

A Computationally Efficient EK-PMBM Filter for Bistatic mmWave Radio SLAM

Yu Ge^{id}, *Student Member, IEEE*, Ossi Kaltiokallio^{id}, Hyowon Kim^{id}, *Member, IEEE*,
 Fan Jiang^{id}, *Member, IEEE*, Jukka Talvitie^{id}, *Member, IEEE*, Mikko Valkama^{id}, *Fellow, IEEE*,
 Lennart Svensson^{id}, *Senior Member, IEEE*, Sunwoo Kim^{id}, *Senior Member, IEEE*,
 and Henk Wymeersch^{id}, *Senior Member, IEEE*

Abstract—Millimeter wave (mmWave) signals are useful for simultaneous localization and mapping (SLAM), due to their inherent geometric connection to the propagation environment and the propagation channel. To solve the SLAM problem, existing approaches rely on sigma-point or particle-based approximations, leading to high computational complexity, precluding real-time execution. We propose a novel low-complexity SLAM filter, based on the Poisson multi-Bernoulli mixture (PMBM) filter. It utilizes the extended Kalman (EK) first-order Taylor series based Gaussian approximation of the filtering distribution, and applies the track-oriented marginal multi-Bernoulli/Poisson (TOMB/P) algorithm to approximate the resulting PMBM as a Poisson multi-Bernoulli (PMB). The filter can account for different landmark types in radio SLAM and multiple data association hypotheses. Hence, it has an adjustable complexity/performance trade-off. Simulation results show that the developed SLAM filter can greatly reduce the computational cost, while it keeps the good performance of mapping and user state estimation.

Index Terms—Bistatic sensing, extended Kalman filter, mmWave sensing, poisson multi-Bernoulli mixture filter, simultaneous localization and mapping.

I. INTRODUCTION

5G AND Beyond 5G systems can provide high-resolution measurements of delays and angles, which make them attractive for localization and sensing applications [1]–[4]. Localization of connected devices is important for autonomous vehicles [5], Vehicle-to-Everything (V2X) [6], [7], and spatial signal design [8]. Sensing of passive objects and connected users is an important part of integrated sensing and

communication (ISAC) [9]–[11], e.g., to support predictive beam tracking, beam alignment, and communication resource allocation. Sensing in ISAC relies on standard communication waveforms, such as orthogonal frequency division multiplexing (OFDM) and can be monostatic or bistatic. In monostatic sensing, the transmitter and receiver are co-located, which brings the advantage of a common clock and perfect knowledge of the transmitted data signal [12], thus allowing a common radio signal for communication and sensing. In bistatic sensing, the transmitter and receiver are spatially separated [3]. Hence, the data symbols are unknown to the receiver, and the transmitter and receiver are not synchronized. The former issue can be resolved by sending predetermined pilot signals within the OFDM frame structure, while the latter issue has the serious implication that only delay-differences among propagation paths bring information. Even in a bistatic setting, there are several benefits and challenges to the integration of sensing and communication, beyond the use of the same waveform type and the aforementioned improvement of communication. In particular, the sensing and communication resources must be multiplexed, which leads to interesting trade-offs at the transmit signal level [13]. At the receiver side, the channel estimation routines used as part of communication can be largely reused for sensing, leading to chip area savings. When the transmitter and receiver have known locations (e.g., base stations (BSs)), the problem is referred to as passive localization or mapping. When the transmitter or the receiver has an unknown location (e.g., a user equipment (UE)), this location must be determined jointly with the map (this also applies to the monostatic case with a UE radar), referred to as radio simultaneous localization and mapping (SLAM) [14], [15]. Here, the UE acts as a *sensor* with an unknown and time-varying state, while the static objects in the propagation environment act as *landmarks* with unknown states and cardinality.

Solving the radio SLAM problem is challenging because i) channel estimation errors, or noise peaks may result in false detections, ii) landmarks in the field-of-view (FoV) of the sensor can be undetected, due to the imperfect detection performance at the sensor, iii) the number of landmarks in the FoV is primarily unknown, as the map is unknown, iv) the source (landmark) of each measurement is unknown at the sensor, so there is an inherent data association problem [16], and v) the signal processing complexity due to accounting for the challenges i)–iv) within a mathematically coherent methodology. Several approaches have been proposed to address

Manuscript received August 20, 2021; revised December 9, 2021; accepted January 14, 2022. Date of publication March 2, 2022; date of current version June 17, 2022. This work was supported in part by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation; in part by the Vinnova 5GPOS Project under Grant 2019-03085; in part by the Swedish Research Council under Grant 2018-03705; in part by the Academy of Finland under Grant 315858, Grant 328214, Grant 319994, Grant 323244, and Grant 346622; and in part by the Ministry of Science and ICT (MSIT), South Korea, under the Information Technology Research Center (ITRC) support program under Grant IITP-2022-2017-0-01637. (*Corresponding author: Yu Ge.*)

Yu Ge, Hyowon Kim, Fan Jiang, Lennart Svensson, and Henk Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Göteborg, Sweden (e-mail: yuge@chalmers.se; hyowon@chalmers.se; fan.jiang@chalmers.se; lennart.svensson@chalmers.se; henkw@chalmers.se).

Ossi Kaltiokallio, Jukka Talvitie, and Mikko Valkama are with the Unit of Electrical Engineering, Tampere University, 33100 Tampere, Finland (e-mail: ossi.kaltiokallio@tuni.fi; jukka.talvitie@tuni.fi; mikko.valkama@tuni.fi).

Sunwoo Kim is with the Department of Electronic Engineering, Hanyang University, Seoul 04763, South Korea (e-mail: remero@hanyang.ac.kr).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JSAC.2022.3155504>.

Digital Object Identifier 10.1109/JSAC.2022.3155504

0733-8716 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

the above-mentioned challenges, including methods based on geometry [17], [18], on belief-propagation (BP) [19]–[21], and on random-finite-set (RFS) theory [22], [23]. Geometry-based methods [17], [18] have low complexity, but cannot inherently deal with the unknown number of targets or the data association problem. BP-SLAM methods [19]–[21] model the landmark state as a vector instead of an RFS, which requires ad-hoc modifications to deal with appearing and disappearing landmarks. Finally, RFS-based methods are particularly attractive for the SLAM problem, because the set of landmarks are modeled as an RFS, where uncertainties on both cardinality and the state of each landmark are considered, and there is no ordering of landmarks [24]. However, they come with a significant computational cost. Among RFS-based methods, the Poisson multi-Bernoulli mixture (PMBM) density is known to be a conjugate prior for most common measurement and dynamic models [25]. The key to the optimality of the PMBM filter lies in keeping track of all possible data association hypotheses over time, conditioned on the unknown sensor state, which renders the PMBM computationally demanding. To reduce the complexity, its simplified version, the Poisson multi-Bernoulli (PMB) filter, can be used, which reduces the number of hypotheses to one after each update. Several methods exist for this reduction, including the track-oriented marginal multi-Bernoulli/Poisson (TOMB/P), the measurement-oriented marginal multi-Bernoulli/Poisson (MOMB/P) [26], and the Kullback-Leibler divergence minimization [27] algorithms. To account for unknown sensor states, a low-complexity PMBM filter for joint target and sensor tracking is proposed in [28] by approximating the joint density with the product of marginal densities and performing separate updates per target.

Several RFS-based methods have been proposed for mmWave radio SLAM [29]–[34], which must account for the specific properties of mmWave sensing, including the highly nonlinear measurement models, time-varying detection probabilities, and the presence of multiple measurement models, due to different landmark types. These approaches mainly differ in terms of the representation of the RFS density and the required approximations. Among those RFS-based SLAM methods, [29] developed a Rao-Blackwellized particle (RBP) probability hypothesis density (PHD) filter. However, the complexity increases exponentially with the state dimension. To reduce the complexity, [30] introduced a cubature Kalman PHD (CK-PHD) filter for radio-SLAM. Although the CK-PHD exhibits a lower computational cost than the RBP-PHD, it relies on the sigma-point approximation, and thus the computational cost can be further reduced by the extended Kalman PHD (EK-PHD) in [33]. These three methods are based on the PHD filter, which considers the data association problem, but there is no explicit enumeration of different data associations. To explicitly consider all possible data associations, the RBP-PMBM SLAM filter is considered in [31], [32]. The PMBM filter has better mapping performance than the PHD filter [34], through exhibiting high computational cost.

In this paper, we address the high complexity of the RBP-PMBM filter by proposing a novel approach that combines the elegant PMBM representation with a low-complexity joint update of the sensor and landmark states. For this update we have selected the extended Kalman (EK) filter,

though other filters such as the cubature or unscented Kalman would also be appropriate. Compared to the RBP-SLAM filters from [29], [31], [32], the resulting filter has very low complexity and can thus be applied to real-time UE localization and environment mapping in bistatic mmWave sensing with a single BS. Compared to the EK-PHD [33] filter, which accounts only for the most likely data association and landmark type, we show how to explicitly account for several data associations and multi-model (MM) implementation. In contrast to [28], the proposed filter can cope with multiple nonlinear measurement models, and applies a joint update to all landmarks and the sensor state. Our main contributions are summarized as follows:

- The derivation of the EK-PMBM SLAM filter, which uses a new and theoretically sound method to jointly update the UE state and landmark states;
- The derivation of a novel algorithm to approximate the resulting PMBM to a PMB, which is based on the TOMB algorithm with a limited number of data associations and can be used to modify the EK-PMBM SLAM filter into the EK-PMB SLAM filter;
- The extension of the EK-PMBM and EK-PMB SLAM filters to the case of multiple landmark types, resulting in a multi-model (MM) implementation with hybrid discrete and continuous landmark states;
- The validation of the proposed SLAM filters, showing it exhibits very low complexity compared to the RBP-PMBM filter, while maintaining comparable SLAM performance.

The remainder of this article is structured as follows. The system models are described in Section II. The PMBM density and the Bayesian recursion of RFS-joint SLAM density are then introduced in Section III. The novel EK-PMBM SLAM filter is derived in Section IV, and its PMB counterpart in Section V. The extension to multiple measurement models is covered in Section VI. Simulation results are presented in Section VII, followed by our conclusions in Section VIII.

Notations: Scalars (e.g., x) are denoted in italic, vectors (e.g., \mathbf{x}) in bold, matrices (e.g., \mathbf{X}) in bold capital letters, sets (e.g., \mathcal{X}) in calligraphic, and its cardinality is denoted as $|\mathcal{X}|$. Transpose is denoted by $(\cdot)^T$. A Gaussian density with mean \mathbf{u} and covariance \mathbf{C} , evaluated in value \mathbf{x} is denoted by $\mathcal{N}(\mathbf{x}; \mathbf{u}, \mathbf{C})$. The union of mutually disjoint sets is denoted by \uplus , and the Kronecker product is denoted by \otimes .

II. SYSTEM MODELS IN MMWAVE BISTATIC SENSING

In this section, we introduce the models for the mobile UE state and the state of the landmarks in the propagation environment of mmWave, as shown in Fig. 1. Then, we provide the measurement model for the mmWave bistatic sensing scenario.

A. State Models

We consider a multi-antenna BS with known location $\mathbf{x}_{\text{BS}} \in \mathbb{R}^3$ and a multi-antenna UE, with sensor state at time step k denoted by \mathbf{s}_k (containing at least the 3D position and clock bias). The user dynamics are given by

$$f(\mathbf{s}_{k+1}|\mathbf{s}_k) = \mathcal{N}(\mathbf{s}_{k+1}; \mathbf{v}(\mathbf{s}_k), \mathbf{Q}_{k+1}), \quad (1)$$

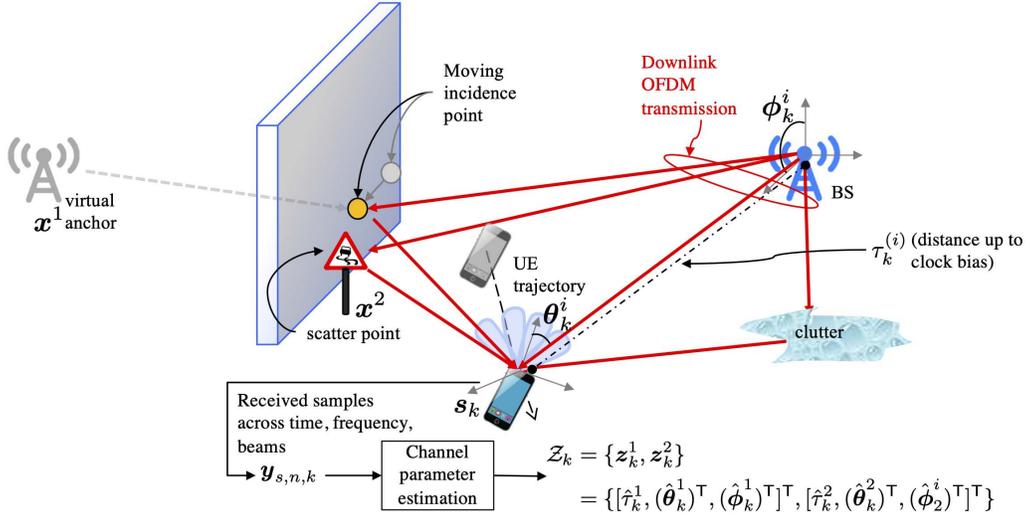


Fig. 1. Bistatic mmWave sensing scenario, where the UE tracks its own state and constructs a map of the environment, by incorporating bistatic measurements from received downlink signals sent by the BS. The downlink measurements per subcarrier and per beam are used to estimate a set of channel parameters \mathcal{Z}_k , which depend on the underlying geometry. Due to missed detections and false alarms, the cardinality of \mathcal{Z}_k (i.e., 2 in the example) may be less or more than the number of landmarks (3 in this case). The landmarks can be of three types: the BS, virtual anchors, or scatter points.

where $\mathbf{v}(\cdot)$ denotes a known transition function and \mathbf{Q}_{k+1} is the process noise covariance. The environment comprises three different types of landmarks, the BS, scattering points (SPs), which could correspond to street lamps, traffic signs, or others types of small objects, and reflecting surfaces, which could correspond to walls, building facades, or other types of large surfaces. Each SP is parameterized by an unknown 3D location $\mathbf{x}_{\text{SP}} \in \mathbb{R}^3$, while each reflecting surface is parameterized by a fixed virtual anchor (VA) with location $\mathbf{x}_{\text{VA}} \in \mathbb{R}^3$. The VA is the reflection of the BS with respect to the reflecting surface [35], [36]

$$\mathbf{x}_{\text{VA}} = (\mathbf{I} - 2\nu\nu^\top)\mathbf{x}_{\text{BS}} + 2\mu^\top\nu\nu, \quad (2)$$

where μ is an arbitrary point on the surface, and ν is the normal to the reflecting surface. The VA is surface-specific; although the incidence point of the downlink signal on the reflecting surface is moving while the UE is moving, the VA remains static.

B. Measurement Models

The measurement process is initiated by the BS every Δ seconds and requires a pilot transmission of duration $T_{\text{meas}} \ll \Delta$. The remaining temporal resources (i.e., $\Delta - T_{\text{meas}}$) are used for communication services. Every time step k (with period Δ), the BS sends downlink signals, which reach the UE via the line-of-sight (LOS) path as well as non-line-of-sight (NLOS) paths, via SPs or reflecting surfaces. We consider the transmission of N_{OFDM} OFDM symbols with S subcarriers. The received signal for OFDM symbol n , at subcarrier s and time step k can be expressed as [37]

$$\mathbf{y}_{s,n,k} = \mathbf{W}_{n,k}^H \sum_{i=0}^{I_k-1} g_k^i \mathbf{a}_R(\theta_k^i) \mathbf{a}_T^\top(\phi_k^i) e^{-j2\pi s \Delta f \tau_k^i} \mathbf{f}_{n,k} + \mathbf{W}_{n,k}^H \mathbf{n}_{s,n,k}, \quad (3)$$

where $\mathbf{f}_{n,k}$ is a precoder (including pilot symbols), $\mathbf{y}_{s,n,k}$ is the received signal; $\mathbf{W}_{n,k}$ is a combining matrix for OFDM symbol n , where the number of columns corresponds to the

number of UE radio frequency chains. The number of visible landmarks is denoted as I_k , while index $i = 0$ corresponds to the BS. We further assume that there is only one path per visible landmark. Moreover, $\mathbf{a}_R(\cdot)$ and $\mathbf{a}_T(\cdot)$ are the steering vectors of the receiver and transmitter antenna arrays, respectively, Δf is the subcarrier spacing, and $\mathbf{n}_{s,n,k}$ is the noise across the UE array. Each path i can be described by a complex gain g_k^i , a time of arrival (TOA) τ_k^i , an angle of arrival (AOA) pair θ_k^i in azimuth and elevation, and an angle of departure (AOD) pair ϕ_k^i in azimuth and elevation. The relations between the channel parameters and the sensor and landmark states can be found, e.g., in [32, Appendix A].¹ These parameters are estimated by a parametric channel estimation algorithm, such as [38]–[42]. We note that the total duration of the pilot transmission is approximately $T_{\text{meas}} \approx N/\Delta f$, ignoring the cyclic prefix overhead.

The channel estimator provides a set of channel parameter estimates \mathcal{Z}_k at time k , with $\mathcal{Z}_k = \{z_k^1, \dots, z_k^{\hat{I}_k}\}$ and \hat{I}_k representing the number of estimated paths, which we use directly as the measurements in the rest of the paper. Each element $z_k^i \in \mathcal{Z}_k$ is a 5-dimensional tuple of the form $z_k^i = [\hat{\tau}_k^i, (\hat{\theta}_k^i)^\top, (\hat{\phi}_k^i)^\top]^\top$. In general, $\hat{I}_k \neq I_k$, since measurements may originate from clutter (e.g., due to transient objects or noise peaks during channel estimation) and landmarks may be misdetections. The clutter can be modeled as a Poisson point process (PPP) with clutter intensity $c(\mathbf{z})$. To account for misdetections, we introduce the detection probability $p_D(\mathbf{x}^i, \mathbf{s}_k) \in [0, 1]$ that landmark \mathbf{x}^i is detected with a measurement when the sensor has state \mathbf{s}_k . The measurement originating from a landmark is characterized by a likelihood function

$$f(z_k^i | \mathbf{x}^i, \mathbf{s}_k) = \mathcal{N}(z_k^i; \mathbf{h}(\mathbf{x}^i, \mathbf{s}_k), \mathbf{R}_k^i), \quad (4)$$

¹The TOA is equal to the path distance over the speed of light, up to the unknown UE clock bias. The AOA is a function of the unknown UE orientation and position, and the locations of the landmarks. Finally, the AOD is a function of the unknown UE and landmark positions, and the known BS orientation.

where $\mathbf{h}(\mathbf{x}^i, \mathbf{s}_k) = [\tau_k^i, (\boldsymbol{\theta}_k^i)^\top, (\boldsymbol{\phi}_k^i)^\top]^\top$ and \mathbf{R}_k^i is the measurement covariance. We also introduce the state variable $\tilde{\mathbf{s}}_k^i = [\mathbf{s}_k^\top, \mathbf{x}_i^\top]^\top$, allowing us to write $\mathbf{h}(\tilde{\mathbf{s}}_k^i)$.

III. BASICS OF THE PMBM SLAM FILTER

In this section, we briefly describe basics of the PMBM and PMB SLAM filters, which will be the benchmark and starting point for the EK-PMB(M) filters in the subsequent sections.

A. PMBM and PMB Densities

An RFS is denoted by $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, where the element \mathbf{x}_i (indexed by i) denotes a random vector, and $|\mathcal{X}| = n$ is the random cardinality, with a set density $f(\mathcal{X})$. As the cardinality and element states are all unknown, the RFS can inherently model the uncertainties of the number of landmarks and the landmark states in an unknown map. When \mathcal{X}_U and \mathcal{X}_D are two independent RFSs, following a PPP and a MBM processes, respectively, we can say that $\mathcal{X} = \mathcal{X}_U \uplus \mathcal{X}_D$ follows a PMBM density [26], [43], [44]. The PPP RFS \mathcal{X}_U is used to model the set of undetected landmarks, which have never been detected before. The MBM RFS \mathcal{X}_D is used to model the set of detected landmarks, which have been detected at least once before. Then, the density of the PMBM RFS is given by

$$f_{\text{PMBM}}(\mathcal{X}) = \sum_{\mathcal{X}_U \uplus \mathcal{X}_D = \mathcal{X}} f_{\text{PPP}}(\mathcal{X}_U) f_{\text{MBM}}(\mathcal{X}_D), \quad (5)$$

where $f_{\text{PPP}}(\cdot)$ is a PPP density, and $f_{\text{MBM}}(\cdot)$ is an MBM density.² The PPP density is given by

$$f_{\text{PPP}}(\mathcal{X}_U) = e^{-\int D_u(\mathbf{x}') d\mathbf{x}'} \prod_{\mathbf{x} \in \mathcal{X}_U} D_u(\mathbf{x}), \quad (6)$$

where $D_u(\cdot)$ is the intensity function, and the MBM density follows

$$f_{\text{MBM}}(\mathcal{X}_D) = \sum_{j \in \mathbb{I}} w^j \sum_{\uplus_{i \in \mathbb{I}} \mathcal{X}^i = \mathcal{X}_D} \prod_{i=1}^n f_B^{j,i}(\mathcal{X}^i), \quad (7)$$

where \mathbb{I} is the index set of the global hypotheses, which in SLAM corresponds to the data associations [26], w^j is the weight for global hypothesis j , satisfying $\sum_{j \in \mathbb{I}} w^j = 1$, n is the number of potentially detected landmarks (for convenience set to the same value for all global hypotheses), \mathbb{I}^j is the index set of landmarks (i.e., the Bernoulli components) under global hypothesis j , and $f_B^{j,i}(\cdot)$ is the Bernoulli density of landmark i under global hypothesis j . Each Bernoulli density follows

$$f_B^{j,i}(\mathcal{X}^i) = \begin{cases} 1 - r^{j,i} & \mathcal{X}^i = \emptyset, \\ r^{j,i} f^{j,i}(\mathbf{x}) & \mathcal{X}^i = \{\mathbf{x}\}, \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

where $r^{j,i} \in [0, 1]$ is the existence probability of the landmark, and $f^{j,i}(\mathbf{x})$ is the probability density function of the vector \mathbf{x} . A larger existence probability means there is more likely a landmark exist. A falsely detected landmark will usually have a low existence probability as time progresses, while a

²To understand the \uplus notation, consider the example where $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2\}$. Then the summation in (5) has four terms: (i) $\mathcal{X}_U = \emptyset$ and $\mathcal{X}_D = \{\mathbf{x}_1, \mathbf{x}_2\}$; (ii) $\mathcal{X}_U = \{\mathbf{x}_1\}$ and $\mathcal{X}_D = \{\mathbf{x}_2\}$; (iii) $\mathcal{X}_U = \{\mathbf{x}_2\}$ and $\mathcal{X}_D = \{\mathbf{x}_1\}$; and (iv) $\mathcal{X}_U = \{\mathbf{x}_1, \mathbf{x}_2\}$ and $\mathcal{X}_D = \emptyset$.

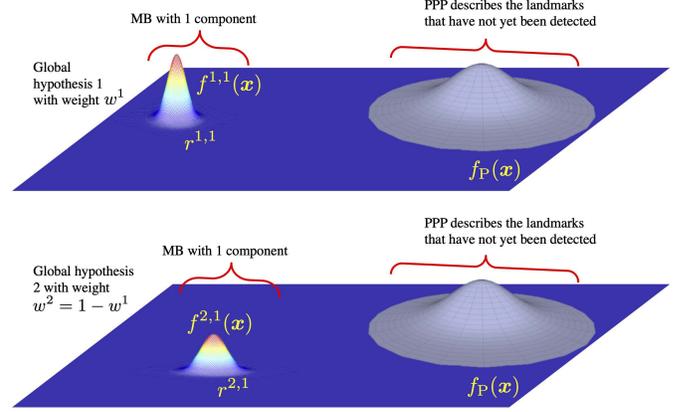


Fig. 2. An example for a PMBM density with 2 mixture components, each comprising a MB and a PPP.

real detected landmark will keep a relatively high existence probability. Moreover, if a landmark is misdetections at some time, the corresponding Bernoulli density can still be kept, with a modified existence probability, based on the associated detection probability (i.e., the misdetections landmark with small detection probability will maintain a high existence probability).

Example 1: Consider a PMBM with $|\mathbb{I}| = 2$ components, each MB contains a single Bernoulli (i.e., $|\mathbb{I}^1| = |\mathbb{I}^2| = n = 1$). Suppose we evaluate the PMBM in $\mathcal{X} = \{\mathbf{x}\}$ (i.e., containing a single element). Then,

$$f_{\text{PMBM}}(\{\mathbf{x}\}) = f_P(\{\mathbf{x}\})(w^1(1 - r^{1,1}) + w^2(1 - r^{2,1})) + f_P(\emptyset)(w^1 r^{1,1} f^{1,1}(\mathbf{x}) + w^2 r^{2,1} f^{2,1}(\mathbf{x}))$$

where $w^2 = 1 - w^1$. Similarly, for $\mathcal{X} = \emptyset$ (i.e., containing no elements), we find that

$$f_{\text{PMBM}}(\emptyset) = f_P(\emptyset)(w^1(1 - r^{1,1}) + w^2(1 - r^{2,1})).$$

Fig. 2 provides a visualization of the PMBM density.

In conclusion, a PMBM is described by a PPP from (6) and a MBM from (7), which is parameterized by $D_u(\mathbf{x})$ and $\{w^j, \{r^{j,i}, f^{j,i}(\mathbf{x})\}_{i \in \mathbb{I}^j}\}_{j \in \mathbb{I}}$. If there is only one mixture component in the MBM (i.e., $|\mathbb{I}| = 1$), then (5) reduces to a PMB.

B. Bayesian Recursion of RFS SLAM

An RFS SLAM filter follows the prediction and update steps of the Bayesian filtering recursion with RFSs. We denote a sensor trajectory at time k by $\mathbf{s}_{0:k}$ and denote a set of landmarks by \mathcal{X} . The joint posterior density of sensor trajectory and set of landmarks can be factorized as

$$f(\mathbf{s}_{0:k}, \mathcal{X} | \mathcal{Z}_{1:k}) = f(\mathbf{s}_{0:k} | \mathcal{Z}_{1:k}) f(\mathcal{X} | \mathbf{s}_{0:k}, \mathcal{Z}_{1:k}), \quad (9)$$

where $f(\mathbf{s}_{0:k} | \mathcal{Z}_{1:k})$ is the density of sensor trajectory, and $f(\mathcal{X} | \mathbf{s}_{0:k}, \mathcal{Z}_{1:k})$ is the set density of landmarks conditioned on the sensor trajectory.

The sensor prediction step is given by

$$f(\mathbf{s}_{0:k+1} | \mathcal{Z}_{1:k}) = f(\mathbf{s}_{0:k} | \mathcal{Z}_{1:k}) f(\mathbf{s}_{k+1} | \mathbf{s}_k), \quad (10)$$

where $f(\mathbf{s}_{k+1} | \mathbf{s}_k)$ is the transition density of the dynamics in (1). We assume landmarks are static and never appear or

disappear, and thus there is no prediction for landmarks. Then, $f(\mathcal{X}|\mathbf{s}_{0:k+1}, \mathcal{Z}_{1:k}) = f(\mathcal{X}|\mathbf{s}_{0:k}, \mathcal{Z}_{1:k})$.

The joint posterior is updated as

$$f(\mathbf{s}_{0:k+1}, \mathcal{X}|\mathcal{Z}_{1:k+1}) = \frac{g(\mathcal{Z}_{k+1}|\mathbf{s}_{k+1}, \mathcal{X})f(\mathbf{s}_{0:k+1}|\mathcal{Z}_{1:k})f(\mathcal{X}|\mathbf{s}_{0:k+1}, \mathcal{Z}_{1:k})}{f(\mathcal{Z}_{k+1}|\mathcal{Z}_{1:k})} \quad (11)$$

where $f(\mathcal{Z}_{k+1}|\mathcal{Z}_{1:k})$ is the normalizing factor, and $g(\mathcal{Z}_{k+1}|\mathbf{s}_{k+1}, \mathcal{X})$ is the RFS likelihood function, given by [43, eqs. (5)–(6)]

$$g(\mathcal{Z}_{k+1}|\mathbf{s}_{k+1}, \{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = e^{-\int c(\mathbf{z})d\mathbf{z}} \sum_{\mathcal{Z}^c \cup \mathcal{Z}^1 \dots \cup \mathcal{Z}^n = \mathcal{Z}_{k+1}} \prod_{\mathbf{z} \in \mathcal{Z}^c} c(\mathbf{z}) \prod_{l=1}^n \ell(\mathcal{Z}^l|\mathbf{s}_{k+1}, \mathbf{x}_l), \quad (12)$$

where \mathcal{Z}^c is the clutter measurement set, and $\ell(\cdot)$ follows

$$\ell(\mathcal{Z}^i|\mathbf{s}_{k+1}, \mathbf{x}_i) = \begin{cases} 1 - p_D(\mathbf{x}^i, \mathbf{s}_{k+1}) & \mathcal{Z}^i = \emptyset, \\ p_D(\mathbf{x}^i, \mathbf{s}_{k+1})f(\mathbf{z}|\mathbf{x}^i, \mathbf{s}_{k+1}) & \mathcal{Z}^i = \{\mathbf{z}\}, \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

C. Overview of PMBM and PMB SLAM Filters

If the PMBM density is the conjugate prior for the transition density and the measurement model, then all subsequent predicted and updated distributions by the Bayesian recursion (see, Section III-B) will preserve the PMBM form with parameters [26]: $D_u(\mathbf{x})$ and $\{w^j, \{r^{j,i}, f^{j,i}(\mathbf{x})\}_{i \in \mathbb{I}^j}\}_{j \in \mathbb{I}}$. The landmark is *conditioned* on the sensor state trajectory, and the set densities $f(\mathcal{X}|\mathbf{s}_{0:k}, \mathcal{Z}_{1:k})$ and $f(\mathcal{X}|\mathbf{s}_{0:k}, \mathcal{Z}_{1:k-1})$ are PMBM.

The PMBM SLAM filter can be implemented using a RBP filter (RBPF), where the sensor trajectory density is represented by particle samples [32]: $f(\mathbf{s}_{0:k+1}|\mathcal{Z}_{1:k+1}) \approx \sum_n^N \mathbf{s}_{0:k+1}^{(n)} w_{k+1}^{(n)}$, where n is the particle index; N is the number of particle samples; $\mathbf{s}_{0:k+1}^{(n)}$ is the particle sample; and $w_{k+1}^{(n)}$ is the particle weight such that $\sum_n w_{k+1}^{(n)} = 1$. The prediction step (10) then follows the standard particle generation of $\mathbf{s}_{0:k+1}^{(n)}$, while the update step (11) requires computation of $f(\mathcal{X}|\mathbf{s}_{0:k+1}^{(n)}, \mathcal{Z}_{1:k+1})$ and $w_{k+1}^{(n)}$ for each particle.

In the PMB SLAM filter, under a RBPF implementation, $f(\mathcal{X}|\mathbf{s}_{0:k}, \mathcal{Z}_{1:k})$ is given by a PMB density. After the update step, $f(\mathcal{X}|\mathbf{s}_{0:k+1}^{(n)}, \mathcal{Z}_{1:k+1})$ is possibly a PMBM since the filter tracks possible association hypotheses to the measurements. We adopt the method of marginal association distribution [26], which enables us to approximate the PMBM to a PMB at the end of time step by marginalizing over the data association.

Complexity: At the end of every time step, each previous global hypothesis considers all possible data associations, which generates a variety of new global hypotheses, and the number of global hypotheses rapidly increases in combinatorial explosion. In particular, the number of global hypotheses per each particle at time step $k+1$ is given by $|\mathbb{I}_{k+1}| = \sum_{j \in \mathbb{I}_k} \sum_{\alpha=0}^{|\mathcal{Z}_{k+1}|} C_{|\mathcal{Z}_{k+1}|}^{\alpha} P_{|\mathbb{I}_k^j|}^{|\mathcal{Z}_{k+1}|-\alpha}$, which grows over time [32, Appendix B]. Here, C and P

denote the combination and permutation operations, respectively. The complexity, at time step $k+1$, scales as $\mathcal{O}(N \sum_{j \in \mathbb{I}_k} |\mathbb{I}_k^j| \sum_{\alpha=0}^{|\mathcal{Z}_{k+1}|} C_{|\mathcal{Z}_{k+1}|}^{\alpha} P_{|\mathbb{I}_k^j|}^{|\mathcal{Z}_{k+1}|-\alpha})$, where number of particles N can be in the order of 1000–10000, depending on the state dimension.

In the PMB filter, the number of global hypotheses, at time step $k+1$, is $\sum_{\alpha=0}^{|\mathcal{Z}_{k+1}|} C_{|\mathcal{Z}_{k+1}|}^{\alpha} P_{|\mathbb{I}_k^1|}^{|\mathcal{Z}_{k+1}|-\alpha}$, which is much smaller than the number of global hypotheses in the PMBM SLAM filter. Then, the marginal association distribution is used to keep $f(\mathcal{X}|\mathbf{s}_{0:k+1}^{(n)}, \mathcal{Z}_{1:k+1})$ as a PMB density. The complexity, at time step $k+1$, scales as $\mathcal{O}(N |\mathbb{I}_k^1| \sum_{\alpha=0}^{|\mathcal{Z}_{k+1}|} C_{|\mathcal{Z}_{k+1}|}^{\alpha} P_{|\mathbb{I}_k^1|}^{|\mathcal{Z}_{k+1}|-\alpha})$.

In summary, both the PMB and PMBM suffer from high complexity, due to both the large number of hypotheses and the number of particles in the RBPF implementation.

IV. PROPOSED EK-PMBM SLAM FILTER

To reduce the computational cost, we motivate the joint sensor and landmark update using the EK filter [45, Ch. 5.2] instead of the RBPF. We will introduce the marginal posterior densities for SLAM and their Bayesian recursion. Then, we describe a novel algorithm of PMBM SLAM filter with the joint update step and its EK filter implementation.³

A. Form of the EK-PMBM Filter

We suppose that at time step k , the sensor state density is a Gaussian distribution $\mathcal{N}(\mathbf{s}_k; \mathbf{m}_{k|k}, \mathbf{P}_{k|k})$, where $\mathbf{m}_{k|k}$ and $\mathbf{P}_{k|k}$ are the mean and the covariance matrix, respectively; the PPP parameter $\lambda_{k|k}(\mathbf{x})$, which can be modeled as a $\eta_{k|k} \mathcal{U}(\mathbf{x})$ with $\mathcal{U}(\mathbf{x})$ representing a uniform distribution over the space, and MBM parameters $\{w_k^j, \{r_{k|k}^{j,i}, f_{k|k}^{j,i}(\mathbf{x})\}_{i \in \mathbb{I}_k^j}\}_{j \in \mathbb{I}_k}$ for the map are also given, where each $f_{k|k}^{j,i}(\mathbf{x})$ is a Gaussian distribution $\mathcal{N}(\mathbf{x}^{j,i}; \mathbf{u}_{k|k}^{j,i}, \mathbf{C}_{k|k}^{j,i})$. Therefore, the MBM parameters can be written as $\{w_k^j, \{r_{k|k}^{j,i}, \mathbf{u}_{k|k}^{j,i}, \mathbf{C}_{k|k}^{j,i}\}_{i \in \mathbb{I}_k^j}\}_{j \in \mathbb{I}_k}$.

B. Marginalization of PMBM SLAM Density

At each time step, rather than keeping the entire state trajectory, we keep track of marginal posteriors $f(\mathcal{X}|\mathcal{Z}_{1:k})$ and $f(\mathbf{s}_k|\mathcal{Z}_{1:k})$ [28], [47]. This implies that this posterior no longer carries the correlation between the sensor state trajectory and the map state, which constitutes an inherent loss of information and is the price to pay for reducing complexity. The prediction step then simplifies to

$$f(\mathbf{s}_{k+1}|\mathcal{Z}_{1:k}) = \int f(\mathbf{s}_k|\mathcal{Z}_{1:k})f(\mathbf{s}_{k+1}|\mathbf{s}_k)d\mathbf{s}_k. \quad (14)$$

The update step for the sensor state becomes

$$f(\mathbf{s}_{k+1}|\mathcal{Z}_{1:k+1}) = \int f(\mathbf{s}_{k+1}, \mathcal{X}|\mathcal{Z}_{1:k+1})\delta\mathcal{X} \quad (15)$$

$$\propto \int f(\mathcal{X}|\mathcal{Z}_{1:k})f(\mathbf{s}_{k+1}|\mathcal{Z}_{1:k})g(\mathcal{Z}_{k+1}|\mathbf{s}_{k+1}, \mathcal{X})\delta\mathcal{X}, \quad (16)$$

³In our EK filter implementation, linearization is performed around the predicted mean. Alternative linearizations (e.g., [46]) can also be considered but are out of the scope of the current work.

whereas for the map state we find that

$$\begin{aligned} f(\mathcal{X}|\mathcal{Z}_{1:k+1}) &= \int f(\mathbf{s}_{k+1}, \mathcal{X}|\mathcal{Z}_{1:k+1}) d\mathbf{s}_{k+1} \\ &\propto \int f(\mathcal{X}|\mathcal{Z}_{1:k}) f(\mathbf{s}_{k+1}|\mathcal{Z}_{1:k}) g(\mathcal{Z}_{k+1}|\mathbf{s}_{k+1}, \mathcal{X}) d\mathbf{s}_{k+1}. \end{aligned} \quad (17)$$

$$\propto \int f(\mathcal{X}|\mathcal{Z}_{1:k}) f(\mathbf{s}_{k+1}|\mathcal{Z}_{1:k}) g(\mathcal{Z}_{k+1}|\mathbf{s}_{k+1}, \mathcal{X}) d\mathbf{s}_{k+1}. \quad (18)$$

In (15), $\int g'(\mathcal{X}) \delta \mathcal{X}$ refers to the set integral [26, eq. (4)].

C. EK-PMBM SLAM Prediction

Following (14), the predicted UE state at the time step $k+1$ can be acquired via first-order EK filter [45, Ch. 5.2]

$$\mathbf{m}_{k+1|k} = \mathbf{v}(\mathbf{m}_{k|k}) \quad (19)$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}_{k|k} \mathbf{P}_{k|k} \mathbf{F}_{k|k}^T + \mathbf{Q}_{k+1}, \quad (20)$$

where $\mathbf{F}_{k|k}$ represents the Jacobian

$$\mathbf{F}_{k|k} = \left. \frac{\partial \mathbf{v}(\mathbf{s}_k)}{\partial \mathbf{s}_k} \right|_{\mathbf{s}_k = \mathbf{m}_{k|k}}. \quad (21)$$

In terms of the map, since the landmarks are assumed to be static, there is no prediction for the landmark states and covariances. Thus, we have $\eta_{k+1|k} = \eta_{k|k}$, $r_{k+1|k}^{j,i} = r_{k|k}^{j,i}$, $\mathbf{u}_{k+1|k}^{j,i} = \mathbf{u}_{k|k}^{j,i}$, $\mathbf{C}_{k+1|k}^{j,i} = \mathbf{C}_{k|k}^{j,i}$.

D. EK-PMBM SLAM Update

The update is more involved and comprises the following steps. First, the data association cost metrics are computed per MBM mixture component j (i.e., per global hypothesis). Second, the top $\gamma \geq 1$ best data associations per global hypothesis are determined from the cost metrics, followed by combining of all best data associations across the global hypotheses. Third, the EK filter joint update is performed for each of the best data associations, and each data association will result in a new global hypothesis in the updated MBM. Finally, the sensor state density is computed, marginalizing out the best data associations. These four steps are now explained in detail.

1) *Computation of Data Association Metric*: The data association cost metric depends on the local hypothesis weights, which are computed from measurements and previously seen landmarks [26]. We distinguish three cases⁴:

- (a) A previously detected landmark i under global hypothesis j is detected again with measurement \mathbf{z}_{k+1}^p . The local association weight is

$$l_{k+1}^{j,i,p} = r_{k+1|k}^{j,i} \rho_{k+1|k+1}^{j,i,p}, \quad (22)$$

where $\rho_{k+1|k+1}^{j,i,p}$ is computed as described in Appendix A-A. The first factor is the existence probability of the Bernoulli, and the second factor accounts for the spatial density and the measurement likelihood.

- (b) A previously detected landmark i under global hypothesis j is not detected at time $k+1$. The local association weight is

$$l_{k+1}^{j,i,0} = (1 - r_{k+1|k}^{j,i}) + (1 - p_{D,k+1}^{j,i}) r_{k+1|k}^{j,i}, \quad (23)$$

⁴The fourth case of an undetected landmark remaining undetected does not affect the cost metric.

where the first term accounts for the landmark not existing in the first place, while the second term accounts for the landmark existing, but leading to a miss-detection. Here, the detection probability is computed per landmark, as detailed in Appendix A-A.

- (c) A previously undetected landmark is detected for the first time with measurement \mathbf{z}_{k+1}^p . The local association weight is

$$l_{B,k+1}^p = c(\mathbf{z}_{k+1}^p) + \rho_{B,k+1|k+1}^p \quad (24)$$

where $\rho_{B,k+1|k+1}^p$ is computed, as described in Appendix A-B. The first term accounts for the fact that the measurement may be due to clutter, while the second term accounts for the fact that the measurement may be due to a new landmark from the PPP.

2) *Computation of Best γ Data Associations*: We construct a cost matrix $\mathbf{L}_{k+1}^j \in \mathbb{R}^{|\mathcal{Z}_{k+1}| \times (|\mathcal{Z}_{k+1}| + |\mathbb{I}_k^j|)}$, using the local association weights [43]

$$\begin{aligned} \mathbf{L}_{k+1}^j &= -\ln \begin{bmatrix} \tilde{l}_{k+1}^{j,1,1} & \cdots & \tilde{l}_{k+1}^{j,|\mathbb{I}_k^j|,1} & | & l_{B,k+1}^1 & \cdots & 0 \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ \tilde{l}_{k+1}^{j,1,|\mathcal{Z}_{k+1}|} & \cdots & \tilde{l}_{k+1}^{j,|\mathbb{I}_k^j|,|\mathcal{Z}_{k+1}|} & | & 0 & \cdots & l_{B,k+1}^{|\mathcal{Z}_{k+1}|} \end{bmatrix}, \end{aligned} \quad (25)$$

where $\tilde{l}_{k+1}^{j,i,p} = l_{k+1}^{j,i,p} / l_{k+1}^{j,i,0}$. The left $|\mathcal{Z}_{k+1}| \times |\mathbb{I}_k^j|$ sub-matrix in \mathbf{L}_{k+1}^j corresponds to previous detections, the right $|\mathcal{Z}_{k+1}| \times |\mathcal{Z}_{k+1}|$ diagonal sub-matrix corresponds to new detections, and the off-diagonal elements of the right sub-matrix are $-\infty$. The γ -best data associations with weights representing the probability per each can be selected out by solving the assignment problem

$$\begin{aligned} &\text{minimize } \text{tr}(\mathbf{A}^T \mathbf{L}_{k+1}^j) \\ &\text{s.t. } [\mathbf{A}]_{\alpha,\beta} \in \{0, 1\} \quad \forall \alpha, \beta \\ &\quad \sum_{\beta=1}^{|\mathbb{I}_k^j| + |\mathcal{Z}_{k+1}|} [\mathbf{A}]_{\alpha,\beta} = 1, \quad \forall \alpha \\ &\quad \sum_{\alpha=1}^{|\mathcal{Z}_{k+1}|} [\mathbf{A}]_{\alpha,\beta} \in \{0, 1\}, \quad \forall \beta \end{aligned} \quad (26)$$

using the Murty's algorithm [48], where $\mathbf{A} \in \mathbb{R}^{|\mathcal{Z}_{k+1}| \times (|\mathcal{Z}_{k+1}| + |\mathbb{I}_k^j|)}$ is the assignment matrix. The solutions are denoted by $\mathbf{A}^{j,h}$, where h is an index in the index set of new data associations under global hypothesis j , denoted as \mathbb{H}_{k+1}^j with $|\mathbb{H}_{k+1}^j| \leq \gamma$. The index set of landmarks under the j, h -th "new data association" is denoted as $\mathbb{I}_{k+1}^{j,h}$, with $|\mathbb{I}_{k+1}^{j,h}| \leq |\mathbb{I}_k^j| + |\mathcal{Z}_{k+1}|$.

For the j, h -th data association, the corresponding assignment matrix $\mathbf{A}^{j,h}$, can be translated to a vector, denoted as $\sigma^{j,h} = [\sigma^{j,h}(1), \dots, \sigma^{j,h}(|\mathbb{I}_k^j| + |\mathcal{Z}_{k+1}|)]$, defined as, for $t \leq |\mathbb{I}_k^j|$

$$\sigma^{j,h}(t) = \begin{cases} p & \exists p : [\mathbf{A}^{j,h}]_{p,t} = 1, \\ 0 & \nexists p : [\mathbf{A}^{j,h}]_{p,t} = 1, \end{cases} \quad (27)$$

and for $t > |\mathbb{I}_k^j|$

$$\sigma^{j,h}(t) = \begin{cases} p & \exists p : [\mathbf{A}^{j,h}]_{p,t} = 1, \\ \emptyset & [\mathbf{A}^{j,h}]_{t-|\mathbb{I}_k^j|,t} = 0. \end{cases} \quad (28)$$

These four cases correspond to a previously detected landmark associated to measurement p , a previously detected landmark being misdeteched, a new landmark being associated to measurement p , and a new landmark being non-existent. Note that p can only be $t - |\mathbb{I}_k^j|$ in the third case.

Each data association has a weight $w_{k+1}^{j,h}$, given by

$$w_{k+1}^{j,h} \propto w_k^j e^{-\text{tr}((\mathbf{A}^{j,h})^\top \mathbf{L}^j)} \quad (29)$$

subject to⁵ $\sum_{j \in \mathbb{I}_k} \sum_{h \in \mathbb{H}_{k+1}^j} w_{k+1}^{j,h} = 1$.

3) *Landmark Update:* Under the j, h -th new data association $\sigma^{j,h}$, we introduce the random variable $\check{\mathbf{s}}_{k+1}^{j,h,i} = [\check{\mathbf{s}}_{k+1}^\top, (\mathbf{x}^{j,h,i})^\top]^\top$ with mean $\check{\mathbf{m}}_{k+1|k}^{j,h,i}$, comprising the predicted sensor state and the state of the i -th landmark under the j, h -th new data association, as well as the random variable $\check{\mathbf{s}}_{k+1}^{j,h} = [\check{\mathbf{s}}_{k+1}^\top, (\mathbf{x}^{j,h,1})^\top, \dots, (\mathbf{x}^{j,h,|\mathbb{I}_k^j|})^\top]^\top$ with mean $\check{\mathbf{m}}_{k+1|k}^{j,h}$ and covariance $\check{\mathbf{P}}_{k+1|k}^{j,h}$. The mean vector, covariance matrix, measurement function and measurement vector for the joint state of the sensor and all previously detected landmarks $i \in \mathbb{I}_{k+1}^j, i \leq |\mathbb{I}_k^j|$, and associated measurement vector for all previously detected landmarks are constructed as

$$\check{\mathbf{m}}_{k+1|k}^{j,h} = [\mathbf{m}_{k+1|k}^\top, (\mathbf{u}_{k+1|k}^{j,h,1})^\top, \dots, (\mathbf{u}_{k+1|k}^{j,h,|\mathbb{I}_k^j|})^\top]^\top, \quad (30)$$

$$\check{\mathbf{P}}_{k+1|k}^{j,h} = \text{blkdiag}(\mathbf{P}_{k+1|k}, \mathbf{C}_{k+1|k}^{j,h,1}, \dots, \mathbf{C}_{k+1|k}^{j,h,|\mathbb{I}_k^j|}), \quad (31)$$

$$\check{\mathbf{h}}(\check{\mathbf{s}}_{k+1}^{j,h}) = [\mathbf{h}(\check{\mathbf{s}}_{k+1}^{j,h,1})^\top, \dots, \mathbf{h}(\check{\mathbf{s}}_{k+1}^{j,h,|\mathbb{I}_k^j|})^\top]^\top, \quad (32)$$

$$\check{\mathbf{z}}_{k+1|k}^{j,h} = \check{\mathbf{h}}(\check{\mathbf{m}}_{k+1|k}^{j,h}) = [\mathbf{h}(\check{\mathbf{m}}_{k+1|k}^{j,h,1})^\top, \dots, \mathbf{h}(\check{\mathbf{m}}_{k+1|k}^{j,h,|\mathbb{I}_k^j|})^\top]^\top, \quad (33)$$

$$\check{\mathbf{z}}_{k+1}^{j,h} = [(\mathbf{z}_{k+1}^{\sigma^{j,h}(1)})^\top, \dots, (\mathbf{z}_{k+1}^{\sigma^{j,h}(|\mathbb{I}_k^j|)})^\top]^\top, \quad (34)$$

where for $i : \sigma^{j,h}(i) = 0$, we set $\mathbf{z}_{k+1}^{\sigma^{j,h}(i)} = \mathbf{0}$ and $\mathbf{h}(\check{\mathbf{m}}_{k+1|k}^{j,h,i}) = \mathbf{0}$.

Then, the posterior mean and covariance under the j, h -th new data association can be determined via the first-order EK filter [45, Ch. 5.2]

$$\check{\mathbf{S}}_{k+1|k+1}^{j,h} = \check{\mathbf{H}}_{k+1|k}^{j,h} \check{\mathbf{P}}_{k+1|k}^{j,h} (\check{\mathbf{H}}_{k+1|k}^{j,h})^\top + \check{\mathbf{R}}_{k+1|k+1}^{j,h}, \quad (35)$$

$$\mathbf{K}_{k+1|k+1}^{j,h} = \check{\mathbf{P}}_{k+1|k}^{j,h} (\check{\mathbf{H}}_{k+1|k}^{j,h})^\top (\check{\mathbf{S}}_{k+1|k+1}^{j,h})^{-1}, \quad (36)$$

$$\check{\mathbf{m}}_{k+1|k+1}^{j,h} = \check{\mathbf{m}}_{k+1|k}^{j,h} + \mathbf{K}_{k+1|k+1}^{j,h} (\check{\mathbf{z}}_{k+1|k+1}^{j,h} - \check{\mathbf{z}}_{k+1|k}^{j,h}), \quad (37)$$

$$\check{\mathbf{P}}_{k+1|k+1}^{j,h} = \check{\mathbf{P}}_{k+1|k}^{j,h} - \mathbf{K}_{k+1|k+1}^{j,h} \check{\mathbf{S}}_{k+1|k+1}^{j,h} (\mathbf{K}_{k+1|k+1}^{j,h})^\top, \quad (38)$$

$$\check{\mathbf{R}}_{k+1|k+1}^{j,h} = \text{blkdiag}(\mathbf{R}_{k+1}^{\sigma^{j,h}(1)}, \dots, \mathbf{R}_{k+1}^{\sigma^{j,h}(|\mathbb{I}_k^j|)}), \quad (39)$$

where we define $\mathbf{R}_{k+1}^0 = \mathbf{I}$. The matrix $\check{\mathbf{H}}_{k+1|k}^{j,h}$ represents the Jacobian of $\check{\mathbf{h}}(\check{\mathbf{s}}_{k+1}^{j,h})$:

$$\check{\mathbf{H}}_{k+1|k}^{j,h} = \left. \frac{\partial \check{\mathbf{h}}(\check{\mathbf{s}}_{k+1}^{j,h})}{\partial \check{\mathbf{s}}_{k+1}^{j,h}} \right|_{\check{\mathbf{s}}_{k+1}^{j,h} = \check{\mathbf{m}}_{k+1|k}^{j,h}}. \quad (40)$$

The mean and the covariance of the j, h, i -th landmark, for $i \in \mathbb{I}_{k+1}^j, i \leq |\mathbb{I}_k^j|$, can be obtained from $\check{\mathbf{m}}_{k+1|k+1}^{j,h}$ and $\check{\mathbf{P}}_{k+1|k+1}^{j,h}$

⁵To further reduce the complexity, we can prune those data associations with weights lower than a threshold, or only keep a certain number of data associations with top weights. If such methods are applied, weights should be renormalized.

by extracting the corresponding parts of the posterior mean $\check{\mathbf{m}}_{k+1|k+1}^{j,h}$ and blocks along the diagonal of $\check{\mathbf{P}}_{k+1|k+1}^{j,h}$, as we drop cross-corrections for simplicity. The updated existence probability is given by

$$r_{k+1|k+1}^{j,h,i} = \begin{cases} 1 & \sigma^{j,h}(i) > 0, \\ \frac{(1 - p_{D,k+1}^j) r_{k+1|k}^{j,i,0}}{1 - r_{k+1|k}^{j,i,0} + (1 - p_{D,k+1}^j) r_{k+1|k}^{j,i,0}} & \sigma^{j,h}(i) = 0, \end{cases} \quad (41)$$

So-far, we have only considered the case of previously detected landmarks $i \in \mathbb{I}_{k+1}^j, i \leq |\mathbb{I}_k^j|$. To account for the newly detected landmarks, we must also consider $i \in \mathbb{I}_{k+1}^j, i > |\mathbb{I}_k^j|$, where $\sigma^{j,h}(\tilde{i}) = p$ for some p , with $\sigma^{j,h}(\tilde{i})$ is the i -th non-empty component⁶ in $\sigma^{j,h}$. The corresponding posterior distributions do not affect the sensor state posterior and are given by

$$\mathbf{u}_{k+1|k+1}^{j,h,i} = \mathbf{u}_{B,k+1|k+1}^p, \quad (42)$$

$$\mathbf{C}_{k+1|k+1}^{j,h,i} = \mathbf{C}_{B,k+1|k+1}^p, \quad (43)$$

$$r_{k+1|k+1}^{j,h,i} = \rho_{B,k+1|k+1}^p / l_{B,k+1}^p, \quad (44)$$

which were already computed in Appendix A-B and Appendix A-C, when determining the local association weight in (24).

4) *Sensor State Update:* The sensor mean and covariance, $\mathbf{m}_{k+1|k+1}$ and $\mathbf{P}_{k+1|k+1}$, can be obtained from $\check{\mathbf{m}}_{k+1|k+1}^{j,h}$ and $\check{\mathbf{P}}_{k+1|k+1}^{j,h}$ by marginalizing the landmark states out over all data associations:

$$\mathbf{m}_{k+1|k+1} = \sum_{j \in \mathbb{I}_k} \sum_{h \in \mathbb{H}_{k+1}^j} w_{k+1}^{j,h} [\check{\mathbf{m}}_{k+1|k+1}^{j,h}]_{1:\nu} \quad (45)$$

$$\begin{aligned} \mathbf{P}_{k+1|k+1} &= \sum_{j \in \mathbb{I}_k} \sum_{h \in \mathbb{H}_{k+1}^j} w_{k+1}^{j,h} ([\check{\mathbf{P}}_{k+1|k+1}^{j,h}]_{1:\nu,1:\nu} \\ &\quad + ([\check{\mathbf{m}}_{k+1|k+1}^{j,h}]_{1:\nu} - \mathbf{m}_{k+1|k+1}) \\ &\quad \times ([\check{\mathbf{m}}_{k+1|k+1}^{j,h}]_{1:\nu} - \mathbf{m}_{k+1|k+1})^\top), \end{aligned} \quad (46)$$

where ν is the length of the sensor state.

5) *Final Form After Update:* After sensor and map update, the sensor posterior distribution is $\mathcal{N}(\mathbf{s}_{k+1}; \mathbf{m}_{k+1|k+1}, \mathbf{P}_{k+1|k+1})$, while the map follows the PMBM format, with MBM components as $\{w_{k+1}^{j,h}, \{r_{k+1|k+1}^{j,h,i}, \mathbf{u}_{k+1|k+1}^{j,h,i}, \mathbf{C}_{k+1|k+1}^{j,h,i}\}_{i \in \mathbb{I}_{k+1}^j}\}_{h \in \mathbb{H}_{k+1}^j}\}_{j \in \mathbb{I}_k}$ and PPP intensity as $\eta_{k+1|k+1} = (1 - \bar{p}_D) \eta_{k+1|k}$, representing the previous undetected landmarks that remain undetected, where \bar{p}_D is a constant that describes the average detection probability over the considered space.

All data associations can be represented by only using one index. Hence, we reorder all data associations using index set $\mathbb{I}_{k+1} = \{1, \dots, \sum_{j \in \mathbb{I}_k} |\mathbb{H}_{k+1}^j|\}$. Then, MBM components can also be written as $\{w_{k+1}^j, \{r_{k+1|k+1}^{j,i}, \mathbf{u}_{k+1|k+1}^{j,i}, \mathbf{C}_{k+1|k+1}^{j,i}\}_{i \in \mathbb{I}_{k+1}^j}\}_{j \in \mathbb{I}_{k+1}}$.

⁶For example, $\sigma^{j,h} = [1, 0, \emptyset, 2]$ can form a MB with 3 Bernoullis, with the first two corresponding to the previously detected landmarks which is detected again with measurement 1 and misdeteched, respectively. The third corresponds to the new Bernoulli detected with measurement 2. The new Bernoulli detected with measurement 1 does not exist. Therefore, $\sigma^{j,h}(\tilde{3})$ is the third non-empty (fourth) component in $\sigma^{j,h}$, which is 2.

V. PROPOSED EK-PMB SLAM FILTER

The EK-PMBM SLAM filter generates γ best global hypotheses for each prior global hypothesis. This means that the complexity of the EK-PMBM filter scales grows exponentially with time k . To avoid this exponential complexity, we propose a variant based on the PMB filter, which only keeps one hypothesis at each time step.

A. Form of the EK-PMB Filter

In a PMB SLAM filter, the RFS landmark density comprises a PPP $\eta_{k|k}$ and a MB $\{r_{k|k}^{1,i}, f_{k|k}^{1,i}(\mathbf{x})\}_{i \in \mathbb{I}_k^1}$, where each $f_{k|k}^{1,i}(\mathbf{x})$ is a Gaussian distribution $\mathcal{N}(\mathbf{x}^{1,i}; \mathbf{u}_{k|k}^{1,i}, \mathbf{C}_{k|k}^{1,i})$. Therefore, the MB components can also be written as $\{r_{k|k}^{1,i}, \mathbf{u}_{k|k}^{1,i}, \mathbf{C}_{k|k}^{1,i}\}_{i \in \mathbb{I}_k^1}$. If we apply the update step from the EK-PMBM filter from Section IV, the resulting map density at time step $k+1$ will be a PMBM with $\{w_{k+1}^j, \{r_{k+1|k+1}^{j,i}, \mathbf{u}_{k+1|k+1}^{j,i}, \mathbf{C}_{k+1|k+1}^{j,i}\}_{i \in \mathbb{I}_{k+1}^j}\}_{j \in \mathbb{I}_{k+1}}$. Hence, to keep the PMB format, we need to approximate the MBM to an MB, which we do through a modified TOMB/P algorithm.

B. Proposed PMB Approximation

We firstly extend all MBs in the posterior PMBM density to the same space, with index set \mathbb{T}_{k+1} , where $|\mathbb{T}_{k+1}| = |\mathbb{I}_k^1| + |\mathcal{Z}_{k+1}|$. Then, we can rewrite the MBM part of the posterior PMBM as

$$\begin{aligned} & f_{k+1|k+1, \text{MBM}}(\mathcal{X} | \mathcal{Z}_{1:k+1}) \\ &= \sum_{j \in \mathbb{I}_{k+1}} w_{k+1}^j \sum_{\Psi_t \in \mathbb{T}_{k+1}} \prod_{t \in \mathbb{T}_{k+1}} f_{k+1|k+1}^{t, \sigma^j(t)}(\mathcal{X}^t | \sigma^j), \quad (47) \end{aligned}$$

where the Bernoulli density $f_{k+1|k+1}^{t, \sigma^j(t)}(\mathcal{X}^t | \sigma^j)$ captures the following cases (i) $t \leq |\mathbb{I}_k^1|$ and $\sigma^j(t) = 0$ correspond to a previously detected landmark that is misdetected, (ii) $t \leq |\mathbb{I}_k^1|$ and $\sigma^j(t) = p$ to a previously detected landmark that is detected with the p -th measurement, (iii) $t > |\mathbb{I}_k^1|$, $\sigma_{k+1}^j(t) = t - |\mathbb{I}_k^1|$ to a newly detected landmark with the $(t - |\mathbb{I}_k^1|)$ -th measurement, and (iv) $t > |\mathbb{I}_k^1|$, $\sigma^j(t) = \emptyset$ to a Bernoulli does not exist (with $f_{k+1|k+1}^{t, \emptyset}(\mathcal{X}^t)$ having 0 existence probability).

The Bernoulli is conditioned on the full association vector σ^j , due to the joint update in (35)–(39). Before we can apply the TOMB/P algorithm, we must remove this conditioning, to obtain Bernoullis that only depend on the local association, i.e., of the form $\hat{f}_{k+1|k+1}^{t, \sigma^j(t)}(\mathcal{X}^t)$. We do this by averaging and representing $\hat{f}_{k+1|k+1}^{t, \sigma^j(t)}(\mathcal{X}^t)$ by the existence probability, mean and covariance, computed as (for $t \leq |\mathbb{I}_k^1|$ and $q \in \{0, 1, \dots, |\mathcal{Z}_{k+1}|\} \cup \emptyset$)

$$r_{k+1|k+1}^{t,q} \propto \sum_{j \in \mathbb{I}_{k+1}: \sigma^j(t)=q} w_{k+1}^j r_{k+1|k+1}^{j, \sigma^j(t)}(\sigma^j) \quad (48)$$

$$\mathbf{u}_{k+1|k+1}^{t,q} \propto \sum_{j \in \mathbb{I}_{k+1}: \sigma^j(t)=q} w_{k+1}^j \mathbf{u}_{k+1|k+1}^{j, \sigma^j(t)}(\sigma^j) \quad (49)$$

$$\begin{aligned} \mathbf{C}_{k+1|k+1}^{t,q} &\propto \sum_{j \in \mathbb{I}_{k+1}: \sigma^j(t)=q} w_{k+1}^j (\mathbf{C}_{k+1|k+1}^{j, \sigma^j(t)}(\sigma^j) \\ &+ (\mathbf{u}_{k+1|k+1}^{j, \sigma^j(t)}(\sigma^j) - \mathbf{u}_{k+1|k+1}^{t,q})(\mathbf{u}_{k+1|k+1}^{j, \sigma^j(t)}(\sigma^j) \\ &- \mathbf{u}_{k+1|k+1}^{t,q})^\top), \quad (50) \end{aligned}$$

Algorithm 1 Proposed EK-PMB(M) SLAM Filter

Input: $\mathbf{m}_{k|k}$, $\mathbf{P}_{k|k}$ and PMB(M) at time k
Output: $\mathbf{m}_{k+1|k+1}$, $\mathbf{P}_{k+1|k+1}$, and updated PMB(M) at time $k+1$

- 1: Sensor state prediction (Section IV-C)
- 2: **for** $j \in \mathbb{I}_k$ **do** ▷ Each previous global hypothesis
- 3: Map prediction (Section IV-C)
- 4: Construct cost matrix \mathbf{L}_{k+1}^j (Section IV-D.1)
- 5: Compute best γ data associations (Section IV-D.2)
- 6: Compute updated PMBM (Section IV-D.3)
- 7: **end for**
- 8: Compute updated sensor density (Section IV-D.4)
- 9: **if** EK-PMB **then** ▷ Convert to PMB
- 10: Express MBM as (47)
- 11: Merge Bernoullis (48)–(50)
- 12: Compute marginal association probabilities (51)–(52)
- 13: Apply the TOMB/P method [26, Fig.10]
- 14: **end if**

where all three terms are normalized by the marginal probability $\beta^{t,q}$ for $q \in \{0, 1, \dots, |\mathcal{Z}_{k+1}|\} \cup \emptyset$, given by

$$\beta^{t,q} = \sum_{j \in \mathbb{I}_{k+1}: \sigma^j(t)=q} w_{k+1}^j. \quad (51)$$

For $t > |\mathbb{I}_k^1|$, newly detected objects already satisfy $\hat{f}_{k+1|k+1}^{t,q}(\mathcal{X}^t) = f_{k+1|k+1}^{t,q}(\mathcal{X}^t | \sigma^j)$ by design, so averaging has no effect. Marginal association probabilities are given by

$$\beta^{t, t-|\mathbb{I}_k^1|} = \sum_{j \in \mathbb{I}_{k+1}: \sigma^j(t)=t-|\mathbb{I}_k^1|} w_{k+1}^j, \quad (52)$$

$$\beta^{t, \emptyset} = 1 - \beta^{t, t-|\mathbb{I}_k^1|}. \quad (53)$$

Note that computing these marginal associations is straightforward, when γ is not too large. Finally, densities $\hat{f}_{k+1|k+1}^{t,q}(\mathcal{X}^t)$ and the marginal probabilities $\beta^{t,q}$ are now be used as an input for the standard TOMB/P method to form the new MB [26, Fig.10].

C. Overview of the EK-PMB(M) SLAM Filters

The proposed EK-PMB(M) SLAM filter is summarized as Algorithm 1 and the corresponding flowchart is shown as Fig. 3. The EK-PMB SLAM filter generates γ best global hypotheses every time step. This means that the complexity of the EK-PMB filter scales as $\mathcal{O}(|\mathbb{I}_k^1|^3 \gamma)$ at time $k+1$, where the power of 3 comes from the complexity of the matrix inverse in (36). To compare the complexities for all above mentioned algorithms, we summarize the complexities of the RBP-PMB(M) SLAM filters with N particles in Section III-C, the proposed EK-PMBM SLAM filter that generates γ best global hypotheses for each prior global hypothesis in Section IV, and the proposed EK-PMB SLAM filter generates γ best global hypotheses every time step in Table I.

VI. EXTENSION TO MULTIPLE MODELS

In the data association, each measurement is associated to a landmark. However, the type of the landmark is still unknown. As mentioned in Section II-A, we consider three different types of landmarks, BS, VA, and SP. Therefore, the SLAM filter should not only be able to figure out the source of each measurement, but also can distinguish the type of the associated landmark.

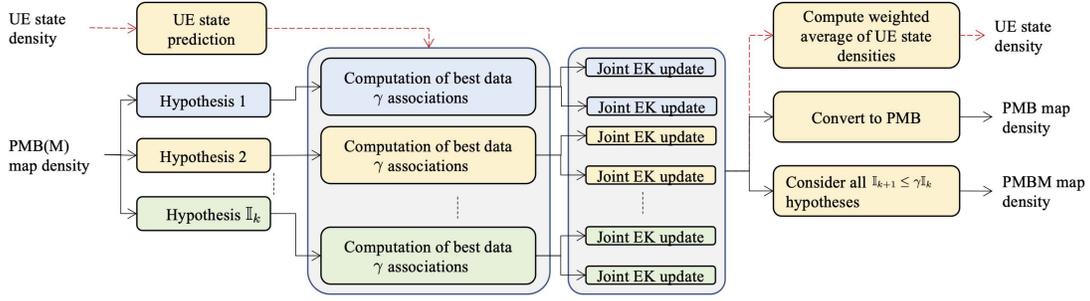


Fig. 3. The flowchart of the EK-PMB(M) SLAM filter for the case of $\gamma = 2$. The γ best data associations are determined per global hypothesis, after which a joint update of the landmark states and sensor state is performed, conditioned on each association within each global hypothesis. Finally, the sensor and landmark states are marginalized out.

TABLE I

COMPLEXITIES OF THE PMBM-BASED SLAM FILTERS AT TIME STEP $k + 1$. HERE, N DENOTES THE NUMBER OF PARTICLES (FOR THE RBP FILTERS), I_k THE NUMBER OF GLOBAL HYPOTHESES AT TIME k , C AND P DENOTE THE COMBINATION AND PERMUTATION OPERATIONS

Filter	Complexity
RBP-PMBM	$\mathcal{O}(N \sum_{j \in I_k} I_k^j \sum_{\alpha=0}^{ Z_{k+1} } C_{ Z_{k+1} }^{\alpha} P_{ I_k^j }^{ Z_{k+1} -\alpha})$
RBP-PMB	$\mathcal{O}(N I_k^1 \sum_{\alpha=0}^{ Z_{k+1} } C_{ Z_{k+1} }^{\alpha} P_{ I_k^1 }^{ Z_{k+1} -\alpha})$
EK-PMBM	$\mathcal{O}(\sum_{j \in I_k} I_k^j ^3 \gamma)$
EK-PMB	$\mathcal{O}(I_k^1 ^3 \gamma)$

A. Problem Description

The presented EK-PMBM and EK-PMB filters were designed only for continuous state spaces. Following [49], we introduce the mixed state space comprising the continuous state \mathbf{x} and the discrete state $\xi \in \mathcal{A} = \{\text{BS}, \text{VA}, \text{SP}\}$. The corresponding Bernoulli densities are given by

$$f_B^{j,i}(\mathcal{X}^i) = \begin{cases} 1 - r^{j,i} & \mathcal{X}^i = \emptyset, \\ r^{j,i} f^{j,i}(\mathbf{x}, \xi) \psi^{j,i,\xi} & \mathcal{X}^i = \{\mathbf{x}, \xi\}, \\ 0 & \text{otherwise,} \end{cases} \quad (54)$$

where $\psi^{j,i,\xi} \in [0, 1]$ represents the probability that the type of the j, i -th landmark is ξ , with $\sum_{\xi \in \mathcal{A}} \psi_{k|k}^{j,i,\xi} = 1$. Furthermore, we set the detection probability model dependent, denoted as $p_D(\mathbf{x}, \mathbf{s}_k, \xi)$.

We now describe how the EK-PMBM and EK-PMB filters should be modified to account for the multiple models.

B. Required Modifications

1) *PMBM and PMB Form*: We can rewrite the PMBM components as $\{\eta_{k|k}^{\xi}\}_{\xi \in \mathcal{A}}$ and MBM parameters $\{w_{k|k}^{j,i}, \{r_{k|k}^{j,i}, \{\psi_{k|k}^{j,i,\xi}, \mathbf{u}_{k|k}^{j,i,\xi}, \mathbf{C}_{k|k}^{j,i,\xi}\}_{\xi \in \mathcal{A}}\}_{i \in I_k^j}\}_{j \in I_k}$.

2) *Prediction Step*: The prediction step can be reformulated as $\eta_{k+1|k}^{\xi} = \eta_{k|k}^{\xi}$, $r_{k+1|k}^{j,i} = r_{k|k}^{j,i}$, $\psi_{k+1|k}^{j,i,\xi} = \psi_{k|k}^{j,i,\xi}$, $\mathbf{u}_{k+1|k}^{j,i,\xi} = \mathbf{u}_{k|k}^{j,i,\xi}$, $\mathbf{C}_{k+1|k}^{j,i,\xi} = \mathbf{C}_{k|k}^{j,i,\xi}$.

3) *Computation of Data Association Metric*: Compared to IV-D.1, there is only a different state definition. Hence, when calculating (22), (23), and (24) in the Appendices, we marginalize over ξ in addition to \mathbf{x} .

4) *Update Step*: To perform the update, we must consider all possible landmark types. Hence, in (30), (31), (32), (33),

$\mathbf{u}_{k+1|k}^{j,h,i}$, $\mathbf{C}_{k+1|k}^{j,h,i}$, $\mathbf{h}(\tilde{\mathbf{s}}_{k+1}^{j,h,i})$, $\mathbf{h}(\tilde{\mathbf{m}}_{k+1|k}^{j,h,i})$ should be replaced by

$$\mathbf{u}_{k+1|k}^{j,h,i} \rightarrow [(\mathbf{u}_{k+1|k}^{j,h,i,1})^T, \dots, (\mathbf{u}_{k+1|k}^{j,h,i,|\mathcal{A}|})^T]^T, \quad (55)$$

$$\mathbf{C}_{k+1|k}^{j,h,i} \rightarrow \text{blkdiag}(\mathbf{C}_{k+1|k}^{j,h,i,1}, \dots, \mathbf{C}_{k+1|k}^{j,h,i,|\mathcal{A}|}), \quad (56)$$

$$\mathbf{h}(\tilde{\mathbf{s}}_{k+1}^{j,h,i}) \rightarrow [\mathbf{h}(\tilde{\mathbf{s}}_{k+1}^{j,h,i,1})^T, \dots, \mathbf{h}(\tilde{\mathbf{s}}_{k+1}^{j,h,i,|\mathcal{A}|})^T]^T, \quad (57)$$

$$\mathbf{h}(\tilde{\mathbf{m}}_{k+1|k}^{j,h,i}) \rightarrow [\mathbf{h}(\tilde{\mathbf{m}}_{k+1|k}^{j,h,i,1})^T, \dots, \mathbf{h}(\tilde{\mathbf{m}}_{k+1|k}^{j,h,i,|\mathcal{A}|})^T]^T, \quad (58)$$

respectively. In (34), $\mathbf{z}_{k+1}^{\sigma^{j,h}(i)}$ is to be replaced by $\mathbf{1}_{|\mathcal{A}| \times 1} \otimes \mathbf{z}_{k+1}^{\sigma^{j,h}(i)}$, and in (39) $\mathbf{R}_{k+1}^{\sigma^{j,h}(t)}$ is replaced by $\mathbf{1}_{|\mathcal{A}| \times |\mathcal{A}|} \otimes \mathbf{R}_{k+1}^{\sigma^{j,h}(t)}$, since replicating the measurements leads to perfect correlation in the covariance matrix. The updates can then be performed as before to recover the joint state posterior. Births should be generated for each landmark type (except $\xi = \text{BS}$).

To account for the type probabilities in the posterior, $\psi_{k+1|k+1}^{j,h,i,\xi}$, we compute, for $i \leq |I_k^j|$

$$\psi_{k+1|k+1}^{j,h,i,\xi} \propto \begin{cases} (1 - p_{D,k+1}^{j,i,\xi}) \psi_{k+1|k}^{j,i,\xi} & \sigma^{j,h}(i) = 0, \\ p_{D,k+1}^{j,i,\xi} \psi_{k+1|k}^{j,i,\xi} \mathcal{N}(\mathbf{z}_{k+1}^p; \mathbf{h}(\tilde{\mathbf{m}}_{k+1|k}^{j,i,\xi}), \mathbf{S}_{k+1|k}^{j,i,\xi}) & \sigma^{j,h}(i) = p, \end{cases} \quad (59)$$

where $p_{D,k+1}^{j,i,\xi} = p_D(\tilde{\mathbf{s}}_{k+1}^{j,i,\xi} = \tilde{\mathbf{m}}_{k+1|k}^{j,i,\xi})$, and for $i > |I_k^j|$, $\sigma^{j,h}(i) \neq \emptyset$

$$\begin{aligned} \psi_{k+1|k+1}^{j,h,i,\xi} &= \psi_{B,k+1|k+1}^{\sigma^{j,h}(i),\xi} \\ &\propto \eta_{k+1|k}^{\xi} p_{D,k+1}^{\sigma^{j,h}(i),\xi} \mathcal{N}(\mathbf{z}^{\sigma^{j,h}(i),\xi}; \mathbf{h}(\tilde{\mathbf{m}}_{B,k+1|k}^{\sigma^{j,h}(i),\xi}), \mathbf{S}_{B,k+1|k}^{\sigma^{j,h}(i),\xi}) \end{aligned} \quad (60)$$

where the proportionality constant can be recovered from $\sum_{\xi} \psi_{k+1|k+1}^{j,h,i,\xi} = 1$. We recall that \tilde{i} is the index of i -th non-empty component in $\sigma^{j,h}$, and the global hypotheses j, h will be finally re-indexed with j .

The PPP is updated by $\eta_{k+1|k+1}^{\xi} = (1 - \tilde{p}_D^{\xi}) \eta_{k+1|k}^{\xi}$, where \tilde{p}_D^{ξ} is a constant that describes the average detection probability over the considered space.

5) *PMB Approximation*: When approximating PMBM to PMB in Section V, the landmark type should also be considered and (49)–(50) should be computed for each landmark type. Similarly, the marginal association probabilities must be computed for each landmark type by marginalizing the probability of the type for each association over data associations,

Algorithm 2 Modified MM-TOMB/P Algorithm

Input: Marginal probabilities $\beta^{t,q}$ and $\beta^{t,q,\xi}$; MBM components;
Output: MB density

$$\{\hat{r}_{k+1|k+1}^{1,i}, \{\hat{\psi}_{k+1|k+1}^{1,i,\xi}, \hat{\mathbf{u}}_{k+1|k+1}^{1,i,\xi}, \hat{\mathbf{C}}_{k+1|k+1}^{1,i,\xi}\}_{\xi \in \mathcal{A}}\}_{i \in \mathbb{T}_{k+1}}$$

- 1: **for** $t \leq \|\mathbb{I}_k^1\|$ **do** ▷ Previously detected tracks
- 2: $\hat{r}_{k+1|k+1}^{1,t} = \sum_q \beta^{t,q} r_{k+1|k+1}^{t,q}$
- 3: **for** $\xi \in \mathcal{A}$ **do**
- 4: $\hat{\psi}_{k+1|k+1}^{1,t,\xi} = \frac{\sum_q \beta^{t,q,\xi} r_{k+1|k+1}^{t,q,\xi}}{\hat{r}_{k+1|k+1}^{1,t}}$
- 5: $\hat{\mathbf{u}}_{k+1|k+1}^{1,t,\xi} = \frac{\sum_q \beta^{t,q,\xi} r_{k+1|k+1}^{t,q,\xi} \mathbf{u}_{k+1|k+1}^{t,q,\xi}}{\sum_q \beta^{t,q,\xi} r_{k+1|k+1}^{t,q,\xi}}$
- 6: $\hat{\mathbf{C}}_{k+1|k+1}^{1,t,\xi} = \frac{\sum_q \beta^{t,q,\xi} r_{k+1|k+1}^{t,q,\xi} (\mathbf{C}_{k+1|k+1}^{t,q,\xi} + (\hat{\mathbf{u}}_{k+1|k+1}^{1,t,\xi} - \mathbf{u}_{k+1|k+1}^{t,q,\xi})(\hat{\mathbf{u}}_{k+1|k+1}^{1,t,\xi} - \mathbf{u}_{k+1|k+1}^{t,q,\xi})^\top)}{\sum_q \beta^{t,q,\xi} r_{k+1|k+1}^{t,q,\xi}}$
- 7: **end for**
- 8: **end for**
- 9: **for** $t > \|\mathbb{I}_k^1\|$ **do** ▷ New tracks
- 10: $\hat{r}_{k+1|k+1}^{1,t} = \beta^{t,t-\|\mathbb{I}_k^1\|} r_{k+1|k+1}^{t,t-\|\mathbb{I}_k^1\|}$
- 11: **for** $\xi \in \mathcal{A}$ **do**
- 12: $\hat{\psi}_{k+1|k+1}^{1,t,\xi} = \frac{\beta^{t,t-\|\mathbb{I}_k^1\|} \psi_{k+1|k+1}^{t,t-\|\mathbb{I}_k^1\|,\xi}}{\hat{r}_{k+1|k+1}^{1,t}}$
- 13: $\hat{\mathbf{u}}_{k+1|k+1}^{1,t,\xi} = \mathbf{u}_{k+1|k+1}^{t,t-\|\mathbb{I}_k^1\|,\xi}$
- 14: $\hat{\mathbf{C}}_{k+1|k+1}^{1,t,\xi} = \mathbf{C}_{k+1|k+1}^{t,t-\|\mathbb{I}_k^1\|,\xi}$
- 15: **end for**
- 16: **end for**

for $t \leq \|\mathbb{I}_k^1\|$, to yield

$$\beta^{t,q,\xi} \propto \sum_{j \in \mathbb{I}_{k+1}^1: \sigma^j(t)=q} w_{k+1}^j \psi_{k+1|k+1}^{j,t,\xi}, \quad (61)$$

and for $t > \|\mathbb{I}_k^1\|$,

$$\beta^{t,t-\|\mathbb{I}_k^1\|,\xi} \propto \sum_{j \in \mathbb{I}_{k+1}^1: \sigma^j(t)=t-\|\mathbb{I}_k^1\|} w_{k+1}^j \psi_{k+1|k+1}^{j,t-\|\mathbb{I}_k^1\|,\xi}, \quad (62)$$

with $\sum_{\xi} \beta^{t,q,\xi} = \beta^{t,q}$. We also have $\psi_{k+1|k+1}^{j,t,\xi} = \beta^{t,q,\xi}$. Then, we modify the TOMB/P algorithm [26, Fig.10] as shown in Algorithm 2.

VII. RESULTS

In this section, the proposed algorithms are evaluated on a vehicular scenario and compared to two benchmarks. We describe the simulation environment, the benchmarks and performance metrics, before discussing the SLAM results in terms of localization and mapping performance.

A. Simulation Environment

We consider a scenario as illustrated in Fig. 4. There is a BS located at $[0 \text{ m}, 0 \text{ m}, 40 \text{ m}]^\top$, 4 reflection surfaces with VAs located at $[200 \text{ m}, 0 \text{ m}, 40 \text{ m}]^\top$, $[-200 \text{ m}, 0 \text{ m}, 40 \text{ m}]^\top$, $[0 \text{ m}, 200 \text{ m}, 40 \text{ m}]^\top$, $[0 \text{ m}, -200 \text{ m}, 40 \text{ m}]^\top$, representing 4 reflection surfaces (wall in the physical environment), and 4 SPs, located at $[99 \text{ m}, 0 \text{ m}, 10 \text{ m}]^\top$, $[-99 \text{ m}, 0 \text{ m}, 10 \text{ m}]^\top$, $[0 \text{ m}, 99 \text{ m}, 10 \text{ m}]^\top$, $[0 \text{ m}, -99 \text{ m}, 10 \text{ m}]^\top$, representing some small landmarks near the walls, for example some street lamps. Compared with the scenarios in some previous work [29]–[31], the SPs are closer to the reflection surfaces. This brings additional challenges into the data association

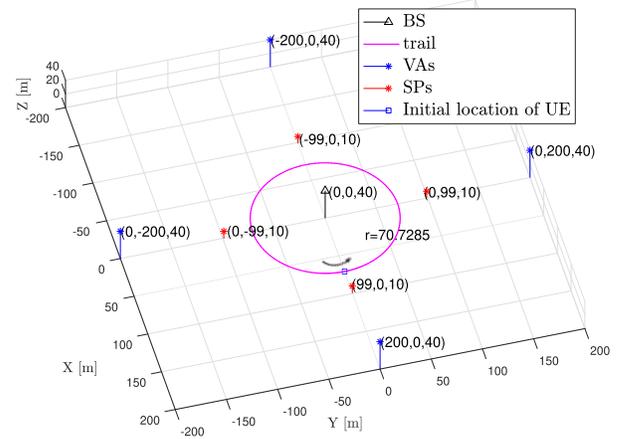


Fig. 4. Scenario with the environment of a BS and 4 VAs and 4 SPs. The UE moves counterclockwise along the trail centered at the BS.

problem, especially when the incidence points on the reflection surfaces are near the SPs. In addition, there is a single sensor doing a counterclockwise constant turn-rate movement around the BS

$$\mathbf{v}(s_{k-1}) = \mathbf{s}_{k-1} + \begin{bmatrix} \frac{\zeta}{\varrho} (\sin(\varpi_{k-1} + \varrho\Delta) - \sin \varpi_{k-1}) \\ \frac{\zeta}{\varrho} (-\cos(\varpi_{k-1} + \varrho\Delta) + \cos \varpi_{k-1}) \\ 0 \\ \varrho\Delta \\ 0 \end{bmatrix}, \quad (63)$$

where the state contains the position of the user $\mathbf{x}_{\text{UE},k} = [x_k, y_k, z_k]^\top$, heading ϖ_k , and clock bias B_k , ζ is the translation speed, set as 22.22 m/s, ϱ is the turn rate, set as $\pi/10$ rad/s, Δ is the sampling time interval, set as 0.5 s. The covariance of the process noise \mathbf{Q} is $\text{diag}[0.2 \text{ m}^2, 0.2 \text{ m}^2, 0 \text{ m}^2, 0.001 \text{ rad}^2, 0.2 \text{ m}^2]$. In practice, a good initial can be acquired by using external sensors, such as GPS or by applying snapshot-based localization algorithms, like [50], [51], which do not require a prior sensor state information. For simplicity, the UE is initialized at $[70.7285 \text{ m}, 0 \text{ m}, 0 \text{ m}, \pi/2 \text{ rad}, 300 \text{ m}]^\top$, and the initial covariance is $\text{diag}[0.3 \text{ m}^2, 0.3 \text{ m}^2, 0 \text{ m}^2, 0.0052 \text{ rad}^2, 0.3 \text{ m}^2]$. The UE has no prior knowledge on the map, apart from the BS location and the PPP intensities, which are initialized with weights $\eta_0^\xi = 1.5 \times 10^{-5}$ for both VA and SP.

To apply the proposed SLAM algorithm to the mmWave scenario, the Fisher information matrix of channel parameters [52] is used to determine the measurement covariances at each time step. The Fisher information matrix is based on the following model. The BS and the user are equipped with square planar arrays, and the numbers of BS and user antennas are 64 and 16, respectively. The operating carrier frequency is 28 GHz. We consider the OFDM pilot signals with $N_p = 16$ symbols, $S = 64$ subcarriers, 200 MHz bandwidth, corresponding to a pilot duration of 5.12 μs . The transmitted power and noise spectral density are set to 5 dBm and -174 dBm/Hz, respectively. Path loss is generated according to [52, eq. (45)], with reflection coefficient of surfaces set as 0.7, and the radar cross-section of SPs set as 50 m^2 .

SPs are only visible when SPs are in the field-of-view of the UE, which is set as 50 m, while the BS and VAs are always visible. The effect of different detection probabilities and clutter intensities can be found in [43, Table II]. The detection probabilities p_D^ξ are set to 0.9 for BS, VAs and visible SPs, and set to 0 for SPs which are out of FoV. The clutter intensity $c(z)$ is $1/(4 \times 200\pi^4)$, with 200 representing the sensing range and 1 representing the average of the number of clutter measurements. Moreover, we utilize pruning, and merging for Bernoullis to decrease the computational complexity, with thresholds set as 10^{-4} and 50, respectively.

B. Studied Methods and Performance Metrics

We evaluate the performance by comparing five algorithms:

- 1) the proposed EK-PMB SLAM filter with $\gamma = 10$;
- 2) the proposed EK-PMB SLAM filter with $\gamma = 1$;
- 3) the proposed EK-PMBM SLAM filter with $\gamma = 10$;
- 4) the RBP-PMBM SLAM filter [32];
- 5) the EK-PHD SLAM filter [33].

The mapping performance is measured by generalized optimal subpattern assignment (GOSPA) distance [53]

$$d_{\text{GOSPA}}(\hat{\mathcal{X}}, \mathcal{X}) = \min_{\gamma \in \Gamma} \left(\sum_{(i,j) \in \gamma} d^{q_p}(\hat{\mathbf{x}}_i, \mathbf{x}_j) + \frac{q_c^{q_p}}{q_a} (N_{\text{miss}} + N_{\text{false}}) \right)^{\frac{1}{q_p}}, \quad (64)$$

where Γ is the set of possible assignment set, N_{miss} is the number of miss detection, N_{false} is the number of false alarm. We set the cut-off distance $q_c = 20$, the cardinality penalty factor $q_a = 2$, the exponent factor $q_p = 2$, and the measure accuracy of the state estimates is evaluated by the root mean squared error (RMSE) over time and the mean absolute error (MAE) changing with time. Overall, 10 Monte Carlo (MC) simulations and 1000 MC simulations are performed for the fourth and the other algorithms, respectively, and the results are obtained by averaging over the different MC simulations. All the codes are written in MATLAB, and the simulations and experiments are run on a MacBook Pro (15-inch, 2019) with a 2.6 GHz 6-Core Intel Core i7 processor and 16 Gb memory.

C. Results and Discussion

Firstly, we study the mapping performance of the different SLAM methods. From Fig. 5 and Fig. 6, we observe that all algorithms can overcome the challenges i)–iv) in the radio SLAM problem, and provide good mapping results, as the GOSPA distances of all algorithms converge to small values in the end, due to the map is modeled as RFSs. We also observe that the first and third algorithms perform slightly better than the second algorithm, as the blue and black solid lines are lower than the red solid lines in Fig. 5 and Fig. 6. The reason is that the second algorithm takes the hard decision for the data association, which may pick up a wrong data association at some time steps, making the second algorithm not very stable and bringing additional error. The first and third algorithms perform similar, while the third algorithm has negligibly better performance. This is because even though the third algorithm keeps the PMBM format, there is usually a dominant MB,

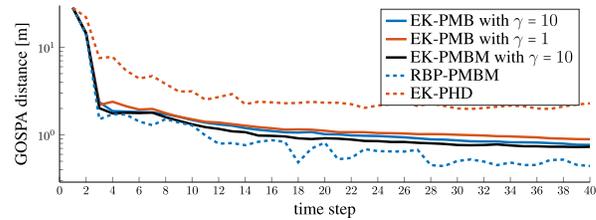


Fig. 5. Comparison of mapping performances for VAs among 5 algorithms.

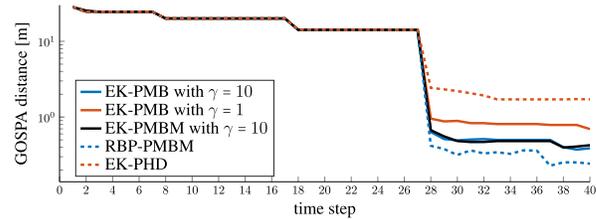


Fig. 6. Comparison of mapping performances for SPs among 5 algorithms.

making the rest MBs unimportant. Moreover, from Fig. 5 and Fig. 6, we can find that the first four algorithms outperform the EK-PHD SLAM filter, as the red dashed lines are higher than the others in both figures, which is because the PHD filter cannot enumerate all possible data associations explicitly. Based on the PMBM filter, the RBP-PMBM filter is slightly better than the first three algorithms, as the nonlinearity is solved by using enough particles, and the density conditioned on each particle keeps the PMBM format and no approximation of MBM to MB is executed.

Next, performance of the proposed SLAM filter in sensor state estimation is studied. Fig. 7 shows the RMSEs of the estimated sensor position, bias, and heading of the five SLAM filters, and Fig. 8 shows the MAE of the estimated position changing with time. Overall, all four PMB(M)-filter-based SLAM filters perform better than the EK-PHD SLAM filter, which is due to the PHD filter do not have explicit enumeration of the different data associations, and approximate the posterior density to a PPP. Within the four PMB(M)-filter-based SLAM filters, the RBP-SLAM filter performs the best, which is due to enough particles are used to solve the nonlinearity, all possible data associations are tracked to keep the PMBM format of the density conditioned on each particle. Because of not doing hard decision in data association and considering more than one MBs, the EK-PMB SLAM filter with $\gamma = 10$ and EK-PMBM SLAM filter perform slightly better than the EK-PMB SLAM filter with $\gamma = 1$. However, these two filters perform similar, as the sensor state prediction and update processes are the same in both algorithms, and there is usually one global hypothesis having the dominant weight every time step.

Another main advantage of the proposed SLAM filters is the low computational cost. In order to show this benefit, we measured the average execution time per time step of five SLAM filters, as shown in Table II. We measured that the proposed EK-PMB SLAM filter with $\gamma = 10$ takes 14.4 ms per time step, with the prediction and the update steps costs 0.34 ms and 14.1 ms, respectively, while the RBP-PMBM SLAM filter takes 71865.9 ms per time step, with the prediction and the update steps costs 581.8 ms and 71284.1 ms, respectively. It is obvious that the proposed EK-PMB algorithm is able

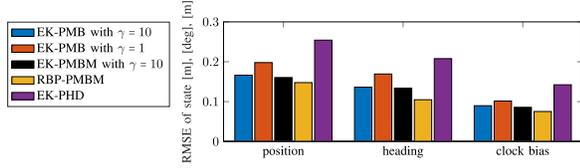


Fig. 7. Comparison of sensor state estimation among 5 algorithms.

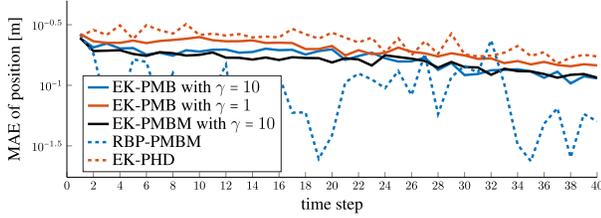


Fig. 8. Comparison of sensor position estimation among 5 algorithms.

TABLE II

AVERAGE COMPUTATION TIME IN MILLISECONDS OF THE PREDICTION AND UPDATE STEPS OF THE SLAM FILTERS

Filter	Prediction [ms]	Update [ms]	Total [ms]
EK-PMB, $\gamma = 10$	0.34	14.1	14.4
EK-PMB, $\gamma = 1$	0.34	11.9	12.3
EK-PMBM	0.34	41.5	41.9
RBP-PMBM	581.8	71284.1	71865.9
EK-PHD	1.2	2.6	3.8

to reduce the complexity, which is 5000 times faster than the RBP-PMBM SLAM filter approximately. One of the reasons is that the later uses 2000 particles to propagate the non-linearity of the UE state and there is a PMBM density presenting the map conditioned on each particle, while the former utilizes the first-order Taylor extension to approximate the nonlinear model and there is only one PMB(M) over the time. Another reason is that the inverse-CKF is used to create new births and CKF is used to update detected landmarks in the RBP-PMBM algorithm. In the other algorithms, the mean of the newly detected landmarks are directly estimated from the measurements, covariance of the birth components is computed from the Jacobians, and the UE and landmark states are jointly updated using the EK filter. The first algorithm takes slightly longer time than the EK-PMB SLAM filter with $\gamma = 1$, which is due to more data associations are considered in the update step. The proposed EK-PMBM SLAM filter takes much longer time than the proposed EK-PMB filter, because the PMBM format is kept and the number of global hypotheses rapidly increases. The EK-PHD SLAM filter takes a short time, that is because of the low complexity of the PHD filter compared to the PMBM filter, and the hard decisions are taken both on the data association and the type of the landmark. Although the EK-PMB does not have the lowest complexity or highest accuracy, online and real-time operation of the filter could still be guaranteed with good accuracy performance. Therefore, it offers a superior overall performance. In the practical usage of the EK-PMB SLAM filter, we could choose adaptive γ to reduce the complexity, keeping a relatively larger number of γ global hypotheses and marginalize over them when the data association problem is more challenging, for example, at the crossing of a street, and choose a relatively smaller γ , when the data association problem is less challenging, for example, if there are only a limited number of geometrically-separated landmarks.

VIII. CONCLUSION

mmWave bistatic sensing is of great relevance to 5G and Beyond 5G systems. When the sensing device is mobile, a SLAM problem needs to be solved to simultaneously localize the sensor and determine the locations of landmarks in the environment. In this paper, we have proposed two novel, low-complexity SLAM filters, based on the PMBM and PMB filters, which utilize an EK filter to perform a joint update of the sensor state and the landmark states. An extension to multiple models, relevant for the mmWave bistatic sensing problem, is also introduced. Via simulation results using reasonable mmWave signal parameters, we demonstrate that the proposed filters can attain good mapping and positioning performance, with very low complexity. Our results also demonstrated that the proposed SLAM filters can not only handle mapping and sensor state estimation simultaneously, but also distinguish the type of landmarks accurately, which is comparable to the performance of the RBP-PMBM SLAM filter. The high mapping and positioning performance and the low computational overhead of the proposed EK-PMB(M) SLAM filter are attractive for real-time execution of 5G and Beyond 5G mmWave SLAM algorithms.

Future work will include comparison with additional SLAM filters, the use of ray tracing data to validate the performance under more realistic operating conditions, the inclusion of high-dimensional channel estimation and optimized signal design (precoding and combining) to boost the localization accuracy, batch processing of measurements over a time window, as well as the extension to a multi-UE and multi-BS setup.

APPENDIX A

COMPUTATION OF DATA ASSOCIATION METRIC

A. Previously Detected Landmark i , Hypothesis j Detected Again

The value of $\rho_{k+1|k+1}^{j,i,p}$ is given by

$$\begin{aligned} \rho_{k+1|k+1}^{j,i,p} &= \int p_D(\tilde{\mathbf{s}}_{k+1}^{j,i}) f(\mathbf{z}_{k+1}^p | \tilde{\mathbf{s}}_{k+1}^{j,i}) f(\tilde{\mathbf{s}}_{k+1}^{j,i}) d\tilde{\mathbf{s}}_{k+1}^{j,i} \\ &= p_{D,k+1}^{j,i} \mathcal{N}(\mathbf{z}^p; \mathbf{h}(\tilde{\mathbf{m}}_{k+1|k}^{j,i}), \mathbf{S}_{k+1|k}^{j,i}), \end{aligned} \quad (65)$$

where $p_D(\tilde{\mathbf{s}}_{k+1}^{j,i})$ is assumed to be a constant over $\tilde{\mathbf{s}}_{k+1}^{j,i}$, with the value $p_{D,k+1}^{j,i} = p_D(\tilde{\mathbf{s}}_{k+1}^{j,i} = \tilde{\mathbf{m}}_{k+1|k}^{j,i})$, $\tilde{\mathbf{m}}_{k+1|k}^{j,i} = [\mathbf{m}_{k+1|k}^T, (\mathbf{u}_{k+1|k}^{j,i})^T]^T$ is the mean of $\tilde{\mathbf{s}}_{k+1}^{j,i}$, which is the joint state of the UE state and j, i th landmark state. The innovation covariance $\mathbf{S}_{k+1|k}^{j,i,p}$ is given by

$$\mathbf{S}_{k+1|k}^{j,i,p} = \mathbf{H}_{k+1|k}^{j,i} \tilde{\mathbf{P}}_{k+1|k}^{j,i} (\mathbf{H}_{k+1|k}^{j,i})^T + \mathbf{R}_{k+1}^p, \quad (66)$$

in which $\mathbf{H}_{k+1|k}^{j,i}$ represents the Jacobian of $\mathbf{h}(\cdot)$ with respect to $\tilde{\mathbf{s}}_{k+1}^{j,i}$, evaluated at $\tilde{\mathbf{s}}_{k+1}^{j,i} = \tilde{\mathbf{m}}_{k+1|k}^{j,i}$, with elements

$$\mathbf{H}_{k|k}^{j,i} = \left. \frac{\partial \mathbf{h}(\tilde{\mathbf{s}}_{k+1}^{j,i})}{\partial \tilde{\mathbf{s}}_{k+1}^{j,i}} \right|_{\tilde{\mathbf{s}}_{k+1}^{j,i} = \tilde{\mathbf{m}}_{k+1|k}^{j,i}}. \quad (67)$$

The covariance of $\tilde{\mathbf{s}}_{k+1}^{j,i}$ is given by $\tilde{\mathbf{P}}_{k+1|k}^{j,i} = \text{blkdiag}(\mathbf{P}_{k+1|k}, \mathbf{C}_{k+1|k}^{j,i})$.

B. Newly Detected Landmark

The value of $\rho_{\mathbf{B},k+1|k+1}^p$ is given by

$$\begin{aligned}\rho_{\mathbf{B},k+1|k+1}^p &= \int p_{\mathbf{D},k+1}^p \eta_{k+1|k} f(\mathbf{z}_{k+1}^p | \tilde{\mathbf{s}}_{k+1}^p) f(\tilde{\mathbf{s}}_{k+1}^p) d\tilde{\mathbf{s}}_{k+1}^p \\ &= p_{\mathbf{D},k+1}^p \eta_{k+1|k} \mathcal{N}(\mathbf{z}^p; \mathbf{h}(\tilde{\mathbf{m}}_{\mathbf{B},k+1|k}^p), \mathbf{S}_{\mathbf{B},k+1|k}^p),\end{aligned}\quad (68)$$

where $p_{\mathbf{D},k+1}^p$ is a constant with $p_{\mathbf{D}}(\tilde{\mathbf{s}}_{k+1}^p = \tilde{\mathbf{m}}_{\mathbf{B},k+1|k}^p)$, $\tilde{\mathbf{m}}_{\mathbf{B},k+1|k}^p = [\mathbf{m}_{k+1|k}^{\top} (\mathbf{u}_{\mathbf{B},k+1|k}^p)^{\top}]^{\top}$ is the mean of $\tilde{\mathbf{s}}_{k+1}^p$, which is the joint state of the UE state and the newly detected landmark state \mathbf{x}_{k+1}^p . The innovation covariance \mathbf{S}_{k+1}^p is given by

$$\mathbf{S}_{\mathbf{B},k+1|k}^p = \mathbf{H}_{\mathbf{B},k+1|k}^p \tilde{\mathbf{P}}_{\mathbf{B},k+1|k}^p (\mathbf{H}_{\mathbf{B},k+1|k}^p)^{\top} + \mathbf{R}_{k+1}^p, \quad (69)$$

where $\mathbf{H}_{\mathbf{B},k+1|k}^p$ is the Jacobian of $\mathbf{h}(\tilde{\mathbf{s}}_{k+1}^p)$, given by

$$\mathbf{H}_{\mathbf{B},k+1|k}^p = \left. \frac{\partial \mathbf{h}(\tilde{\mathbf{s}}_{k+1}^p)}{\partial \tilde{\mathbf{s}}_{k+1}^p} \right|_{\tilde{\mathbf{s}}_{k+1}^p = \tilde{\mathbf{m}}_{\mathbf{B},k+1|k}^p}. \quad (70)$$

The covariance of $\tilde{\mathbf{s}}_{k+1}^p$ is given by $\tilde{\mathbf{P}}_{\mathbf{B},k+1|k}^p = \text{blkdiag}(\mathbf{P}_{k+1|k}, \mathbf{C}_{\mathbf{B},k+1|k}^p)$. Generation of the mean and covariance of a new landmark is explained in Appendix A-C.

C. Birth Generation

The mean of newly detected landmark $\mathbf{u}_{\mathbf{B},k+1|k+1}^p$ can be estimated by [29, Appendix B]. The covariance of the newly detected landmark can be computed as follows. Consider a joint prior distribution of the sensor state and the newly detected landmark of the form

$$\mathcal{N}(\tilde{\mathbf{s}}_{k+1}; [\mathbf{m}_{k+1|k}^{\top}, (\mathbf{u}_{\mathbf{B},k+1|k+1}^p)^{\top}]^{\top}, \text{blkdiag}(\mathbf{P}_{k+1|k}, \mathbf{C}^{\text{prior}})),$$

with $\mathbf{C}^{\text{prior}} \rightarrow \infty \mathbf{I}$. Then, after applying an EK filter update, the marginal posterior covariance of the landmark is given by

$$\begin{aligned}\mathbf{C}_{\mathbf{B},k+1|k+1}^p &= ((\dot{\mathbf{H}}_{\mathbf{B},k+1|k}^p)^{\top} \\ &\quad \times (\dot{\mathbf{H}}_{\mathbf{B},k+1|k}^p \mathbf{P}_{k+1|k} (\dot{\mathbf{H}}_{\mathbf{B},k+1|k}^p)^{\top} + \mathbf{R}_{k+1}^p)^{-1} \dot{\mathbf{H}}_{\mathbf{B},k+1|k}^p)^{-1},\end{aligned}\quad (71)$$

with the Jacobians of $\mathbf{h}(\cdot)$

$$\begin{aligned}\dot{\mathbf{H}}_{\mathbf{B},k+1|k}^p &= \left. \frac{\partial \mathbf{h}(\mathbf{s}_{k+1} = \mathbf{m}_{k+1|k}, \mathbf{x}_{k+1}^p)}{\partial \mathbf{x}_{k+1}^p} \right|_{\mathbf{x}_{k+1}^p = \mathbf{u}_{\mathbf{B},k+1|k+1}^p}, \\ \dot{\mathbf{H}}_{\mathbf{B},k+1|k}^p &= \left. \frac{\partial \mathbf{h}(\mathbf{s}_{k+1}, \mathbf{x}_{k+1}^p = \mathbf{u}_{\mathbf{B},k+1|k}^p)}{\partial \mathbf{s}_{k+1}} \right|_{\mathbf{s}_{k+1} = \mathbf{m}_{k+1|k}}.\end{aligned}$$

REFERENCES

- [1] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmWave positioning for vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, Dec. 2017.
- [2] K. Witralsal *et al.*, "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 59–70, Mar. 2016.
- [3] O. Kanhere, S. Goyal, M. Beluri, and T. S. Rappaport, "Target localization using bistatic and multistatic radar with 5G NR waveform," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–7.
- [4] J. Andrew Zhang *et al.*, "An overview of signal processing techniques for joint communication and radar sensing," 2021, *arXiv:2102.12780*.
- [5] T. G. Reid *et al.*, "Localization requirements for autonomous vehicles," 2021, *arXiv:1906.01061*.
- [6] S.-W. Ko, H. Chae, K. Han, S. Lee, D.-W. Seo, and K. Huang, "V2X-based vehicular positioning: Opportunities, challenges, and future directions," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 144–151, Dec. 2021.
- [7] Z. Zhang, S.-W. Ko, R. Wang, and K. Huang, "Cooperative multi-point vehicular positioning using millimeter-wave surface reflection," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2221–2236, Apr. 2021.
- [8] M. Furkan Keskin, F. Jiang, F. Munier, G. Seco-Granados, and H. Wymeersch, "Optimal spatial signal design for mmWave positioning under imperfect synchronization," 2021, *arXiv:2105.07664*.
- [9] N. C. Luong, X. Lu, D. T. Hoang, D. Niyato, and D. I. Kim, "Radio resource management in joint radar and communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 780–814, 2nd Quart., 2021.
- [10] P. Kumari, A. Mezghani, and R. W. Heath, "JCR70: A low-complexity millimeter-wave Proof-of-Concept platform for a fully-digital SIMO joint communication-radar," *IEEE Open J. Veh. Technol.*, vol. 2, pp. 218–234, 2021.
- [11] L. Han and K. Wu, "24-GHz integrated radio and radar system capable of time-agile wireless communication and sensing," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 619–631, Mar. 2012.
- [12] L. Reichardt, C. Sturm, F. Grunhaupt, and T. Zwick, "Demonstrating the use of the IEEE 802.11 P Car-to-Car communication standard for automotive radar," in *Proc. 6th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2012, pp. 1576–1580.
- [13] S. H. Dokhanchi, M. R. B. Shankar, T. Stifter, and B. Ottersten, "OFDM-based automotive joint radar-communication system," in *Proc. IEEE Radar Conf.*, Apr. 2018, pp. 0902–0907.
- [14] 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.
- [15] 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.
- [16] Y. Bar-Shalom, *Tracking and Data Association*. New York, NY, USA: Academic, 1987.
- [17] A. Yassin, Y. Nasser, A. Y. Al-Dubai, and M. Awad, "MOSAIC: Simultaneous localization and environment mapping using mmWave without priori knowledge," *IEEE Access*, vol. 6, pp. 68932–68947, 2018.
- [18] M. Aladsani, A. Alkhateeb, and G. C. Trichopoulos, "Leveraging mmWave imaging and communications for simultaneous localization and mapping," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 4539–4543.
- [19] R. Mendrzik, F. Meyer, G. Bauch, and M. Z. Win, "Enabling situational awareness in millimeter wave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1196–1211, Sep. 2019.
- [20] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witralsal, 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, Dec. 2019.
- [21] E. Leitinger, S. Grebien, and K. Witralsal, "Multipath-based SLAM exploiting AOA and amplitude information," in *Proc. IEEE ICC Workshops*, Jul. 2019, pp. 1–7.
- [22] J. Mullane, B.-N. Vo, M. D. Adams, and B.-T. Vo, "A random-finite-set approach to Bayesian SLAM," *IEEE Trans. Robot.*, vol. 27, no. 2, pp. 268–282, Apr. 2011.
- [23] R. P. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2014.
- [24] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.
- [25] A. Garcia-Fernandez, J. Williams, L. Svensson, and Y. Xia, "A Poisson multi-Bernoulli mixture filter for coexisting point and extended targets," *IEEE Trans. Signal Process.*, vol. 69, pp. 2600–2610, 2021.
- [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] J. L. Williams, "An efficient, variational approximation of the best fitting multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 258–273, 2014.
- [28] M. Frohle, C. Lindberg, K. Granstrom, and H. Wymeersch, "Multisensor Poisson multi-Bernoulli filter for joint target-sensor state tracking," *IEEE Trans. Intell. Vehicles*, vol. 4, no. 4, pp. 609–621, Dec. 2019.

- [29] H. Kim, K. Granström, L. Gao, G. Battistelli, S. Kim, and H. Wymeersch, "5G mmWave cooperative positioning and mapping using multi-model PHD filter and map fusion," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3782–3795, Jun. 2020.
- [30] H. Kim, K. Granström, L. Gao, G. Battistelli, S. Kim, and H. Wymeersch, "Joint CKF-PHD Filter and map fusion for 5G multi-cell SLAM," in *Proc. Int. Conf. Commun.*, 2020, pp. 1–6.
- [31] Y. Ge, H. Kim, F. Wen, L. Svensson, S. Kim, and H. Wymeersch, "Exploiting diffuse multipath in 5G SLAM," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.
- [32] Y. Ge *et al.*, "5G SLAM using the clustering and assignment approach with diffuse multipath," *Sensors*, vol. 20, no. 16, p. 4656, Aug. 2020, doi: [10.3390/s20164656](https://doi.org/10.3390/s20164656).
- [33] O. Kaltiokallio, Y. Ge, J. Talvitie, H. Wymeersch, and M. Valkama, "MmWave simultaneous localization and mapping using a computationally efficient EK-PHD filter," in *Proc. Int. Conf. Inf. Fusion*, 2021, pp. 1–6.
- [34] H. Kim, H. Wymeersch, and S. Kim, "Multiple model Poisson multi-Bernoulli mixture for 5G mapping," in *Proc. Conf. Korean Commun. Soc.*, 2020, pp. 1–4.
- [35] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78729–78757, 2019.
- [36] J. Palacios, G. Bielsa, P. Casari, and J. Widmer, "Single and multiple-access point indoor localization for millimeter-wave networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1927–1942, Mar. 2019.
- [37] R. W. Heath, Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2017.
- [38] A. Richter, "Estimation of radio channel parameters: Models and algorithms," Ph.D. dissertation, Dept. Elect. Eng. Inf. Technol., Ilmenau Univ. Technol., Ilmenau, Germany, 2005.
- [39] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
- [40] K. Venugopal, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, Jr., "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.
- [41] A. B. Gershman, M. Rübbsamen, and M. Pesavento, "One- and two-dimensional direction-of-arrival estimation: An overview of search-free techniques," *Signal Process.*, vol. 90, no. 5, pp. 1338–1349, 2010.
- [42] F. Jiang, Y. Ge, M. Zhu, and H. Wymeersch, "High-dimensional channel estimation for simultaneous localization and communications," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.
- [43] Á. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: Direct derivation and implementation," *IEEE Trans. Aersp. Electron. Syst.*, vol. 54, no. 4, pp. 1883–1901, Aug. 2018.
- [44] M. Fatemi, K. Granström, L. Svensson, F. J. R. Ruiz, and L. Hammarstrand, "Poisson multi-Bernoulli mapping using Gibbs sampling," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2814–2827, Jun. 2017.
- [45] S. Särkkä, *Bayesian Filtering Smoothing*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
- [46] Á. F. García-Fernández, L. Svensson, M. R. Morelande, and S. Särkkä, "Posterior linearization filter: Principles and implementation using sigma points," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5561–5573, Oct. 2015.
- [47] H. Kim, K. Granstrom, S. Kim, and H. Wymeersch, "Low-complexity 5G SLAM with CKF-PHD filter," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 5220–5224.
- [48] K. G. Murty, "Letter to the editor—an algorithm for ranking all the assignments in order of increasing cost," *Oper. Res.*, vol. 16, no. 3, pp. 682–687, 1968.
- [49] G. Li, L. Kong, W. Yi, and X. Li, "Multiple model Poisson multi-Bernoulli mixture filter for maneuvering targets," *IEEE Sensors J.*, vol. 21, no. 3, pp. 3143–3154, Feb. 2021.
- [50] F. Wen and H. Wymeersch, "5G synchronization, positioning, and mapping from diffuse multipath," *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 43–47, Jan. 2021.
- [51] A. Kakkavas, H. Wymeersch, G. Seco-Granados, M. H. C. Garcia, R. A. Stirling-Gallacher, and J. A. Nossek, "Power allocation and parameter estimation for multipath-based 5G positioning," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7302–7316, Nov. 2021.
- [52] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados, and H. Wymeersch, "Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4939–4954, Aug. 2018.
- [53] A. S. Rahmathullah, A. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Proc. 20th IEEE Int. Conf. Inf. Fusion (Fusion)*, Jul. 2017, pp. 1–8.

Yu Ge (Student Member, IEEE) received the B.E. degree from Zhejiang University, Hangzhou, China, in 2017, and the M.Sc. degree from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2019. He is currently pursuing the Ph.D. degree with the Department of Electrical and Engineering, Chalmers University of Technology, Sweden. His research interests include wireless localization systems, simultaneous localization and mapping, joint communication and sensing, and multi-object tracking, particularly in 5G and beyond 5G scenarios.

Ossi Kaltiokallio is currently a Post-Doctoral Researcher with the Unit of Electrical Engineering, Tampere University, Finland. His research interests lie at the intersection of statistical signal processing and wireless networking for improving radio-based positioning and sensing technologies.

Hyowon Kim (Member, IEEE) is currently a Post-Doctoral Scholar with the Department of Electrical Engineering, Chalmers University of Technology, Sweden. His research interests include wireless positioning systems, simultaneous localization and mapping, and situational-aware communications.

Fan Jiang (Member, IEEE) is currently a Post-Doctoral Researcher with the Chalmers University of Technology, Sweden. His research interests include signal processing for wireless communications, massive multiple-input multiple-output (MIMO) and millimeter wave systems, and localization, tracking, and navigation framework.

Jukka Talvitie (Member, IEEE) is currently a University Lecturer with the Unit of Electrical Engineering, Tampere University, Finland. His research interests include signal processing for radio communications, positioning, and sensing, particularly concerning 5G and beyond mobile technologies.

Mikko Valkama (Fellow, IEEE) is currently a Full Professor and the Head of the Unit of Electrical Engineering, Tampere University, Finland. His general research interests include radio communications, radio localization, and radio-based sensing, with particular emphasis on 5G and 6G mobile radio networks.

Lennart Svensson (Senior Member, IEEE) is currently a Professor of signal processing with the Department of Electrical Engineering, Chalmers University of Technology, Sweden. His research interests include multi-object tracking and deep machine learning for medical and autonomous vehicle applications.

Sunwoo Kim (Senior Member, IEEE) is currently a Professor at Hanyang University, South Korea. He is also the Director of the 5G/Unmanned Vehicle Research Center, funded by the Ministry of Science and ICT of Korea. His research interests include wireless communication/positioning/localization, signal processing, vehicular networks, and location-aware communications.

Henk Wymeersch (Senior Member, IEEE) is currently a Professor of communication systems with the Department of Electrical Engineering, Chalmers University of Technology, Sweden. His current research interests include the convergence of communication and sensing in a 5G and beyond 5G context.