimally Markovian proposal density as proposed in [10], except the enumerated label set is now $\{\ell_{1:P}\}=\mathbb{B}_+\uplus I\uplus\mathcal{T}_+(I)$ and equation (39) in [10] is replaced by

$$\eta_n^{(v)}(j^{(v)})=\begin{cases}(1-P_B(\ell_n))^{\delta_1[v]}, & \ell_n\in\mathbb{B}_+,j^{(v)}=-1\\ (P_B(\ell_n))^{\delta_1[v]}\bar{\psi}_{Z_+^{(v)}}^{(\xi,v,j^{(v)})}(\ell_n), & \ell_n\in\mathbb{B}_+,j^{(v)}\in\Lambda^{(v)}\\ (1-\bar{P}_S(\ell_n))^{\delta_1[v]}, & \ell_n\in I,j^{(v)}=-1\\ (\bar{P}_S(\ell_n))^{\delta_1[v]}\bar{\psi}_{Z_+^{(v)}}^{(\xi,v,j^{(v)})}(\ell_n), & \ell_n\in I,j^{(v)}\in\Lambda^{(v)}\\ (1-\bar{P}_T(\ell_n))^{\delta_1[v]}, & \ell_n\in\mathcal{T}(I),j^{(v)}=-1\\ (\bar{P}_T(\ell_n))^{\delta_1[v]}\bar{\psi}_{Z_+^{(v)}}^{(\xi,v,j^{(v)})}(\ell_n), & \ell_n\in\mathcal{T}(I),j^{(v)}\in\Lambda^{(v)}\end{cases}.$$

where $\bar{\psi}_{Z_+^{(v)}}^{(\xi,v,j^{(v)})}(\ell)=\langle\tilde{p}^{(I,\xi,I_+)},\psi_{Z_+^{(v)}}^{(v,j^{(v)})}\rangle(\ell)$ (note that $\tilde{p}^{(I,\xi,I_+)}$ does not depend on I_+ in our implementation as discussed above). We also assume the detection probability depends only on object labels, i.e., $P_D^{(v)}(x,\ell)=\bar{P}_D^{(v)}(\ell)$ and the single-object likelihood function is linear Gaussian[1], i.e., $g^{(v)}(\cdot|x,\ell)=\mathcal{N}(\cdot|H^{(v)}x,R^{(v)})$. Thus, the Bayes evidence terms $\bar{\psi}_{Z_+}^{(\xi,v,j^{(v)})}(\ell)$ can be computed analytically. After obtaining the multi-sensor data association samples from the Gibbs sampler (using Algorithm 2 in [10]), we can analytically compute the updated weights and the single-object densities according to (18) and (19), respectively.

V. Experimental Results

To evaluate the performance of the proposed method, we construct a 3D tracking scenario with object spawning. The total number of trajectories is 22 and the tracking volume is $[-1000\text{m},1000\text{m}]^3$. The lineage information of the ground truth objects are given in Table I. We test three multi-sensor algorithms: multi-sensor GLMB with the minimally Markovian proposal with object spawning (MM-Spawn) and without object spawning (MM-No Spawn); and the iterated corrector multi-sensor GLMB with object spawning (IT-Spawn).

We model the 3D state vector as $[p_x,\dot{p}_x,p_y,\dot{p}_y,p_z,\dot{p}_z]$. There are four birth regions modeled by an LMB with four components, each has the existence (birth) probability of 0.03. The state probability density functions of the components are:

- Region 1: $\mathcal{N}(\cdot,[0.1,0,0.1,0,0.1,0]^T,100\otimes\mathbf{I}_6)$;
- Region 2: $\mathcal{N}(\cdot,[400,0,-600,0,200,0]^T,100\otimes\mathbf{I}_6)$;
- Region 3: $\mathcal{N}(\cdot,[-800,0,-200,0,-400,0]^T,100\otimes\mathbf{I}_6)$;
- Region 4: $\mathcal{N}(\cdot,[-200,0,800,0,600,0]^T,100\otimes\mathbf{I}_6)$;

where $\mathbf{I}_n$ is the n-D identity matrix, and $\otimes$ denotes the Kronecker multiplication. The survival state transition matrix and the process noise covariance are given respectively as

$$F_S=\begin{bmatrix}1&1\\0&1\end{bmatrix}\otimes\mathbf{I}_3\text{ and }Q_S=25\times\begin{bmatrix}1/4&1/2\\1/3&1\end{bmatrix}\otimes\mathbf{I}_3.$$

The spawning state transition matrix and the process noise covariance are given respectively as

$$F_T^{(i)}=\begin{bmatrix}1&0\\0&0\end{bmatrix}\otimes\mathbf{I}_3\text{ and }Q_T=50\otimes\mathbf{I}_6,$$

[1]Although we assume linear Gaussian in this work for efficiency, more complicated models can also be handled, e.g., using particle representation.

Table I: Ground truth birth and spawning events. The 'Parent' column indicates the ID of the parents. Linages with more than one object are bolded and highlighted in different colors.

ID	Birth Region	Start Time	Parent	Generation
1	**1**	**1**	**-**	**0**
2	**1**	**10**	**1**	**1**
3	**1**	**10**	**1**	**1**
4	**1**	**30**	**2**	**2**
5	**2**	**1**	**-**	**0**
6	**2**	**15**	**5**	**1**
7	**2**	**15**	**5**	**1**
8	**2**	**15**	**5**	**1**
9	**2**	**35**	**6**	**2**
10	**3**	**1**	**-**	**0**
11	**3**	**20**	**10**	**1**
12	**3**	**35**	**11**	**2**
13	**4**	**40**	**-**	**0**
14	**4**	**50**	**13**	**1**
15	2	20	-	0
16	2	20	-	0
17	1	20	-	0
18	3	40	-	0
19	3	60	-	0
20	4	60	-	0
21	1	80	-	0
22	4	80	-	0

for $i \in \{1:5\}$ (i.e., each object can spawn up to 5 objects at each time step). The survival and spawn probabilities are constants of 0.99 and 0.03, respectively. The single-sensor observation matrix and covariance of the noise are

$$H^{(v)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } R^{(v)} = 100 \otimes \mathbf{I}_3,$$

respectively, for $v \in \{1:4\}$. The detection probability of the sensors is 0.8 and the (uniform) clutter rate is 20. The experiment is run over 100 Monte Carlo (MC) trials.

We use OSPA$^{(2)}$ metric [11], [37] to evaluate the distance (with localization and cardinality components, average across MC runs) between the estimates and the ground truth. Figure 1 indicates that MM-Spawn exhibits the lowest OSPA$^{(2)}$ error among the tested filters. IT-Spawn shows the worst cardinality estimation while its localization error is similar to one of the MM-No Spawn. The inferior performance of the IT-Spawn filter is due to the inaccurate update step. The MM-No Spawn does not perform well since it cannot predict the existence of spawned objects. To investigate the lineage estimation performance, we plot the average numbers (across all MC runs) of birth and spawning events from 4 birth regions in Figure 2. To avoid reporting false tracks, we only count an event if the corresponding object survives for more than 5 time steps. Compare to the ground truth in Table I, there are one or two time steps delays for the filter to initialize new tracks. Nevertheless, the overall lineage information is correct across MC runs, verifying the efficacy of our technique in estimating lineage information. On average, for the tested scenario and our

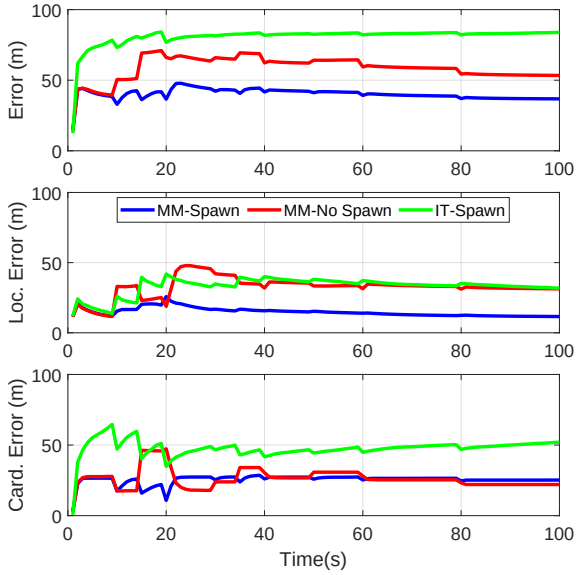

Figure 1: Mean OSPA$^{(2)}$ errors for different algorithms.

implementations, MM-No-Spawn has the fastest computation time, averaging 7.81 seconds per time step. The proposed technique takes approximately four times as long as MM-No-Spawn, while IT-Spawn takes about 1.4 times as long. The runtime increase is due to the higher number of tracks and hypotheses required for object spawning, and the multi-dimensional assignment problem in MM-Spawn.

VI. Conclusion

We have proposed an algorithm to estimate object trajectories and lineage information leveraging the advancement in LRFS. Our technique approximates the predicted multi-object density by a GLMB to achieve closed-form filtering recursion. To handle multi-sensor observation, we apply the multi-sensor GLMB update step using Gibbs sampling with a minimally Markovian proposal for efficiency. Experimental results demonstrate the efficacy of our technique in estimate 3D trajectories and their lineage. Future research will focus on different approximation techniques as well as extension to more complicated models.

References

[1] D. S. Bryant, B.-T. Vo, B.-N. Vo, and B. A. Jones, "A generalized labeled multi-Bernoulli filter with object spawning," *IEEE Trans. Signal Process.*, 66(23):6177–6189, 2018.

[2] T. T. D. Nguyen and D. Y. Kim, "On-line tracking of cells and their lineage from time lapse video data," in *Int. Conf. Control Automat. Inf. Sci.*, pp. 291–296, 2018.

[3] B.-N. Vo, B.-T. Vo, T. T. D. Nguyen, and C. Shim, "An overview of multi-object estimation via labeled random finite set," *IEEE Trans. Signal Process.*, 72:4888–4917, 2024.

[4] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, "The labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, 62(12):3246–3260, 2014.

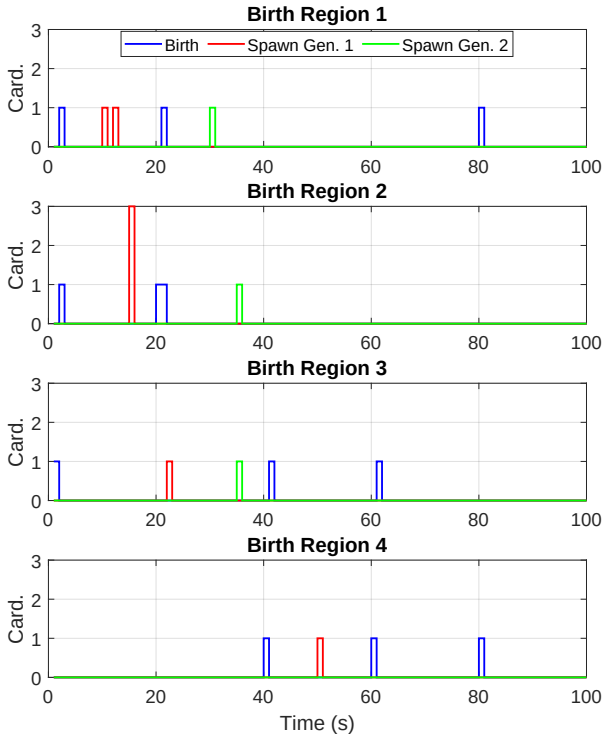

Figure 2: Mean numbers of birth and spawning events from MM-Spawn algorithm.

[5] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Trans. Signal Process.*, 61(13):3460–3475, 2013.

[6] B.-N. Vo and B.-T. Vo, "A multi-scan labeled random finite set model for multi-object state estimation," *IEEE Trans. Signal Process.*, 67(19):4948–4963, 2019.

[7] B.-N. Vo, B.-T. Vo, and H. G. Hoang, "An efficient implementation of the generalized labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, 65(8):1975–1987, 2017.

[8] W. Yi, M. Jiang, and R. Hoseinnezhad, "The multiple model Vo-Vo filter," *IEEE Trans. Aerosp. Electron. Syst.*, 53(2):1045–1054, 2017.

[9] D. Moratuwage, B.-N. Vo, B.-T. Vo, and C. Shim, "Multi-scan multi-sensor multi-object state estimation," *IEEE Trans. Signal Process.*, 70:5429–5442, 2022.

[10] B.-N. Vo, B.-T. Vo, and M. Beard, "Multi-sensor multi-object tracking with the generalized labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, 67(23):5952–5967, 2019.

[11] M. Beard, B. T. Vo, and B.-N. Vo, "A solution for large-scale multi-object tracking," *IEEE Trans. Signal Process.*, 68:2754–2769, 2020.

[12] C. Shim, B.-T. Vo, B.-N. Vo, J. Ong, and D. Moratuwage, "Linear complexity Gibbs sampling for generalized labeled multi-Bernoulli filtering," *IEEE Trans. Signal Process.*, 71:1981–1994, 2023.

[13] J. Ong, B.-T. Vo, B.-N. Vo, D. Y. Kim, and S. Nordholm, "A Bayesian filter for multi-view 3D multi-object tracking with occlusion handling," *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(5):2246-2263, 2022.

[14] L. Van Ma, T. T. D. Nguyen, B.-N. Vo, H. Jang, and M. Jeon, "Track initialization and re-identification for 3D multi-view multi-object tracking," *Inf. Fusion*, 111:102496, 2024.

[15] F. Papi, B.-N. Vo, B.-T. Vo, C. Fantacci, and M. Beard, "Generalized labeled multi-Bernoulli approximation of multi-object densities," *IEEE Trans. Signal Process.*, 63(20)5487–5497, 2015.

[16] B.-N. Vo, B.-T. Vo, and D. Phung, "Labeled random finite sets and the Bayes multi-target tracking filter," *IEEE Trans. Signal Process.*, 62(24):6554–6567, 2014.

[17] C.-T. Do, T. T. D. Nguyen, and H. V. Nguyen, "Robust multi-sensor generalized labeled multi-Bernoulli filter," *Signal Process.*, 192:108368, 2022.

[18] F. Papi, B.-T. Vo, M. Bocquel, and B.-N. Vo, "Multi-target track-before-detect using labeled random finite set," in *Int. Conf. Control Automat. Inf. Sci.*, pp. 116–121, 2013.

[19] M. Beard, B.-T. Vo, and B.-N. Vo, "Bayesian multi-target tracking with merged measurements using labelled random finite sets," *IEEE Trans. Signal Process.*, 63(6):1433–1447, 2015.

[20] M. Beard, S. Reuter, K. Granström, B.-T. Vo, B.-N. Vo, and A. Scheel, "Multiple extended target tracking with labeled random finite sets," *IEEE Trans. Signal Process.*, 64(7):1638–1653, 2015.

[21] N. Ishtiaq, A. K. Gostar, A. Bab-Hadiashar, and R. Hoseinnezhad, "Interaction-aware labeled multi-Bernoulli filter," *IEEE Trans. Intell. Transp. Syst.*, 24(11):11668–11681, 2023.

[22] L. Van Ma, T. T. D. Nguyen, C. Shim, D. Y. Kim, N. Ha, and M. Jeon, "Visual multi-object tracking with re-identification and occlusion handling using labeled random finite sets," *Pattern Recognit.*, 156:110785, 2024.

[23] A. Scheel and K. Dietmayer, "Tracking multiple vehicles using a variational radar model," *IEEE Trans. Intell. Transp. Syst.*, 20(10):3721–3736, 2019.

[24] D. Moratuwage, M. Adams, and L. Cament, "Space object tracking using a jump Markov system based δ-glmb filter for space situational awareness," in *Adv. Maui Opt. Space Surveillance Technol. Conf.*, p. 32, 2019.

[25] D. Y. Kim, B.-N. Vo, A. Thian, and Y. S. Choi, "A generalized labeled multi-Bernoulli tracker for time lapse cell migration," in *Int. Conf. Control Automat. Inf. Sci.*, pp. 20–25, 2017.

[26] T. T. D. Nguyen, B.-N. Vo, B.-T. Vo, D. Y. Kim, and Y. S. Choi, "Tracking cells and their lineages via labeled random finite sets," *IEEE Trans. Signal Process.*, 69:5611–5626, 2021.

[27] D. Moratuwage, M. Adams, and F. Inostroza, "δ-generalised labelled multi-Bernoulli simultaneous localisation and mapping," in *Int. Conf. Control Automat. Inf. Sci.*, pp. 175–182, 2018.

[28] S. Panicker, A. K. Gostar, A. Bab-Hadiashar, and R. Hoseinnezhad, "Tracking of targets of interest using labeled multi-Bernoulli filter with multi-sensor control," *Signal Process.*, 171:107451, 2020.

[29] A. K. Gostar, R. Hoseinnezhad, T. Rathnayake, X. Wang, and A. Bab-Hadiashar, "Constrained sensor control for labeled multi-Bernoulli filter using Cauchy-Schwarz divergence," *IEEE Signal Process. Lett.*, 24(9):1313–1317, 2017.

[30] H. Cai, S. Gehly, Y. Yang, R. Hoseinnezhad, R. Norman, and K. Zhang, "Multisensor tasking using analytical Rényi divergence in labeled multi-Bernoulli filtering," *J. Guid. Control Dyn.*, 42(9):2078–2085, 2019.

[31] C. Fantacci, B.-N. Vo, B.-T. Vo, G. Battistelli, and L. Chisci, "Robust fusion for multisensor multiobject tracking," *IEEE Signal Process. Lett.*, 25(5):640–644, 2018.

[32] L. Gao, G. Battistelli, and L. Chisci, "Resilient labeled multi-Bernoulli fusion with peer-to-peer sensor network," *Inf. Fusion*, 100:101965, 2023.

[33] T. Li, K. Da, H. Fan, and B. Yu, "Multisensor random finite set information fusion: Advances, challenges, and opportunities," in *Secure and Digitalized Future Mobility*, pp. 33–64, CRC Press, 2022.

[34] S. Li, G. Battistelli, L. Chisci, W. Yi, B. Wang, and L. Kong, "Computationally efficient multi-agent multi-object tracking with labeled random finite sets," *IEEE Trans. Signal Process.*, 67(1):260–275, 2018.

[35] S. Li, W. Yi, R. Hoseinnezhad, G. Battistelli, B. Wang, and L. Kong, "Robust distributed fusion with labeled random finite sets," *IEEE Trans. Signal Process.*, 66(2):278–293, 2017.

[36] T. Li, "Arithmetic average density fusion - part II: Unified derivation for (labeled) RFS fusion," *IEEE Trans. Aerosp. Electron. Syst.*, 2024.

[37] T. T. D. Nguyen, H. Rezatofighi, B.-N. Vo, B.-T. Vo, S. Savarese, and I. Reid, "How trustworthy are performance evaluations for basic vision tasks?," *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(7):8538–8552, 2023.