

Set-Type Belief Propagation with Applications to Poisson Multi-Bernoulli SLAM

Hyowon Kim, *Member, IEEE*, Ángel F. García-Fernández, Yu Ge, *Student Member, IEEE*,
Yuxuan Xia, *Member, IEEE*, Lennart Svensson, *Senior Member, IEEE*, and Henk Wymeersch, *Fellow, IEEE*

Abstract—Belief propagation (BP) is a useful probabilistic inference algorithm for efficiently computing approximate marginal probability densities of random variables. However, in its standard form, BP is only applicable to the vector-type random variables with a fixed and known number of vector elements, while certain applications rely on random finite sets (RFSs) with an unknown number of vector elements. In this paper, we develop BP rules for factor graphs defined on sequences of RFSs where each RFS has an unknown number of elements, with the intention of deriving novel inference methods for RFSs. Furthermore, we show that vector-type BP is a special case of set-type BP, where each RFS follows the Bernoulli process. To demonstrate the validity of developed set-type BP, we apply it to the Poisson multi-Bernoulli (PMB) filter for simultaneous localization and mapping (SLAM), which naturally leads to a set-type BP PMB-SLAM method, which is analogous to a vector type SLAM method, subject to minor modifications.

Index Terms—Belief propagation, multi-target tracking, Poisson multi-Bernoulli filter, random finite sets, simultaneous localization and mapping.

I. INTRODUCTION

Belief propagation (BP) is a powerful statistical inference algorithm, applicable to different fields including signal processing, robotics, autonomous vehicles, image processing, and artificial intelligence. In particular, BP is attractive for computing probability densities in Bayesian networks [1]–[4], where one can efficiently compute the marginal probability densities of random variables of interest given the observed data. BP has been actively applied in localization [5], [6], mapping [7], multiple-target tracking (MTT) [8], [9], simultaneous localization and mapping (SLAM) [10]–[14], and simultaneous localization and tracking (SLAT) [15]. Mapping and MTT respectively involve mapping landmarks and tracking targets, based on the noisy sensor measurements. Simultaneously estimating the unknown sensor state as well as landmarks

and targets leads to SLAM and SLAT, respectively. The main challenge in this research stems from the data association uncertainty between the unknown number of targets and imperfect measurements by missed detections and clutter. To effectively address this challenge, several approaches have been developed as follows. Classical SLAM methods [16]–[20] consider the data associations outside the Bayesian filter, while the filter operates on vector random variables. With the introduction of random finite sets (RFSs) in [21], [22], a theoretically sound tool for modeling the unknown number of targets and for handling the data association between targets and measurements became available. By modeling targets with vector-type random variables, the marginal densities for targets are efficiently computed by BP, relying on several ad-hoc modifications [12]–[14]. Between them, the RFS- and BP-based methods have rigorously handled the main challenge of MTT and SLAM by starting with the formulation of joint posterior density of targets and data association.

There are several RFS-based methods for MTT and SLAM. The probability hypothesis density (PHD) filter [21]–[24] propagates the first-order statistical moment of the RFS, which is its intensity. The cardinalized probability hypothesis density (CPHD) filter propagates the intensity and the cardinality distribution [21], [22], [25]. A more accurate solution can be obtained with RFS filters based on multi-object conjugate priors, which are inherently closed under prediction and update [26]–[29]. The Poisson multi-Bernoulli mixture (PMBM) [27], [28] adopts the Poisson birth model and multi-Bernoulli mixture (MBM). For standard multi-object models with Poisson birth [21], the Bayesian optimal solutions to MTT and SLAM are given by the PMBM conjugate prior [26], [28]. With the multi-Bernoulli (MB) birth model instead of Poisson, the conjugate prior is an MBM [27]. The MBM filter corresponds to a PMBM filter where the Poisson intensity equals to zero, and the birth Bernoulli components are added into the MBM in the prediction step. If we expand the number of global hypotheses in the MBM filter such that each Bernoulli has deterministic target existence, we obtain the MBM₀₁ filter [27]. Both MBM and MBM₀₁ filters can be labeled without changing the filtering recursion. The MBM₀₁ filter recursion is analogous to the δ -generalized labeled multi-Bernoulli (δ -GLMB) filtering recursion [30]–[34]. Some RFS-based methods rely on BP to compute marginal association probabilities [26], which converts the PMBM [27], [28] to a Poisson multi-Bernoulli (PMB) [26], [28], [29], [35] after every update, by marginalizing out the association variables.

In addition to RFS-based methods, vector-type BP methods

H. Kim is with the Department of Electronics Engineering, Chungnam National University, 34134 Daejeon, South Korea (email: hyowon.kim@cnu.ac.kr).

Á. F. García-Fernández is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, United Kingdom, and also with the ARIES Research Centre, Universidad Antonio de Nebrija, Madrid, Spain (e-mail: angel.garcia-fernandez@liverpool.ac.uk).

Y. Ge, L. Svensson, and H. Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, 412 58 Gothenburg, Sweden (email: {yuge; lennart.svensson; and henkw}@chalmers.se).

Y. Xia is with the Department of Electrical Engineering, Linköping University, 581 83 Linköping, Sweden (email: yuxuan.xia@liu.se).

This work was supported in part by the Swedish Research Council (VR grant 2022-03007), the Chalmers Area of Advance Transport 6G-Cities project, Basic Science Research Program through the NRF of Korea (2022R1A6A3A03068510), and MSCA-IF grant 101065422 (6G-ISLAC).

have also been applied to tracking and mapping problems. In [8], [9], the MTT problem is tackled by running vector-type BP on the factor graph representation of the joint distribution of target states and data association variables. Here targets are modeled by augmented vectors including the binary target existence indicators, and message passing between targets and data association variables is adopted from [9]. The SLAM problems with radio signal propagation are formulated in [10]–[14]. The authors of [10], [11] formulate the factor graph and introduce an efficient message passing scheduling, among the unknown sensor variables consisting of position, orientation, and clock bias, and multiple-types of targets are further studied in [11]. With the augmented target vectors and the factorized joint distribution of targets and data association variables from [9], the message passing methods for joint SLAM and data association are developed [12]–[14]. These vector-type BP-based approaches [12]–[14] can be explained as the track-oriented PMB filter [28] and yield competitive performance. To handle undetected targets and newly detected targets, the auxiliary PHD is adopted in [12]–[14], which was initially considered in [36] for vector-type MTT and in [37] for set-type MTT. The propagation of the auxiliary PHD filter is explicitly introduced in [13], but external to the factor graph representation. Therefore, undetected targets and their connections to newly detected targets are not represented in the formulated factor graph. The modeling of undetected targets in vector-based methods is treated in [38] but the distinction between eventually and never detected targets is required. Moreover, it is not explicitly revealed how to apply BP in the multiple hypotheses tracking formalism. Hence, while BP-based methods have benefits in algorithm implementation, they require certain ad-hoc modifications.

In this paper, we aim to bridge the gap between RFS theory and BP, by developing a novel BP algorithm, running on factor graphs defined on the sequence of RFSs. Like conventional BP, this opens the door to automated inference over factor graphs, once the RFS density is factorized. This approach can avoid the need for heuristics and approximations. Related work has been done in [39], where BP is applied to only the update step, without a systematic treatment to the entire filtering recursion, except for the modeling of undetected targets and newly detected targets. The goal of this paper is thus to formalize set-type BP and the corresponding factor graphs from the RFS densities. From the newly proposed set-type BP, we derive the PMB filter using the developed set-type BP. The contributions of this paper are summarized as follows:

- **The specification of set-type BP:** We derive set-type BP and demonstrate that vector-type BP is a special case of set-type BP. We also show that as in vector-type BP, the interior stationary points of the constrained Bethe free energy are set-type BP fixed points.
- **The introduction of novel factors for set-type BP:** We devise a partition and merging factor, which partitions a single set into multiple sets and merges multiple sets into a single set, useful for handling sets with unknown cardinalities. We also propose a conversion factor for sets augmented with auxiliary vectors such as unique marks.

- **Derivation of PMB and MB filters with set-type BP for the related problems of mapping, MTT, SLAM, SLAT:** With the developed set-type BP, we revisit the PMB- and MB-SLAM filters. We first introduce auxiliary variables to factorize their joint SLAM and data association distribution, formulating a factor graph from the factorized density, and running set-type BP on the factor graph. This work can also lead to set-type BP PMB-mapping, MTT, and SLAT filters. The resulting method bears a close resemblance to the PMB-SLAM filter [28] that computes the updated sensor state density by ignoring the sensor state information generated from newly detected targets and selecting the most likely association determined by the Murty’s algorithm [40].
- **Relation to vector-type BP-SLAM:** We clearly show the connections between the proposed set-type BP PMB-SLAM and vector-type BP-SLAM filters [12]–[14]. The methods are algorithmically equivalent, except for minor details described in Section IV-F, even though the proposed filter is derived directly from the RFS extension and developed set-type BP framework. However, vector-type BP-SLAM filters [12]–[14] requires certain heuristics as part of the algorithm development, which are avoided in the proposed set-type BP PMB-SLAM filter. The simulation results show that the proposed set-type BP PMB-SLAM filter outperforms the vector-type BP-SLAM filter [12]–[14], in scenarios with informative Poisson point process (PPP) birth.

This paper is organized as follows. Section II provides the background of vector-type BP and RFSs. In Section III, the set-type BP rules and set-type factor nodes are proposed. Proposed set-type BP is applied to the PMB filter, and the connections between set-type and vector-type BP-SLAM filters are analyzed in Section IV. The numerical results and discussions are reported in Section V, and conclusions are drawn in Section VI.

II. BACKGROUND

In this section, we review factor graphs and belief propagation, and we recall the RFS approaches.

Notation: Scalars are denoted by italic font, vectors and matrices are respectively indicated by bold lowercase and uppercase letters, and sets are displayed in calligraphic font, e.g., x , \mathbf{x} , \mathbf{X} , and \mathcal{X} . The set of finite subsets of a space \mathbb{R} is denoted by $\mathcal{F}(\mathbb{R})$. The vector consisting of a sequence of vectors \mathbf{x}^i is denoted by $\underline{\mathbf{x}}$, and the sequence of multiple sets \mathcal{X} is denoted by $\underline{\mathcal{X}}$.

A. Vector-Type Factor Graph and Belief Propagation

1) *Joint Density Factorization and Factor Graph:* Let $\mathbf{x}^i \in \mathbb{R}^{n_x}$ denote a single state vector and $\underline{\mathbf{x}} = [(\mathbf{x}^1)^\top, \dots, (\mathbf{x}^N)^\top]^\top$ denote augment single state vectors representing N states. Then, we denote a joint probability density of the hidden variables $\underline{\mathbf{x}}$ by $f(\underline{\mathbf{x}})$, which can be factorized as [1], [4]

$$f(\underline{\mathbf{x}}) \propto \prod_a f_a(\underline{\mathbf{x}}^a), \quad (1)$$

where $f_a(\underline{\mathbf{x}}^a)$ denotes a nonnegative function, and $\underline{\mathbf{x}}^a$ denotes the argument vector of the function $f_a(\cdot)$.

The factorized functions in (1) and corresponding argument vectors can be represented by a factor graph. The factor graph with the general graphical model [1] consists of nodes for the different factors and variables, illustrated by squares for the factors, $f_a(\cdot)$, and circles for the variables, \mathbf{x}^i , respectively, with edge connections between the factors and their argument variables.

Example 1. Suppose we have a probability density $f(\underline{\mathbf{x}})$, such that $\underline{\mathbf{x}} = [(\mathbf{x}^1)^\top, (\mathbf{x}^2)^\top]^\top$, $\underline{\mathbf{x}}^A = \mathbf{x}^1$, $\underline{\mathbf{x}}^B = \mathbf{x}^2$, $\underline{\mathbf{x}}^C = [(\mathbf{x}^1)^\top, (\mathbf{x}^2)^\top]^\top$, which can be factorized as

$$f(\underline{\mathbf{x}}) \propto f_A(\mathbf{x}^1) f_B(\mathbf{x}^2) f_C(\mathbf{x}^1, \mathbf{x}^2). \quad (2)$$

The corresponding factor graph is illustrated in Fig. 1.

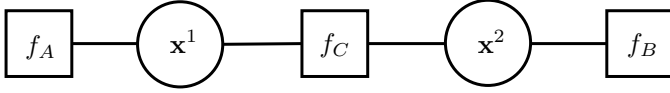


Fig. 1: Factor graph representation of the factorized vector density (2).

2) *Belief Propagation:* BP is an efficient approach for estimating the marginal densities of the variables from a joint probability density function [1]. By running BP on the factor graph, we compute the messages passed between factors and variables for all links. We denote the propagated messages from factor a to variable i and from variable i to factor a by $\mathfrak{m}_{a \rightarrow i}(\mathbf{x}^i)$ and $\mathfrak{n}_{i \rightarrow a}(\mathbf{x}^i)$, respectively. The messages are updated with the following rules:

$$\mathfrak{m}_{a \rightarrow i}(\mathbf{x}^i) = \int f_a(\underline{\mathbf{x}}^a) \prod_{j \in \mathcal{N}(a) \setminus \{i\}} \mathfrak{n}_{j \rightarrow a}(\mathbf{x}^j) d\mathbf{x}^{\sim i}, \quad (3)$$

$$\mathfrak{n}_{i \rightarrow a}(\mathbf{x}^i) = \prod_{b \in \mathcal{M}(i) \setminus \{a\}} \mathfrak{m}_{b \rightarrow i}(\mathbf{x}^i), \quad (4)$$

where $\mathcal{N}(a)$ denotes the set of indices i of neighboring vectors linked to the factor $f_a(\underline{\mathbf{x}}^a)$, such that $i \in \mathcal{N}(a)$ if and only if \mathbf{x}^i is an argument of $f_a(\cdot)$, $\mathcal{M}(i)$ denotes the set of indices a of neighboring factors linked to the variable \mathbf{x}^i , such that $a \in \mathcal{M}(i)$ if and only if \mathbf{x}^i is an argument of $f_a(\cdot)$, and $\int \dots d\mathbf{x}^{\sim i}$ denotes integration with respect to all vectors \mathbf{x}^j except \mathbf{x}^i . The messages can be used to compute beliefs that approximate the marginalized posterior densities. The beliefs at variable i and factor a are denoted $\mathfrak{b}(\mathbf{x}^i)$ and $\mathfrak{b}(\underline{\mathbf{x}}^a)$, respectively, and are updated using the following rules:

$$\mathfrak{b}(\mathbf{x}^i) \propto \prod_{a \in \mathcal{M}(i)} \mathfrak{m}_{a \rightarrow i}(\mathbf{x}^i), \quad (5)$$

$$\mathfrak{b}(\underline{\mathbf{x}}^a) \propto f_a(\underline{\mathbf{x}}^a) \prod_{i \in \mathcal{N}(a)} \mathfrak{n}_{i \rightarrow a}(\mathbf{x}^i). \quad (6)$$

B. Random Finite Sets

1) *Set-Variables, Density, and Integral:* Let us denote an RFS by $\mathcal{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^n\} \in \mathcal{F}(\mathbb{R}^{n \times \mathbf{x}})$, where both vector \mathbf{x}^i ,

$i \in \{1, \dots, n\}$ and cardinality $n = |\mathcal{X}|$ are random. We define a set-density $f(\mathcal{X})$ as [26]

$$f(\mathcal{X}) = p(n) \sum_{\pi} f^n(\mathbf{x}^{\pi(1)}, \dots, \mathbf{x}^{\pi(n)}), \quad (7)$$

in which $p(n) = \Pr(|\mathcal{X}| = n)$ denotes the probability mass function of the set cardinality, π denotes a possible permutation of the set $\mathcal{N} = \{1, \dots, n\}$ with $\pi(i) \in \mathcal{N}$, and $f^n(\cdot)$ is the joint probability density function of the vector with n elements, evaluated for permutation π . Given the set-valued function $g(\cdot)$, we define the set integral as [22, eq. (3.11)]

$$\int g(\mathcal{X}) \delta \mathcal{X} = g(\emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int g(\{\mathbf{x}^1, \dots, \mathbf{x}^n\}) d\mathbf{x}^1 \dots d\mathbf{x}^n. \quad (8)$$

2) *Poisson, Bernoulli, and PMB Densities:* We will follow the definitions from [21]. Suppose we have a set \mathcal{X}^U that follows the Poisson process. The density is given by [26]

$$f^{\text{PPP}}(\mathcal{X}^U) = e^{-\int \lambda(\mathbf{x}) d\mathbf{x}} \prod_{\mathbf{x} \in \mathcal{X}^U} \lambda(\mathbf{x}), \quad (9)$$

where $\lambda(\mathbf{x})$ denotes the Poisson intensity function. A Bernoulli density $f^{\text{B}}(\mathcal{X})$ is given by

$$f^{\text{B}}(\mathcal{X}) = \begin{cases} 1 - r, & \mathcal{X} = \emptyset \\ r f(\mathbf{x}), & \mathcal{X} = \{\mathbf{x}\} \\ 0, & |\mathcal{X}| > 1, \end{cases} \quad (10)$$

where $f(\mathbf{x})$ and $r \in [0, 1]$ denote the spatial density and the existence probability, respectively. An MB density with n Bernoulli components is given by

$$f^{\text{MB}}(\mathcal{X}^{\text{D}}) = \sum_{\uplus_{i=1}^n \mathcal{X}^i = \mathcal{X}^{\text{D}}} \prod_{i=1}^n f^i(\mathcal{X}^i), \quad (11)$$

where $f^i(\mathcal{X}^i)$ is a Bernoulli density, and \uplus stands for disjoint set union [41, pp. 24].

We are now ready to introduce a PMB density, defined as follows. Suppose we have two independent RFSs \mathcal{X}^U and \mathcal{X}^{D} such that $\mathcal{X} = \mathcal{X}^U \uplus \mathcal{X}^{\text{D}}$, where \mathcal{X}^{D} follows a MB process and \mathcal{X}^U follows a PPP. Using the convolution formula for independent RFSs [21], a PMB density $f(\mathcal{X})$ is

$$f^{\text{PMB}}(\mathcal{X}) \propto \sum_{\uplus_{i=1}^n \mathcal{X}^i \uplus \mathcal{X}^U = \mathcal{X}} \prod_{\mathbf{x} \in \mathcal{X}^U} \lambda(\mathbf{x}) \prod_{i=1}^n f^i(\mathcal{X}^i). \quad (12)$$

3) *Auxiliary Variables:* The derivation of set-type BP PMB filters will require us to introduce auxiliary variables to remove the summation in (12). In particular, we introduce $u \in \mathbb{U}$ in the PMB density (12), where $\mathbb{U} = \{0, 1, \dots, n\}$ [39], [42]. We thus extend the single state space, such that $(u, \mathbf{x}) \in \mathbb{U} \times \mathbb{R}^{n \times \mathbf{x}}$, and denote a set of target states with auxiliary variables by $\tilde{\mathcal{X}} \in \mathcal{F}(\mathbb{U} \times \mathbb{R}^{n \times \mathbf{x}})$. The set with the auxiliary variable $u = 0$ follows a PPP and indicates that the targets have not been detected, denoted by $\tilde{\mathcal{X}}^U = \{(u, \mathbf{x}) \in \tilde{\mathcal{X}} : u = 0\}$. Similarly, the set with $u = i$ follows a Bernoulli process and indicates that the single target has previously been detected, denoted by $\tilde{\mathcal{X}}^i = \{(u, \mathbf{x}) \in \tilde{\mathcal{X}} : u = i\}$. For the set of targets with

auxiliary variables $\tilde{\mathcal{X}}$, the PMB density is [42, Definition 1]

$$\tilde{f}^{\text{PMB}}(\tilde{\mathcal{X}}) = \tilde{f}^{\text{U}}(\tilde{\mathcal{X}}^{\text{U}}) \prod_{i=1}^n \tilde{f}^i(\tilde{\mathcal{X}}^i). \quad (13)$$

Here $\tilde{f}^{\text{U}}(\tilde{\mathcal{X}}^{\text{U}})$ and $\tilde{f}^i(\tilde{\mathcal{X}}^i)$ are given by

$$\tilde{f}^{\text{U}}(\tilde{\mathcal{X}}^{\text{U}}) = \exp\left(-\int \lambda(\mathbf{x})d\mathbf{x}\right) \prod_{(u,\mathbf{x}) \in \tilde{\mathcal{X}}^{\text{U}}} \delta_0[u]\lambda(\mathbf{x}), \quad (14)$$

$$\tilde{f}^i(\tilde{\mathcal{X}}^i) = \begin{cases} 1 - r^i, & \tilde{\mathcal{X}}^i = \emptyset \\ r^i f^i(\mathbf{x})\delta_i[u], & \tilde{\mathcal{X}}^i = (u, \mathbf{x}), \\ 0, & \text{otherwise} \end{cases}, \quad (15)$$

where $\delta_i[\cdot]$ denotes the Kronecker delta function. For notational simplicity, $\tilde{\cdot}$ on sets with auxiliary variables will be omitted, when possible.

III. FACTOR GRAPH AND BELIEF PROPAGATION FOR RANDOM FINITE SETS

We describe the proposed set-type BP update rules and special factors for RFSs. We reveal that set-type BP is a generalization of standard vector-type BP since a vector can be represented as a set with a single element $p(n=1) = 1$ of (7).

A. Factor Graphs and BP over a Sequence of RFSs

Suppose we have n RFSs $\mathcal{X}^1, \dots, \mathcal{X}^n$, with the joint density $f(\mathcal{X}^1, \dots, \mathcal{X}^n)$. In the general formulation of set-type BP, $\mathcal{X}^1, \dots, \mathcal{X}^n$ may or may not contain auxiliary variables, and the number of RFSs in the sequence, n , is known.

Definition 1 (Factorization of Set-Density and Factor Graph). Let us denote by $f_a(\cdot)$ the set-factor a , which is a nonnegative function; by $\mathcal{N}(a)$ the set of neighboring set-variable indices linked to the set-factor a ; by \mathcal{X}^i the set-variable i ; by $\mathcal{M}(i)$ the set of neighboring set-factor indices linked to the set-variable \mathcal{X}^i ; and by $\underline{\mathcal{X}}^a$ the arguments of the set-factor a , represented by the sequence of all RFSs \mathcal{X}^i for $i \in \mathcal{N}(a)$. Suppose the joint density $f(\mathcal{X}^1, \dots, \mathcal{X}^n)$ is factorized as follows:

$$f(\mathcal{X}^1, \dots, \mathcal{X}^n) \propto \prod_a f_a(\underline{\mathcal{X}}^a). \quad (16)$$

The factorized density can then be represented by a factor graph, consisting of the set-variables \mathcal{X}^i and set-factors $f_a(\underline{\mathcal{X}}^a)$, which can be represented by circles and squares, and edge connections between \mathcal{X}^i and $f_a(\underline{\mathcal{X}}^a)$. It should be noted that $f_a(\underline{\mathcal{X}}^a)$ has adequate units for each cardinality of $\underline{\mathcal{X}}^a$ such that we can integrate (16) using set integrals over $\mathcal{X}^1, \dots, \mathcal{X}^n$ [22, Sec. 3.2.4].

Example 2. Consider a set-density $f(\mathcal{X}^1, \mathcal{X}^2)$, such that $\underline{\mathcal{X}}^A = \mathcal{X}^1$, $\underline{\mathcal{X}}^B = \mathcal{X}^2$, and $\underline{\mathcal{X}}^C = (\mathcal{X}^1, \mathcal{X}^2)$, which can be factorized as

$$f(\mathcal{X}^1, \mathcal{X}^2) \propto f_A(\mathcal{X}^1)f_B(\mathcal{X}^2)f_C(\mathcal{X}^1, \mathcal{X}^2). \quad (17)$$

Using the set-variables and set-factors, the factor graph corresponding to (17) is shown in Fig. 2.

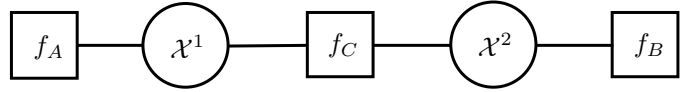


Fig. 2: Factor graph representation of set-density of (17).

Definition 2 (Set-Type BP Update Rules). The set-messages from the set-factor a to the set-variable i are denoted by $\mathfrak{m}_{a \rightarrow i}(\mathcal{X}^i)$, while those from the set-variable i to the set-factor a are denoted by $\mathfrak{n}_{i \rightarrow a}(\mathcal{X}^i)$. The beliefs at the set-variable i and set-factor a are denoted by $\mathfrak{b}(\mathcal{X}^i)$ and $\mathfrak{b}(\underline{\mathcal{X}}^a)$, respectively. The set-messages are updated with the following rules:

$$\mathfrak{m}_{a \rightarrow i}(\mathcal{X}^i) = \int f_a(\underline{\mathcal{X}}^a) \prod_{j \in \mathcal{N}(a) \setminus \{i\}} \mathfrak{n}_{j \rightarrow a}(\mathcal{X}^j) \delta \mathcal{X}^{\sim i}, \quad (18)$$

$$\mathfrak{n}_{i \rightarrow a}(\mathcal{X}^i) = \prod_{b \in \mathcal{M}(i) \setminus \{a\}} \mathfrak{m}_{b \rightarrow i}(\mathcal{X}^i), \quad (19)$$

where \setminus denotes the set difference, and $\int \dots \delta \mathcal{X}^{\sim i}$ denotes integration with respect to all sets \mathcal{X}^j except \mathcal{X}^i , i.e., with respect to all \mathcal{X}^j for which $j \in \mathcal{N}(a) \setminus \{i\}$. The beliefs at the set-variable i and set-factor a are updated with the following rules:

$$\mathfrak{b}(\mathcal{X}^i) \propto \prod_{a \in \mathcal{M}(i)} \mathfrak{m}_{a \rightarrow i}(\mathcal{X}^i) \quad (20)$$

$$\mathfrak{b}(\underline{\mathcal{X}}^a) \propto f_a(\underline{\mathcal{X}}^a) \prod_{i \in \mathcal{N}(a)} \prod_{b \in \mathcal{M}(i) \setminus \{a\}} \mathfrak{m}_{b \rightarrow i}(\mathcal{X}^i). \quad (21)$$

The optimality of the set-type BP update rules in Definition 1 is described by Theorem 1 and Corollary 1, provided next.

Theorem 1. The interior stationary points of the constrained Bethe free energy are set-type BP fixed points with positive set-beliefs and vice versa.

Proof. See Appendix A. \square

Corollary 1. The set-beliefs obtained by running set-type BP on a factor graph that has no cycles, represent the exact marginal probability densities.

Proof. See Appendix B. \square

Example 3. Consider the factor graph in Fig. 2, representing the factorized density in Example 2. Using set-type BP, the beliefs at set-variable 1 and factor C are obtained as follows:

$$\mathfrak{m}_{A \rightarrow 1}(\mathcal{X}^1) = f_A(\mathcal{X}^1), \quad \mathfrak{m}_{B \rightarrow 2}(\mathcal{X}^2) = f_B(\mathcal{X}^2), \quad \mathfrak{n}_{1 \rightarrow C}(\mathcal{X}^1) = \mathfrak{m}_{A \rightarrow 1}(\mathcal{X}^1), \quad \mathfrak{n}_{2 \rightarrow C}(\mathcal{X}^2) = \mathfrak{m}_{B \rightarrow 2}(\mathcal{X}^2),$$

$$\mathfrak{m}_{C \rightarrow 1}(\mathcal{X}^1) = \int \mathfrak{n}_{2 \rightarrow C}(\mathcal{X}^2) f_C(\mathcal{X}^1, \mathcal{X}^2) \delta \mathcal{X}^2, \quad (22a)$$

$$\mathfrak{b}(\mathcal{X}^1) \propto \mathfrak{m}_{A \rightarrow 1}(\mathcal{X}^1) \mathfrak{m}_{C \rightarrow 1}(\mathcal{X}^1), \quad (22b)$$

$$\mathfrak{b}(\mathcal{X}^C) \propto f_C(\mathcal{X}^1, \mathcal{X}^2) \mathfrak{m}_{A \rightarrow 1}(\mathcal{X}^1) \mathfrak{m}_{B \rightarrow 2}(\mathcal{X}^2). \quad (22c)$$

Remark 1 (Vector-Type BP is a Special Case of Set-Type BP). Note that the vector-type BP message passing rules and beliefs can be obtained from the set-type BP expressions when considering sets whose cardinality is 1 with probability 1, i.e.,

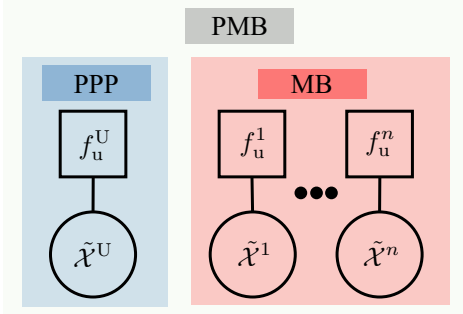


Fig. 3: Factor graph of a PMB density with auxiliary variables, see (13), the product of a PPP and n Bernoulli densities.

$p(|\mathcal{X}^i| = 1) = 1, \forall i.$ ¹ Note that this property implies that we can directly apply BP to joint set and vector densities using the set- and vector-integrals for set- and vector-variables, respectively. A vector-type factor graph (1) can be written as a set-type factor graph in (16) with the property that each factor requires sets with cardinality 1. Then, for factors that only consider sets with cardinality 1, the set integral is equivalent to a vector integral, and the set-type BP message (20)-(21) becomes equivalent to those in (5)-(6).

B. Examples and Special Factors for Set Densities

1) *Factor Graph of a PMB Density:* We have so far explained how set-type BP can be applied to a joint density over a sequence of RFSs. We now proceed to explain how to obtain this type of density from a PMB density (12).

A set-density defined in a single-target space that is the disjoint union of different sub-spaces can be used to define a density over a sequence of sets [22, Eq. 3.52]. This type of single-target space was obtained when we introduced auxiliary variables to the PMB in (12), resulting in (13). Therefore, applying this result to the PMB density of the form (13) yields

$$\tilde{f}^{\text{PMB}}(\tilde{\mathcal{X}}^U, \tilde{\mathcal{X}}^1, \dots, \tilde{\mathcal{X}}^n) = \tilde{f}^{\text{PMB}}(\tilde{\mathcal{X}}^U \uplus \tilde{\mathcal{X}}^1 \uplus \dots \uplus \tilde{\mathcal{X}}^n). \quad (23)$$

We can now use this joint density over a sequence of RFSs to apply set-type BP. The factor graph of a PMB density (13) with auxiliary variables is then shown in Fig. 3. Since the density is fully factorized, it appears as a collection of disjoint factors. We can also directly obtain the factor graph of a PPP and an MB density with auxiliary variables by removing the required factors and variables in Fig. 3.

2) *Partitioning and Merging Factor:* Unions of RFSs are common in the literature [22, Sec. 3.5.3]. To represent unions of RFSs in a factor graph, we introduce what we refer to as partitioning and merging factors, defined as follows.

Definition 3 (Partitioning and Merging Factor). We define a partitioning and merging set-factor as

$$f_a(\underline{\mathcal{X}}^a) = \delta_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^i}(\mathcal{X}^j), \quad (24)$$

where $\delta_{\mathcal{X}}(\cdot)$ denotes a set Dirac delta centered at set \mathcal{X} , defined in [21, Sec. 11.3.4.3] and \mathcal{X}^j is the union $\mathcal{X}^j =$

¹The number of elements of the sets is set deterministically to 1. For example, suppose we have $\mathcal{X}^1 = \{\mathbf{x}^1\}, \mathcal{X}^2 = \{\mathbf{x}^2\}, \mathcal{X}^3 = \{\mathbf{x}^3\}$ in the RFS representation and $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$ in the vector representation. Then, both representations are equivalent.

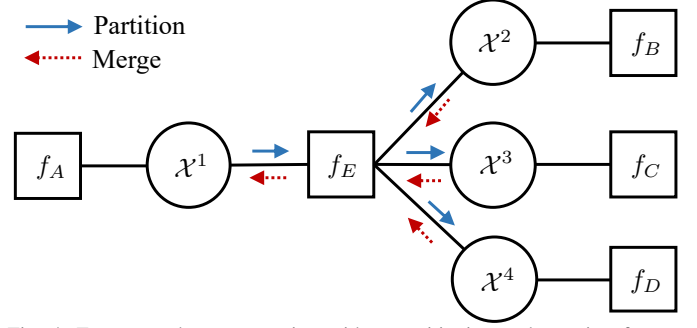


Fig. 4: Factor graph representation with a partitioning and merging factor: (left-right) one incoming message is partitioned into multiple messages; and (right-left) multiple messages are merged into one message.

$\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^i$. This factor partitions a single set \mathcal{X}^j into $|\mathcal{N}(a)| - 1$ subsets, i.e., \mathcal{X}^i for $i \in \mathcal{N}(a) \setminus \{j\}$, and merges $|\mathcal{N}(a)| - 1$ sets, i.e., \mathcal{X}^i for $i \in \mathcal{N}(a) \setminus \{j\}$, into a single set \mathcal{X}^j .

It is useful to understand how this factor affects the set-messages. Suppose we have incoming messages $\mathbf{n}_{j \rightarrow a}(\mathcal{X}^j)$ and $\mathbf{n}_{q \rightarrow a}(\mathcal{X}^q) = 1$ for $q \in \mathcal{N}(a) \setminus \{j, i\}$. By following the set-type BP update rules, the outgoing messages from the set-factor a to set-variable i for $i \in \mathcal{N}(a) \setminus \{j\}$ are

$$\mathbf{m}_{a \rightarrow i}(\mathcal{X}^i) = \int \mathbf{n}_{j \rightarrow a}(\mathcal{X}^j) \delta_{\uplus_{q \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^q}(\mathcal{X}^j) \delta \mathcal{X}^{\sim i} \quad (25)$$

$$= \int \mathbf{n}_{j \rightarrow a}(\uplus_{q \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^q) \delta \mathcal{X}^{\sim(i,j)} \quad (26)$$

$$= \int \mathbf{n}_{j \rightarrow a}(\mathcal{X}^i \uplus \mathcal{X}) \delta \mathcal{X}, \quad (27)$$

where $\int \delta \mathcal{X}^{\sim(i,j)}$ indicates the integration with respect to all sets \mathcal{X}^q except \mathcal{X}^i and \mathcal{X}^j , and \mathcal{X} is a dummy variable that is used to integrate over all possible sets. It indicates that $\mathcal{X} = \uplus_{q \in \mathcal{N}(a) \setminus \{j, i\}} \mathcal{X}^q$, and the single set \mathcal{X}^j with the incoming message $\mathbf{n}_{j \rightarrow a}(\mathcal{X}^j)$ is partitioned into $|\mathcal{N}(a)| - 1$ subsets \mathcal{X}^i with the outgoing messages $\mathbf{m}_{a \rightarrow i}(\mathcal{X}^i)$, for $i \in \mathcal{N}(a) \setminus \{j\}$.

Conversely, suppose we have incoming messages $\mathbf{n}_{i \rightarrow a}(\mathcal{X}^i)$ for $i \in \mathcal{N}(a) \setminus \{j\}$, then the outgoing message from the set-factor a to the set-variable j is

$$\mathbf{m}_{a \rightarrow j}(\mathcal{X}^j) = \int \prod_{i \in \mathcal{N}(a) \setminus \{j\}} \mathbf{n}_{i \rightarrow a}(\mathcal{X}^i) \delta_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^i}(\mathcal{X}^j) \delta \mathcal{X}^{\sim j} \quad (28)$$

$$= \int \sum_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{W}^i = \mathcal{X}^j} \prod_{i \in \mathcal{N}(a) \setminus \{j\}} \mathbf{n}_{i \rightarrow a}(\mathcal{X}^i) \delta_{\mathcal{W}^i}(\mathcal{X}^i) \delta \mathcal{X}^{\sim j} \quad (29)$$

$$= \sum_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{W}^i = \mathcal{X}^j} \prod_{i \in \mathcal{N}(a) \setminus \{j\}} \mathbf{n}_{i \rightarrow a}(\mathcal{W}^i), \quad (30)$$

where in step (28), we have used that $\delta_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^i}(\mathcal{X}^j) = \sum_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{W}^i = \mathcal{X}^j} \prod_{i \in \mathcal{N}(a) \setminus \{j\}} \delta_{\mathcal{W}^i}(\mathcal{X}^i)$ by the convolution formula. Then, (30) indicates that the sets \mathcal{X}^i with the incoming messages $\mathbf{n}_{i \rightarrow a}(\mathcal{X}^i)$ for $i \in \mathcal{N}(a) \setminus \{j\}$ are merged into the single set \mathcal{X}^j with the outgoing message $\mathbf{m}_{a \rightarrow j}(\mathcal{X}^j)$.

Example 4. Given the factor graph shown in Fig. 4 such that $\mathcal{X}^1 = \mathcal{X}^2 \uplus \mathcal{X}^3 \uplus \mathcal{X}^4$, the argument at the set-factor E is rep-

resented as the sequence of sets as $\underline{\mathcal{X}}^E = (\mathcal{X}^1, \mathcal{X}^2, \mathcal{X}^3, \mathcal{X}^4)$ by Definition 1. By Definition 3, the partitioning and merging factor is given by $f_E(\underline{\mathcal{X}}^E) = \delta_{\mathcal{X}^2 \uplus \mathcal{X}^3 \uplus \mathcal{X}^4}(\mathcal{X}^1)$. Suppose we have incoming messages $\mathbf{n}_{1 \rightarrow E}(\mathcal{X}^1)$, $\mathbf{n}_{i \rightarrow E}(\mathcal{X}^i) = 1$, for $i \in \{2, 3, 4\}$. From (26), the partitioned message $\mathbf{m}_{E \rightarrow i}(\mathcal{X}^i)$ is computed as

$$\mathbf{m}_{E \rightarrow i}(\mathcal{X}^i) = \int \mathbf{n}_{1 \rightarrow E}(\mathcal{X}^2 \uplus \mathcal{X}^3 \uplus \mathcal{X}^4) \delta_{\mathcal{X}^{\sim(i,1)}}, \quad (31)$$

for $i = 2, 3, 4$. Conversely, suppose we have incoming messages $\mathbf{n}_{j \rightarrow E}(\mathcal{X}^j)$ for $j \in \{2, 3, 4\}$, then the outgoing message from the set-factor E to set-variable 1 is computed as

$$\begin{aligned} \mathbf{m}_{E \rightarrow 1}(\mathcal{X}^1) &= \sum_{\mathcal{W}^2 \uplus \mathcal{W}^3 \uplus \mathcal{W}^4 = \mathcal{X}^1} \mathbf{n}_{2 \rightarrow E}(\mathcal{W}^2) \mathbf{n}_{3 \rightarrow E}(\mathcal{W}^3) \mathbf{n}_{4 \rightarrow E}(\mathcal{W}^4). \end{aligned} \quad (32)$$

That is, the outgoing message is the convolution of the three incoming messages, representing the union of three independent RFSs.

Proposition 1 (PPP Partitioning and Merging). *Suppose we have a set factor $f_a(\underline{\mathcal{X}}^a)$ and an RFS $\mathcal{X}^j = \uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^i$. Let the RFSs \mathcal{X}^i for $i \in \mathcal{N}(a)$ follow a Poisson process. From (27), the partitioning messages from $f_a(\underline{\mathcal{X}}^a)$ to \mathcal{X}^i with $i \in \mathcal{N}(a) \setminus \{j\}$ are*

$$\mathbf{m}_{a \rightarrow i}(\mathcal{X}^i) \propto f^{\text{PPP}}(\mathcal{X}^i), \quad (33)$$

From (30), the merging messages from $f_a(\underline{\mathcal{X}}^a)$ to \mathcal{X}^j is

$$\mathbf{m}_{a \rightarrow j}(\mathcal{X}^j) = \sum_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{W}^i = \mathcal{X}^j} \prod_{i \in \mathcal{N}(a) \setminus \{j\}} f^{\text{PPP}}(\mathcal{W}^i), \quad (34)$$

which follows a PPP.

Proof. See Appendix C. \square

3) *Auxiliary Variable Shifting Factor:* We introduce a factor to change the auxiliary variables.

Definition 4 (Auxiliary Variable Shifting Function). *Given an arbitrary integer L and a single-target space with auxiliary variables, such that $(u, \mathbf{x}) \in \mathbb{U} \times \mathbb{R}^{n_x}$, we define the function*

$$\mathbf{h}_L(u, \mathbf{x}) = (u + L, \mathbf{x}), \quad (35)$$

which shifts the auxiliary variables units by L . The function $\mathbf{h}_L(\cdot)$ can be extended to a set $\tilde{\mathcal{X}} \in \mathcal{F}(\mathbb{U} \times \mathbb{R}^{n_x})$ such that

$$\mathbf{h}_L(\tilde{\mathcal{X}}) = \uplus_{(u, \mathbf{x}) \in \tilde{\mathcal{X}}} \{(u + L, \mathbf{x})\}. \quad (36)$$

Definition 5 (Conversion Factor for Auxiliary Variables). *We have two sets $\tilde{\mathcal{X}}^i = \{(u^1, \mathbf{x}^1), \dots, (u^n, \mathbf{x}^n)\}$ and $\tilde{\mathcal{X}}^j = \{(u^1 + L, \mathbf{x}^1), \dots, (u^n + L, \mathbf{x}^n)\}$, where L is an integer. By Definition 4, we define the conversion factor for auxiliary variables as*

$$f_a(\tilde{\mathcal{X}}^i, \tilde{\mathcal{X}}^j) = \delta_{\mathbf{h}_{-L}(\tilde{\mathcal{X}}^j)}(\tilde{\mathcal{X}}^i), \quad (37)$$

where $\mathbf{h}_{-L}(\cdot)$ is given by (36).

It is useful to understand how this factor affects the set-messages. Suppose we have an incoming message $\mathbf{n}_{i \rightarrow a}(\tilde{\mathcal{X}}^i)$.

Then, the outgoing message from $f_a(\tilde{\mathcal{X}}^i, \tilde{\mathcal{X}}^j)$ to $\tilde{\mathcal{X}}^j$ is

$$\mathbf{m}_{a \rightarrow j}(\tilde{\mathcal{X}}^j) = \int \mathbf{n}_{i \rightarrow a}(\tilde{\mathcal{X}}^i) \delta_{\mathbf{h}_{-L}(\tilde{\mathcal{X}}^j)}(\tilde{\mathcal{X}}^i) \delta_{\tilde{\mathcal{X}}^j} \quad (38)$$

$$= \mathbf{n}_{i \rightarrow a}(\mathbf{h}_{-L}(\tilde{\mathcal{X}}^j)) = \mathbf{n}_{i \rightarrow a}(\tilde{\mathcal{X}}^i). \quad (39)$$

That is, the outgoing message has the same form as the incoming message with the difference that the auxiliary variables of the incoming message have been converted by the auxiliary variable shifting functions.

IV. APPLICATION OF SET-TYPE BELIEF PROPAGATION

The aim of this section is to propose an application of the developed set-type BP update rules and special factors for RFSs. In particular, we derive set-type BP PMB and set-type MB filters for SLAM, where the targets and measurements are modeled by RFSs.

A. Problem Formulation

1) *Objective:* Our objective is to compute the marginal densities $f(\mathbf{s}_k | \mathcal{Z}_{1:k})$ and $f(\tilde{\mathcal{X}}_k | \mathcal{Z}_{1:k})$ at discrete time k , where the random vector \mathbf{s}_k denotes a sensor state, the RFS $\tilde{\mathcal{X}}_k$ denotes the set of target states with auxiliary variables, modeled by a PMB. To compute the marginal densities, we adopt the sequential Bayesian framework consisting of prediction and update steps, which will be detailed in Section IV-B and IV-C.

2) *Multi-Target Dynamics:* Each target $\mathbf{x}_{k-1} \in \tilde{\mathcal{X}}_{k-1}$ at time $k-1$ survives with probability $p_S(\mathbf{x}_{k-1})$ or dies with probability $1 - p_S(\mathbf{x}_{k-1})$. The surviving targets evolve with a transition density $f(\cdot | \mathbf{x}_{k-1})$ but may be static (in which case they are landmarks) or mobile (in this case they are targets, which is the terminology we will adopt here). The set of targets at time step k , $\tilde{\mathcal{X}}_k$, is the union of surviving and evolving targets and new targets, where target birth follows a PPP with the intensity $\lambda_B(\cdot)$. The sensor may have an unknown state (in the case of SLAT and SLAM) or a known state (in the case of MTT and mapping).

3) *Measurements:* The targets \mathcal{X}_k are observed at the sensor state \mathbf{s}_k , and the observations are denoted using a measurement set \mathcal{Z}_k . Each target $\mathbf{x}_k \in \tilde{\mathcal{X}}_k$ is detected with probability $p_D(\mathbf{s}_k, \mathbf{x}_k)$, and if detected, it generates a single measurement with the single target measurement likelihood function $g(\cdot | \mathbf{s}_k, \mathbf{x}_k)$. The measurement \mathcal{Z}_k is the union of target measurements and PPP with the intensity $c(\cdot)$.

B. Prediction with Joint Density and Factor Graph

Without loss of generality, we consider two time steps, $k-1$ and k , as part of an iterative Bayesian filter. For the factor graph formulation, the joint density for all variables (including auxiliary variables) in the prediction is factorized as

$$\begin{aligned} & f(\mathbf{s}_{k-1:k}, \tilde{\mathcal{X}}_{k-1}^U, \tilde{\mathcal{X}}_k^S, \tilde{\mathcal{X}}_k^B, \tilde{\mathcal{P}}_k^U, \tilde{\mathcal{X}}_{k-1:k}^{1:I_{k-1}}) \\ & \propto f_u(\mathbf{s}_{k-1}) f_u^U(\tilde{\mathcal{X}}_{k-1}^U) \prod_{i \in \mathcal{I}_{k-1}} f_u^i(\tilde{\mathcal{X}}_{k-1}^i) \end{aligned} \quad (40a)$$

$$\times f(\mathbf{s}_k | \mathbf{s}_{k-1}) f^P(\tilde{\mathcal{X}}_k^S | \tilde{\mathcal{X}}_{k-1}^U) f(\tilde{\mathcal{X}}_k^i | \tilde{\mathcal{X}}_{k-1}^i) \quad (40b)$$

$$\times f^U(\tilde{\mathcal{X}}_k^B) \delta_{\tilde{\mathcal{X}}_k^S \uplus \tilde{\mathcal{X}}_k^B}(\tilde{\mathcal{P}}_k^U), \quad (40c)$$

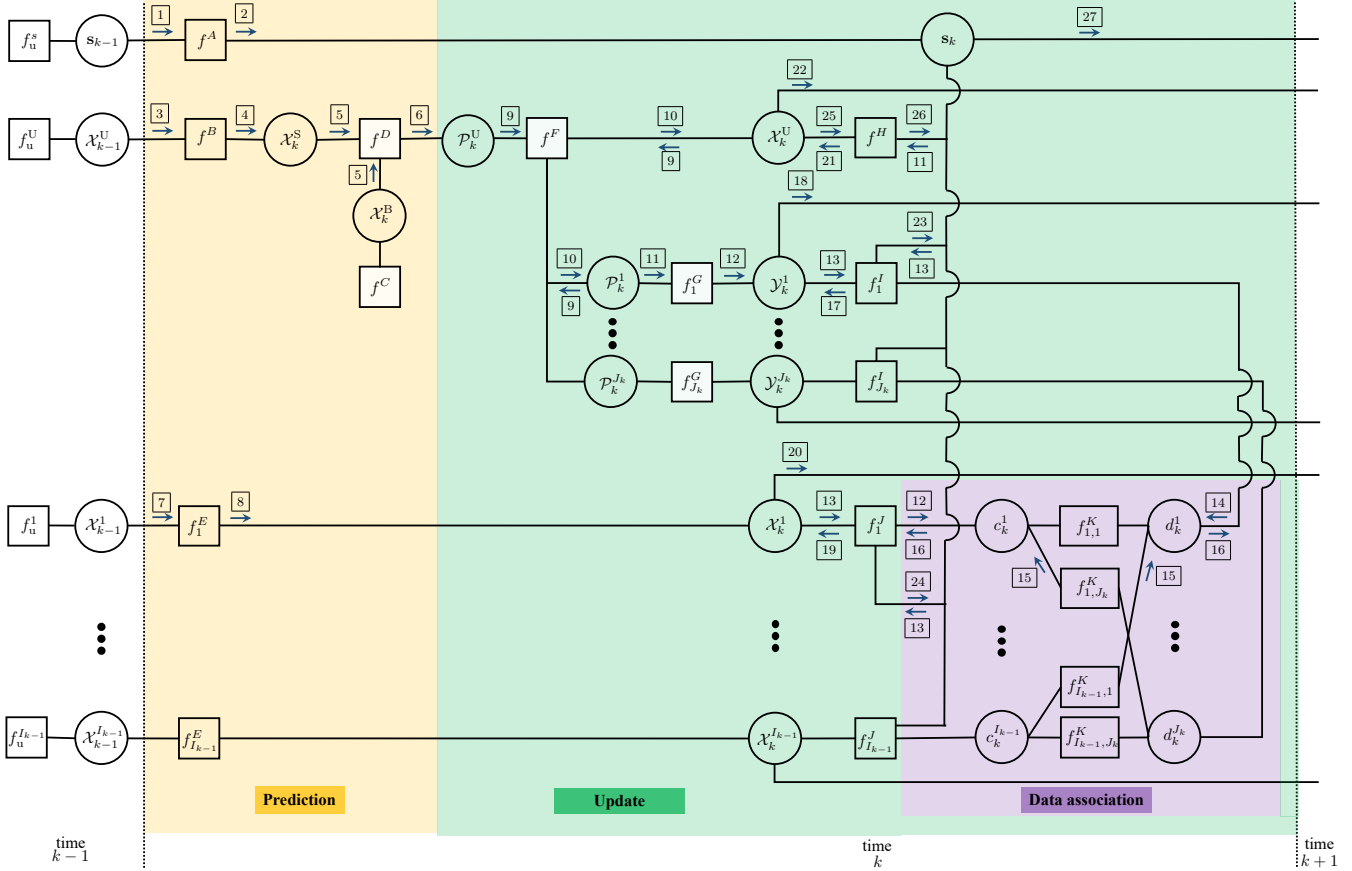


Fig. 5: Concatenated factor graph of the joint densities of (40) and (46). The special factors for set densities, discussed in Section III-B, are represented by white squares. The abbreviated notations for the factor nodes are represented by dropping the arguments and (\cdot) on sets, as follows: $f_u^S \triangleq \mathbf{b}(\mathbf{s}_{k-1})$, $f_u^U \triangleq \mathbf{b}(\mathcal{X}_{k-1}^U)$, $f_u^D \triangleq \mathbf{b}(\mathcal{X}_{k-1}^D)$, $f^A \triangleq f^A(\mathbf{s}_k, \mathbf{s}_{k-1})$, $f^B \triangleq f^B(\mathcal{X}_k^S, \mathcal{X}_{k-1}^U)$, $f^C \triangleq f^C(\mathcal{X}_k^B)$, $f^D \triangleq f^D(\mathcal{X}_k^S, \mathcal{X}_k^B, \mathcal{P}_k^U)$, $f^E \triangleq f^E(\mathcal{X}_k^U, \mathcal{X}_{k-1}^U)$, $f^F \triangleq f^F(\mathcal{X}_k^U, \mathcal{P}_k^1, \dots, \mathcal{P}_k^{J_k})$, $f_j^G \triangleq f_j^G(\mathcal{P}_k^j, \mathcal{Y}_k^j)$, $f_j^H \triangleq f_j^H(\mathbf{s}_k, \mathcal{X}_k^U)$, $f_j^I \triangleq f_j^I(\mathbf{s}_k, \mathcal{Y}_k^j, d_k^j)$, $f_i^J \triangleq f_i^J(\mathbf{s}_k, \mathcal{X}_k^i, c_k^i)$, and $f_{i,j}^K \triangleq f_{i,j}^K(c_k^i, d_k^j)$.

where $\tilde{\mathcal{X}}_{k-1:1:k}^{1:I_{k-1}}$ represents the sequence of sets $\tilde{\mathcal{X}}^i$ for $i = 1, \dots, I_{k-1}$ at time k and $k-1$, and I_{k-1} is the number of Bernoullis at time $k-1$. For completeness, the meaning of each line in (40) is described as follows.

- *Posterior at time $k-1$ (40a):* This line describes the posterior at time step $k-1$, which is assumed to be $f_u(\mathbf{s}_{k-1}, \tilde{\mathcal{X}}_{k-1}) = f_u(\mathbf{s}_{k-1})f_u(\tilde{\mathcal{X}}_{k-1})$, where $f_u(\mathbf{s}_{k-1})$ is the sensor state posterior, and $f_u(\tilde{\mathcal{X}}_{k-1})$ is the target set posterior. Here $\tilde{\mathcal{X}}_{k-1}$ follows a PMB, endowed with auxiliary variables (see Section II-B3). Due to the auxiliary variables, the PMB posterior at time $k-1$, $f_u(\tilde{\mathcal{X}}_{k-1})$, can be factorized as the product of $f_u(\tilde{\mathcal{X}}_{k-1}^U)$ and $f_u(\mathcal{X}_{k-1}^i)$ for $i \in \mathcal{I}_{k-1}$, where $\mathcal{I}_{k-1} = \{1, \dots, I_{k-1}\}$. Here $\tilde{\mathcal{X}}_{k-1}^U$ is the set of undetected targets (modeled as a PPP), $\tilde{\mathcal{X}}_{k-1}^i$ is the set of detected target i (each modeled as a Bernoulli). The subsets $\tilde{\mathcal{X}}_{k-1}^U$ and $\tilde{\mathcal{X}}_{k-1}^i$ of $\tilde{\mathcal{X}}_{k-1}^U$ are defined as

$$\tilde{\mathcal{X}}_{k-1}^U = \{(u, \mathbf{x}) \in \tilde{\mathcal{X}}_{k-1} : u = 0\}, \quad (41)$$

$$\tilde{\mathcal{X}}_{k-1}^i = \{(u, \mathbf{x}) \in \tilde{\mathcal{X}}_{k-1} : u = i\}, \quad i \in \mathcal{I}_{k-1}. \quad (42)$$

The set densities for targets are given by the form of (13)–(15).

- *Transition densities (40b):* This line describes the transition densities, where $f(\mathbf{s}_k|\mathbf{s}_{k-1})$, $f(\tilde{\mathcal{X}}_k^i|\tilde{\mathcal{X}}_{k-1}^i)$, and $f^P(\tilde{\mathcal{X}}_k^S|\tilde{\mathcal{X}}_{k-1}^U)$ correspond to the sensor state, previously

detected targets, and undetected targets. Here $\tilde{\mathcal{X}}_k^S$ denotes a surviving set from $\tilde{\mathcal{X}}_{k-1}^U$, modeled as a PPP with auxiliary variable $u = 0$. The transition density for Bernoullis with auxiliary variables is

$$f(\tilde{\mathcal{X}}_k|\tilde{\mathcal{Y}}_k) = \begin{cases} p_S(\mathbf{y}_k)f(\mathbf{x}_k|\mathbf{y}_k)\delta_{u'_k}[u_k], & \tilde{\mathcal{X}}_k = \{(u_k, \mathbf{x}_k)\}, \\ 1 - p_S(\mathbf{y}_k), & \tilde{\mathcal{Y}}_k = \{(u'_k, \mathbf{y}_k)\} \\ 1, & \tilde{\mathcal{X}}_k = \emptyset, \\ 0, & \tilde{\mathcal{Y}}_k = \{(u'_k, \mathbf{y}_k)\} \\ & \tilde{\mathcal{X}}_k = \emptyset, \tilde{\mathcal{Y}}_k = \emptyset \\ & \text{otherwise} \end{cases}, \quad (43)$$

and the transition density for the PPP is

$$f^P(\tilde{\mathcal{X}}_k|\tilde{\mathcal{Y}}_k) = \begin{cases} \sum_{\cup_{i=1}^n \tilde{\mathcal{X}}_k^i = \tilde{\mathcal{X}}_k} \prod_{i=1}^n f(\tilde{\mathcal{X}}_k^i|\{\tilde{\mathbf{y}}_k^i\}), & \tilde{\mathcal{Y}}_k = \{\tilde{\mathbf{y}}_k^1, \dots, \tilde{\mathbf{y}}_k^n\} \\ 1, & \tilde{\mathcal{X}}_k = \emptyset, \tilde{\mathcal{Y}}_k = \emptyset \\ 0, & \text{otherwise} \end{cases}.$$

where $f(\tilde{\mathcal{X}}_k^i|\{\tilde{\mathbf{y}}_k^i\})$ follows (43) and $\tilde{\mathbf{y}}_k^i = (0, \mathbf{y}_k^i)$.

- *Undetected targets (40c):* This last line describes the set-density of newborn targets at time k , represented by

$f(\tilde{\mathcal{X}}_k^B)$, where $\tilde{\mathcal{X}}_k^B$ denotes a set of newborn targets that follows a PPP, and the set-factor $\delta_{\tilde{\mathcal{X}}_k^S \uplus \tilde{\mathcal{X}}_k^B}(\tilde{\mathcal{P}}_k^U)$ merges the set of surviving targets $\tilde{\mathcal{X}}_k^S$ and the set of newborn targets $\tilde{\mathcal{X}}_k^B$ into the set $\tilde{\mathcal{P}}_k^U$ (see also Definition 3). Both $\tilde{\mathcal{X}}_k^B$ and $\tilde{\mathcal{P}}_k^U$ have auxiliary variable $u = 0$.

Using the factorized density in (40), we depict the factor graph for prediction, as shown in Fig. 5 (yellow area). Here we use the following notations for the factor nodes:

$$f^A(\mathbf{s}_k, \mathbf{s}_{k-1}) \triangleq f(\mathbf{s}_k | \mathbf{s}_{k-1}), \quad (45a)$$

$$f^B(\tilde{\mathcal{X}}_k^S, \tilde{\mathcal{X}}_{k-1}^U) \triangleq f(\tilde{\mathcal{X}}_k^S | \tilde{\mathcal{X}}_{k-1}^U), \quad (45b)$$

$$f^C(\tilde{\mathcal{X}}_k^B) \triangleq f^U(\tilde{\mathcal{X}}_k^B) \quad (45c)$$

$$f^D(\tilde{\mathcal{X}}_k^S, \tilde{\mathcal{X}}_k^B, \tilde{\mathcal{P}}_k^U) \triangleq \delta_{\tilde{\mathcal{X}}_k^S \uplus \tilde{\mathcal{X}}_k^B}(\tilde{\mathcal{P}}_k^U), \quad (45d)$$

$$f_i^E(\tilde{\mathcal{X}}_k^i, \tilde{\mathcal{X}}_{k-1}^i) \triangleq f(\tilde{\mathcal{X}}_k^i | \tilde{\mathcal{X}}_{k-1}^i). \quad (45e)$$

We compute the messages on the prediction factor graph in Fig. 5, which has no cycle, by running developed set-type BP.² Note that the messages [2], [6], and [8] are equivalent to the marginal densities in prediction of \mathbf{s}_k , $\tilde{\mathcal{P}}_k^U$, and $\tilde{\mathcal{X}}_k^{1:I_{k-1}}$, corresponding to prediction of the standard PMB filter [26], represented as $f_p(\mathbf{s}_k)$, $f_p^U(\tilde{\mathcal{P}}_k^U)$, and $f_p^i(\tilde{\mathcal{X}}_k^i)$ for $i \in \mathcal{I}_{k-1}$.

C. Update with Joint Density and its Factor Graph

Without any loss of generality, we consider the update step at time k after the prediction step. The joint density for all variables in the update is

$$f(\mathbf{s}_k, \tilde{\mathcal{P}}_k^U, \tilde{\mathcal{P}}_k^{1:J_k}, \tilde{\mathcal{X}}_k^U, \tilde{\mathcal{Y}}_k^{1:J_k}, \tilde{\mathcal{X}}_k^{1:I_{k-1}}, \mathbf{c}_k, \mathbf{d}_k | \mathcal{Z}_k) \propto f_p(\mathbf{s}_k) f_p^U(\tilde{\mathcal{X}}_k^U) \prod_{i \in \mathcal{I}_{k-1}} f_p^i(\tilde{\mathcal{X}}_k^i) \quad (46a)$$

$$\times \delta_{\tilde{\mathcal{X}}_k^U \uplus \tilde{\mathcal{P}}_k^U}(\tilde{\mathcal{P}}_k^U) \prod_{j \in \mathcal{J}_k} \psi(c_k^j, d_k^j) \delta_{\mathbf{h}_{-(I_{k-1}+j)}}(\tilde{\mathcal{Y}}_k^j) (\tilde{\mathcal{P}}_k^j) \quad (46b)$$

$$\times \tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \tilde{\mathcal{Y}}_k^j, d_k^j) t(\mathcal{Z}_k^i | \mathbf{s}_k, \tilde{\mathcal{X}}_k^i, c_k^i) [1 - p_D(\mathbf{s}_k, \cdot)]^{\tilde{\mathcal{X}}_k^U}. \quad (46c)$$

The proof of (46) is an extension of Appendix E in [28] and [27], found in Appendix E. For completeness, we describe the meaning of each line in (46) as follows.

- *Prediction at time $k - 1$* (46a): This line describes the predicted densities at time step k , discussed at the end of Section IV-B.
- *Consistency constraints* (46b): This line ensures consistency among the introduced hidden variables. A set of measurements \mathcal{Z}_k is provided at time k with index set $\mathcal{J}_k = \{1, \dots, J_k\}$. We introduce data association variables $\mathbf{c}_k \in \mathbb{N}^{J_k-1}$, $c_k^i \in \{0, \dots, J_k\}$, and $\mathbf{d}_k \in \mathbb{N}^{J_k-1}$, $d_k^j \in \{0, \dots, I_{k-1}\}$, where, in target-oriented data association, $c_k^i = j$ implies that target i is associated with measurement j , and $c_k^i = 0$ implies that target i is not detected. In measurement-oriented data association, $d_k^j = i > 0$ implies that measurement j is associated with target i , and $d_k^j = 0$ implies that measurement j originates from either a target that has never been detected or clutter. In particular, we introduced

$\Psi(\mathbf{c}_k, \mathbf{d}_k) = \prod_{i=1}^{I_{k-1}} \prod_{j=1}^{J_k} \psi(c_k^i, d_k^j)$, which ensures mutual consistency between \mathbf{c}_k and \mathbf{d}_k , with $\psi(c_k^i, d_k^j) = 0$ when $c_k^i = j, d_k^j \neq i$ or $c_k^i \neq j, d_k^j = i$, and 1 otherwise. The factor $\delta_{\tilde{\mathcal{X}}_k^U \uplus \tilde{\mathcal{P}}_k^U}(\tilde{\mathcal{P}}_k^U)$ partitions the set $\tilde{\mathcal{P}}_k^U$ into $J_k + 1$ sets, such that $\tilde{\mathcal{P}}_k^U = \tilde{\mathcal{X}}_k^U \uplus \tilde{\mathcal{P}}_k^j$, where $\tilde{\mathcal{X}}_k^U$ is the set of targets that remain undetected (as indicated by $[1 - p_D(\mathbf{s}_k, \cdot)]^{\tilde{\mathcal{X}}_k^U}$ in the likelihood part), and $\tilde{\mathcal{P}}_k^j$ for $j \in \mathcal{J}_k$ represent surviving undetected or newborn targets that are first detected at time k . The factor $\delta_{\mathbf{h}_{-(I_{k-1}+j)}}(\tilde{\mathcal{Y}}_k^j)$ (see Definition 5) converts the auxiliary variable of $\tilde{\mathcal{P}}_k^j$ from 0 to $I_{k-1} + j$, turning the set $\tilde{\mathcal{P}}_k^j$ into a set $\tilde{\mathcal{Y}}_k^j$ corresponding to a newly detected target (or clutter) arising from measurement j . The subsets $\tilde{\mathcal{X}}_k^U$, $\tilde{\mathcal{X}}_k^i$, and $\tilde{\mathcal{Y}}_k^j$ are defined as

$$\tilde{\mathcal{X}}_k^U = \{(u, \mathbf{x}) \in \tilde{\mathcal{X}}_k : u = 0\}, \quad (47)$$

$$\tilde{\mathcal{X}}_k^i = \{(u, \mathbf{x}) \in \tilde{\mathcal{X}}_k : u = i\}, \quad i \in \mathcal{I}_{k-1}, \quad (48)$$

$$\tilde{\mathcal{Y}}_k^j = \{(u, \mathbf{x}) \in \tilde{\mathcal{X}}_k : u = I_{k-1} + j\}, \quad j \in \mathcal{J}_k. \quad (49)$$

- *Set-likelihoods* (46c): This last line describes the likelihood, conditioned on the data association \mathbf{c}_k (or equivalently \mathbf{d}_k). The factor $\tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \tilde{\mathcal{Y}}_k^j, d_k^j)$ considers potential new targets or clutter, while $t(\mathcal{Z}_k^i | \mathbf{s}_k, \tilde{\mathcal{X}}_k^i, c_k^i)$ considers detections and missed detections of the previously detected targets. Both likelihoods are defined as follows. We consider the likelihood when the j -th measurement \mathbf{z}_k^j is associated with a newly detected target or with clutter conditioned on measurement-oriented data association variable d_k^j : $\tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \tilde{\mathcal{Y}}_k^j, d_k^j)$ is given by

$$\tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \tilde{\mathcal{Y}}_k^j, d_k^j) = \begin{cases} p_D(\mathbf{s}_k, \mathbf{x}_k) g(\mathbf{z}_k^j | \mathbf{s}_k, \mathbf{x}_k) \delta_0[u_k], & \tilde{\mathcal{Y}}_k^j = \{(u_k, \mathbf{x}_k)\}, \\ c(\mathbf{z}_k), & d_k^j = 0 \\ 1, & \tilde{\mathcal{Y}}_k^j = \emptyset, d_k^j = 0 \\ 0, & \tilde{\mathcal{Y}}_k^j = \emptyset, d_k^j \neq 0 \\ & \text{otherwise.} \end{cases} \quad (50)$$

Here $p_D(\cdot)$ accounts for the fact that a target may be detected, while $c(\mathbf{z}_k)$ represents the clutter intensity (when a measurement is generated by clutter and not a newly detected target). The function $g(\mathbf{z}_k^j | \mathbf{s}_k, \mathbf{x}_k)$ is a classical likelihood function. Then, the likelihood $t(\mathcal{Z}_k^i | \mathbf{s}_k, \tilde{\mathcal{X}}_k^i, c_k^i)$ considers the case when a set \mathcal{Z}_k^i is associated to the i -th previously detected target conditioned on target-oriented data association variable c_k^i :

$$t(\mathcal{Z}_k^i | \mathbf{s}_k, \tilde{\mathcal{X}}_k^i, c_k^i) = \begin{cases} p_D(\mathbf{s}_k, \mathbf{x}_k^i) g(\mathbf{z}_k^i | \mathbf{s}_k, \mathbf{x}_k^i) \delta_i[u_k], & \mathcal{Z}_k^i = \{\mathbf{z}_k^i\}, c_k^i = j, \\ (1 - p_D(\mathbf{s}_k, \mathbf{x}_k^i)) \delta_i[u_k], & \tilde{\mathcal{X}}_k^i = \{(u_k, \mathbf{x}_k^i)\} \\ 1, & \mathcal{Z}_k^i = \emptyset, c_k^i = 0, \\ & \tilde{\mathcal{X}}_k^i = \{(u_k, \mathbf{x}_k^i)\} \\ 0, & \mathcal{Z}_k^i = \emptyset, c_k^i = 0, \\ & \tilde{\mathcal{X}}_k^i = \emptyset \\ & \text{otherwise.} \end{cases} \quad (51)$$

²This step is a straightforward application of the message passing rules introduced in Section III. For completeness, the messages are detailed in Appendix F of the supplementary material.

The factor $[1 - p_D(\mathbf{s}_k, \cdot)]^{\tilde{\mathcal{X}}_k^U}$ describes the set of targets that remain undetected.

Using the factorized density in (46), we depict the factor graph for update, as shown in Fig. 5 (green area). Here we use the following notations for the factor nodes:

$$f^F(\tilde{\mathcal{X}}_k^U, \tilde{\mathcal{P}}_k^1, \dots, \tilde{\mathcal{P}}_k^{J_k}) \triangleq \delta_{\tilde{\mathcal{X}}_k^U \uplus_{j=1}^{J_k} \tilde{\mathcal{P}}_k^j}(\tilde{\mathcal{P}}_k^U), \quad (52a)$$

$$f_j^G(\tilde{\mathcal{P}}_k^j, \tilde{\mathcal{Y}}_k^j) \triangleq \delta_{\mathbf{h}_{-(I_{k-1}+j)}}(\tilde{\mathcal{Y}}_k^j)(\tilde{\mathcal{P}}_k^j), \quad (52b)$$

$$f^H(\mathbf{s}_k, \tilde{\mathcal{X}}_k^U) \triangleq [1 - p_D(\mathbf{s}_k, \cdot)]^{\tilde{\mathcal{X}}_k^U}, \quad (52c)$$

$$f_j^I(\mathbf{s}_k, \tilde{\mathcal{Y}}_k^j, d_k^j) \triangleq \tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \tilde{\mathcal{Y}}_k^j, d_k^j), \quad (52d)$$

$$f_i^J(\mathbf{s}_k, \tilde{\mathcal{X}}_k^i, c_k^i) \triangleq t(\mathcal{Z}_k^i | \mathbf{s}_k, \tilde{\mathcal{X}}_k^i, c_k^i), \quad (52e)$$

$$f_{i,j}^K(c_k^i, d_k^j) \triangleq \psi(c_k^i, d_k^j). \quad (52f)$$

We compute the messages and beliefs on the update factor graph in Fig. 5 by running developed set-type BP, detailed in Appendix F of the supplementary material. Messages are not sent backward in time. The iterative message is only performed for the data association step, whereas for other steps a single message passing is performed. As a final result, the message [27] represents the posterior of the sensor state at the end of time step k , [22] represents the posterior PPP of undetected targets, [18] represents the posteriors of the Bernoullis of each newly detected target, and [20] represents the posteriors of the Bernoullis of each previously detected target. These posteriors will be used when creating the factor graph connecting time step k with time step $k+1$. Note that the number of detected targets grows over time.

D. Connecting the Factor Graphs in Prediction and Update

So far, we have discussed the prediction and update factor graphs independently. We now proceed to explain how these factor graphs are connected. The factor graph in the prediction (see, yellow part in Fig. 5) includes variable nodes $\mathbf{s}_k, \tilde{\mathcal{P}}_k^U, \tilde{\mathcal{X}}_k^1, \dots, \tilde{\mathcal{X}}_k^{I_{k-1}}$ while the factor graph in the update (see, green and purple parts in Fig. 5) contains variables $\mathbf{s}_k, \tilde{\mathcal{X}}_k^U, \tilde{\mathcal{P}}_k^1, \dots, \tilde{\mathcal{P}}_k^{J_k}, \tilde{\mathcal{X}}_k^1, \dots, \tilde{\mathcal{X}}_k^{I_{k-1}}$. We note that the variables $\mathbf{s}_k, \tilde{\mathcal{X}}_k^1, \dots, \tilde{\mathcal{X}}_k^{I_{k-1}}$ are the same. To connect the other variable nodes in the prediction and update factor graphs, we can use a merging and partitioning factor (see, Definition 3) as well as a conversion factor for auxiliary variables (see, Definition 5) for newly detected Bernoullis. That is, we first adopt the factor $f^F(\tilde{\mathcal{P}}_k^U, \tilde{\mathcal{X}}_k^U, \tilde{\mathcal{P}}_k^1, \dots, \tilde{\mathcal{P}}_k^{J_k})$, which partitions the set $\tilde{\mathcal{P}}_k^U$ into subsets $\tilde{\mathcal{X}}_k^U, \tilde{\mathcal{Y}}_k^1, \dots, \tilde{\mathcal{Y}}_k^{J_k}$ (all with auxiliary variable 0). Then, to each of the newly detected Bernoullis, we apply the conversion factor for auxiliary variables $f_j^G(\tilde{\mathcal{P}}_k^j, \tilde{\mathcal{Y}}_k^j)$. Considering the set-variables that are linked to the set-factors $f^F(\cdot)$ and $f_j^G(\cdot)$, see Fig. 5, we know that the set-variable $\tilde{\mathcal{Y}}_k^j$ and $\tilde{\mathcal{X}}_k^U$ follow a Poisson process with $u = I_{k-1} + j$ and $u = 0$, respectively. Finally, the factor graphs for prediction and update are connected by the factors ($f^A(\cdot)$, $f^D(\cdot)$, and $f_i^E(\cdot)$) and their linked variables.

Remark 2 (Factor Graphs for Smoothing at the Previous Time Step). *Fig. 5 shows the concatenation of factor graphs for the prediction and update steps at time $k-1$ and k . It is possible to derive a joint prediction and update factor graph (instead*

of its concatenation). This factor graph would require the use of PMBs for sets of trajectories between time $k-1$ and k [42], [43]. Furthermore, this factor graph would enable us to infer the state at time $k-1$ of a Bernoulli created at time k , via smoothing.

Remark 3 (Special Cases and Generalizations). *The factorization ((40) and (46)) and the corresponding factor graph in Fig. 5 can be specialized and generalized to cover a variety of applications. Some special cases include:*

- *Mapping and SLAM are obtained when the targets have no mobility, which is obtained when the transition density is $f(\mathbf{x}_k | \mathbf{x}_{k-1}) = \delta(\mathbf{x}_k - \mathbf{x}_{k-1})$.*
- *Mapping and MTT are obtained when the sensor state is known at all times, so that $f_u(\mathbf{s}_{k-1})$ and $f(\mathbf{s}_k | \mathbf{s}_{k-1})$ are removed and \mathbf{s}_{k-1} and \mathbf{s}_k no longer explicitly appear as vertices in the factor graph (instead they are absorbed as parameters in the likelihood functions).*

Remark 4 (Set-Type BP MB Filter). *To obtain a set-type BP MB filter (whether labeled or not), we use the same modeling assumptions as in the set-type BP PMB filter except that the birth model is MB instead of Poisson. Then, we set the intensity of PPP to zero and add the Bernoulli components of the birth process in the prediction step [44]. The corresponding factor graph is equivalent to the one in Fig. 5 but removing all the PPP variables ($\mathcal{X}_{k-1}^U, \mathcal{X}_k^S, \mathcal{X}_k^B, \mathcal{P}_k^U, \mathcal{P}_k^1, \dots, \mathcal{P}_k^{J_k}, \mathcal{X}_k^U$), their connected factors ($f_u^U, f^B, f^C, f^D, f^F, f_1^G, \dots, f_{J_k}^G$), and newly detected targets ($\mathcal{Y}_k^1, \dots, \mathcal{Y}_k^{J_k}$), and also adding the new birth Bernoulli variables and factors in the prediction step.*

Remark 5 (On the Use of Set-Type BP with Other RFS Filters). *It should be noted that set-type BP can be used to derive RFS filters based on computing marginal distributions for targets, such as PMB and MB filters, for general measurement and clutter models, including coexisting point and extended targets [45]. It is not suitable to derive PHD and CPHD filters as these filters do not calculate marginals but approximate the posterior as a PPP or as an independent and identically distributed cluster process [22]. Set-type BP is also not suitable for implementing filters based on conjugate priors such as PMBM and δ -GLMB filters.*

E. Approximate KLD Minimization of Set-Type BP PMB for SLAM

Assuming that the prior is a PMB, the prediction of the set-type BP is represented in closed-form. As a direct application of the PMBM update [28, Sec. IV-B], the updated density is $\tilde{f}_u^{\text{PMBM}}(\tilde{\mathcal{X}}_k | \mathbf{s}_k) f_u(\mathbf{s}_k)$, where $\tilde{f}_u^{\text{PMBM}}(\tilde{\mathcal{X}}_k | \mathbf{s}_k)$ corresponds to the PMBM with the auxiliary variables [43, Sec. III-A], given the sensor state \mathbf{s}_k , where $\tilde{\mathcal{X}}_k = \tilde{\mathcal{X}}_k^U \uplus \tilde{\mathcal{X}}_k^{1:I_{k-1}} \uplus \tilde{\mathcal{Y}}_k^{1:J_k}$ (see, (46)).

Lemma 1. *The marginals of $\mathbf{s}_k, \tilde{\mathcal{X}}_k^U, \tilde{\mathcal{X}}_k^i, \tilde{\mathcal{Y}}_k^j$ of (46) represent a PMB that is an optimal approximation of the corresponding PMBM posterior in the sense that it minimizes the Kullback-Leibler divergence*

(KLD) $D(\tilde{f}_u^{\text{PMBM}}(\tilde{\mathcal{X}}_k | \mathbf{s}_k) f_u(\mathbf{s}_k) || \tilde{q}_u^{\text{PMB}}(\tilde{\mathcal{X}}) q_u(\mathbf{s}_k))$. The marginal densities of \mathbf{s}_k , $\tilde{\mathcal{X}}_k^U$, $\tilde{\mathcal{X}}_k^i$, $\tilde{\mathcal{Y}}_k^j$ are given by

$$\tilde{q}_u^{\text{PMB}}(\tilde{\mathcal{X}}) = \tilde{q}_u^U(\tilde{\mathcal{X}}^U) \prod_{i=1}^n \tilde{q}_u^i(\tilde{\mathcal{X}}^i), \quad (53)$$

$$q_u(\mathbf{s}_k) = f_u(\mathbf{s}_k), \quad (54)$$

$$q_u(\tilde{\mathcal{X}}_k^U) = \int f_u(\tilde{\mathcal{X}}_k^U | \mathbf{s}_k) f_u(\mathbf{s}_k) d\mathbf{s}_k, \quad (55)$$

$$q_u(\tilde{\mathcal{X}}_k^i) = \int f_u(\tilde{\mathcal{X}}_k^i | \mathbf{s}_k) f_u(\mathbf{s}_k) d\mathbf{s}_k, \quad (56)$$

$$q_u(\tilde{\mathcal{Y}}_k^j) = \int f_u(\tilde{\mathcal{Y}}_k^j | \mathbf{s}_k) f_u(\mathbf{s}_k) d\mathbf{s}_k. \quad (57)$$

Then, upon convergence, the set-type BP applied to (46) approximates these marginal densities by producing the interior points of the constrained Bethe free energy on the factor graph, see Theorem 1.

Proof. The proof of the KLD minimization is an extension of Proposition 1 of [42], showing the optimality of the solution. The Bethe free minimization is a result of Theorem 1. \square

F. Exploring the Relationships: Set-Type and Vector-Type BP PMB-SLAM Filters

We now reveal connections between the proposed set-type BP filter to the vector-type BP filters [13]. The proposed set-type BP PMB-SLAM filter turned out to be algorithmically identical to the vector-type BP-SLAM filters [13], except for minor details.³ We note that the connection between vector-type and RFS-based method was partially discussed from the perspective of the expression of target densities and data association in [9, Sec. XIII-A] and [28, Sec. IV-E].

1) *Previously Detected Targets:* Bernoulli densities in the proposed set-type BP-SLAM filter can be parameterized using a state $\mathbf{y} = [\mathbf{x}^\top, \epsilon]^\top$, where ϵ is the existence variable such that $\epsilon = 1$ if the target exists and $\epsilon = 0$ if the target does not exist. A Bernoulli density in vector representation is $f(\mathbf{x}, \epsilon) = f(\epsilon) f(\mathbf{x} | \epsilon)$ where $f(\mathbf{x}, \epsilon = 1) = f(\{\mathbf{x}\})$, and $f(\mathbf{x}, \epsilon = 0) = f(\emptyset)$. The vector representation has the theoretical drawback of requiring a dummy density $f(\mathbf{x} | \epsilon = 0)$ which does not exist, though this does not affect practical results. The same procedure can be followed to define equivalent conditional densities (factors) that consider sets of at maximum one element. With this in mind, we can write the part of the factor graph in Fig. 5 which has Bernoulli densities using the existence variable representation. It is also straightforward to express the parts that involve variables that have vector densities (such as \mathbf{s}_k).

2) *Undetected Targets:* As the undetected targets are PPP distributed, this part of the factor graph, including the factors f^B (for prediction) and f^H (for update) cannot be written with existence variables. Therefore, the analogous factor graph with existence variables has to drop this source of information. One solution to compensate for this loss of information in

vector-type BP is to use an external PHD filter to model undetected features, as in [13]. The prediction step of the external PHD filter is designed to match the solution obtained by the (prediction) factor graph in Fig. 5.

3) *Newly Detected Targets:* In the update, the external PHD filter can be used to describe the density of the newly detected Bernoullis. In order to link this PHD filter with the factor graph, reference [13] requires direct modeling of the density of newly detected features. This density corresponds to the predicted intensity of the PHD filter multiplied by the probability of detection, normalized to integrate to one in [13, eq. (16)].

4) *Detection Probability:* Vector-type BP SLAM requires that the detection probability for undetected targets does not depend on the sensor state, such that $p_D(\mathbf{s}_k, \mathbf{x}_k) = p_D(\mathbf{x}_k)$, as in [13]. The reason why this is required is that if $p_D(\mathbf{s}_k, \mathbf{x}_k)$ depends on \mathbf{s}_k and \mathbf{x}_k , the detection probability should be part of the factor that connects each newly detected feature and the sensor state, as in Fig. 5. This could be fixed in vector-type BP by defining the density of undetected features (as the normalized PHD intensity) and then considering $p_D(\mathbf{s}_k, \mathbf{x}_k)$ in the factor, as in Fig. 5.

5) *Sensor State Update:* It is also relevant to mention that, while the factor graphs in [12], [13] consider messages from the sensor state to the newly detected features, the belief update equation of the sensor state does not have these messages, see [12, eq. (37)] and [13, Sec. III-G]. The equations with the messages from newly detected features to the sensor state are not considered in [12], [13] either. These equations can be corrected to obtain a vector-type BP-SLAM algorithm similar to the set-type BP-SLAM algorithm presented in this paper.

6) *Derivations:* It is also important to realize that this paper provides an alternative derivation of the (slightly modified) vector-type BP-SLAM algorithm in [13] obtained from first principles using RFSs. That is, vector-type BP-SLAM is obtained assuming models for single-feature dynamics, single-feature death, sensor dynamics, and measurement models, but no feature births. Instead, it is assumed that, at the current time step, we know the multi-Bernoulli density of previous features and the multi-Bernoulli density of newly detected features. Then, the factor graph is obtained. Afterwards, an external PHD filter can be used in vector-based SLAM to model undetected features.

Instead, the set-type BP PMB-SLAM derivation works as follows. We require models for feature birth, single-feature dynamics, single-feature death, sensor dynamics, and measurements, similar to Bayesian multi-target tracking algorithms. Then, auxiliary variables are used to obtain the associated factor graph, and then the factor graph is solved using set-BP. The resulting algorithm, with the minor modifications explained above, is similar to vector-BP SLAM with the external PHD filter in [13]. The connection is clearer by writing each Bernoulli component in terms of existence variables. This shows that the vector-BP SLAM algorithm with the external PHD filter, with the minor modifications above, can be derived from RFS theory, which is an important result for the literature. It is also relevant to note that the set-BP MB-SLAM algorithm

³Note that the vector-type BP-SLAM filter in [12] mentions the possibility of using an external PHD filter, but it is not explained in detail, so we focus on the comparison with [13], as it has an explicit description of the external PHD filter.

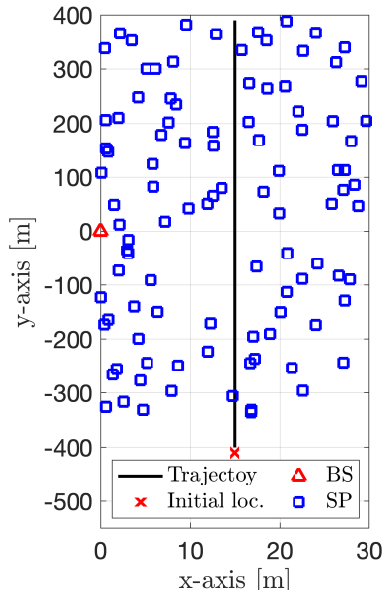


Fig. 6: Bistatic radio SLAM scenario, where 176 SPs are uniformly distributed. A single BS transmits the signal, and the sensor receives the two types of measurements: BS-sensor; and BS-SPs-sensor.

that we also propose could also be written with existence variables.

V. NUMERICAL RESULTS

In this section, we analyze the proposed set-type BP PMB-SLAM and BP MB-SLAM filters, abbreviated as SBP-PMB-SLAM and SBP-MB-SLAM, respectively, in comparison with BP-SLAM [13], abbreviated as VBP-SLAM1. Note that the vector-type BP-SLAM filters [12], [13] with the minor modifications, abbreviated as VBP-SLAM2, is equivalent to the SBP-PMB-SLAM, as discussed in Sec. IV-F. We introduce the simulation setup for evaluating the SLAM filters, and subsequently the results are discussed.

A. Simulation Setup

1) *Sensor and Landmarks*: We consider a propagation environment of bistatic radio SLAM, where a single base station (BS) transmits the pilot signals, and scattering points (SPs) (i.e., landmarks) are uniformly distributed, as shown in Fig. 6. A single sensor can receive two types of measurements: one from BS-sensor path; and others from BS-SPs-sensor paths. We denote the locations of the BS and SP j by \mathbf{x}_{BS} and \mathbf{x}_{SP}^j . The sensor state at time k is denoted by $\mathbf{s}_k = [\mathbf{x}_{k,s}^\top, \dot{\mathbf{x}}_{k,s}^\top]^\top$, where $\mathbf{x}_{k,s} = [x_{k,s}, y_{k,s}]^\top$ denotes the location, and $\dot{\mathbf{x}}_{k,s} = [\dot{x}_{k,s}, \dot{y}_{k,s}]^\top$ denotes the velocity.

2) *Dynamics*: For the sensor dynamics during K time steps, we adopt constant-velocity motion model [46, Sec. 6.3.2] with the known transition density $f(\mathbf{s}_k|\mathbf{s}_{k-1})$, and the sensor state evolution follows

$$\mathbf{s}_k = \mathbf{F}\mathbf{s}_{k-1} + \mathbf{B}\mathbf{q}_k, \quad (58)$$

\mathbf{F} and \mathbf{B} are defined as [46, Sec. 6.3.2]

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_{2 \times 2} & \Delta t \mathbf{I}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \mathbf{I}_{2 \times 2} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0.5 \Delta t^2 \mathbf{I}_{2 \times 2} \\ \Delta t \mathbf{I}_{2 \times 2} \end{bmatrix}. \quad (59)$$

Here Δt is the sampling time, and \mathbf{q}_k is the driving process that follows zero-mean Gaussian distribution with the covariance matrix $\sigma_w^2 \mathbf{I}_{2 \times 2}$, where σ_w is the standard deviation, and $\mathbf{I}_{2 \times 2}$ is the 2 by 2 identity matrix.

We set the parameters as follows: $K = 80$ time steps, $\sigma_w = 0.1$ m/s², and $\Delta t = 0.5$ s. We set the prior density of the sensor state to $f(\mathbf{s}_0) = \mathcal{N}(\mathbf{s}_0; \mathbf{s}_{0|0}, \mathbf{P}_{0|0})$, where $\mathbf{s}_{0|0}$ is sampled from $\mathcal{N}(\bar{\mathbf{s}}_0, \mathbf{P}_{0|0})$ for each simulation run. Here $\bar{\mathbf{s}}_0 = [15, -420, 0, 20]^\top$ is the ground truth of the initial sensor state, with the units of m, m, m/s, and m/s, and $\mathbf{P}_{0|0} = \text{blkdiag}[0.5 \mathbf{I}_{2 \times 2}, 0.005 \mathbf{I}_{2 \times 2}]$, and the units of the diagonal term of \mathbf{P} are m², m², m²/s², and m²/s².

3) *Measurements*: The sensor obtains the measurement set $\mathcal{Z}_k = \{\mathbf{z}_k^0, \mathbf{z}_k^1, \dots, \mathbf{z}_k^j\}$, where \mathbf{z}_k^0 and \mathbf{z}_k^j for $j > 0$ are the measurements corresponding to the BS and different SPs, respectively. Let $\mathbf{t}_k^j = [(\mathbf{x}_{\text{SP}}^j)^\top, \mathbf{x}_{k,s}^\top]^\top$ be the augmented vector of SP location \mathbf{x}_{SP}^j and the sensor location $\mathbf{x}_{k,s}$. With the known BS location, the measurements \mathbf{z}_k^0 and \mathbf{z}_k^j for $j > 0$ are modeled as

$$\mathbf{z}_k^0 = \mathbf{x}_{k,s} + \mathbf{r}_k, \quad (60)$$

$$\mathbf{z}_k^j = \mathbf{H}\mathbf{t}_k^j + \mathbf{r}_k. \quad (61)$$

Here $\mathbf{H} = [\mathbf{I}_{2 \times 2}, -\mathbf{I}_{2 \times 2}]$, and \mathbf{r}_k is the measurement noise that follows zero-mean Gaussian distribution with the covariance $\sigma_r^2 \mathbf{I}_{2 \times 2}$, where σ_r is the standard deviation. We regard the false alarms and shortly visible SPs as clutter, modeled as $\mathbf{z}_k^C \in \mathcal{Z}_k$. The number of clutter measurements, which are observed by the sensor, is modeled by the Poisson distribution with mean μ_c [42], and the clutter intensity is denoted by $c(\mathbf{z})$.

The parameters are set as follows: $\mu_c = 1$, $c(\mathbf{z}) = 1.6 \cdot 10^{-4}$ in the area of interest, and $\sigma_r = 0.707$ m. The number of SPs that are uniformly distributed in a network is denoted by N_{SP} . The network size is $[0, 30]$ m \times $[-400, 470]$ m, $N_{\text{SP}} = 176$, and the BS location is $\mathbf{x}_{\text{BS}} = [0, 0]^\top$. The field-of-view of the sensor is determined by

$$p_{\text{D}}(\mathbf{s}_k, \mathbf{x}_k) = \begin{cases} p_{\text{D}} & \|\mathbf{x}_k - \mathbf{s}_k\| < r \\ 0 & \text{otherwise} \end{cases} \quad (62)$$

where $r = 20$ m and $p_{\text{D}} = 0.95$. The landmarks are observable starting from $k = 5$, and additional landmarks can be observable at each subsequent time step. For sensor localization in both scenarios, multipath is employed from $k = 5$, whereas for $k < 5$ the measurement for BS-sensor path is exploited.⁴

4) *SLAM Filter*: To investigate the influence of newly detected landmarks on sensor state updates, we consider K time steps with two scenarios: one with an uninformative birth model, and the other with an informative birth model in modeling undetected landmarks. In realistic environments, the uninformative birth model represents cases where we lack prior map knowledge, whereas under the informative birth model the assumption is that we have a previously available map.

⁴By the ellipsoidal gating method [47], the measurement for BS-sensor path is determined as $\hat{\mathbf{z}} = \min_{\mathbf{z}_k^j \in \mathcal{Z}_k} (\mathbf{z}_k^j - \mathbf{x}_{\text{BS}})^\top \mathbf{R}^{-1} (\mathbf{z}_k^j - \mathbf{x}_{\text{BS}})$.

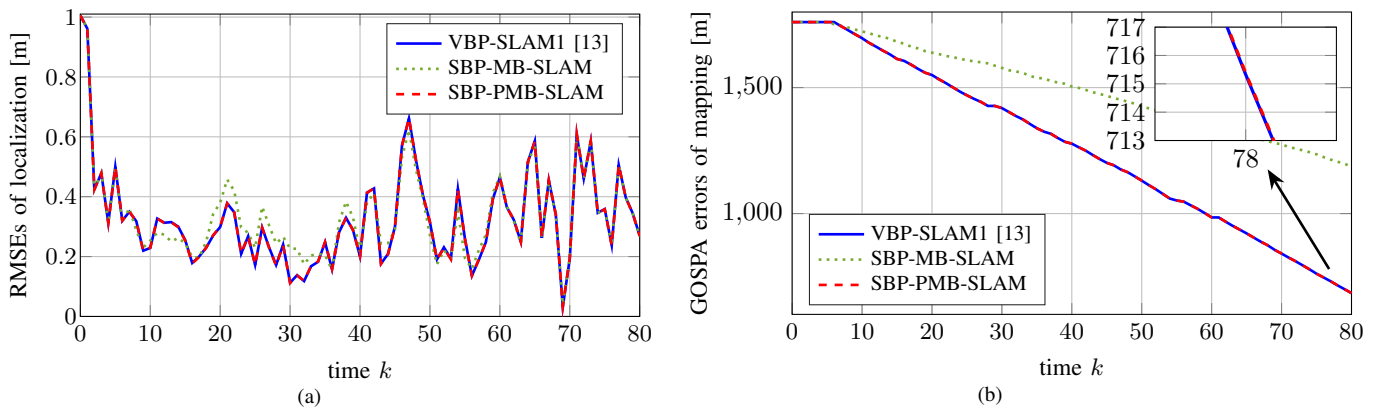


Fig. 7: Uninformative birth case. Comparison of the proposed set-type BP PMB-SLAM SBP-PMB-SLAM (i.e., VBP-SLAM2) with VBP-SLAM1 [13] and SBP-MB-SLAM: (a) RMSEs of sensor localization and (b) GOSPA errors of landmark mapping.

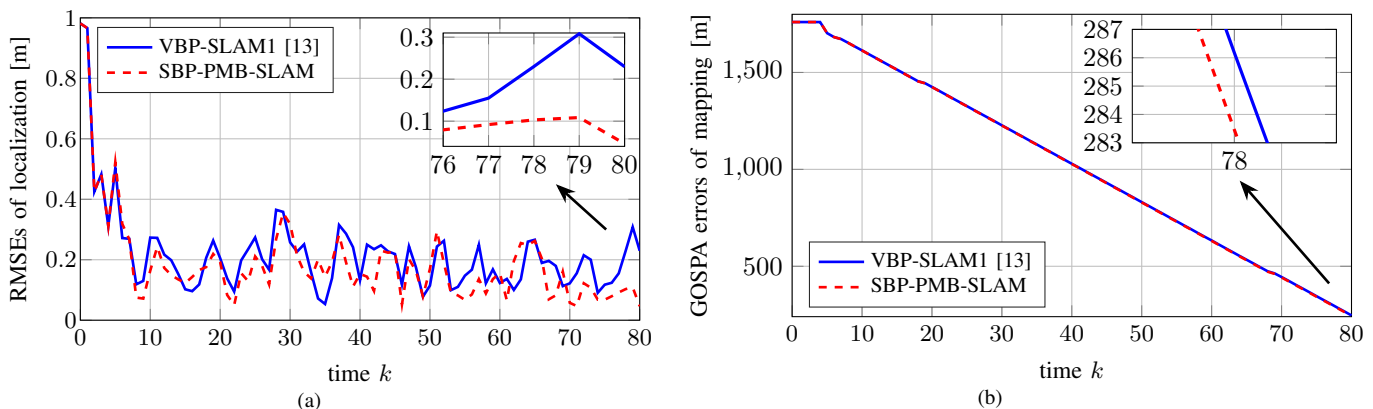


Fig. 8: Informative birth case. Comparison of the proposed SBP-PMB-SLAM (i.e., VBP-SLAM2) and VBP-SLAM1 [13]: (a) RMSEs of sensor localization and (b) GOSPA errors of landmark mapping. The performance of SBP-MB-SLAM is identical to SBP-PMB-SLAM, and the plots are omitted.

Undetected landmarks and the birth for both SBP-PMB-SLAM and VBP-PMB1 [13] are modeled by the PPP, implemented by the intensity functions. The intensities for undetected targets and the birth are represented by $\lambda(\mathbf{x}) = \sum_q \eta^q \mathcal{N}(\mathbf{x}; \mathbf{x}^q, \mathbf{U}^q)$ and $\lambda^B(\mathbf{x}) = \sum_q \eta^{B,q} \mathcal{N}(\mathbf{x}; \mathbf{x}^{B,q}, \mathbf{U}^{B,q})$, where $\mathcal{N}(\mathbf{x}; \mathbf{x}, \mathbf{U})$ is the Gaussian density, and η is the Gaussian weight. For SBP-MB-SLAM, the birth is modeled by the MB process. The MB birth density is the form of (11), where the q -th Bernoulli has an existence probability $r^{B,q}$ and Gaussian density $\mathcal{N}(\mathbf{x}; \mathbf{x}^{B,q}, \mathbf{U}^{B,q})$. For the informative birth case, we set the number of births to the number of newly detected landmarks, $\mathbf{x}^{B,q} \sim \mathcal{N}(\mathbf{x}; \mathbf{x}_{\text{SP}}^{j(q)}, \mathbf{U}^{B,q})$, $\eta^{B,q} = 1$, and the small Gaussian covariance $\mathbf{U}^{B,q} = 0.01 \mathbf{I}_{2 \times 2}$, where $j(q)$ is the SP index corresponding to the q -th Gaussian. For the uninformative birth case, we set the number of Gaussians for births to the same as the number of measurements, birth weight $\eta^{B,q} = 10^{-3}$, and the large Gaussian covariance $\mathbf{U}^{B,q} = 10^6 \times \mathbf{I}_{2 \times 2}$.

Using the Kalman filter [48], we implement the messages and beliefs corresponding to landmarks and data associations. The belief of the sensor state of [27] is intractable by the Kalman filter, and thus implemented by the particle filter [4], [49] with N_P samples. After the belief computation at each time step, we declare that a landmark is detected when $r_{k|k}^i >$

Γ_D . The previously detected landmarks with $r_{k|k}^i < \Gamma_{\text{Ber}}$ for $i = 1, \dots, I_{k-1} + J_k$ are removed, and the Gaussians in the PPP with $\eta_{k|k}^q < \Gamma_{\text{Poi}}$ are eliminated.

The rest of the algorithm parameters are set as follows: $N_P = 10^4$, $\Gamma_D = 0.4$; $\Gamma_{\text{Ber}} = 10^{-5}$; $\Gamma_{\text{Poi}} = 5 \cdot 10^{-10}$; $p_S = 0.99$. The simulation results are averaged over 500 Monte Carlo trials. The performance of sensor localization and landmark mapping are evaluated by root mean square error (RMSE) and generalized optimal subpattern assignment (GOSPA) [50, eq. (1)] with parameters $G_p = 1$, $G_c = 2$, and $G_\alpha = 2$, respectively.

B. Results and Discussions

1) *Uninformative Birth*: Fig. 7 shows the SLAM performance against time, under the scenario that the PPP for undetected landmarks and birth in SBP-PMB-SLAM and the MB for birth in SBP-MB-SLAM are uninformative. During $K = 80$, the SLAM performance of the SBP-PMB-SLAM is identical to that of the VBP-BP-SLAM filter even though the messages [23] are employed for sensor belief computation. This happens because the messages [23] corresponding to newly detected landmarks or clutter are not as informative as the messages [20] corresponding to previously detected landmarks. The mapping performance of the SBP-MB-SLAM

TABLE I: Informative birth case: RMSEs of sensor localization and GOSPA errors of mapping with different N_{SP} and μ_c .

		VBP-SLAM1 [13]	SBP-PMB-SLAM	
RMSEs [m]	$N_{\text{SP}} = 176$	$\mu_c = 1$	0.1651	0.1306
		$\mu_c = 5$	0.1881	0.1334
		$\mu_c = 10$	0.2014	0.1401
	$N_{\text{SP}} = 352$	$\mu_c = 1$	0.1321	0.1125
		$\mu_c = 5$	0.1361	0.1137
		$\mu_c = 10$	0.1559	0.1195
GOSPA Errors [m]	$N_{\text{SP}} = 176$	$\mu_c = 1$	638.48	636.42
		$\mu_c = 5$	979.74	977.31
		$\mu_c = 10$	1162.34	1160.52
	$N_{\text{SP}} = 352$	$\mu_c = 1$	622.52	620.99
		$\mu_c = 5$	794.65	793.17
		$\mu_c = 10$	881.29	880.04

filter is inferior to the other two filters. This is because the MB birth model is limited in the number of Bernoulli sets, where each is available for modeling a single landmark, whereas the set that follows PPP captures multiple landmarks.

2) *Informative Birth*: Fig. 8 shows the SLAM performance against time, under the scenario that the PPP for undetected landmarks and birth in SBP-PMB-SLAM and MB birth in SBP-MB-SLAM are informative. When using the informative PPP, we achieve performance improvement in set-type BP PMB-SLAM SBP-PMB-SLAM compared to that of VBP-SLAM1. This is because new landmarks appear and the PPP for undetected landmarks is informative. The sensor localization gap is clearly visible, and the landmarks are determined, starting from $k = 5$ since the landmarks begin to be observable from $k = 5$. The gap of GOSPAs increases over time steps because there exist newly detected landmarks at every time step. We obtain that the performance of SBP-MB-SLAM is identical to that of SBP-PMB-SLAM, and thus the results are omitted. This is because the clutter part of the messages [23]⁵ is much smaller than the newly detected landmark part, and the implementation of MB birth for SBP-MB-SLAM is equivalent to that of SBP-PMB-SLAM, under the informative birth case.

Table I present the RMSEs of sensor localization and GOSPA errors of landmark mapping respectively, with the different setup: $\mu_c = 1$ with $c(\mathbf{z}) = 1.6 \cdot 10^{-4}$, $\mu_c = 5$ with $c(\mathbf{z}) = 8 \cdot 10^{-4}$, and $\mu_c = 10$ with $c(\mathbf{z}) = 1.6 \cdot 10^{-3}$, for $N_{\text{SP}} = 176$, $N_{\text{SP}} = 352$. The results are averaged over all Monte Carlo simulation runs during the steady-state operation after $k > 40$. Table I show the performance of both methods deteriorates progressively as the clutter Poisson mean μ_c increases ($\mu_c = 1, \mu_c = 5, \mu_c = 10$). The sensor localization accuracy with $N_{\text{D}} = 4$ is improved, compared to with $N_{\text{D}} = 2$. This improvement is attributed to the fact that sensor localization accuracy increases as either mapping accuracy or the number of detected targets increases. The performance gap between SBP-PMB-SLAM and VBP-SLAM1 arises from the usage of newly detected targets in sensor state updates.

VI. CONCLUSIONS

In this paper, we have developed the general framework of set-type BP, which can serve as a fundamental tool for computing either the marginal (or its approximate density) of an RFS. From the framework, we derived the PMB and MB

⁵This message comprises the sum of two parts: clutter and newly detected landmark.

filters with the developed set-type BP, applicable to the related problems of mapping, MTT, SLAM, and SLAT. Under the densities that follow the Bernoulli process, we demonstrated that vector-type BP is the special case of set-type BP. To handle the unknown set cardinality in the factor graph, we developed the set-factor nodes for set partitioning, set merging, and shifting set auxiliary variables. We applied the proposed set-type BP to PMB-SLAM filter and explored the relations between the set-type BP PMB and vector-type BP-SLAM [12], [13] filters. Our results demonstrated that the proposed set-type BP-SLAM filter (and the equivalent vector type BP algorithm) outperforms vector-type BP from [12], [13], under the informative PPP for undetected landmarks, and equivalent for the uninformative PPP. Vector-type BP can be readily modified, through various heuristics, to match the performance and operation of set-type BP.

Other applications where set-type BP can be applied include smoothing (obtained by computing messages backward in time), extended targets (where each target may generate a set of measurements), cooperative processing (obtained by linking the factor graphs related to two different sensors), general target-generated measurements and arbitrary clutter, and positioning (by adding factors related to fixed anchors in the environment). Applications of the set-type BP framework to other families of RFS filters, along with their BP counterparts, deserve further study. Finally, we hope that the proposed framework can find application in problems outside the sphere of mapping and tracking.

APPENDIX A

PROOF OF THEOREM 1

We prove Theorem 1 in a similar fashion as in the vector-type BP [3, Theorem 2].

Proof. Using the set beliefs of (20) and (21), we introduce the Bethe free energy [3]: $F_{\text{Bethe}} = U_{\text{Bethe}} - H_{\text{Bethe}}$, where U_{Bethe} is the Bethe average energy, given by⁶

$$U_{\text{Bethe}} = - \sum_a \int \mathbf{b}(\underline{\mathcal{X}}^a) \ln f_a(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a, \quad (63)$$

and H_{Bethe} is the Bethe entropy, given by

$$H_{\text{Bethe}} = - \sum_a \int \mathbf{b}(\underline{\mathcal{X}}^a) \ln \mathbf{b}(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a + \sum_i (|\mathcal{M}(i)| - 1) \int \mathbf{b}(\mathcal{X}^i) \ln \mathbf{b}(\mathcal{X}^i) \delta \mathcal{X}^i. \quad (64)$$

The constraints of the Bethe free energy, the average energy, and the entropy are all functions of set-beliefs. They are introduced as follows. The normalization constraints are $\int \mathbf{b}(\mathcal{X}^i) \delta \mathcal{X}^i = 1$ for all set-variable i , $\int \mathbf{b}(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a = 1$ for all set-factor a with $|\mathcal{M}(i)| \geq 2$. The marginalization constraints are $\int \mathbf{b}(\underline{\mathcal{X}}^a) \delta \mathcal{X}^{\sim i} = \mathbf{b}(\mathcal{X}^i)$ such that $i \in \mathcal{N}(a)$.

⁶The set integral such as $\int \mathbf{b}(\mathcal{X}) \ln \mathbf{b}(\mathcal{X}) \delta \mathcal{X}$ and $\int \mathbf{b}(\mathcal{X}) \ln f(\mathcal{X}) \delta \mathcal{X}$ requires the use of the measure-theoretic integral [51], due to the units of the standard set integral and densities [22, Sec. 3.2.4]. In this sense, the integral is then $\int \mathbf{b}(\mathcal{X}) \ln(K^{|\mathcal{X}|} \mathbf{b}(\mathcal{X})) \delta \mathcal{X}$ and $\int \mathbf{b}(\mathcal{X}) \ln(K^{|\mathcal{X}|} f(\mathcal{X})) \delta \mathcal{X}$, where K is the unit of the hypervolume of the single state $\mathbf{x} \in \mathcal{X}$. For notational simplicity, we omit the unit $K^{|\mathcal{X}|}$ in the integration.

The inequality constraints are $0 \leq \mathbf{b}(\underline{\mathcal{X}}^a) \leq 1$ for all set-factor a and $0 \leq \mathbf{b}(\mathcal{X}^i) \leq 1$ for all set-variable i . Due to the assumption of the interior stationary point, the inequality constraints will be inactive. Thus, we enforce the equality constraints with the Lagrange multipliers, denoted by γ_a , γ_i , and $\lambda_{a,i}(\mathcal{X}^i)$, respectively, and the Lagrangian is formulated as

$$\begin{aligned} L = & - \sum_a \int \mathbf{b}(\underline{\mathcal{X}}^a) \ln f_a(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a \\ & + \sum_a \int \mathbf{b}(\underline{\mathcal{X}}^a) \ln \mathbf{b}(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a \\ & - \sum_i (|\mathcal{M}(i)| - 1) \int \mathbf{b}(\mathcal{X}^i) \ln \mathbf{b}(\mathcal{X}^i) \delta \mathcal{X}^i \\ & + \sum_a \gamma_a \left(\int \mathbf{b}(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a - 1 \right) + \sum_i \gamma_i \left(\int \mathbf{b}(\mathcal{X}^i) \delta \mathcal{X}^i - 1 \right) \\ & + \sum_i \sum_{a \in \mathcal{M}(i)} \int \lambda_{a,i}(\mathcal{X}^i) (\mathbf{b}(\mathcal{X}^i) - \int \mathbf{b}(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a) \delta \mathcal{X}^i. \end{aligned} \quad (65)$$

By the derivatives of the Lagrangian with respect to $\mathbf{b}(\underline{\mathcal{X}}^a)$ and $\mathbf{b}(\mathcal{X}^i)$, we can obtain the interior stationary points as follows:

$$\hat{\mathbf{b}}(\mathcal{X}^i) = \frac{1}{Z_i} \exp \left(\frac{1}{|\mathcal{M}(i)| - 1} \sum_{a \in \mathcal{M}(i)} \lambda_{a,i}(\mathcal{X}^i) \right), \quad (66)$$

$$\hat{\mathbf{b}}(\underline{\mathcal{X}}^a) = \frac{1}{Z_a} f_a(\underline{\mathcal{X}}^a) \exp \left(\sum_{i \in \mathcal{N}(a)} \lambda_{a,i}(\mathcal{X}^i) \right), \quad (67)$$

where Z_a and Z_i are the normalization constants. Making the identification

$$\lambda_{a,i}(\mathcal{X}^i) = \ln n_{i \rightarrow a}(\mathcal{X}^i) = \ln \prod_{b \in \mathcal{M}(i) \setminus \{a\}} m_{b \rightarrow i}(\mathcal{X}^i), \quad (68)$$

and substitute (68) into (66) and (67), then we recover the set-type BP fixed points of (20) and (21) as follows:

$$\hat{\mathbf{b}}(\mathcal{X}^i) \propto \prod_{a \in \mathcal{M}(i)} m_{a \rightarrow i}(\mathcal{X}^i), \quad (69)$$

$$\hat{\mathbf{b}}(\underline{\mathcal{X}}^a) \propto f_a(\underline{\mathcal{X}}^a) \prod_{i \in \mathcal{N}(a)} \prod_{b \in \mathcal{M}(i) \setminus \{a\}} m_{b \rightarrow i}(\mathcal{X}^i). \quad (70)$$

To derive this theorem conversely, we introduce $m_{a \rightarrow i}$

$$m_{a \rightarrow i} = \exp \left(\frac{2 - |\mathcal{M}(i)|}{|\mathcal{M}(i)| - 1} \lambda_{a,i}(\mathcal{X}^i) + \frac{1}{|\mathcal{M}(i)| - 1} \lambda_{b,i}(\mathcal{X}^i) \right), \quad (71)$$

where $\lambda_{b,i}(\mathcal{X}^i) = \ln n_{i \rightarrow b}(\mathcal{X}^i)$, which can be obtained from (68) and set-message update rules of (18) and (19). Substituting $\prod_{b \in \mathcal{M}(i) \setminus \{a\}} m_{b \rightarrow i}$ of (68) and $m_{a \rightarrow i}$ of (71) into the set-type BP fixed points of (20) and (21), the reverse of this theorem can be shown.

We omitted a single variable that is only connected to a single factor (i.e., $|\mathcal{M}(i)| = 1$), called a dead-end variable, in the Lagrangian since the dead-end variable does not contribute to the Bethe free energy and the beliefs. The beliefs at dead-end variables are not required but it can be easily computed from the belief $\int \mathbf{b}(\underline{\mathcal{X}}^a) \delta \underline{\mathcal{X}}^a$ as needed. \square

APPENDIX B PROOF OF COROLLARY 1

We prove Corollary 1 by showing that the set-belief $\mathbf{b}(\mathcal{X}^i)$ obtained after running set-type BP in a factor graph with no cycles corresponds to the marginal probability density of \mathcal{X}^i .

Proof. Note that $m_{a \rightarrow i}(\mathcal{X}^i) = f_a(\underline{\mathcal{X}}^a)$ if $|\mathcal{N}(a)| = 1$. By the chain rule with (18) and (19), the set-belief $\mathbf{b}(\mathcal{X}^i)$ of (18) is represented by the integration of the product of all set-factors $f_a(\underline{\mathcal{X}}^a)$. It corresponds to $\int f(\mathcal{X}^1, \dots, \mathcal{X}^n) \delta \mathcal{X}^{\sim i}$ of (16). \square

APPENDIX C PROOF OF PROPOSITION 1

We prove Proposition 1 with the partitioning and merging factor of Definition 3.

Proof. Using the partitioning and merging factor of Definition 3, we can then partition the Poisson set \mathcal{X}^j with the incoming Poisson message $\mathbf{n}_{j \rightarrow a}(\mathcal{X}^j)$ into $|\mathcal{N}(a)| - 1$ Poisson sets \mathcal{X}^i with the Poisson outgoing messages $m_{a \rightarrow i}(\mathcal{X}^i)$ for $i \in \mathcal{N}(a) \setminus \{j\}$. Suppose we have incoming messages $\mathbf{n}_{j \rightarrow a}(\mathcal{X}^j) = f^{\text{PPP}}(\mathcal{X}^j)$ that follows the PPP density, $\mathbf{n}_{q \rightarrow a}(\mathcal{X}^q) = 1$, for $q \in \mathcal{N}(a) \setminus \{j, i\}$. From (26), the partitioning messages from $f_a(\underline{\mathcal{X}}^a)$ to \mathcal{X}^i with $i \in \mathcal{N}(a) \setminus \{j\}$ are

$$m_{a \rightarrow i}(\mathcal{X}^i) = \int \mathbf{n}_{j \rightarrow a}(\uplus_{q \in \mathcal{N}(a) \setminus \{j\}} \mathcal{X}^q) \delta \mathcal{X}^{\sim i} \quad (72)$$

$$= \int \mathbf{n}_{j \rightarrow a}(\mathcal{X}^i \uplus \mathcal{X}) \delta \mathcal{X} \quad (73)$$

$$\propto f^{\text{PPP}}(\mathcal{X}^i), \quad (74)$$

the same PPP intensity function. It indicates that the Poisson set is partitioned into multiple Poisson sets with the same PPP intensity function, which is derived in Appendix D.

Conversely, the $|\mathcal{N}(a)| - 1$ Poisson sets \mathcal{X}^i with the incoming Poisson messages $\mathbf{n}_{i \rightarrow a}(\mathcal{X}^i)$ for $i \in \mathcal{N}(a) \setminus \{j\}$ are merged into a single Poisson set \mathcal{X}^j with the Poisson outgoing message $m_{a \rightarrow j}(\mathcal{X}^j)$. Suppose we have incoming messages $\mathbf{n}_{i \rightarrow a}(\mathcal{X}^i) = f^{\text{PPP}}(\mathcal{X}^i)$, for $i \in \mathcal{N}(a) \setminus \{j\}$ that follow PPP densities. The message of (30) is

$$m_{a \rightarrow j}(\mathcal{X}^j) = \sum_{\uplus_{i \in \mathcal{N}(a) \setminus \{j\}} \mathcal{W}^i = \mathcal{X}^j} \prod_{i \in \mathcal{N}(a) \setminus \{j\}} f^{\text{PPP}}(\mathcal{W}^i). \quad (75)$$

It indicates that the outgoing message is the convolution of all incoming PPP messages, representing the union of multiple Poisson sets. \square

APPENDIX D PROOF OF POISSON SET PARTITION

We find that $f^{\text{PPP}}(\mathcal{X}) = e^{-\bar{\lambda}} \prod_{\mathbf{x} \in \mathcal{X}} \lambda(\mathbf{x})$ and $\bar{\lambda} = \int \lambda(\mathbf{x}) d\mathbf{x}$ from (9), and then (72) is expressed as

$$m_{a \rightarrow i}(\mathcal{X}^i) \quad (76)$$

$$= \int e^{-\bar{\lambda}} \prod_{q \in \mathcal{N}(a) \setminus \{j\}} \prod_{\mathbf{x} \in \mathcal{X}^q} \lambda(\mathbf{x}) \delta \mathcal{X}^{\sim i} \quad (77)$$

$$\propto e^{-\bar{\lambda}} \prod_{\mathbf{x} \in \mathcal{X}^i} \lambda(\mathbf{x}) \prod_{q \in \mathcal{N}(a) \setminus \{j\}} \int e^{-\bar{\lambda}} \prod_{\mathbf{x} \in \mathcal{X}^q} \lambda(\mathbf{x}) \delta \mathcal{X}^{\sim (i,j)} \quad (78)$$

$$= e^{-\bar{\lambda}} \prod_{\mathbf{x} \in \mathcal{X}^i} \lambda(\mathbf{x}) = f^{\text{PPP}}(\mathcal{X}^i). \quad (79)$$

Finally, we find that $m_{a \rightarrow i}(\mathcal{X}^i) \propto f^{\text{PPP}}(\mathcal{X}^i)$.

APPENDIX E JOINT UPDATE DENSITY

We prove the joint update density of (46) with introducing $\mathcal{X}_k^1 \uplus \dots \uplus \mathcal{X}_k^{I_{k-1}} = \bar{\mathcal{X}}_k^{I_{k-1}}$, $\mathcal{Z}_k^1 \uplus \dots \uplus \mathcal{Z}_k^{I_{k-1}} = \bar{\mathcal{Z}}_k^{I_{k-1}}$, and $\mathcal{Y}_k^1 \uplus \dots \uplus \mathcal{Y}_k^{J_k} = \bar{\mathcal{Y}}_k^{J_k}$. We start with the prior at time step k (without auxiliary variables) such that

$$\begin{aligned} f_{\text{p}}(\mathbf{s}_k, \mathcal{X}_k) &= f_{\text{p}}(\mathbf{s}_k) f_{\text{p}}(\mathcal{X}_k) \\ &= f_{\text{p}}(\mathbf{s}_k) \sum_{\mathcal{P}_k^{\text{U}} \uplus \bar{\mathcal{X}}_k^{I_{k-1}} = \mathcal{X}_k} f_{\text{p}}^{\text{U}}(\mathcal{P}_k^{\text{U}}) \prod_{i=1}^{I_{k-1}} f_{\text{p}}^i(\mathcal{X}_k^i), \end{aligned} \quad (80)$$

where the prior assumes that the sensor state and the set of targets are independent. The set of measurements received at time step k is $\mathcal{Z}_k = \{z_k^1, \dots, z_k^{J_k}\}$.

A. Likelihood

For any sets $\mathcal{P}_k^{\text{U}}, \mathcal{X}_k^1, \dots, \mathcal{X}_k^{I_{k-1}}$ such that $|\mathcal{X}_k^i| \leq 1$ for $i = 1, \dots, I_{k-1}$, we introduce the likelihood functions [27, Eq. (25)]

$$\begin{aligned} l_o(\mathcal{Z}_k | \mathcal{P}_k^{\text{U}}, \mathcal{X}_k^1, \dots, \mathcal{X}_k^{I_{k-1}}, \mathbf{s}_k) \\ = \sum_{\mathcal{Z}_k^{\text{U}} \uplus \bar{\mathcal{Z}}_k^{I_{k-1}} = \mathcal{Z}_k} l(\mathcal{Z}_k^{\text{U}} | \mathcal{P}_k^{\text{U}}, \mathbf{s}_k) \prod_{i=1}^{I_{k-1}} t(\mathcal{Z}_k^i | \mathcal{X}_k^i, \mathbf{s}_k), \end{aligned} \quad (82)$$

where \mathcal{Z}_k^{U} represents a measurement set including measurement elements that are generated from both targets in \mathcal{X}_k^{U} and clutter, and $t(\mathcal{Z}_k^i | \mathcal{X}_k^i)$ is the likelihood for a set with zero or one measurement element without clutter, given by

$$\begin{aligned} t(\mathcal{Z}_k^i | \mathcal{X}_k^i, \mathbf{s}_k) \\ = \begin{cases} p_{\text{D}}(\mathbf{x}_k, \mathbf{s}_k) l(\mathbf{z}_k | \mathbf{x}_k, \mathbf{s}_k), & \mathcal{Z}_k^i = \{\mathbf{z}_k\}, \mathcal{X}_k^i = \{\mathbf{x}_k\} \\ 1 - p_{\text{D}}(\mathbf{x}_k, \mathbf{s}_k), & \mathcal{Z}_k^i = \emptyset, \mathcal{X}_k^i = \{\mathbf{x}_k\} \\ 1, & \mathcal{Z}_k^i = \emptyset, \mathcal{X}_k^i = \emptyset \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (83)$$

In addition, we know that, for $\mathcal{P}_k^{\text{U}} = \mathcal{X}_k^{\text{U}} \uplus \mathcal{Y}_k^1 \uplus \dots \uplus \mathcal{Y}_k^{J_k}$, the multi-target likelihood meets

$$l(\mathcal{Z}_k | \mathcal{P}_k^{\text{U}}, \mathbf{s}_k) = l_o(\mathcal{Z}_k | \mathcal{X}_k^{\text{U}}, \mathcal{Y}_k^1, \dots, \mathcal{Y}_k^{J_k}, \mathbf{s}_k). \quad (84)$$

B. Posterior

From [27, Eq. (34)], the posterior density is then

$$f_{\text{u}}(\mathbf{s}_k, \mathcal{X}_k) \propto l(\mathcal{Z}_k | \mathcal{X}_k, \mathbf{s}_k) f_{\text{p}}(\mathbf{s}_k, \mathcal{X}_k) \quad (85)$$

$$\begin{aligned} &= f_{\text{p}}(\mathbf{s}_k) \sum_{\mathcal{Z}_k^{\text{U}} \uplus \bar{\mathcal{Z}}_k^{I_{k-1}} = \mathcal{Z}_k} \sum_{\mathcal{P}_k^{\text{U}} \uplus \bar{\mathcal{X}}_k^{I_{k-1}} = \mathcal{X}_k} l(\mathcal{Z}_k^{\text{U}} | \mathcal{P}_k^{\text{U}}, \mathbf{s}_k) f_{\text{p}}^{\text{U}}(\mathcal{P}_k^{\text{U}}) \\ &\times \prod_{i=1}^{I_{k-1}} t(\mathcal{Z}_k^i | \mathcal{X}_k^i, \mathbf{s}_k) f_{\text{p}}^i(\mathcal{X}_k^i) \end{aligned} \quad (86)$$

$$\begin{aligned} &= f_{\text{p}}(\mathbf{s}_k) \sum_{\mathcal{Z}_k^{\text{U}} \uplus \bar{\mathcal{Z}}_k^{I_{k-1}} = \mathcal{Z}_k} \sum_{\mathcal{P}_k^{\text{U}} \uplus \bar{\mathcal{X}}_k^{I_{k-1}} = \mathcal{X}_k} \sum_{\mathcal{X}_k^{\text{U}} \uplus \bar{\mathcal{Y}}_k^{J_k} = \mathcal{P}_k^{\text{U}}} \\ &\times f_{\text{p}}^{\text{U}}(\mathcal{X}_k^{\text{U}}) [1 - p_{\text{D}}(\cdot, \mathbf{s}_k)]^{\mathcal{X}_k^{\text{U}}} \\ &\times \prod_{j=1}^{J_k} [\chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j) \tilde{l}(\mathbf{z}_k^j | \mathcal{Y}_k^j, \mathbf{s}_k) f_{\text{p}}^{\text{U}}(\mathcal{Y}_k^j) \\ &+ (1 - \chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j)) \delta_{\emptyset}(\mathcal{Y}_k^j)] \prod_{i=1}^{I_{k-1}} t(\mathcal{Z}_k^i | \mathcal{X}_k^i, \mathbf{s}_k) f_{\text{p}}^i(\mathcal{X}_k^i), \end{aligned} \quad (87)$$

where $\chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j)$ is defined to be 1 if $\mathbf{z}_k^j \in \mathcal{Z}_k^{\text{U}}$, and to be 0 otherwise. Combining the last two convolution sums into one, we have

$$\begin{aligned} f_{\text{u}}(\mathbf{s}_k, \mathcal{X}_k) \\ \propto f_{\text{p}}(\mathbf{s}_k) \sum_{\mathcal{Z}_k^{\text{U}} \uplus \bar{\mathcal{Z}}_k^{I_{k-1}} = \mathcal{Z}_k} \sum_{\mathcal{X}_k^{\text{U}} \uplus \bar{\mathcal{Y}}_k^{J_k} \uplus \bar{\mathcal{X}}_k^{I_{k-1}} = \mathcal{X}_k} f_{\text{p}}^{\text{U}}(\mathcal{X}_k^{\text{U}}) \\ \times [1 - p_{\text{D}}(\cdot, \mathbf{s}_k)]^{\mathcal{X}_k^{\text{U}}} \prod_{j=1}^{J_k} [\chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j) \tilde{l}(\mathbf{z}_k^j | \mathcal{Y}_k^j, \mathbf{s}_k) f_{\text{p}}^{\text{U}}(\mathcal{Y}_k^j) \\ + (1 - \chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j)) \delta_{\emptyset}(\mathcal{Y}_k^j)] \prod_{i=1}^{I_{k-1}} t(\mathcal{Z}_k^i | \mathcal{X}_k^i, \mathbf{s}_k) f_{\text{p}}^i(\mathcal{X}_k^i). \end{aligned} \quad (88)$$

We can now add auxiliary variables for the previous targets $\mathcal{X}_k^1, \dots, \mathcal{X}_k^{I_{k-1}}$ and also for the newly detected targets $\mathcal{Y}_k^1, \dots, \mathcal{Y}_k^{J_k}$. The auxiliary variables for the newly detected targets will start from $I_{k-1} + 1$. Then, we have

$$\begin{aligned} \tilde{f}_{\text{u}}(\mathbf{s}_k, \tilde{\mathcal{X}}_k) \propto f_{\text{p}}(\mathbf{s}) \sum_{\mathcal{Z}_k^{\text{U}} \uplus \bar{\mathcal{Z}}_k^{I_{k-1}} = \mathcal{Z}_k} \tilde{f}_{\text{p}}^{\text{U}}(\tilde{\mathcal{X}}_k^{\text{U}}) [1 - p_{\text{D}}(\cdot, \mathbf{s}_k)]^{\tilde{\mathcal{X}}_k^{\text{U}}} \\ \times \prod_{j=1}^{J_k} [\chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j) \tilde{l}(\mathbf{z}_k^j | \tilde{\mathcal{Y}}_k^j, \mathbf{s}_k) \tilde{f}_{\text{p}}^{\text{U}}(\tilde{\mathcal{Y}}_k^j) \\ + (1 - \chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j)) \delta_{\emptyset}(\tilde{\mathcal{Y}}_k^j)] \\ \times \prod_{i=1}^{I_{k-1}} t(\mathcal{Z}_k^i | \tilde{\mathcal{X}}_k^i, \mathbf{s}_k) \tilde{f}_{\text{p}}^i(\tilde{\mathcal{X}}_k^i). \end{aligned} \quad (89)$$

By jointly considering the data associations \mathbf{c}_k and \mathbf{d}_k introduced in (46b), we can introduce the measurement sets as follows:

$$\mathcal{Z}_k^i = \begin{cases} \{\mathbf{z}_k^{c_k^i}\} & c_k^i > 0 \\ \emptyset & c_k^i = 0 \end{cases}, \quad (90)$$

$$\mathcal{Z}_k^{\text{U}} = \{\mathbf{z}_k^j : d_k^j = 0\}. \quad (91)$$

With the association function $\psi(\mathbf{c}_k^i, \mathbf{d}_k^j)$ introduced in (46b), we can rewrite (89) as

$$\begin{aligned} \tilde{f}_{\text{u}}(\mathbf{s}_k, \tilde{\mathcal{X}}_k) \propto f_{\text{p}}(\mathbf{s}) \sum_{\mathbf{c}_k, \mathbf{d}_k} \tilde{f}_{\text{p}}^{\text{U}}(\tilde{\mathcal{X}}_k^{\text{U}}) [1 - p_{\text{D}}(\cdot, \mathbf{s}_k)]^{\tilde{\mathcal{X}}_k^{\text{U}}} \\ \times \prod_{j=1}^{J_k} [\chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j) \tilde{l}(\mathbf{z}_k^j | \tilde{\mathcal{Y}}_k^j, \mathbf{s}_k, \mathbf{d}_k^j) \tilde{f}_{\text{p}}^{\text{U}}(\tilde{\mathcal{Y}}_k^j) \\ + (1 - \chi_{\mathcal{Z}_k^{\text{U}}}(\mathbf{z}_k^j)) \delta_{\emptyset}(\tilde{\mathcal{Y}}_k^j)] \end{aligned} \quad (92)$$

$$\times \prod_{i=1}^{I_{k-1}} t(\mathbf{z}_k^i | \tilde{\mathcal{X}}_k^i, \mathbf{s}_k, c_k^i) \tilde{f}_p^i(\tilde{\mathcal{X}}_k^i),$$

where $\tilde{l}(\mathbf{z}_k^j | \tilde{\mathcal{Y}}_k^j, \mathbf{s}_k, d_k^j)$ and $t(\mathbf{z}_k^i | \tilde{\mathcal{X}}_k^i, \mathbf{s}_k, c_k^i)$ were defined in (50) and (51), respectively. We now make association variables \mathbf{c}_k and \mathbf{d}_k explicit in the posterior to define the density. We can then define the joint update density as

$$\begin{aligned} \tilde{f}_u(\mathbf{s}_k, \tilde{\mathcal{X}}_k, \mathbf{c}_k, \mathbf{d}_k) &\propto f_p(\mathbf{s}_k) \tilde{f}_p^U(\tilde{\mathcal{X}}_k^U) [1 - p_D(\cdot, \mathbf{s}_k)]^{\tilde{\mathcal{X}}_k^U} \quad (93) \\ &\times \prod_{j=1}^{J_k} \tilde{l}(\mathbf{z}_k^j | \tilde{\mathcal{Y}}_k^j, \mathbf{s}_k, d_k^j) \tilde{f}_p^U(\tilde{\mathcal{Y}}_k^j) \\ &\times \prod_{i=1}^{I_{k-1}} t(\mathbf{z}_k^i | \tilde{\mathcal{X}}_k^i, \mathbf{s}_k, c_k^i) \tilde{f}_p^i(\tilde{\mathcal{X}}_k^i) \psi(c_k^i, d_k^j), \end{aligned}$$

REFERENCES

- [1] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [2] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.
- [3] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.
- [4] H. Wymeersch, *Iterative Receiver Design*. Cambridge University Press, 2007.
- [5] A. Ihler, J. Fisher, R. Moses, and A. Willsky, "Nonparametric belief propagation for sensor network self-calibration," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 809–819, Apr. 2005.
- [6] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, vol. 97, no. 2, pp. 427–450, Feb. 2009.
- [7] C. Rhodes, C. Liu, and W.-H. Chen, "Scalable probabilistic gas distribution mapping using gaussian belief propagation," in *2022 IEEE/RSJ Int. Conf. Intell. Robots and Syst. (IROS)*, Kyoto, Japan, Oct. 2022, pp. 9459–9466.
- [8] F. Meyer, P. Braca, P. Willett, and F. Hlawatsch, "A scalable algorithm for tracking an unknown number of targets using multiple sensors," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3478–3493, Jul. 2017.
- [9] F. Meyer, T. Kropfreiter, J. L. Williams, R. Lau, F. Hlawatsch, P. Braca, and M. Z. Win, "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.
- [10] H. Wymeersch, N. Garcia, H. Kim, G. Seco-Granados, S. Kim, F. Wen, and M. Fröhle, "5G mmWave downlink vehicular positioning," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 206–212.
- [11] H. Kim, H. Wymeersch, N. Garcia, G. Seco-Granados, and S. Kim, "5G mmWave vehicular tracking," in *Proc. IEEE 52nd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2018, pp. 541–547.
- [12] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, "A belief propagation algorithm for multipath-based SLAM," *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 5613–5629, Sep. 2019.
- [13] E. Leitinger, S. Grebien, and K. Witrisal, "Multipath-based SLAM exploiting AoA and amplitude information," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Shanghai, China, May 2019.
- [14] E. Leitinger, A. Venus, B. Teague, and F. Meyer, "Data fusion for multipath-based slam: Combining information from multiple propagation paths," *IEEE Trans. Signal Process.*, vol. 71, Sep. 2023.
- [15] F. Meyer, O. Hlinka, H. Wymeersch, E. Riegler, and F. Hlawatsch, "Distributed localization and tracking of mobile networks including noncooperative objects," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 2, no. 1, pp. 57–71, Mar. 2015.
- [16] A. Stentz, D. Fox, and M. Montemerlo, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," in *Proc. Nat. Conf. Artif. Intell.*, 2002, pp. 593–598.
- [17] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A tutorial on graph-based SLAM," *IEEE Intell. Transp. Syst. Mag.*, vol. 2, no. 4, pp. 31–43, 2010.
- [18] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robot. Autom. Mag.*, vol. 13, no. 2, pp. 99–110, Jun. 2006.
- [19] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *IEEE Robot. Autom. Mag.*, vol. 13, no. 3, pp. 108–117, Sep. 2006.
- [20] R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *Int. J. Robot. Res.*, vol. 5, no. 4, pp. 56–68, 1986.
- [21] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.
- [22] —, *Advances Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.
- [23] —, "Multitarget Bayes filtering via first-order multi target moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.
- [24] H. Kim, K. Granström, L. Gao, G. Battistelli, S. Kim, and H. Wymeersch, "5G mmWave cooperative positioning and mapping using multi-model PHD," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3782–3795, Mar. 2020.
- [25] A. F. Garcia-Fernandez and B.-N. Vo, "Derivation of the PHD and CPHD filters based on direct Kullback-Leibler divergence minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 21, pp. 5812–5820, Nov. 2015.
- [26] J. L. Williams, "Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA, and association-based MeMBer," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 1664–1687, Jul. 2015.
- [27] Á. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: Direct derivation and implementation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1883–1901, Aug. 2018.
- [28] H. Kim, K. Granström, L. Svensson, S. Kim, and H. Wymeersch, "PMBM-based SLAM filters in 5G mmWave vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8646–8661, Aug. 2022.
- [29] M. Fröhle, C. Lindberg, K. Granström, and H. Wymeersch, "Multisensor Poisson multi-Bernoulli filter for joint target–sensor state tracking," *IEEE Trans. Int. Veh.*, vol. 4, no. 4, pp. 609–621, Dec. 2019.
- [30] D. Moratuwage, M. Adams, and F. Inostroza, "δ-generalised labelled multi-Bernoulli simultaneous localisation and mapping," in *Int. Conf. Control, Autom. Inf. Sci. (ICCAIS)*, Pyeong Chang, South Korea, Oct. 2018, pp. 175–182.
- [31] —, "δ-generalized labeled multi-Bernoulli simultaneous localization and mapping with an optimal kernel-based particle filtering approach," *Sensors*, vol. 19, no. 10, May 2019.
- [32] H. Deusch, "Random finite set-based localization and SLAM for highly automated vehicles," Ph.D. dissertation, driveU/Institute of Measurement, Control and Microtechnology, Universität Ulm, Germany, 2015.
- [33] H. Deusch, S. Reuter, and K. Dietmayer, "The labeled multi-Bernoulli SLAM filter," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1561–1565, Oct. 2015.
- [34] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3460–3475, Jul. 2013.
- [35] Y. Ge, O. Kaitiokallio, 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 J. Sel. Areas Commun.*, Jul. 2022.
- [36] P. Horridge and S. Maskell, "Using a probabilistic hypothesis density filter to confirm tracks in a multi-target environment," in *Proc. Workshop on Sensor Data Fusion*, Berlin, Germany, Oct. 2011, pp. 1103–1110.
- [37] J. L. Williams, "Hybrid Poisson and multi-Bernoulli filters," in *Proc. Int. Conf. Inf. Fusion*, Singapore, Jul. 2012, pp. 1103–1110.
- [38] 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, Jul. 2014.
- [39] 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 Trans. Aerosp. Electron. Syst.*, 2023.
- [40] K. G. Murthy, "An algorithm for ranking all the assignments in order of increasing costs," *Oper. Res.*, vol. 16, no. 3, pp. 682–687, 1968.
- [41] M. A. Armstrong, *Basic topology*. Springer Science & Business Media, 2013.
- [42] Á. F. García-Fernández, L. Svensson, J. L. Williams, Y. Xia, and K. Granström, "Trajectory Poisson multi-Bernoulli filters," *IEEE Trans. Signal Process.*, vol. 68, pp. 4933–4945, Aug. 2020.
- [43] Á. F. García-Fernández, L. Svensson, and M. R. Morelande, "Multiple target tracking based on sets of trajectories," *IEEE Trans. Aerosp. and Electron. Syst.*, vol. 56, no. 3, pp. 1685–1707, Jun. 2019.

- [44] Á. F. García-Fernández, Y. Xia, K. Granström, L. Svensson, and J. L. Williams, "Gaussian implementation of the multi-Bernoulli mixture filter," in *22th Int. Conf. Inf. Fusion (FUSION)*, Ottawa, Canada, Jul. 2019, pp. 1–8.
- [45] Á. F. García-Fernández, Y. Xia, and L. Svensson, "Poisson multi-Bernoulli mixture filter with general target-generated measurements and arbitrary clutter," *IEEE Trans. Signal Process.*, vol. 71, pp. 1895–1906, 2023.
- [46] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Hoboken, NJ, USA: Wiley, 2002.
- [47] K. Panta, B.-N. Vo, and S. Singh, "Novel data association schemes for the probability hypothesis density filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 2, pp. 556–570, Apr. 2007.
- [48] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.
- [49] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
- [50] A. S. Rahmathullah, A. F. García Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Proc. 20th Int. Conf. Inf. Fusion (FUSION)*, Xian, China, Jul. 2017, pp. 1–8.
- [51] H. G. Hoang, B.-N. Vo, B.-T. Vo, and R. Mahler, "The Cauchy-Schwarz divergence for Poisson point processes," *IEEE Trans. Inf. Theory*, vol. 61, no. 8, pp. 4475–4485, Aug. 2015.



Hyowon Kim received the Ph.D. degree from the Department of Electronic Engineering, Hanyang University, Seoul, Korea in 2021. He is currently an Assistant Professor in the Department of Electronics Engineering at Chungnam National University, Daejeon, Korea. He was a Marie Skłodowska-Curie Fellow/Postdoctoral Researcher in the Department of Electrical Engineering at Chalmers University of Technology, Sweden, from 2021 to 2023. He was a Visiting Researcher with the Department of Electrical Engineering, Chalmers University of Technology,

Sweden, from 2019 to 2020. His main research interests include wireless communications and integrated sensing/localization/communications, in 5G and Beyond 5G communication systems.



Ángel F. García-Fernández received the telecommunication engineering degree and the Ph.D. degree from Universidad Politécnica de Madrid, Madrid, Spain, in 2007 and 2011, respectively. He is currently a Senior Lecturer in the Department of Electrical Engineering and Electronics at the University of Liverpool, Liverpool, UK. He previously held postdoctoral positions at Universidad Politécnica de Madrid, Chalmers University of Technology, Gothenburg, Sweden, Curtin University, Perth, Australia, and Aalto University, Espoo, Finland. His main research activities and interests are in the area of Bayesian inference, with emphasis on dynamic systems and multiple target tracking.



Yu Ge (S'20) received his B.E. degree from Zhejiang University, Hangzhou, China, in 2017, and M.Sc. degree from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2019. He is currently a Ph.D. candidate in the Department of Electrical and Engineering at Chalmers University of Technology, Sweden. His research interests include integrated communication and sensing, wireless positioning systems, simultaneous localization and mapping, and multi-object tracking, particularly in 5G and Beyond 5G scenarios.



multi-object tracking and SLAM, especially for automotive applications.

Yuxuan Xia received his M.Sc. degree in communication engineering and Ph.D. degree in signal processing both from Chalmers University of Technology, Gothenburg, Sweden, in 2017 and 2022, respectively. After obtaining his Ph.D. degree, he stayed at the same research group as a postdoctoral researcher for a year. He is currently an industrial postdoctoral researcher with Zenseact AB, Gothenburg, Sweden and also affiliated with the Division of Automatic Control, Linköping University, Linköping, Sweden. His main research interests include sensor fusion,



Lennart Svensson (M'09-SM'10) is a Professor of Signal Processing with the Chalmers University of Technology. His main research interests include machine learning and Bayesian inference in general, and nonlinear filtering, deep learning, and tracking in particular. He has organized a massive open online course on multiple object tracking, available on edX and YouTube, and received paper awards at the International Conference on Information Fusion in 2009, 2010, 2017, and 2019.



Henk Wymeersch (S'01, M'05, SM'19, F'24) obtained the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium. He is currently a Professor of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Prof. Wymeersch served as Associate Editor for *IEEE Communication Letters* (2009-2013), *IEEE Transactions on Wireless Communications* (since 2013), and *IEEE Transactions on Communications* (2016-2018) and is currently Senior Member of the *IEEE Signal Processing Magazine* Editorial Board. During 2019-2021, he was an *IEEE Distinguished Lecturer* with the Vehicular Technology Society. His current research interests include the convergence of communication and sensing, in a 5G and Beyond 5G context.

Set-Type Belief Propagation with Applications to Poisson Multi-Bernoulli SLAM: Supplementary Material

APPENDIX F

SET-TYPE BP MESSAGES AND BELIEFS

The messages and beliefs on the factor graph (Fig. 5) are computed as follows.

A. Prediction

The prediction step is cascaded to the updated step of the previous time step, where we have a sensor state vector and $I_{k-1} + 1$ sets: 1 for undetected targets that have never been detected; I_{k-1} for detected targets that have been previously detected.

1) *Sensor*: We compute the message of the predicted density for the sensor state. The belief of the sensor variable \mathbf{s}_{k-1} is denoted by $\mathbf{b}(\mathbf{s}_{k-1})$, obtained from message [27] at the previous time step.

- [1] A message from the the sensor variable \mathbf{s}_{k-1} to linked factor A , i.e., $f^A(\mathbf{s}_k, \mathbf{s}_{k-1}) = f(\mathbf{s}_k | \mathbf{s}_{k-1})$, is denoted by $\mathbf{n}_{\mathbf{s}_{k-1} \rightarrow A}(\mathbf{s}_{k-1})$, and $\mathbf{n}_{\mathbf{s}_{k-1} \rightarrow A}(\mathbf{s}_{k-1}) = \mathbf{b}(\mathbf{s}_{k-1})$.
- [2] A message from factor A to the variable \mathbf{s}_k is denoted by $\mathbf{m}_{A \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$, given by

$$\mathbf{m}_{A \rightarrow \mathbf{s}_k}(\mathbf{s}_k) = \int \mathbf{n}_{\mathbf{s}_{k-1} \rightarrow A}(\mathbf{s}_{k-1}) f(\mathbf{s}_k | \mathbf{s}_{k-1}) d\mathbf{s}_{k-1}. \quad (\text{S1})$$

2) *Undetected Targets*: The undetected targets will either remain undetected again or be detected for the first time. Thus, we compute the messages of the predicted densities for the undetected targets and newly detected targets. A belief of undetected target set \mathcal{X}_{k-1}^U is denoted by $\mathbf{b}(\mathcal{X}_{k-1}^U)$, obtained from message [22] at the previous time step.

- [3] A message from the set-variable \mathcal{X}_{k-1}^U to the linked factor B , i.e., $f^B(\mathcal{X}_k^S, \mathcal{X}_{k-1}^U) = f(\mathcal{X}_k^S | \mathcal{X}_{k-1}^U)$, is denoted by $\mathbf{n}_{\mathcal{X}_{k-1}^U \rightarrow B}(\mathcal{X}_{k-1}^U) = \mathbf{b}(\mathcal{X}_{k-1}^U)$ that follows the PPP distribution.
- [4] A message from the factor B to \mathcal{X}_k^S is denoted by $\mathbf{m}_{B \rightarrow \mathcal{X}_k^S}(\mathcal{X}_k^S)$, given by

$$\mathbf{m}_{B \rightarrow \mathcal{X}_k^S}(\mathcal{X}_k^S) = \int \mathbf{n}_{\mathcal{X}_{k-1}^U \rightarrow B}(\mathcal{X}_{k-1}^U) f(\mathcal{X}_k^S | \mathcal{X}_{k-1}^U) \delta \mathcal{X}_{k-1}^U, \quad (\text{S2})$$

which follows the PPP distribution.

- [5] Messages from \mathcal{X}_k^S and \mathcal{X}_k^B to the linked factor D , i.e., $f^D(\mathcal{X}_k^S, \mathcal{X}_k^B, \mathcal{P}_k^U) = \delta_{\mathcal{X}_k^S \cup \mathcal{X}_k^B}(\mathcal{P}_k^U)$, are denoted by $\mathbf{n}_{\mathcal{X}_k^S \rightarrow D}(\mathcal{X}_k^S)$, respectively, and $\mathbf{n}_{\mathcal{X}_k^B \rightarrow D}(\mathcal{X}_k^B)$, where $\mathbf{n}_{\mathcal{X}_k^B \rightarrow D}(\mathcal{X}_k^B) = f^C(\mathcal{X}_k^B) = f^{\text{Poi}}(\mathcal{X}_k^B)$ follows the PPP distribution.
- [6] A message from the factor D to \mathcal{P}_k^U is denoted by $\mathbf{m}_{D \rightarrow \mathcal{P}_k^U}(\mathcal{P}_k^U)$, given by

$$\mathbf{m}_{D \rightarrow \mathcal{P}_k^U}(\mathcal{P}_k^U) = \iint \mathbf{n}_{\mathcal{P}_k^S \rightarrow D}(\mathcal{X}_k^S) \mathbf{n}_{\mathcal{X}_k^B \rightarrow D}(\mathcal{X}_k^B) \times \delta_{\mathcal{X}_k^S \cup \mathcal{X}_k^B}(\mathcal{P}_k^U) \delta \mathcal{X}_k^S \delta \mathcal{X}_k^B, \quad (\text{S3})$$

which follows the PPP distribution.

3) *Previously Detected Targets*: We compute the messages of the predicted densities for the previously detected targets. Beliefs of previously detected target set-variables \mathcal{X}_{k-1}^i are denoted by $\mathbf{b}(\mathcal{X}_{k-1}^i)$, for $i = \{1, \dots, I_{k-1}\}$, obtained from message [20] at the previous time step.

- [7] Messages from the set-variable \mathcal{X}_{k-1}^i to the the linked factors E^i , i.e., $f_i^E(\mathcal{X}_k^i, \mathcal{X}_{k-1}^i) = f(\mathcal{X}_k^i | \mathcal{X}_{k-1}^i)$, are denoted by $\mathbf{n}_{\mathcal{X}_{k-1}^i \rightarrow E^i}(\mathcal{X}_{k-1}^i)$, and $\mathbf{n}_{\mathcal{X}_{k-1}^i \rightarrow E^i}(\mathcal{X}_{k-1}^i) = \mathbf{b}(\mathcal{X}_{k-1}^i)$ for $i \in \mathcal{I}_{k-1}$.
- [8] Messages from factors E^i to the linked detected set-variables \mathcal{X}_k^i are denoted by $\mathbf{m}_{E^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i)$, given by

$$\mathbf{m}_{E^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i) = \int \mathbf{n}_{\mathcal{X}_{k-1}^i \rightarrow E^i}(\mathcal{X}_{k-1}^i) f(\mathcal{X}_k^i | \mathcal{X}_{k-1}^i) \delta \mathcal{X}_{k-1}^i \quad (\text{S4})$$

B. Update

The update step is cascaded to the prediction step from which the messages [2], [6], and [8] are obtained.

1) *Separation of Undetected Targets*: The set of undetected targets is partitioned into 1 set of targets that remains undetected and J_k sets of newly detected targets.

- [9] A message from the set-variable \mathcal{P}_k^U to factor F , i.e., $f^F(\mathcal{P}_k^U, \mathcal{X}_k^U, \mathcal{P}_k^1, \dots, \mathcal{P}_k^{J_k}) = \delta_{\mathcal{X}_k^U \cup \mathcal{P}_k}(\mathcal{P}_k^U)$, is denoted by $\mathbf{n}_{\mathcal{P}_k^U \rightarrow F}(\mathcal{P}_k^U)$, where $\mathbf{n}_{\mathcal{P}_k^U \rightarrow F}(\mathcal{P}_k^U) = \mathbf{m}_{D \rightarrow \mathcal{P}_k^U}(\mathcal{P}_k^U)$. Messages from the sets $\mathcal{X}_k^U, \mathcal{P}_k^1, \dots, \mathcal{P}_k^{J_k}$ to factor F are $\mathbf{n}_{\mathcal{X}_k^U \rightarrow F}(\mathcal{X}_k^U) = 1, \mathbf{n}_{\mathcal{P}_k^j \rightarrow F}(\mathcal{P}_k^j) = 1$, for $j \in \{1, \dots, J_k\}$.
- [10] A message from factor F to \mathcal{X}_k^U is denoted by $\mathbf{m}_{F \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U)$. With the Proposition 1, $\mathbf{m}_{F \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U)$ is given by

$$\mathbf{m}_{F \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U) = \int \mathbf{n}_{\mathcal{P}_k^U \rightarrow F}(\mathcal{P}_k^U) \delta_{\mathcal{X}_k^U \cup \mathcal{P}_k}(\mathcal{P}_k^U) \delta \mathcal{X}_k^{\sim U} \propto f^{\text{Poi}}(\mathcal{X}_k^U), \quad (\text{S5})$$

$$\propto f^{\text{Poi}}(\mathcal{X}_k^U), \quad (\text{S6})$$

proportional to the PPP distribution. In similarly, messages from factor F to \mathcal{P}_k^j , for $j \in \{1, \dots, J_k\}$, are denoted by $\mathbf{m}_{F \rightarrow \mathcal{P}_k^j}(\mathcal{P}_k^j)$, and with Proposition 1, $\mathbf{m}_{F \rightarrow \mathcal{P}_k^j}(\mathcal{P}_k^j)$ for $j \in \{1, \dots, J_k\}$ are given by

$$\mathbf{m}_{F \rightarrow \mathcal{P}_k^j}(\mathcal{P}_k^j) = \int \mathbf{n}_{\mathcal{P}_k^U \rightarrow F}(\mathcal{P}_k^U) \delta_{\mathcal{X}_k^U \cup \mathcal{P}_k}(\mathcal{P}_k^U) \delta \mathcal{P}_k^{\sim j}, \quad (\text{S7})$$

$$\propto f^{\text{Poi}}(\mathcal{P}_k^j) \quad (\text{S8})$$

also proportional to the PPP distribution.

- [11] Messages from the set-variable \mathcal{P}_k^j to linked factor G^j , i.e., $f_j^G(\mathcal{P}_k^j, \mathcal{Y}_k^j) = \delta_{\mathbf{h}_{-(I_{k-1}+j)}}(\mathcal{Y}_k^j)(\mathcal{P}_k^j)$, are denoted by $\mathbf{n}_{\mathcal{P}_k^j \rightarrow G^j}(\mathcal{P}_k^j) = \mathbf{m}_{F \rightarrow \mathcal{P}_k^j}(\mathcal{P}_k^j)$, for $j = \{1, \dots, J_k\}$.
- [12] Messages from factor G^j to the linked set-variable \mathcal{Y}_k^j are denoted by $\mathbf{m}_{G^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j)$. With the Definition 5, the messages $\mathbf{m}_{G^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j)$ are given by

$$\mathbf{m}_{G^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j) = \int \mathbf{n}_{F \rightarrow \mathcal{P}_k^j}(\mathcal{P}_k^j) \delta_{\mathbf{h}_{-(I_{k-1}+j)}}(\mathcal{Y}_k^j)(\mathcal{P}_k^j), \quad (\text{S9})$$

which follows the PPP distribution with auxiliary variables $u = I_{k-1} + j$.

2) *Data Association*: We compute the messages of marginal probabilities for c_k^i and d_k^j by running loopy BP [26] on the factor graph with cycle. Initial association probabilities are determined by the predicted messages and their linked likelihood factors.

13) Messages from the sensor state variable \mathbf{s}_k to linked factors I^j , i.e., $f_{I^j}^I(\mathbf{s}_k, \mathcal{Y}_k^j, d_k^j) = \tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \mathcal{Y}_k^j, d_k^j)$ for $j \in \mathcal{J}_k$, and J^i , i.e., $f_{J^i}^J(\mathbf{s}_k, \mathcal{X}_k^i, c_k^i) = t(\mathbf{Z}_k^i | \mathbf{s}_k, \mathcal{X}_k^i, c_k^i)$ for $i \in \mathcal{I}_{k-1}$, are respectively denoted by $\mathbf{n}_{\mathbf{s}_k \rightarrow I^j}(\mathbf{s}_k)$, and $\mathbf{n}_{\mathbf{s}_k \rightarrow J^i}(\mathbf{s}_k)$, given by $\mathbf{n}_{\mathbf{s}_k \rightarrow I^j}(\mathbf{s}_k) = \mathbf{n}_{\mathbf{s}_k \rightarrow J^i}(\mathbf{s}_k) = \mathbf{m}_{A \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$. Messages from the set-variables \mathcal{Y}_k^j to factor I^j are denoted by $\mathbf{n}_{\mathcal{Y}_k^j \rightarrow I^j}(\mathcal{Y}_k^j)$, and $\mathbf{n}_{\mathcal{Y}_k^j \rightarrow J^i}(\mathcal{Y}_k^j) = \mathbf{m}_{G^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j)$, for $j \in \{1, \dots, J_k\}$. Messages from the set-variable \mathcal{X}_k^i to linked factors J^i are denoted by $\mathbf{n}_{\mathcal{X}_k^i \rightarrow J^i}(\mathcal{X}_k^i)$, and $\mathbf{n}_{\mathcal{X}_k^i \rightarrow J^i}(\mathcal{X}_k^i) = \mathbf{m}_{E^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i)$, for $i \in \mathcal{I}_{k-1}$.

14) Messages from the linked factors I^j to the the linked variables d_k^j are denoted by $\mathbf{m}_{I^j \rightarrow d_k^j}(d_k^j)$. The messages $\mathbf{m}_{I^j \rightarrow d_k^j}(d_k^j)$, for $j \in \mathcal{J}_k$, are given by

$$\begin{aligned} & \mathbf{m}_{I^j \rightarrow d_k^j}(d_k^j) \\ &= \iint \mathbf{n}_{\mathbf{s}_k \rightarrow I^j}(\mathbf{s}_k) \mathbf{n}_{\mathcal{Y}_k^j \rightarrow I^j}(\mathcal{Y}_k^j) \tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \mathcal{Y}_k^j, d_k^j) d\mathbf{s}_k \delta \mathcal{Y}_k^j. \end{aligned} \quad (S10)$$

Messages from the linked factors J^i to the linked variables c_k^i are denoted by $\mathbf{m}_{J^i \rightarrow c_k^i}(c_k^i)$. The messages $\mathbf{m}_{J^i \rightarrow c_k^i}(c_k^i)$, for $i \in \mathcal{I}_{k-1}$, are given by

$$\begin{aligned} & \mathbf{m}_{J^i \rightarrow c_k^i}(c_k^i) \\ &= \iint \mathbf{n}_{\mathbf{s}_k \rightarrow J^i}(\mathbf{s}_k) \mathbf{n}_{\mathcal{X}_k^i \rightarrow J^i}(\mathcal{X}_k^i) t(\mathbf{Z}_k^i | \mathbf{s}_k, \mathcal{X}_k^i, c_k^i) d\mathbf{s}_k \delta \mathcal{X}_k^i. \end{aligned} \quad (S11)$$

15) During L iteration, loopy BP [26] between target-oriented data association variables c_k^i and measurement-oriented data association variables d_k^j with the factors $K^{i,j}$, i.e., $f_{K^{i,j}}^K(c_k^i, d_k^j) = \psi(c_k^i, d_k^j)$, is performed. Messages from the variables c_k^i to d_k^j and from the variables d_k^j to c_k^i at the l -th iteration are respectively denoted by $\mathbf{m}_{c_k^i \rightarrow d_k^j}^{(l)}(d_k^j)$ and $\mathbf{m}_{d_k^j \rightarrow c_k^i}^{(l)}(c_k^i)$, given by

$$\begin{aligned} & \mathbf{m}_{d_k^j \rightarrow c_k^i}^{(l)}(c_k^i) \\ &= \sum_{j=0}^{J_k} \mathbf{m}_{J^i \rightarrow c_k^i}(c_k^i) \psi(c_k^i, d_k^j) \prod_{i' \in \mathcal{I}_{k-1} \setminus \{i\}} \mathbf{m}_{c_k^{i'} \rightarrow d_k^j}^{(l-1)}(d_k^j) \end{aligned} \quad (S12)$$

$$\begin{aligned} & \mathbf{m}_{c_k^i \rightarrow d_k^j}^{(l)}(d_k^j) \\ &= \sum_{i=0}^{I_{k-1}} \mathbf{m}_{I^j \rightarrow d_k^j}(d_k^j) \psi(c_k^i, d_k^j) \prod_{j' \in \mathcal{J}_k \setminus \{j\}} \mathbf{m}_{d_k^{j'} \rightarrow c_k^i}^{(l-1)}(c_k^i). \end{aligned} \quad (S13)$$

Their initial messages are given by $\mathbf{m}_{d_k^j \rightarrow c_k^i}^{(0)}(c_k^i) = \sum_{j=0}^{J_k} \mathbf{m}_{J^i \rightarrow c_k^i}(c_k^i) \psi(c_k^i, d_k^j)$ and $\mathbf{m}_{c_k^i \rightarrow d_k^j}^{(0)}(d_k^j) = \sum_{i=0}^{I_{k-1}} \mathbf{m}_{I^j \rightarrow d_k^j}(d_k^j) \psi(c_k^i, d_k^j)$.

16) Messages from the variables c_k^i to factor J^i are denoted by $\mathbf{n}_{c_k^i \rightarrow J^i}(c_k^i)$. The messages $\mathbf{n}_{c_k^i \rightarrow J^i}(c_k^i)$, for $i \in \mathcal{I}_{k-1}$, are given by

$$\mathbf{n}_{c_k^i \rightarrow J^i}(c_k^i) = \prod_{j \in \mathcal{J}_k} \mathbf{m}_{d_k^j \rightarrow c_k^i}^{(L)}(c_k^i). \quad (S14)$$

Messages from the variables d_k^j to linked factors I^j are denoted by $\mathbf{n}_{d_k^j \rightarrow I^j}(d_k^j)$. The messages $\mathbf{n}_{d_k^j \rightarrow I^j}(d_k^j)$, for $j \in \mathcal{J}_k$, are given by

$$\mathbf{n}_{d_k^j \rightarrow I^j}(d_k^j) = \prod_{i \in \mathcal{I}_{k-1}} \mathbf{m}_{c_k^i \rightarrow d_k^j}^{(L)}(d_k^j). \quad (S15)$$

3) *Belief Computation*: We compute beliefs of previously detected targets, newly detected targets, undetected targets, and sensor state. The total number of beliefs is $2 + I_{k-1} + J_k$: 1 for the sensor state vector; 1 for undetected targets; and $I_{k-1} + J_k$ for detected targets.

a) *Newly Detected Targets*: We compute J_k beliefs of the set-variables \mathcal{X}_k^j for $j \in \mathcal{J}_k$. Each set-variable \mathcal{X}_k^j represents a target, newly detected for the first time or clutter, obtained as follows.

17) Messages from factors I^j to the linked set-variables \mathcal{Y}_k^j are denoted by $\mathbf{m}_{I^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j)$. The messages $\mathbf{m}_{I^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j)$, for $j \in \mathcal{J}_k$, are given by

$$\begin{aligned} \mathbf{m}_{I^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j) &= \int \sum_{i=0}^{I_{k-1}} \mathbf{n}_{d_k^i \rightarrow I^j}(d_k^i = i) \mathbf{n}_{\mathbf{s}_k \rightarrow I^j}(\mathbf{s}_k) \\ &\quad \times \tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \mathcal{Y}_k^j, d_k^i = i) d\mathbf{s}_k. \end{aligned} \quad (S16)$$

18) Beliefs of the set-variables \mathcal{Y}_k^j are denoted by $\mathbf{b}(\mathcal{Y}_k^j)$, for $j \in \mathcal{J}_k$. The beliefs $\mathbf{b}(\mathcal{Y}_k^j)$ are obtained by

$$\mathbf{b}(\mathcal{Y}_k^j) \propto \mathbf{m}_{G^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j) \mathbf{m}_{I^j \rightarrow \mathcal{Y}_k^j}(\mathcal{Y}_k^j), \quad (S17)$$

which follow the Bernoulli distribution.

b) *Previously Detected targets*: We compute I_{k-1} beliefs of the set-variables \mathcal{X}_k^i for $i \in \mathcal{I}_{k-1}$. Each set-variable \mathcal{X}_k^i represents the target that had been previously detected, obtained as follows.

19) Messages from factors J^i to the set-variables \mathcal{X}_k^i are denoted by $\mathbf{m}_{J^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i)$. The messages $\mathbf{m}_{J^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i)$, for $i \in \mathcal{I}_{k-1}$, are given by

$$\begin{aligned} \mathbf{m}_{J^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i) &= \int \sum_{j=0}^{J_k} \mathbf{n}_{c_k^i \rightarrow J^i}(c_k^i = j) \mathbf{n}_{\mathbf{s}_k \rightarrow J^i}(\mathbf{s}_k) \\ &\quad \times t(\mathbf{Z}_k^i | \mathbf{s}_k, \mathcal{X}_k^i, c_k^i = j) d\mathbf{s}_k. \end{aligned} \quad (S18)$$

20) Beliefs of the set-variables are denoted by $\mathbf{b}(\mathcal{X}_k^i)$, for $i \in \mathcal{I}_{k-1}$. The beliefs $\mathbf{b}(\mathcal{X}_k^i)$ are obtained by

$$\mathbf{b}(\mathcal{X}_k^i) \propto \mathbf{m}_{E^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i) \mathbf{m}_{J^i \rightarrow \mathcal{X}_k^i}(\mathcal{X}_k^i), \quad (S19)$$

which follows the Bernoulli distribution.

c) *Undetected Targets*: We compute 1 belief of the set-variable \mathcal{X}_k^U representing the targets that have never been detected and thus remain undetected again, obtained as follows.

- [21] A message from the sensor state \mathbf{s}_k to factor H , i.e., $f^H(\mathbf{s}_k, \mathcal{X}_k^U) = [1 - \text{pD}(\mathbf{s}_k, \mathcal{X}_k^U)]^{\mathcal{X}_k^U}$, is denoted by $\mathbf{n}_{\mathbf{s}_k \rightarrow I}(\mathbf{s}_k)$, given by $\mathbf{n}_{\mathbf{s}_k \rightarrow H}(\mathbf{s}_k) = \mathbf{n}_{A \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$. A message from factor H to the linked set-variable \mathcal{X}_k^U is denoted by $\mathbf{m}_{H \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U)$, given by

$$\mathbf{m}_{H \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U) = \int \mathbf{n}_{\mathbf{s}_k \rightarrow H}(\mathbf{s}_k) [1 - \text{pD}(\mathbf{s}_k, \mathcal{X}_k^U)]^{\mathcal{X}_k^U} d\mathbf{s}_k. \quad (\text{S20})$$

- [22] The belief of the set-variables \mathcal{X}_k^U is denoted by $\mathbf{b}(\mathcal{Y}_k^j)$, computed by

$$\mathbf{b}(\mathcal{X}_k^U) \propto \mathbf{m}_{F \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U) \mathbf{m}_{H \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U), \quad (\text{S21})$$

which follows the Bernoulli distribution.

d) *Sensor*: We compute 1 belief of the sensor state \mathbf{s}_k using the messages from the predicted sensor state, previously detected targets, and newly detected targets.

- [23] Messages from factor I^j to the linked vector variable \mathbf{s}_k are denoted by $\mathbf{m}_{I^j \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$. The messages $\mathbf{m}_{I^j \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$, for $j \in \mathcal{J}_k$, are given by

$$\begin{aligned} \mathbf{m}_{I^j \rightarrow \mathbf{s}_k}(\mathbf{s}_k) &= \int \sum_{i=0}^{I_k-1} \mathbf{n}_{d_k^j \rightarrow I^j}(d_k^j = i) \mathbf{n}_{\mathcal{Y}_k^j \rightarrow I^j}(\mathcal{Y}_k^j) \\ &\quad \times \tilde{l}(\mathbf{z}_k^j | \mathbf{s}_k, \mathcal{Y}_k^j, d_k^j = i) \delta \mathcal{Y}_k^j. \end{aligned} \quad (\text{S22})$$

- [24] Messages from factors J^i to the linked vector variable \mathbf{s}_k are denoted by $\mathbf{m}_{J^i \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$. The messages $\mathbf{m}_{J^i \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$, for $i \in \mathcal{I}_{k-1}$, are given by

$$\begin{aligned} \mathbf{m}_{J^i \rightarrow \mathbf{s}_k}(\mathbf{s}_k) &= \int \sum_{j=0}^{J_k} \mathbf{n}_{c_k^i \rightarrow J^i}(c_k^i = j) \mathbf{n}_{\mathcal{X}_k^i \rightarrow J^i}(\mathcal{X}_k^i) \\ &\quad \times t(\mathcal{Z}_k^i | \mathbf{s}_k, \mathcal{X}_k^i, c_k^i = j) \delta \mathcal{X}_k^i. \end{aligned} \quad (\text{S23})$$

- [25] A message from \mathcal{X}_k^U to factor H is denoted by $\mathbf{n}_{\mathcal{X}_k^U \rightarrow H}(\mathcal{X}_k^U)$, and $\mathbf{n}_{\mathcal{X}_k^U \rightarrow H}(\mathcal{X}_k^U) = \mathbf{m}_{F \rightarrow \mathcal{X}_k^U}(\mathcal{X}_k^U)$.

- [26] A message from factor H to the sensor state \mathbf{s}_k is denoted by $\mathbf{m}_{H \rightarrow \mathbf{s}_k}(\mathbf{s}_k)$, computed by

$$\mathbf{m}_{H \rightarrow \mathbf{s}_k}(\mathbf{s}_k) = \int \mathbf{n}_{\mathcal{X}_k^U \rightarrow H}(\mathcal{X}_k^U) [1 - \text{pD}(\mathbf{s}_k, \mathcal{X}_k^U)]^{\mathcal{X}_k^U} \delta \mathcal{X}_k^U \quad (\text{S24})$$

- [27] A belief of the sensor variable \mathbf{s}_k is denoted by $\mathbf{b}(\mathbf{s}_k)$, computed by

$$\begin{aligned} \mathbf{b}(\mathbf{s}_k) &\propto \mathbf{m}_{A \rightarrow \mathbf{s}_k}(\mathbf{s}_k) \mathbf{m}_{H \rightarrow \mathbf{s}_k}(\mathbf{s}_k) \\ &\quad \times \prod_{j \in \mathcal{J}_k} \mathbf{m}_{I^j \rightarrow \mathbf{s}_k}(\mathbf{s}_k) \prod_{i \in \mathcal{I}_{k-1}} \mathbf{m}_{J^i \rightarrow \mathbf{s}_k}(\mathbf{s}_k). \end{aligned} \quad (\text{S25})$$

Finally, we obtain the marginal posterior densities, $f_u(\mathbf{s}_k) = \mathbf{b}(\mathbf{s}_k)$, $f_u^U(\mathcal{X}_k^U) = \mathbf{b}(\mathcal{X}_k^U)$, $f_u^i(\mathcal{X}_k^i) = \mathbf{b}(\mathcal{X}_k^i)$ for $i \in \mathcal{I}_{k-1}$, $f_u^i(\mathcal{X}_k^{I_{k-1}+J}) = \mathbf{b}(\mathcal{X}_k^j)$ for $j \in \mathcal{J}_k$.