



## **5G mmWave Cooperative Positioning and Mapping Using Multi-Model PHD Filter and Map Fusion**

Downloaded from: <https://research.chalmers.se>, 2022-08-27 21:01 UTC

Citation for the original published paper (version of record):

Kim, H., Granström, K., Gao, L. et al (2020). 5G mmWave Cooperative Positioning and Mapping Using Multi-Model PHD Filter and Map Fusion. *IEEE Transactions on Wireless Communications*, 19(6): 3782-3795. <http://dx.doi.org/10.1109/TWC.2020.2978479>

N.B. When citing this work, cite the original published paper.

©2020 IEEE. Personal use of this material is permitted.

However, permission to reprint/republish this material for advertising or promotional purposes

# 5G mmWave Cooperative Positioning and Mapping using Multi-Model PHD Filter and Map Fusion

Hyowon Kim, *Student Member, IEEE*, Karl Granström, *Member, IEEE*, Lin Gao, Giorgio Battistelli, Sunwoo Kim, *Senior Member, IEEE*, and Henk Wymeersch, *Senior Member, IEEE*

**Abstract**—5G millimeter wave (mmWave) signals can enable accurate positioning in vehicular networks when the base station and vehicles are equipped with large antenna arrays. However, radio-based positioning suffers from multipath signals generated by different types of objects in the physical environment. Multipath can be turned into a benefit, by building up a radio map (comprising the number of objects, object type, and object state) and using this map to exploit all available signal paths for positioning. We propose a new method for cooperative vehicle positioning and mapping of the radio environment, comprising a multiple-model probability hypothesis density filter and a map fusion routine, which is able to consider different types of objects and different fields of views. Simulation results demonstrate the performance of the proposed method.

**Index Terms**—5G millimeter-wave, cooperative positioning and mapping, map fusion, probability hypothesis density, vehicular networks.

## I. INTRODUCTION

5G millimeter wave (mmWave) considers the potential of large bandwidths and large antenna arrays at the user and base station (BS), which enable accurate ranging, and angle of arrival (AOA) and angle of departure (AOD) estimation [1]. Thus, 5G mmWave positioning is expected to be the next-generation cellular positioning framework [2]. Thanks to the aforementioned advantages of 5G mmWave, measurements of multipath components can be related to objects in the physical environment [2]–[4]. Therefore, it is possible to build up a map, which, e.g., can be reused by other users to cooperatively improve their position estimates. Such maps can also remove the need for a priori synchronization of the users, and support single BS localization [5]. Mapping and positioning using 5G can be categorized as a simultaneous localization and mapping (SLAM) problem (for SLAM basics see, e.g., [6], [7]). In addition, 5G communication links can be used to share measurements, map, or location information, leading to

cooperative positioning and mapping. In 5G mmWave cooperative positioning and mapping (i.e., positioning and mapping based on measurements from 5G mmWave communication signals), there are three main tasks: (i) *Vehicle positioning*: determine the states (position, velocity, heading, clock bias) of the vehicles; (ii) *Environment mapping*: estimate the number of objects, as well as each object’s type and position; and (iii) *Cooperation*: fusing the collected the map information from the vehicles, and relay it to each vehicle. Even with the advantages of 5G mmWave, there are several challenges. First of all, due to the imperfect detection process at the receiver, there is a chance that objects that are inside the field-of-view (FoV) of vehicles are undetected. Secondly, the measurements may include false detections because of clutter, channel estimation error, and objects that are only visible in a short time. Third, since there are no origin-related tags on measurements, the data association must be addressed, in either in an explicit or an implicit manner.

To solve these general challenges of SLAM, a variety of approaches have been developed. They can be coarsely divided into three classes of methods (elaborated further in Section II): geometry-based [8]–[11], based on message passing [5], [12], [13], and based on random finite sets (RFSs) [14]–[20]. Among RFS-based methods for tracking and mapping, probability hypothesis density (PHD) filters [14] are widely used because they are computationally efficient alternatives that avoids explicit enumeration of the different data associations. Finally, in the considered SLAM problem, there are different types of measurements that are received by the vehicle, specifically measurements from the BS, scattering points, and reflecting surfaces [5], [13]. These different types of measurements should be handled in the mapping using a multiple model approach, with one model for each measurement type. In addition, each vehicle has a limited FoV, and has thus only observed the parts of the map space that has been inside the FoV. It follows that the multi-object densities only contain information for a subset of the map space, specifically the parts that have been inside the FoVs. For this reason, a direct application of the standard fusion [21]–[23] is not possible.

In this paper, we address the aforementioned challenges and propose a new method for 5G mmWave cooperative positioning and mapping that is based on RFS theory. The proposed method comprises a Rao-Blackwellized (RB) representation of the joint vehicle-map density, with particle filters for the vehicle location and a multiple-model PHD filter for the map (i.e., PHD-SLAM), similar to [24]. To update the particle weight, we theoretically derive a form of the set likelihood

H. Kim and S. Kim are with the Department of Electronics and Computer Engineering, Hanyang University, 04763 Seoul, South Korea (email: khw870511@hanyang.ac.kr; remero@hanyang.ac.kr).

K. Granström and H. Wymeersch are with the Department of Electrical Engineering, Chalmers University of Technology, 412 58 Gothenburg, Sweden (email: henkw@chalmers.se; karl.granstrom@chalmers.se).

L. Gao and G. Battistelli are with the Dipartimento di Ingegneria dell’Informazione (DINFO), Università degli Studi di Firenze, 50139 Firenze, Italy (email: lin.gao@unifi.it; giorgio.battistelli@unifi.it).

This work was supported, in part, by the Swedish Research Council, under grant 2018-03701, by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2017-0-01637) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), and by Samsung Research Funding & Incubation Center of Samsung Electronics under Project Number SRFC-IT-1601-09.

calculation. The PHD-SLAM filter is implemented by the Rao-Blackwellized particle filter (RBPF). Cooperation is handled using asynchronous map fusion through a modified arithmetic average (AA) approach, taking into account the different FoVs of the vehicles by designing fusion weights for AA map fusion by first decomposing the map space into mutually disjoint subsets, and then selecting weights for each subset. The BS performs map fusion with one vehicle at a time, so that through multiple interactions, each vehicle contributes to, and has access to, the global BS map. The main contributions of the paper are summarized as follows:

- For the propagation environment with multiple objects, multiple object types, and multipath measurements, we present and evaluate a novel solution to the cooperative positioning and mapping problem, based on a RB representation of the joint position-map density and a multiple-model PHD representation of the map, as well as a novel AA fusion rule.
- We derive a multiple-model PHD, which considers different measurements models, rather than different mobility models, generally considered in the literature.
- A new and theoretically sound method to update the vehicle state is provided, by deriving the closed form of the RFS-likelihood.
- By decomposing the overall source features into three mutually disjoint subsets, the fusion of each subset can be carried out independently. With this decomposition, we are able to flexibly design fusion weights for the different subsets and deal with non-overlapping sensor FoV.
- Through a Gaussian mixture implementation with online source code, the efficacy of the proposed filter and fusion approach is demonstrated in a two-vehicle scenario with 5G mmWave communication links, where all propagation paths are exploited and vehicles cooperatively map the environment, which is shown to speed up the mapping process.

The rest of the paper is organized as follows. In Section II, we discuss related works, and how our work compares to previous work. Section III describes the considered vehicular networks with 5G mmWave communication links and a problem formulation. Section IV introduces the multiple-model PHD-SLAM at the vehicle. In Section V, asynchronous map fusion is presented. Numerical results and discussions are reported in Section VI. Finally, Section VII concludes the paper.

*Notation:* Throughout this paper, we will use the following basic notations. Scalars are denoted by italic, e.g.,  $x$ . Vectors are indicated by the bold lower-case letters, e.g.,  $\mathbf{x}$ , and matrices are denoted by the bold upper-case letters, e.g.,  $\mathbf{X}$ . Transpose of both vector and matrix is represented by superscript T, e.g.,  $\mathbf{x}^T$  and  $\mathbf{X}^T$ . Random sets are denoted by calligraphic, e.g.,  $\mathcal{X}$ . We denote probability density functions (pdfs) and probability mass functions (pmfs) by  $f(\cdot)$  and  $p(\cdot)$ , respectively. We will use the following indexing: vehicle  $n$ , time step  $k$ , particle  $i$ , source type  $m$ , Gaussian mixture component  $j$ .

## II. RELATED WORKS

In this section, we introduce the previous works for handling the aforementioned challenges in SLAM, multiple-model object tracking, and map fusion, considering methods based on geometry, methods based on message passing, and methods based on RFS theory.

In the geometry-based SLAM methods, [8] formulates the SLAM problem using the geometric relation between observations, and a non-Bayesian estimator for the user location and extended Kalman filter for mapping are introduced. MmWave imaging for one single reflected path is utilized in [9]. Neither [8] nor [9] considered the unknown number of objects or the data association uncertainty; in our paper we handle both an unknown number of objects and unknown data association. The authors in [10], [11] develop SLAM methods that are applicable when the BS location is unknown. Their solutions require four anchors to localize a node regardless of whether the four anchors are physical BSs or VAs. Therefore, at least one physical anchor is required with its corresponding VA mirrored through reflective surfaces for SLAM and obstacle detection.

In the second category, a message passing-based estimator for position and orientation of the vehicle, as well as mapping objects is introduced in [12]. In [5], the clock bias of the vehicle is considered as an additional unknown, and scheduling method for effective message passing is introduced. In [5], [25], only reflecting surfaces are regarded as objects generating multipath signals and small objects are ignored. The authors in [13] consider scatterers as well as reflection surfaces. However, these message passing-based SLAM filters [5], [13], [25] do not include the data association uncertainty as part of the message passing problem, which we do in this paper. For effective data association in message passing-based SLAM, the joint probability data association scheme is dealt with in [26].

The third approach involves RFS theory, which is a powerful tool for probabilistic modelling of a set of objects with uncertainties on both cardinality and object states. RFSs have been used for SLAM problems, see [14]–[20]; these approaches mainly differ in terms of their representation of the object RFS and the required approximations. When using an RB SLAM density with an RFS based map, the position particle weights must be updated. In previous work, the RFS-likelihood for vehicle position update was approximated by using a dummy map, where the dummy map was either an empty map or a map with a single feature [15], [16]. In comparison, in this paper we use the theoretically exact RFS-likelihood. Unlike this work, where a PHD representation of the map RFS density is used, the authors of [18]–[20] represent the RFS density different types of densities (e.g., multi-Bernoulli (MB) and Poisson MB). However, those RFS densities require explicit data association, which is less computationally efficient. Further, in [20] the authors only consider the mapping problem, which is a simpler problem compared to SLAM.

In object tracking, multiple models are commonly used to handle maneuvering targets that switch between different types of motion, e.g., going straight forward or turning, see,

e.g., [27], however it is also possible to use multiple models to handle different types of measurements, see, e.g., [28], [29]. The objects can transition from one type of motion to another type, and this is commonly modelled using a jump Markov system, which can be handled, e.g., using the interactive multiple model (IMM) estimator [27]. However, in the considered 5G SLAM application, the objects do not transition from one type to another, and subsequently jump Markov system modelling is not necessary. There are multiple ways in which one can integrate multiple models into a PHD filter; an overview of different approaches to multiple model PHD filters was given in [30]. In this paper we adopt the proper jump Markov chain model in [24] which is based on an augmented object state consisting of the object state and the object type.

Fusion of the different map PHDs from different vehicles defined in the different and limited FoVs brings considerable challenges. Generally speaking, two frameworks can be employed in this situation: (i) centralized methods, where each vehicle directly sends the raw measurements to the fusion center to perform SLAM; (ii) decentralized methods, where each vehicle process the measurements and then share their posteriors with each other (or a fusion center) to perform density fusion. The centralized method is computationally intensive for the BS and treats the vehicles as decentralized sensors. To spread out the complexity over the network, the focus has been on decentralized methods. The most prevalent methods for multi-object density fusion consider generalized covariance intersection (GCI) [31], [32], which amounts to computing the intersection of information among densities. GCI cannot be directly applied in 5G cooperative SLAM because of the multi-object densities are defined for different FoVs. This difficulty is overcome in [17], where the PHD of each vehicle is initialized as a non-zero constant throughout the whole area of interest. Though it works well in fusing maps with different FoVs, it becomes troublesome when applied to large-scale scenarios since the storage and the propagation of the resulting PHD, which is non-zero everywhere, require both large memory and computational resources. In addition, GCI in [31], [32] extracts minimum information in fusing the maps. Thus we adopt AA which takes the union of involved densities and leads to minimum information loss [21]–[23]. However, there is a challenge of selecting the fusion weight in our scenario.

None of the above methods have been applied to the problem of 5G mmWave cooperative positioning and mapping.

### III. MODEL

In this section, we describe a vehicle, environment, and measurement models for the considered propagation environment with 5G mmWave communication links.

#### A. Vehicle Model

We consider a set of  $N$  vehicles, traversing a common environment, in communication with a common BS. The BS has a known and fixed location  $\mathbf{x}_{\text{BS}} \in \mathbb{R}^3$ . Each vehicle  $n$  has a dynamic state  $\mathbf{s}_k^{(n)}$  at time  $k$ . Time is discrete with sampling

interval  $\Delta$ . The state comprises the three-dimensional position  $\mathbf{v}_k^{(n)} = [x_k^{(n)}, y_k^{(n)}, z_k^{(n)}]^\top$ , heading  $\alpha_k^{(n)}$ , translation speed  $\zeta_k^{(n)}$ , turn-rate  $\rho_k^{(n)}$ , and clock bias  $B_k^{(n)}$ . Vehicle  $n$  has a known dynamic model with the transition density  $f(\mathbf{s}_k^{(n)} | \mathbf{s}_{k-1}^{(n)})$ . The vehicle dynamics follow a velocity motion model

$$\mathbf{s}_k^{(n)} = \mathbf{v}(\mathbf{s}_{k-1}^{(n)}) + \mathbf{q}_k, \quad (1)$$

where  $\mathbf{v}(\cdot)$  is a known transition function (see [33, Chapter 5], [34] and Section VI) and  $\mathbf{q}_k$  denotes the process noise, modeled as zero-mean Gaussian with known covariance  $\mathbf{Q}$ .

#### B. Environment Model

The environment is characterized by scattering points (SPs) and reflecting surfaces. A scattering point has an unknown three-dimensional location  $\mathbf{x}_{\text{SP}}$ , while a reflecting surface can be parameterized by a fixed virtual anchor (VA) location  $\mathbf{x}_{\text{VA}}$ , obtained by mirroring<sup>1</sup> the BS with respect to the surface. The details of geometric relation to the propagation environment are described in Appendix B.

#### C. Observation Model

A common model of a 5G mmWave received signal from the BS to vehicle  $n$  at time  $k$  is [35]

$$\mathbf{y}_k^{(n)}(t) = (\mathbf{W}_k^{(n)})^H \sum_{l=0}^{L_{k,l}^{(n)}} h_{k,l}^{(n)} \mathbf{a}_R(\boldsymbol{\theta}_{k,l}^{(n)}) \mathbf{a}_T^H(\boldsymbol{\phi}_{k,l}^{(n)}) \mathbf{b}_k(t - \tau_{k,l}^{(n)}) + \mathbf{n}_k^{(n)}(t), \quad (2)$$

where  $\mathbf{b}_k(t)$  is a transmitted signal (possibly precoded) to all the users,  $\mathbf{W}_k^{(n)}$  is a combining matrix,  $h_{k,l}^{(n)}$  is a complex path gain,  $\boldsymbol{\theta}_{k,l}^{(n)}$  is the AOA (also denoted as direction of arrival - DOA) in azimuth and elevation,  $\boldsymbol{\phi}_{k,l}^{(n)}$  is the AOD (also denoted as direction of departure - DOD) in azimuth and elevation,  $\tau_{k,l}^{(n)}$  is the time of arrival (TOA), and  $\mathbf{n}_k^{(n)}(t)$  is (possibly colored) noise. The vectors  $\mathbf{a}_T^H(\boldsymbol{\phi})$  and  $\mathbf{a}_R^H(\boldsymbol{\theta})$  are the steering vectors of the transmit and receive array, respectively. The AOA and TOA are measured in the frame of reference of the receiver, while the AOD is measured in the frame of reference of the transmitter. The path index  $l = 0$  is the line-of-sight (LOS) path, while the  $L_{k,l}^{(n)}$  remaining paths are non-LOS (NLOS) paths. The AOA, TOA, and AOD of each path has a geometric meaning, which depends on the location of the transmitter and receiver, as well as the points of incidence of the NLOS paths in the environment (see further). We further assume a channel estimation routine is present at the receiver, which provides, at time  $k$ , a set  $\mathcal{Z}_k^{(n)}$  of measurements with elements

$$\mathcal{Z}_{k,l}^{(n)} = \mathbf{h}(\mathbf{x}, \mathbf{s}_k^{(n)}, m) + \mathbf{w}_{k,l}^{(n)}, \quad (3)$$

where

$$\mathbf{h}(\mathbf{x}, \mathbf{s}_k^{(n)}, m) = [\tau_{k,l}^{(n)}, (\boldsymbol{\theta}_{k,l}^{(n)})^\top, (\boldsymbol{\phi}_{k,l}^{(n)})^\top]^\top \quad (4)$$

<sup>1</sup>Mathematically, the reflecting surface can be described by a point  $\mathbf{f}$  and a normal vector  $\mathbf{u}$ . With each reflecting surface we can associate a virtual anchor location  $\mathbf{x}_{\text{VA}} = \mathbf{P}\mathbf{x}_{\text{BS}} + \mathbf{t}$ , where  $\mathbf{P} = \mathbf{I}_3 - 2\mathbf{u}\mathbf{u}^\top$  is a Householder matrix and  $\mathbf{t} = 2\mathbf{f}^\top\mathbf{u}\mathbf{u}$  is a translation vector.

and  $\mathbf{w}_{k,l} \sim \mathcal{N}(\mathbf{0}, \Sigma_{k,l}^{(n)})$  for a certain number of paths  $l = 0, \dots, L_k^{(n)}$ . Here,  $m$  denotes the source type and  $\mathbf{x}$  the source location. We distinguish between three different sources: the BS, a VA, a SP and correspondingly have  $m \in \{\text{BS}, \text{SP}, \text{VA}\}$ . Both the source type and source location are unknown. We define  $\mathcal{X}$  as a random set of sources with entries  $(\mathbf{x}, m)$  with density  $f(\mathcal{X})$ . The functional form of  $\mathbf{h}(\mathbf{x}, \mathbf{s}_k^{(n)}, m)$  and of the corresponding likelihood function  $f(\mathbf{z}_l^{(n)} | \mathbf{x}, \mathbf{s}_k^{(n)}, m)$  is described in detail in Appendix B.

Finally, not all sources give rise to measurements and some measurements don't correspond to any fixed source. This is described as follows:

- *Missed detections:* A vehicle may only be able to detect a source if it is within the field of view. Hence, we introduce  $p_{D,k}(\mathbf{x}, \mathbf{s}_k^{(n)}, m) \in [0, 1]$  as the probability that a source of type  $m$  with location  $\mathbf{x}$  can give rise to a measurements when the vehicle is in state  $\mathbf{s}_k^{(n)}$ .
- *False alarms:* Some measurements in  $\mathcal{Z}_k^{(n)}$  may correspond to clutter (e.g., due to noise peaks that are detected as paths during channel estimation). We model this through the clutter intensity  $c(\mathbf{z})$ , which assumes that clutter is generated according to a Poisson point process.
- *Transient sources:* Measurements may also correspond to transient physical objects in the environment (e.g., a vehicle that moves). The corresponding measurements can be seen as a landmark that is visible only for a short time (a few seconds) and will be treated as a transient SP, meaning that it will appear and then disappear from the map.

We assume that  $\mathbf{h}(\mathbf{x}, \mathbf{s}_k^{(n)}, m)$ ,  $\Sigma_{k,l}^{(n)}$ ,  $p_{D,k}(\mathbf{x}, \mathbf{s}_k^{(n)}, m)$  and  $c(\mathbf{z})$  are known to vehicle  $n$ .

#### D. Problem Formulation

Given a certain prior  $f(\mathbf{s}_0^{(n)})$ , our goal is to track the state of the vehicles' states and build a common map of the environment (VAs and SPs). To solve this problem, we first detail the SLAM algorithm running locally on each vehicle and then go on to detail the map fusion at the BS.

### IV. LOCAL PROCESSING: MULTIPLE-MODEL PHD-SLAM

In this section, we describe a local multiple-model PHD filter at each vehicle. We will consider a single vehicle and thus drop the vehicle index  $n$ .

#### A. Approach

The map state will be modeled as a multi-object Poisson process (MPP), which is fully characterized by its PHD (first-order statistical moment), hence the conditional map PHD is propagated rather than its density. Further, in order to distinguish the type of each source, the discrete state  $m$  is also included in conditional map PHD. We rely on the standardized RB approach, whereby the vehicle state trajectory is represented by particles, and PHDs conditioned on each particle are maintained. Hence, the data structure at the end of time  $k-1$  consists of (i) a list of  $I$  particles  $\mathbf{s}_{0:k-1}^i$  with particle

weights  $\omega_{k-1|k-1}^i \geq 0$ ,  $\sum_i \omega_{k-1}^i = 1$ ; (ii) for each particle, the PHD  $D_{k-1|k-1}(\mathbf{x}, m | \mathbf{s}_{0:k-1}^i)$ ,  $m \in \{\text{BS}, \text{SP}, \text{VA}\}$ . We initialize  $D_{0|0}(\mathbf{x}, \text{VA} | \mathbf{s}_0^i) = D_{0|0}(\mathbf{x}, \text{SP} | \mathbf{s}_0^i) = 0$  and  $D_{0|0}(\mathbf{x}, \text{BS} | \mathbf{s}_0^i) = \delta(\mathbf{x} - \mathbf{x}_{\text{BS}})$ . As a shorthand, we will denote  $D_{k-1|k-1}(\mathbf{x}, m | \mathbf{s}_{0:k-1}^i)$  as  $D_{k-1|k-1}^i(\mathbf{x}, m)$ ,  $p_{D,k}(\mathbf{x}, \mathbf{s}_k^{(n,i)}, m)$  as  $p_{D,k}^i(\mathbf{x}, m)$ , and  $\mathbf{h}(\mathbf{x}, \mathbf{s}_k^{(n,i)}, m)$  as  $\mathbf{h}^i(\mathbf{x}, m)$ .

We are now ready to describe the recursive formulation of the PHD filter (Section IV-C), followed by a practical Gaussian mixture implementation (Section IV-D).

#### B. Basics on PHDs

An RFS  $\mathcal{X}$  is characterized by its set density  $f(\mathcal{X})$ , which in turn depends on the cardinality distribution and the cardinality-conditioned joint distributions [36]

$$f(\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}) = p(n) \sum_{\pi} f_n(\mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(n)}), \quad (5)$$

where  $p(n)$  is the cardinality distribution evaluated in  $n$ , the sum goes over all permutations  $\pi$  of the set  $\{1, \dots, n\}$ , and  $f_n(\cdot)$  is standard vector density of  $n$  elements. The set integral is defined as

$$\begin{aligned} & \int g(\mathcal{X}) \delta \mathcal{X} \\ &= g(\emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int g(\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \dots d\mathbf{x}_n. \end{aligned} \quad (6)$$

If  $\delta_{\mathcal{X}}(\mathbf{x}) = \sum_{\mathbf{w} \in \mathcal{X}} \delta_{\mathbf{w}}(\mathbf{x})$ , where  $\delta(\cdot)$  indicates the delta Dirac function, the PHD associated with  $\mathcal{X}$  is the function [14]

$$D(\mathbf{x}) = \int f(\mathbf{x} \cup \mathcal{W}) \delta \mathcal{W} = \int \delta_{\mathcal{X}}(\mathbf{x}) f(\mathcal{X}) \delta \mathcal{X}, \quad (7)$$

which has as property that for any region  $S$  in the underlying state space,  $\int_S D(\mathbf{x}) d\mathbf{x}$  is the expected number of elements in  $S$ . Note that  $D(\mathbf{x}) \geq 0$  is generally not normalized, and generally does not provide a unique representation of an RFS density (multiple RFS densities may have the same PHD). One exception is the Poisson Point Process (PPP) RFS, which has a single parameter, called the PPP intensity, which is equal to the PPP PHD. In this case the RFS density is defined as follows, see, e.g., [15],

$$f(\mathcal{X}) = \frac{\prod_{\mathbf{x} \in \mathcal{X}} D(\mathbf{x})}{\exp(\int D(\mathbf{x}) d\mathbf{x})}. \quad (8)$$

A common representation of a PHD is through a Gaussian mixture (GM)

$$D(\mathbf{x}) \approx \sum_{j=1}^J \gamma^j \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}^j, \boldsymbol{\Sigma}^j), \quad (9)$$

where  $\sum_{j=1}^J \gamma^j$  represents the expected number of elements, with locations  $\boldsymbol{\mu}^j$ . The GM representation allows closed form computation of the PHD mapping filter under certain conditions.

#### C. General Formulation

The filter comprises two steps: the prediction step, which accounts for the motion model (1), and the update step, which accounts for the measurement set  $\mathcal{Z}_k$ .

1) *Prediction*: The PHD prediction is [37]

$$D_{k|k-1}^i(\mathbf{x}, m) = D_{k-1|k-1}^i(\mathbf{x}, m) + b^i(\mathbf{x}, m), \quad (10)$$

where  $b^i(\mathbf{x}, m)$  is a birth process, indicating where and with which intensities we expect sources of type  $m$  to appear. Note that  $b^i(\mathbf{x}, \text{BS}) = 0$  since the BS location is already known. For the vehicle state prediction, we use the process model (1) to generate predicted trajectories,  $\mathbf{s}_{0:k}^i$ , where  $\mathbf{s}_k^i \sim f(\mathbf{s}_k | \mathbf{s}_{k-1}^i)$ , with  $\omega_{k|k-1}^i = \omega_{k-1|k-1}^i$ .

2) *Measurement Update*: Given the measurement set  $\mathcal{Z}_k$  at time  $k$ , we update the 3 PHDs for each particle as follows: for the BS PHD,  $D_{k|k}^i(\mathbf{x}, \text{BS}) = D_{k|k-1}^i(\mathbf{x}, \text{BS})$ , which for the VA and SP PHDs [37],

$$D_{k|k}^i(\mathbf{x}, m) = D_{k|k-1}^i(\mathbf{x}, m)(1 - p_D^i(\mathbf{x}, m)) + \sum_{\mathbf{z} \in \mathcal{Z}_k} \frac{\nu^i(\mathbf{z}, \mathbf{x}, m)}{c(\mathbf{z}) + \sum_{m'} \int \nu^i(\mathbf{z}, \mathbf{x}', m') d\mathbf{x}'}, \quad (11)$$

where we recall that  $c(\mathbf{z})$  is the clutter intensity,  $p_D^i(\mathbf{x}, m)$  is shorthand for the detection probability of a source of type  $m$  at location  $\mathbf{x}$  (given the current vehicle state  $\mathbf{s}_k^i$ ) and

$$\nu^i(\mathbf{z}, \mathbf{x}, m) = p_D^i(\mathbf{x}, m) f(\mathbf{z} | \mathbf{x}, \mathbf{s}_k^i, m) D_{k|k-1}^i(\mathbf{x}, m). \quad (12)$$

The first term in (11) corresponds to the update when no measurement comes from the source at location  $\mathbf{x}$  (as it is out of the field of view), while the second term corresponds to the update when there is a measurement. In the latter case, the measurement can come from clutter, which is accounted for in the denominator.

In parallel, using the same measurement set  $\mathcal{Z}_k$ , we update the vehicle state distribution, but updating the weights

$$\begin{aligned} \omega_{k|k}^i &= \omega_{k|k-1}^i f(\mathcal{Z}_k | \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) \\ &= \omega_{k|k-1}^i \int f(\mathcal{Z}_k | \mathbf{s}_{0:k}^i, \mathcal{X}, \mathcal{Z}_{1:k-1}) f(\mathcal{X} | \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) \delta \mathcal{X}, \end{aligned} \quad (13)$$

where  $\delta \mathcal{X}$  refers to a set integral. To avoid numerical problems, rather than working with the particle weights  $\omega_{k|k}^i$ , we work with the log-weights  $\ell_{k|k}^i = \log \omega_{k|k}^i$ . The log-weight update is  $\ell_{k|k}^i = \ell_{k|k-1}^i + \log f(\mathcal{Z}_k | \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1})$ .

In previous work on PHD-SLAM [15], [16], the integral in the weight update (13),

$$\int f(\mathcal{Z}_k | \mathbf{s}_{0:k}^i, \mathcal{X}, \mathcal{Z}_{1:k-1}) f(\mathcal{X} | \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) \delta \mathcal{X} \quad (14)$$

was approximated using a ‘‘dummy’’ map  $\hat{\mathcal{X}}$ ; in [15, Sec. 4.E] it is proposed to use either an empty map or a map with a single feature, in [16, Sec. 3.C] a map with multiple features is used. In this paper, we use the exact expression for the integral in (13). With a PPP prior  $f(\mathcal{X} | \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1})$  and a point object measurement model,  $f(\mathcal{Z}_k | \mathbf{s}_{0:k}^i, \mathcal{X}, \mathcal{Z}_{1:k-1})$ , the solution to the integral is

$$f(\mathcal{Z}_k | \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) = \prod_{\mathbf{z} \in \mathcal{Z}_k} \left( c(\mathbf{z}) + \sum_{m'} \int \nu^i(\mathbf{z}, \mathbf{x}', m') d\mathbf{x}' \right). \quad (15)$$

This result follows as a special case of the more general PMBM update, see details in [38, Sec. 3.B.2], as derived in Appendix C. Note that (15) is easily evaluated during the map update step.

## D. Gaussian Mixture Implementation

While the expression above provide a solution to the SLAM problem, considering multiple source types and limited field of view, a practical implementation requires several choices and approximations to be made. In this section, we provide a GM implementation, inspired by [15]. The proposed implementation has a complexity cost that scales as  $\mathcal{O}(I \times N_{\text{model}} \times J \times |\mathcal{Z}|)$  per vehicle and per time step. Here,  $I$  denotes the number of particles,  $N_{\text{model}}$  is the number of models,  $J$  is the number of Gaussian mixture components per model, and  $|\mathcal{Z}|$  is the number of measurements per time step. Note that for single-model SLAM  $N_{\text{model}} = 1$ , for mapping only  $I = 1$ , and for localization only,  $N_{\text{model}} \times J = 1$ .

Using a set of  $I$  particles, the multiple-model PHD-SLAM density at time  $k - 1$  is expressed as

$$\{\mathbf{s}_{k-1|k-1}^i, \omega_{k-1|k-1}^i, D_{k-1|k-1}^i(\mathbf{x}, m)\}_{i=1}^I, \quad (16)$$

where  $D_{k-1|k-1}^i(\mathbf{x}, m)$  will be described by a GM. In this section, we will detail the implementation of map prediction (10), map update (11), and vehicle state update (13).

1) *Map Prediction (10)*: The map PHD  $D_{k-1|k-1}^i(\mathbf{x}, m)$  at the end of time step  $k - 1$  is assumed to be of a GM

$$\begin{aligned} D_{k-1|k-1}^i(\mathbf{x}, m) &= \sum_{j=1}^{J_{k-1|k-1}^i(m)} \gamma_{k-1|k-1}^{i,j}(\mathbf{x}, m) \\ &\quad \times \mathcal{N}(\mathbf{x}; \mathbf{x}_{k-1|k-1}^{i,j}(\mathbf{x}, m), \mathbf{P}_{k-1|k-1}^{i,j}(\mathbf{x}, m)), \end{aligned} \quad (17)$$

where  $J_{k-1|k-1}^i(m)$  is the number of Gaussians in the map PHD for the source type  $m$ , and  $\gamma_{k-1|k-1}^{i,j}(\mathbf{x}, m) \geq 0$ ,  $\mathbf{x}_{k-1|k-1}^{i,j}(\mathbf{x}, m)$ , and  $\mathbf{P}_{k-1|k-1}^{i,j}(\mathbf{x}, m)$  are respectively the weight, mean, and covariance of  $j$ -th Gaussian. Note that  $\sum_j \gamma_{k-1|k-1}^{i,j}(\mathbf{x}, m)$  is not necessary to be equal to 1. Similarly, the birth process PHD  $b_k^i(\mathbf{x}, m)$ , which is determined as the measurement arrives, is also represented as a GM

$$b_k^i(\mathbf{x}, m) = \sum_{j(\mathbf{z})=1}^{J_{b,k}^i(m)} \gamma_{b,k}^{i,j}(\mathbf{x}, m) \mathcal{N}(\mathbf{x}; \mathbf{x}_{b,k}^{i,j}(\mathbf{z}, m), \mathbf{P}_{b,k}^{i,j}(\mathbf{z}, m)), \quad (18)$$

where  $j(\mathbf{z})$  denote the measurement index corresponding to the measurement  $\mathbf{z}$ , and  $J_{b,k}^i(m)$  is the number of Gaussians in the birth process PHD, which is equal to the number of elements in measurement  $\mathcal{Z}_k$ . Hence,  $\mathbf{x}_{b,k}^{i,j}(\mathbf{z}, m)$  and  $\mathbf{P}_{b,k}^{i,j}(\mathbf{z}, m)$  are respectively the mean and covariance of Gaussians which indicate the statistics of the birth location. Hence, the prediction map PHD  $D_{k|k-1}^i(\mathbf{x}, m)$  in (10) is given by the sum of  $D_{k-1|k-1}^i(\mathbf{x}, m)$  and birth process PHD  $b_k^i(\mathbf{x}, m)$ , which is a new GM, denoted by

$$\begin{aligned} D_{k|k-1}^i(\mathbf{x}, m) &= \sum_{j=1}^{J_{k|k-1}^i(m)} \gamma_{k|k-1}^{i,j}(\mathbf{x}, m) \mathcal{N}(\mathbf{x}; \mathbf{x}_{k|k-1}^{i,j}(\mathbf{x}, m), \mathbf{P}_{k|k-1}^{i,j}(\mathbf{x}, m)), \end{aligned} \quad (19)$$

where  $J_{k|k-1}^i(m) = J_{k-1|k-1}^i(m) + J_{b,k}^i(m)$ .

An important practical consideration is how to set the weights, means and covariances of the birth process. We have

found that in order to have an implementation that is able to successfully incorporate new information, it was crucial to let these depend on the measurements at time  $k$ , so that  $\nu^i(\mathbf{z}, \mathbf{x}, m)$  in (12) takes on significant values [39]. The main idea is, for each measurement  $\mathbf{z}$ , to generate a birth for each source type  $m$ . Using the inverse sigma point of the cubature Kalman filter (CKF) [40], the mean  $\mathbf{x}_{b,k}^{i,j}(m)$  and covariance  $\mathbf{P}_{b,k}^{i,j}(m)$  of these births can be determined with respect to the measurement  $\mathbf{z}$  and source type  $m$ , details of which are described in Appendix D. The weight  $\gamma_{k|k-1}^{i,j}(m)$  is set to a low constant value, depending on the application. Complexity can be reduced by not generating sources with low likelihood (e.g., when the generated source location is out of the field of view so that  $p_{\mathcal{D}}^i(\mathbf{x}_{k|k-1}^{i,j}(m), m)$  close to zero).

2) *Map Update (11)*: In order to evaluate the update in closed form, we utilize two approximations: the first approximation involves the detection probability and the second approximating the Bayes update. We note that since the births are generated from the measurements, their detection probability should be 1 and they should not be updated with their corresponding measurements (i.e., the likelihood for a birth and its corresponding measurement is set to 1) [39]. For the existing targets, on the other hand, we consider an adaptive detection probability  $p_{\mathcal{D},k}^{i,j}(m)$ . We may set this adaptive detection probability to the expected value (i.e.,  $p_{\mathcal{D},k}^{i,j}(m) = \mathbb{E}\{p_{\mathcal{D},k}^{i,j}(\mathbf{x}, m)\}$  where the expectation is over  $\mathbf{x}$  with density  $\mathcal{N}(\mathbf{x}; \mathbf{x}_{k|k-1}^{i,j}(m), \mathbf{P}_{k|k-1}^{i,j}(m))$ ) or to a robust value to avoid weight decrease of objects that were previously detected (i.e.,  $p_{\mathcal{D},k}^{i,j}(m) = \min_{\mathbf{x} \in \mathcal{S}} p_{\mathcal{D},k}^{i,j}(\mathbf{x}, m)$ , where  $\mathcal{S}$  could be the highest density region of  $\mathcal{N}(\mathbf{x}; \mathbf{x}_{k|k-1}^{i,j}(m), \mathbf{P}_{k|k-1}^{i,j}(m))$  containing a large fraction (e.g., 95%) PHD filters are known for being sensitive to both missed detections and false alarms, due to the approximation of the multi-object density as a Poisson RFS. The Poisson cardinality has high variance, so a missed detection leads to a drastic decrease in the landmark weight (except when the detection probability is very low), while clutter often leads to false landmarks. Hence, if we don't want to lose the sources due to missed detections, we must set the detection probability to low values, at a cost of a higher sensitivity to clutter (false landmarks).

The second approximation is related to the Bayes update, and allows a closed-form evaluation of (12)

$$\nu^i(\mathbf{z}, \mathbf{x}, m) = \sum_{j=1}^{J_{k|k-1}^i(m)} \nu^{i,j}(\mathbf{z}, \mathbf{x}, m). \quad (20)$$

We will denote  $j(\mathbf{z})$  the birth index  $j$  corresponding to measurement  $\mathbf{z}$ . Considering a particular measurement  $\mathbf{z}$ , then when where for a birth  $j = j(\mathbf{z})$ , [39]

$$\nu^{i,j}(\mathbf{z}, \mathbf{x}, m) = \gamma_{k|k-1}^{i,j}(m) \mathcal{N}(\mathbf{x}; \mathbf{x}_{k|k-1}^{i,j}(m), \mathbf{P}_{k|k-1}^{i,j}(m)), \quad (21)$$

while for any  $j \neq j(\mathbf{z})$

$$\begin{aligned} & \nu^{i,j}(\mathbf{z}, \mathbf{x}, m) \\ &= p_{\mathcal{D},k}^{i,j}(m) \gamma_{k|k-1}^{i,j}(m) \mathcal{N}(\mathbf{x}; \mathbf{x}_{k|k-1}^{i,j}(m), \mathbf{P}_{k|k-1}^{i,j}(m)) \\ & \times \mathcal{N}(\mathbf{z}; \mathbf{h}^i(\mathbf{x}, m), \mathbf{\Sigma}_k(\mathbf{z})) \\ & \approx p_{\mathcal{D},k}^{i,j}(m) \gamma_{k|k-1}^{i,j}(m) \mathcal{N}(\mathbf{x}; \mathbf{x}_{k|k}^{i,j}(m), \mathbf{P}_{k|k}^{i,j}(m)) \\ & \times \mathcal{N}(\mathbf{z}; \mathbf{h}^i(\mathbf{x}_{k|k-1}^{i,j}(m), m), \mathbf{S}_{\mathbf{z}\mathbf{z},k}^{i,j}(m)), \end{aligned} \quad (22)$$

where  $\mathbf{\Sigma}_k(\mathbf{z})$  is the measurement covariance of measurement  $\mathbf{z}$ . The approximation in (22) follows from the CKF, described in CKF update of Algorithm 2, 3 in Appendix A.

3) *Vehicle Update (13)*: Computing (15) in the log-domain, log-weight update is<sup>2</sup>

$$\ell_{k|k}^i = \ell_{k|k-1}^i + \sum_{\mathbf{z} \in \mathcal{Z}} \log(c(\mathbf{z})) + \sum_{m'} \int \nu^i(\mathbf{z}, \mathbf{x}', m') d\mathbf{x}'. \quad (23)$$

We note that the closed form evaluation in (20)–(22) is used for evaluating (15).

Finally, we denote the estimated vehicle state and estimated vehicle location by  $\hat{\mathbf{s}}_k$  and  $\hat{\mathbf{v}}_k$ , respectively. The vehicle state is estimated by the sample mean,  $\hat{\mathbf{s}}_k = \sum_{i=1}^I \omega_{k|k}^i \mathbf{s}_k^i$ , and the estimated vehicle location  $\hat{\mathbf{v}}_k \in \mathbb{R}^3$  is extracted from  $\hat{\mathbf{s}}_k \in \mathbb{R}^7$ . We denote the resampled particle set by  $\{\bar{\mathbf{s}}_k^i, \bar{\omega}_{k|k}^i\}_{i=1}^I$ ,  $\bar{\omega}_{k|k}^i = 1/I \quad \forall i$ .

## V. GLOBAL PROCESSING: MAP FUSION

In this section, we consider fusion of information from different vehicles. As mentioned in Section I, we aim to leverage the local processing capabilities of each vehicle, as described in Section IV. To allow simple processing, we consider the case where vehicles asynchronously communicate with the BS, where each communication involves an uplink transmission and a downlink transmission. Hence, a vehicle may only sporadically communicate with the BS. At the beginning of a time slot  $k$ , the BS maintains maps  $D_{k|k-1}^{\text{BS}}(\mathbf{x}, m)$  in GM form, for  $m \in \{\text{SP}, \text{VA}\}$ .

### A. Uplink Transmission

At time  $k$ , a certain vehicle  $n$  determines a particle average PHD

$$D_{k|k}^{(n)}(\mathbf{x}, m) = \sum_{i=1}^I \omega_{k|k}^{i,(n)} D_{k|k}^{i,(n)}(\mathbf{x}, m), \quad (24)$$

to which we apply pruning and merging<sup>3</sup> for implementation, described in [37, Table II]. The vehicle sends the average PHD

<sup>2</sup>The logarithm term  $\log(c(\mathbf{z})) + \sum_{m'} \int \nu^i(\mathbf{z}, \mathbf{x}', m') d\mathbf{x}'$  is implemented by first introducing  $W^{i,j}(\mathbf{z}) = c(\mathbf{z}) + \sum_{m'} \int \nu^{i,j}(\mathbf{z}, \mathbf{x}', m') d\mathbf{x}'$ , sorting these (for a given  $i$ ) from large to small and re-indexing. Then  $\log(c(\mathbf{z})) + \sum_{m'} \int \nu^i(\mathbf{z}, \mathbf{x}', m') d\mathbf{x}' = \log W^{i,(1)}(\mathbf{z}) + \log(1 + \sum_{l>1} W^{i,(l)}(\mathbf{z})/W^{i,(1)}(\mathbf{z}))$ .

<sup>3</sup>Gaussian components (mean, covariance, normalized weight) for all particles are imported as the input since our PHD uses the particle approach.

as well as a representation of the accumulated FoV since the last communication instant  $k^\dagger$

$$\mathcal{F}_k^{(n)}(m) = \{\mathbf{x} : \exists k' \in (k^\dagger, k], \max_i \{p_{D,k'}(\mathbf{x}, \hat{\mathbf{s}}_{k'}^i, m)\} \geq \gamma_D\}, \quad (25)$$

where  $\gamma_D$  is a detection threshold (close to 1).

### B. Map Fusion at the BS

The BS receives  $D_{k|k}^{(n)}(\mathbf{x}, m)$  and fuses with the local map  $D_{k|k-1}^{BS}(\mathbf{x}, m)$ . There are two common approaches for fusing two PHDs  $D_1(\mathbf{x})$  and  $D_2(\mathbf{x})$

$$D_\beta(\mathbf{x}) = \begin{cases} (D_1(\mathbf{x}))^{\beta_1} (D_2(\mathbf{x}))^{\beta_2} & \text{GCI fusion [41]} \\ \beta_1 D_1(\mathbf{x}) + \beta_2 D_2(\mathbf{x}) & \text{AA fusion [21], [22]} \end{cases} \quad (26)$$

where  $\beta_1, \beta_2$  are the fusion weights which satisfy  $\beta_1, \beta_2 \geq 0$  and  $\beta_1 + \beta_2 = 1$ . The values of  $\beta_1$  and  $\beta_2$  are set to reflect the relative contributions of  $D_1$  and  $D_2$ . From the information-theoretic point of view, both approaches lead to a fused PHD that can be interpreted as the (respectively left- and right-) centroid of the PHDs to be fused when the Kullback-Leibler divergence is used as discrepancy measure [23]. However, the two fusion rules have different characteristics. For instance, due to its multiplicative nature, GCI tends to preserve only objects present in all the PHDs to be fused and, hence, is preferable when the PHDs to be fused originate from sensors having a high clutter rate. On the other hand, AA is more suitable for higher rates of missed detections since it tends to preserve all the detected objects. Thus, it is clear that GCI fusion is hard to combine with sensors that have limited FoVs, since, by definition, the probability of detection is equal to zero outside each vehicle FoV [15]. For this reason, we choose to use AA.

Let  $\mathcal{I}_k^{(n)}(m)$  denote the region of the map space where vehicle  $n$  has information. Notice that such a region includes the accumulated FoV  $\mathcal{F}_k^{(n)}(m)$  but also all the regions containing the components of the map  $D_{k|k}^{(n)}(\mathbf{x}, m)$ . In fact, each vehicle can have information also outside its own FoV thanks to the downlink transmission from the BS to the vehicles. Similarly, let  $\mathcal{I}_k^{BS}(m)$  denote the region of the map space where the BS has information. Accordingly, the source sets  $\mathcal{X}_k$  is divided into three mutually disjoint sets:

- (a) the set  $\bar{\mathcal{X}}_k$  of sources on which both the BS and vehicle  $n$  have information, belonging to the intersection  $\mathcal{I}_k^{(n)}(m) \cap \mathcal{I}_k^{BS}(m)$ ;
- (b) the set  $\tilde{\mathcal{X}}_k$  of sources on which only vehicle  $n$  has information, belonging to the relative complement  $\mathcal{I}_k^{(n)}(m) \setminus \mathcal{I}_k^{BS}(m)$ ; and
- (c) the set  $\hat{\mathcal{X}}_k$  of sources on which only the BS has information, belonging to the relative complement  $\mathcal{I}_k^{BS}(m) \setminus \mathcal{I}_k^{(n)}(m)$ .

It is clear that the three sets  $\bar{\mathcal{X}}_k$ ,  $\hat{\mathcal{X}}_k$ , and  $\tilde{\mathcal{X}}_k$  have to be considered separately because an actual information fusion is possible only for the sources belonging to  $\bar{\mathcal{X}}_k$ . To this end,

we can exploit the property that the PHD of the union of independent Poisson RFSs is the sum of the PHDs and write

$$D_{k|k-1}^{BS}(\mathbf{x}, m) = \bar{D}_{k|k-1}^{BS}(\mathbf{x}, m) + \tilde{D}_{k|k-1}^{BS}(\mathbf{x}, m) + \hat{D}_{k|k-1}^{BS}(\mathbf{x}, m), \quad (27)$$

$$D_{k|k}^{(n)}(\mathbf{x}, m) = \bar{D}_{k|k}^{(n)}(\mathbf{x}, m) + \tilde{D}_{k|k}^{(n)}(\mathbf{x}, m) + \hat{D}_{k|k}^{(n)}(\mathbf{x}, m), \quad (28)$$

where  $\bar{D}_{k|k-1}^{BS}(\mathbf{x}, m)$  and  $\bar{D}_{k|k}^{(n)}(\mathbf{x}, m)$  refer to  $\bar{\mathcal{X}}_k$ ,  $\tilde{D}_{k|k-1}^{BS}(\mathbf{x}, m)$  and  $\tilde{D}_{k|k}^{(n)}(\mathbf{x}, m)$  refer to  $\tilde{\mathcal{X}}_k$ , and finally  $\hat{D}_{k|k-1}^{BS}(\mathbf{x}, m)$  and  $\hat{D}_{k|k}^{(n)}(\mathbf{x}, m)$  refer to  $\hat{\mathcal{X}}_k$ . As can be concluded immediately, by construction, we have  $\tilde{D}_{k|k}^{(n)}(\mathbf{x}, m) = 0$  and  $\hat{D}_{k|k-1}^{BS}(\mathbf{x}, m) = 0$ . Then the idea is to carry out the AA fusion independently on the three sets  $\bar{\mathcal{X}}_k$ ,  $\tilde{\mathcal{X}}_k$ , and  $\hat{\mathcal{X}}_k$  and, accordingly, the fusion result at the BS takes the form

$$D_{k|k}^{BS}(\mathbf{x}, m) = \bar{D}_{k|k}^{BS}(\mathbf{x}, m) + \tilde{D}_{k|k}^{BS}(\mathbf{x}, m) + \hat{D}_{k|k}^{BS}(\mathbf{x}, m), \quad (29)$$

where

$$\bar{D}_{k|k}^{BS}(\mathbf{x}, m) = \bar{\beta}_p \bar{D}_{k|k-1}^{BS}(\mathbf{x}, m) + \bar{\beta}_a \bar{D}_{k|k}^{(n)}(\mathbf{x}, m), \quad (30)$$

$$\tilde{D}_{k|k}^{BS}(\mathbf{x}, m) = \tilde{\beta}_p \tilde{D}_{k|k-1}^{BS}(\mathbf{x}, m) + \tilde{\beta}_a \tilde{D}_{k|k}^{(n)}(\mathbf{x}, m). \quad (31)$$

$$\hat{D}_{k|k}^{BS}(\mathbf{x}, m) = \hat{\beta}_p \hat{D}_{k|k-1}^{BS}(\mathbf{x}, m) + \hat{\beta}_a \hat{D}_{k|k}^{(n)}(\mathbf{x}, m). \quad (32)$$

Here different fusion weights have been assigned for the three disjoint sets since each fusion is supposed to be independently carried out. Such an additional flexibility allows us to take into account directly in the fusion rule the decomposition of the source sets. Due to the fact that both the BS and the vehicle have information on the common set  $\bar{\mathcal{X}}_k$ , the uniform weights  $\bar{\beta}_p = \bar{\beta}_a = 1/2$  can be adopted for  $\bar{\mathcal{X}}_k$ , while for  $\tilde{\mathcal{X}}_k$  (the source set for which the vehicle has no information) the weight can be set to  $\tilde{\beta}_a = 0$ ,  $\tilde{\beta}_p = 1$ , and for  $\hat{\mathcal{X}}_k$  (the source set for which the BS has no information) the weight can be set to  $\hat{\beta}_a = 1$ ,  $\hat{\beta}_p = 0$ .<sup>4</sup>

In practice, the decompositions in (27) and (28) are not known and have to be approximately determined directly from the densities  $D_{k|k}^{(n)}(\mathbf{x}, m)$  and  $D_{k|k-1}^{BS}(\mathbf{x}, m)$ . We conclude this section by presenting a procedure for deriving such decompositions when the average vehicle map and the BS map are expressed as GMs

$$D_{k|k}^{(n)}(\mathbf{x}, m) = \sum_{j_a=1}^{J_k} \hat{\gamma}_k^{j_a} \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_k^{j_a}, \hat{\mathbf{P}}_k^{j_a}), \quad (33)$$

$$D_{k|k-1}^{BS}(\mathbf{x}, m) = \sum_{j_p=1}^{J_{p,k}} \hat{\gamma}_k^{j_p} \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_k^{j_p}, \hat{\mathbf{P}}_k^{j_p}). \quad (34)$$

<sup>4</sup>Setting  $\hat{\beta}_a = 1/2$  would lead to a reduction by a half of the weights of the Gaussian components related to the sources belonging to  $\hat{\mathcal{X}}_k$ , which hence would become lower than the threshold for declaring the presence of the source (i.e. 0.5). Then, if these sources are not detected again, the fused PHD will not be able to declare them anymore. Such a situation can happen to newly detected landmarks located at the border of the FoV or when the vehicles move fast. It has been pointed out in [42] that the fusion weight should be selected according to the probability that the next prediction made using its corresponding density outperform predictions made from all other individual densities.



We notice preliminarily that, in this case, the fusion rule (29) can be rewritten as

$$D_{k|k}^{\text{BS}}(\mathbf{x}, m) = \sum_{j_a=1}^{\hat{J}_k} \beta_{a,j_a} \hat{\gamma}_k^{j_a} \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_k^{j_a}, \hat{\mathbf{P}}_k^{j_a}) + \sum_{j_p=1}^{\hat{J}_{p,k}} \beta_{p,j_p} \hat{\gamma}_k^{j_p} \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_k^{j_p}, \hat{\mathbf{P}}_k^{j_p}), \quad (35)$$

where  $\beta_{a,j_a}$  takes value 1/2 when the component  $j_a$  is assigned to  $\overline{D}_{k|k}^{(n)}(\mathbf{x}, m)$  or value 1 when it is assigned to  $\widehat{D}_{k|k}^{(n)}(\mathbf{x}, m)$  and, similarly,  $\beta_{p,j_p}$  takes value 1/2 when the component  $j_p$  is assigned to  $\overline{D}_{k|k-1}^{\text{BS}}(\mathbf{x}, m)$  or value 1 when it is assigned to  $\widetilde{D}_{k|k-1}^{\text{BS}}(\mathbf{x}, m)$ .

To set the values of  $\beta_{a,j_a}$  and  $\beta_{p,j_p}$  (i.e. to approximately determine the decompositions in (27) and (28)), we use the Mahalanobis cost metric to compute the distance between the components of the two PHDs. Specifically, we introduce two distance metrics

$$\mathcal{D}_{j_a \rightarrow j_p} = (\hat{\mathbf{x}}_k^{j_a} - \hat{\mathbf{x}}_k^{j_p})^{\text{T}} (\hat{\mathbf{P}}_k^{j_a})^{-1} (\hat{\mathbf{x}}_k^{j_a} - \hat{\mathbf{x}}_k^{j_p}), \quad (36)$$

$$\mathcal{D}_{j_p \rightarrow j_a} = (\hat{\mathbf{x}}_k^{j_p} - \hat{\mathbf{x}}_k^{j_a})^{\text{T}} (\hat{\mathbf{P}}_k^{j_p})^{-1} (\hat{\mathbf{x}}_k^{j_p} - \hat{\mathbf{x}}_k^{j_a}), \quad (37)$$

where  $\mathcal{D}_{j_a \rightarrow j_p}$  is the Mahalanobis distance between  $\hat{\mathbf{x}}_k^{j_p}$  and the distribution  $\mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_k^{j_a}, \hat{\mathbf{P}}_k^{j_a})$  while  $\mathcal{D}_{j_p \rightarrow j_a}$  is the Mahalanobis distance between  $\hat{\mathbf{x}}_k^{j_a}$  and  $\mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}_k^{j_p}, \hat{\mathbf{P}}_k^{j_p})$ . With these metrics, we compute binary proximity matrices  $\mathbf{C}^a \in \mathbb{B}^{\hat{J}_k} \times \mathbb{B}^{\hat{J}_{p,k}}$  and  $\mathbf{C}^p \in \mathbb{B}^{\hat{J}_k} \times \mathbb{B}^{\hat{J}_{p,k}}$ , initialized as zeros. Then, we cycle through all pairs  $(j_a, j_p)$ : if  $\mathcal{D}_{j_p \rightarrow j_a} < \gamma_{\text{UP}}$ , then we set  $C_{j_a, j_p}^p = 1$ . If  $\mathcal{D}_{j_a \rightarrow j_p} < \gamma_{\text{UP}}$ , then we set  $C_{j_a, j_p}^a = 1$ . Here,  $\gamma_{\text{UP}}$  is a threshold on the Mahalanobis distances. Finally, we determine the  $\beta$  values for each component. We initialize  $D_{k|k}^{\text{BS}}(\mathbf{x}, m) = 0$  and add entries as follows:

- 1) *Assign equal weights for matches:* If  $\max(C_{j_a, j_p}^p, C_{j_a, j_p}^a) = 1$ , the components  $j_a$  of  $D_{k|k}^{(n)}(\mathbf{x}, m)$  and  $j_p$  of  $D_{k|k-1}^{\text{BS}}(\mathbf{x}, m)$  are deemed to belong to the region of common information  $\mathcal{I}_k^{(n)}(m) \cap \mathcal{I}_k^{\text{BS}}(m)$ , since both of them can find their respective correspondences in the other map PHDs. Thus we set  $\beta_{p, j_p} = \beta_{a, j_a} = 1/2$ . Note that a source  $j_p$  could be matched with multiple sources  $j_a$  and vice versa.
- 2) *Find unmatched sources in the BS map:* If  $\sum_{j_a} C_{j_a, j_p}^p = 0$ , then source  $j_p$  in the BS map could not be associated with any entry in the vehicle map. Then, recalling that  $\mathcal{I}_k^{(n)}(m)$  always contains the accumulated FoV  $\mathcal{F}_k^{(n)}(m)$ , we add the source  $j_p$  to  $D_{k|k}^{\text{BS}}(\mathbf{x}, m)$  with weight

$$\beta_{p, j_p} = \begin{cases} 1 & \hat{\mathbf{x}}_k^{j_p} \notin \mathcal{F}_k^{(n)}(m), \\ 1/2 & \hat{\mathbf{x}}_k^{j_p} \in \mathcal{F}_k^{(n)}(m). \end{cases} \quad (38)$$

This ensures that sources outside the FoV are kept. However, sources that suddenly appear could possibly be false alarms, therefore sources in the field of view that were not seen by vehicle  $n$  are reduced in weight and will gradually disappear from the BS map.

- 3) *Find unmatched sources in the vehicle map:* If  $\sum_{j_p} C_{j_a, j_p}^p = 0$ , then source  $j_a$  in the vehicle map could

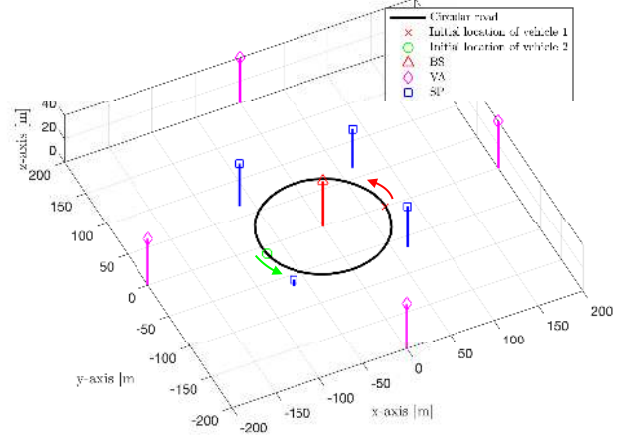


Fig. 1. Vehicle trajectories and common map of the environment (1 BS, 4 VAs, and 4 SPs). Two vehicles are initially located at opposite sides of  $[0, 0, 0]$  m, and move along with counterclockwise. Two vehicles have observed same objects (i.e., explored area) when  $k$  is about 40.

not be associated with any entry in the BS map. Hence, the component  $j_a$  is deemed to belong to the relative complement  $\mathcal{I}_k^{(n)}(m) \setminus \mathcal{I}_k^{\text{BS}}(m)$  and we set  $\beta_{a, j_a} = 1$ .

The BS map is then found by adding all these sources with their corresponding weights as in (35) and by applying pruning and merging so as to keep the number of components limited. Clearly, at the beginning, when the BS map is empty, instead of applying (35) the BS map is simply overwritten with the vehicle map.

### C. Downlink Transmission

The BS sends the computed  $D_{k|k}^{\text{BS}}(\mathbf{x}, m)$  to vehicle  $n$ . This map can contain new information for the vehicle as it contains all the information provided by *other* vehicles between times  $k^\dagger$  and  $k$ . Hence, the vehicle overwrites the fused map to  $D_{k|k}^i(\mathbf{x}, m)$  as follows

$$D_{k|k}^i(\mathbf{x}, m) = D_{k|k}^{\text{BS}}(\mathbf{x}, m) \quad \forall i. \quad (39)$$

While this leads to a lack of diversity among the maps across the particles, it has the distinct benefit of being a low-complexity solution.

## VI. NUMERICAL RESULTS

### A. Simulation Setup

We consider a three dimensional (3D) vehicular network where two moving vehicles are on a circular road with a BS, four VAs, and four SPs as shown in Fig. 1. Such a scenario is sufficiently representative to demonstrate the efficacy of the proposed PHD filter and map fusion, though actual performance in localization and mapping will depend on the chosen scenario. The details of the scenario are shown and available in [43]. During  $K = 40$  time steps, the vehicle

states are evolved with the dynamics model (1) as discussed in Section III-A, with

$$\mathbf{v}(\mathbf{s}_{k-1}^{(n)}) \quad (40)$$

$$= \mathbf{s}_{k-1}^{(n)} + \begin{bmatrix} \frac{\zeta_{k-1}^{(n)}}{\rho_{k-1}^{(n)}} \left( \sin(\alpha_{k-1}^{(n)} + \rho_{k-1}^{(n)} \Delta) - \sin \alpha_{k-1}^{(n)} \right) \\ \frac{\zeta_{k-1}^{(n)}}{\rho_{k-1}^{(n)}} \left( -\cos(\alpha_{k-1}^{(n)} + \rho_{k-1}^{(n)} \Delta) + \cos \alpha_{k-1}^{(n)} \right) \\ 0 \\ \rho_{k-1}^{(n)} \Delta \\ \mathbf{0}_3 \end{bmatrix},$$

where  $\mathbf{0}_\psi$  denotes a column vector of  $\psi$  zeros,  $\Delta$  is the sampling time and  $\mathbf{q}_k$  denotes the process noise, modeled as zero-mean Gaussian with covariance  $\mathbf{Q} = \text{diag}[\sigma_x^2, \sigma_y^2, 0, \sigma_\alpha^2, \mathbf{0}_2, \sigma_B^2]$ . The vehicle states are initialized as  $\mathbf{s}_0^{(1)} = [70.7285, 0, 0, \pi/2, 22.22, \pi/10, 300]^T$  and  $\mathbf{s}_0^{(2)} = [-70.7285, 0, 0, \pi/2, -22.22, \pi/10, 300]^T$ , with units m, m, m, rad, m/s, rad/s, and m. The time interval  $\Delta$  is set to 0.5 s. The process noise standard deviations are set to  $\sigma_x = 0.2$  m,  $\sigma_y = 0.2$  m,  $\sigma_\alpha = 0.001$  rad, and  $\sigma_B = 0.2$  m. The initial prior of the vehicle state follows zero-mean Gaussian distribution with 0.3 m standard deviation for both x and y location, 0.3 rad for the vehicle heading, 0.3 m the bias. The longitudinal velocity  $\zeta_k^{(n)}$ , rotational velocity  $\rho_k^{(n)}$  are assumed to be known. The measurement covariance matrix  $\Sigma_{k,l}^{(n)}$  is diagonal, and is set to  $\text{diag}(10^{-2} \text{ m}^2, 10^{-4} \text{ rad}^2, 10^{-4} \text{ rad}^2, 10^{-4} \text{ rad}^2)$ . To mitigate the effect of the errors in the CKF (due to the non-invertible nonlinearity), we replace the measurement  $\Sigma_{k,l}^{(n)}$  with  $\Sigma_{\text{PHD}} = 9 \times \Sigma_{k,l}^{(n)}$  in (47) for the birth process, and in the CKF update of Algorithm 1 for the map correction. A BS is located at  $[0, 0, 40]^T$  m. Four VAs are located at  $[200, 0, 40]^T$  m,  $[-200, 0, 40]^T$  m,  $[0, 200, 40]^T$  m,  $[0, -200, 40]^T$  m. Four SPs are located at  $[65, 65, z_{\text{SP}}]^T$  m,  $[-65, 65, z_{\text{SP}}]^T$  m,  $[-65, -65, z_{\text{SP}}]^T$  m, and  $[65, -65, z_{\text{SP}}]^T$  m, where  $z_{\text{SP}} \sim \mathcal{U}(0, 40)$ . The SPs are only visible when the distance between the SP and vehicle is within the FoV range  $r_{\text{FoV}} = 50$  m, while VAs are always visible. The detection probability  $p_D$  is set to 0.9 within the FOV. In the Gaussian representation of the birth process (18), we consider the birth weight  $\gamma_{b,k}^{i,j(\mathbf{z})}(m) = 1.5 \times 10^{-5}$  for  $m = \{\text{VA}, \text{SP}\}$ . For the clutter intensity  $c(\mathbf{z})$ , we consider the average of the number of clutter measurements (following Poisson distribution)  $\lambda = 1$ , and the maximum sensing range  $R_{\text{max}} = 200$  m, so  $c(\mathbf{z}) = \lambda / (4R_{\text{max}}\pi^4)$ . We utilized the pruning and merging in [37, Table II], and also used its parameter notations as follows: truncation threshold  $T$ ; merging threshold  $U$ ; and maximum allowable number of Gaussians  $J_{\text{max}}$ . We considered  $T = 10^{-4}$ ,  $U = 49$ , and  $J_{\text{max}} = 50$ . The object detection parameters are set to as follows: the VA detection threshold  $T_{\text{VA}} = 0.7$ ; the SP detection threshold  $T_{\text{SP}} = 0.55$ . We consider an asynchronous map fusion where each vehicle communicates with the BS every 4 time steps, with vehicle 1 starting at time 10 and vehicle 2 at time 12. Each vehicle's state was represented by  $I = 2000$  particles, and simulation results were obtained by averaging over  $N_{\text{mc}} = 20$  Monte Carlo runs (we observed no significant performance differences when increasing  $N_{\text{mc}}$  from

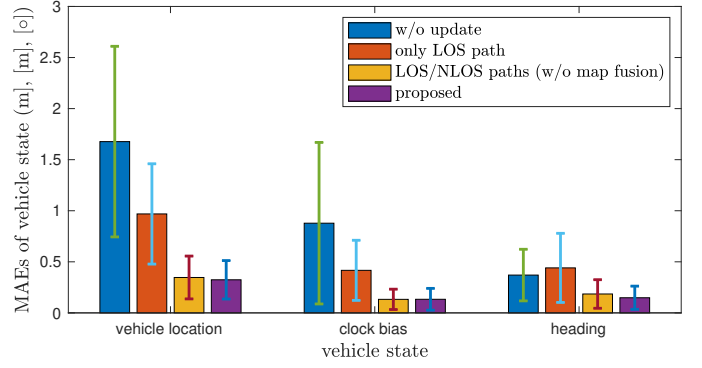


Fig. 2. MAEs and RMSE bars of the vehicle states (vehicle location, clock bias, and heading).

10 to 20). Complete source code is available at <https://github.com/HyowonKim-P1/5GmmWavePHDFilterMapFusion>.

### B. Performance Metric

To demonstrate the efficacy of the method and support the contributions of this paper described in Section I. The performances of the vehicle state estimation and the mapping of the environment are evaluated, over all Monte Carlo runs during the steady-state operation, which was determined to be after  $k = 20$ . For the vehicle state estimation, we compute the mean absolute error (MAE) on each component (location, clock bias, heading), along with root mean square error (RMSE) bars. For the mapping, we compute the average of the generalized optimal subpattern assignment (GOSPA) distance [44], as follows (removing the time index  $k$  and the source type index  $m$ ). We denote  $\mathcal{L} = \{(\mathbf{x}_1, m_1), \dots, (\mathbf{x}_M, m_M)\}$  and  $\hat{\mathcal{L}} = \{(\hat{\mathbf{x}}_1, \hat{m}_1), \dots, (\hat{\mathbf{x}}_{\hat{M}}, \hat{m}_{\hat{M}})\}$  by the set of sources (of type  $m$  at time  $k$ ) and its estimated set, respectively. The FoV was not considered in the GOSPA distance metric in order to evaluate the map fusion performance. Then, the GOSPA is defined as

$$d_{\text{GOSPA}}^{(q_c, q_a, q_p)} \quad (41)$$

$$= \left( \min_{\mathcal{M} \in \Pi_{|\hat{\mathcal{L}}|}} \sum_{e=1}^M d^{q_c}(\hat{\mathbf{x}}_e, \hat{\mathbf{x}}_{\mathcal{M}(e)}) + \frac{q_c}{q_a} (\hat{M} - M) \right)^{\frac{1}{q_p}},$$

where  $\Pi_*$  indicates the permutations of set  $*$ , cut off distance  $q_c = 20$ ,  $q_a = 2$ , power parameter  $q_p = 2$ , and  $d^{q_c}(\mathbf{x}_e, \hat{\mathbf{x}}_{\mathcal{M}(e)}) = \min(\|\hat{\mathbf{x}}_e - \hat{\mathbf{x}}_{\mathcal{M}(e)}\|, q_c)$ .

### C. Results and Discussions

1) *Vehicle tracking*: Fig. 2 shows the MAE and RMSE bars of the estimated vehicle location, clock bias, and heading with respect to (w.r.t.) the four cases as follows: i) only performing the vehicle state prediction without the update step; ii) using only the measurement from the LOS path; iii) proposed PHD filter for positioning and mapping per vehicle from Section IV; and iv) proposed PHD filtering and map fusion from Section V. In case i), the accuracy of the estimated vehicle state gradually increases and demonstrates the need for measurements in the considered scenario. Case ii) can be

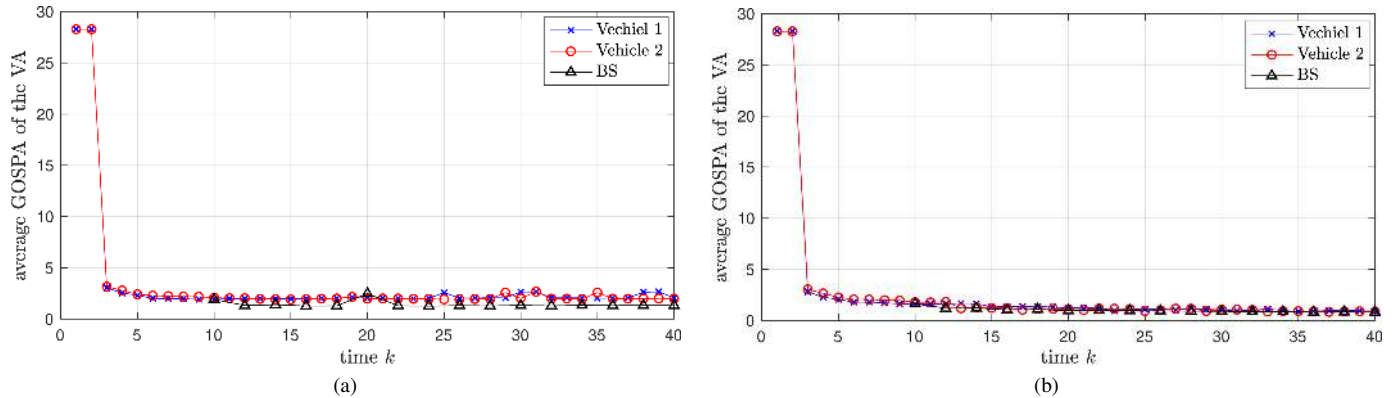


Fig. 3. Average GOSPA of the VA for (a) map fusion per vehicle and uplink transmission to the BS (b) map fusion per vehicle and uplink transmission to the BS as well as downlink transmission to the vehicles.

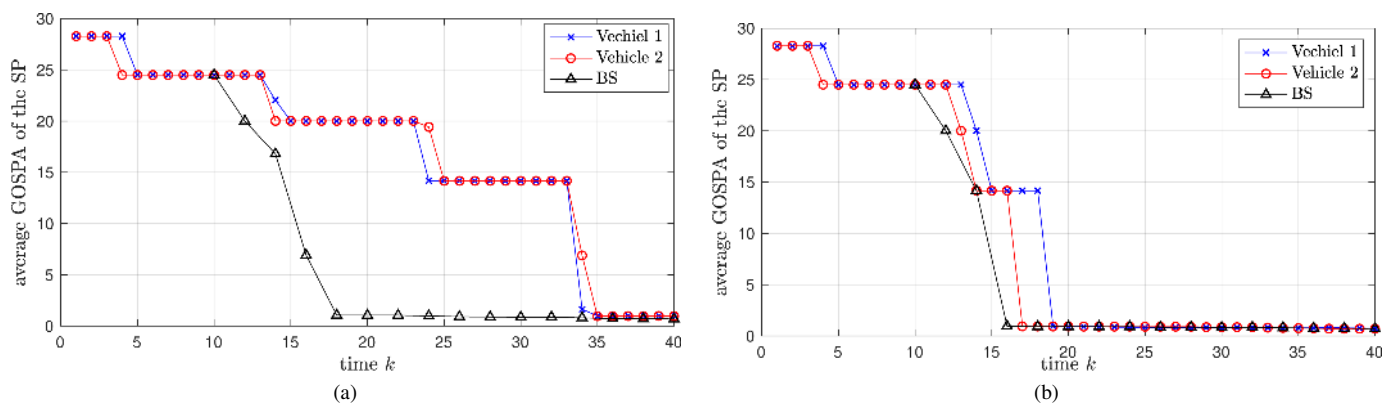


Fig. 4. Average GOSPA of the SP for (a) map fusion per vehicle and uplink transmission to the BS (b) map fusion per vehicle and uplink transmission to the BS as well as downlink transmission to the vehicles.

considered a best-case, without any objects in the environment and a clear LOS at all times. We see that the performance is significantly improved compared to case i). In case iii), the performance is much better than case ii) showing the benefit of NLOS information, even with unknown source association. In case iv), despite the reduced map diversity (see Section V-C) the performance is not reduced compared to case iii), but there are only marginal performance gains either. This is due to the specific scenario, where the vehicles independently are able to localize themselves well. In addition, since the VAs are always visible for both vehicles and can be mapped accurately, the main cooperative localization gain can come from the SPs. Since SPs have variable detection probability, they provide only limited information for the vehicles' positions.

2) *Mapping*: Fig. 3 shows the average GOSPA of the VA map, with Fig. 3a considering the case of the local PHD filter and map fusion at the BS, but no downlink transmission, while Fig. 3b presents the performance of the proposed PHD filter and map fusion with downlink transmission. Comparing Fig. 3a and Fig. 3b, there is only a small, little or no benefit the downlink transmissions, as both vehicles have all VAs in their FoV at all time. This is in contrast to the SP map, where Fig. 4 reports the GOSPA results. In Fig. 4a, we see that the GOSPA per vehicle goes down as they move in the environment. The GOSPA at the BS is reduced faster, as it can benefit from

the information of all vehicles. In Fig. 4b, we note that when the BS sends back the map to the vehicles over the downlink, each vehicle can benefit from the measurements of the other vehicle, so that the GOSPA is reduced faster on the vehicle maps as well.

## VII. CONCLUSIONS

It is expected that the framework of cooperative positioning and mapping with the proposed multiple-model PHD filter and map fusion is essential for vehicular positioning. In this paper, we proposed a multiple-model PHD filter and map fusion for cooperative positioning and mapping in vehicular networks with 5G mmWave communication links. The environment comprises a single BS, multiple vehicles, and different types of objects (small scattering objects and large reflecting surfaces). The challenges of the mapping such as the number of objects, object type, and their position were dealt with the proposed PHD filter. In addition, asynchronous map transmission to the BS is solved by the proposed map fusion method. From the results, it is confirmed that our PHD filter can handle the challenges of the mapping and vehicle state estimation simultaneously. We also confirmed that the proposed map fusion using map information of other vehicles significantly improves the mapping performance.

APPENDIX A  
PSEUDO-CODE FOR MAP UPDATE

The pseudo-code is provided in Algorithms 1–3.

---

**Algorithm 1: Map Correction**

---

**input :**  $\{D_{k|k-1}^i(\mathbf{x}, m)\}_{i=1}^I \forall m$ , where  $D_{k|k-1}^i(\mathbf{x}, m) = \{\gamma_{k|k-1}^{i,j}(m), \mathbf{x}_{k|k-1}^{i,j}(m), \mathbf{P}_{k|k-1}^{i,j}(m)\}_{j=1}^{J_{k|k-1}^{i,(m)}}$ ,  $\{\mathbf{S}_{k|k-1}^i, \omega_{k|k-1}^i\}_{i=1}^I, \mathcal{Z}_k$

**output:**  $\{D_{k|k}^i(\mathbf{x}, m)\}_{i=1}^I \forall m$ , where  $D_{k|k}^i(\mathbf{x}, m) = \{\gamma_{k|k}^{i,j}(m), \mathbf{x}_{k|k}^{i,j}(m), \mathbf{P}_{k|k}^{i,j}(m)\}_{j=1}^{J_{k|k}^{i,(m)}}$

**for** (particle sample)  $i = 1$  to  $I$  **do**  
 | Compute PHD update components (Algorithm 2);  
 | PHD update (Algorithm 3);  
**end**

---



---

**Algorithm 2: Compute PHD update components by CKF**

---

**for**  $m = \{\text{BS, VA, SP}\}$  **do**  
 | **for**  $j = 1$  to  $J_{k|k-1}^i(m)$  **do**  
 | |  $\mathbf{P}_{k|k-1}^{i,j}(m) = \mathbf{G}_k^{i,j}(m)\mathbf{G}_k^{i,j}(m)^\top$ ;  
 | | **for**  $c = 1$  to  $2d_{\mathbf{x}}$  **do**  
 | | |  $\mathbf{X}_{c,k}^{i,j}(m) = \mathbf{x}_{k|k-1}^{i,q}(m) + \mathbf{G}_k^{i,j}(m)\delta_{\mathbf{x},c}$ ;  
 | | |  $\mathbf{Z}_{c,k}^{i,j}(m) = \mathbf{h}_k(\mathbf{x}_{k|k-1}^{i,j}(m), \mathbf{s}_k^i, m)$ ;  
 | | **end**  
 | |  $\mathbf{Z}_k^{i,j}(m) = \frac{1}{2d_{\mathbf{x}}} \sum_{c=1}^{2d_{\mathbf{x}}} \mathbf{Z}_{c,k}^{i,j}(m)$ ;  
 | |  $\mathbf{S}_{\mathbf{z}\mathbf{z},k}^{i,j}(m) = \frac{1}{2d_{\mathbf{x}}} \sum_{c=1}^{2d_{\mathbf{x}}} \mathbf{Z}_{c,k}^{i,j}(m)\mathbf{Z}_{c,k}^{i,j}(m)^\top - \mathbf{Z}_k^{i,j}(m)\mathbf{Z}_k^{i,j}(m)^\top + \Sigma_{k,l}$ ;  
 | | **if**  $m = \{\text{VA, SP}\}$  **then**  
 | | |  $\mathbf{X}_k^{i,j}(m) = \frac{1}{2d_{\mathbf{x}}} \sum_{c=1}^{2d_{\mathbf{x}}} \mathbf{X}_{c,k}^{i,j}(m)$ ;  
 | | |  $\mathbf{S}_{\mathbf{z}\mathbf{z},k}^{i,j}(m) = \frac{1}{2d_{\mathbf{x}}} \sum_{c=1}^{2d_{\mathbf{x}}} \mathbf{X}_{c,k}^{i,j}(m)\mathbf{Z}_{c,k}^{i,j}(m)^\top - \mathbf{X}_k^{i,j}(m)\mathbf{Z}_k^{i,j}(m)^\top$ ;  
 | | |  $\mathbf{K}_k^{i,j}(m) = \mathbf{S}_{\mathbf{z}\mathbf{z},k}^{i,j}(m)\mathbf{S}_{\mathbf{z}\mathbf{z},k}^{i,j}(m)^{-1}$ ;  
 | | |  $\mathbf{P}_{k|k}^{i,j}(m) = \mathbf{P}_{k|k-1}^{i,j}(m) - \mathbf{K}_k^{i,j}(m)\mathbf{S}_{\mathbf{z}\mathbf{z},k}^{i,j}(m)\mathbf{K}_k^{i,j}(m)^\top$ ;  
 | | **end**  
 | |  $q = 0$ ;  
 | | **for**  $\mathbf{z} \in \mathcal{Z}_k$  **do**  
 | | |  $q = q + 1$ ;  
 | | | **for**  $j = 1$  to  $J_{k|k-1}^i(m)$  **do**  
 | | | | **if**  $j = j(\mathbf{z})$  **then**  
 | | | | |  $\mu_k^{i,qJ_{k|k-1}^i(m)+j}(m) = \gamma_{k|k-1}^{i,j}(m)$ ;  
 | | | | **else**  
 | | | | |  $\mu_k^{i,qJ_{k|k-1}^i(m)+j}(m) = \mathcal{N}(\mathbf{z}; \mathbf{h}_k(\mathbf{x}_{k|k-1}^{i,j}(m), \mathbf{s}_k^i, m), \mathbf{S}_{\mathbf{z}\mathbf{z},k}^{i,j}(m)) \times p_{D,k}^{i,j}(m)\gamma_{k|k-1}^{i,j}(m)$ ;  
 | | | | **end**  
 | | | **end**  
 | | **end**  
 | **end**  
**end**

---



---

**Algorithm 3: PHD update**

---

**for**  $m = \{\text{VA, SP}\}$  **do**  
 | **for**  $j = 1$  to  $J_{k|k-1}^i(m)$  (missed detections) **do**  
 | |  $\mathbf{x}_{k|k}^{i,j}(m) = \mathbf{x}_{k|k-1}^{i,j}(m)$ ;  
 | |  $\mathbf{P}_{k|k}^{i,j}(m) = \mathbf{P}_{k|k-1}^{i,j}(m)$ ;  
 | | **if**  $j$  is a birth **then**  
 | | |  $\gamma_{k|k}^{i,j}(m) = 0$ ;  
 | | **else**  
 | | |  $\gamma_{k|k}^{i,j}(m) = \{1 - p_{D,k}^{i,j}(m)\}\gamma_{k|k-1}^{i,j}(m)$ ;  
 | | **end**  
 | **end**  
 $q = 0$ ;  
 | **for**  $\mathbf{z} \in \mathcal{Z}_k$  (detections) **do**  
 | |  $q = q + 1$ ;  
 | | **for**  $j = 1$  to  $J_{k|k-1}^i(m)$  **do**  
 | | | **if**  $j = j(\mathbf{z})$  **then**  
 | | | |  $\mathbf{x}_{k|k}^{i,qJ_{k|k-1}^i(m)+j}(m) = \mathbf{x}_{k|k-1}^{i,j}(m)$ ;  
 | | | |  $\mathbf{P}_{k|k}^{i,qJ_{k|k-1}^i(m)+j}(m) = \mathbf{P}_{k|k-1}^{i,j}(m)$ ;  
 | | | **else**  
 | | | |  $\mathbf{x}_{k|k}^{i,qJ_{k|k-1}^i(m)+j}(m) = \mathbf{x}_{k|k-1}^{i,j}(m) + \mathbf{K}_k^{i,j}(m)(\mathbf{z} - \mathbf{Z}_k^{i,j}(m))$ ;  
 | | | |  $\mathbf{P}_{k|k}^{i,qJ_{k|k-1}^i(m)+j}(m) = \mathbf{P}_{k|k}^{i,j}(m)$ ;  
 | | | **end**  
 | | |  $\gamma_{k|k}^{i,qJ_{k|k-1}^i(m)+j}(m) = \mu_k^{i,qJ_{k|k-1}^i(m)+j}(m) / (c(\mathbf{z}) + \sum_{m'=\{\text{BS, VA, SP}\}} \sum_{j'=1}^{J_{k|k-1}^i(m')} \times \mu_k^{i,qJ_{k|k-1}^i(m')+j'}(m'))$ ;  
 | | | **end**  
 | | |  $J_{k|k}^i(m) = (q + 1)J_{k|k-1}^i(m)$ ;  
 | | **end**  
 |  $\mathbf{x}_{k|k}^{i,j}(\text{BS}) = \mathbf{x}_{k|k-1}^{i,j}(\text{BS}), \mathbf{P}_{k|k}^{i,j}(\text{BS}) = \mathbf{P}_{k|k-1}^{i,j}(\text{BS})$ ;  
 |  $\gamma_{k|k}^{i,j}(\text{BS}) = 1, J_{k|k}^i(\text{BS}) = 1$ ;  
**end**

---

APPENDIX B  
GEOMETRIC RELATIONS

In the relation between the observations (3), the state of the vehicle and the map depend on the origin of the measurement. We distinguish between 3 different cases.

*A. Source is the BS*

For the LOS path between BS and vehicle, we have the following relations:  $\tau = \|\mathbf{x}_{\text{BS}} - \mathbf{v}_k\|/c + B$ , where  $c$  denotes the speed of light;  $\phi_0^{\text{az}} = \arctan(y_k/x_k)$ ,  $\phi_0^{\text{el}} = \arcsin((z_k - z_{\text{BS}})/(\|\mathbf{v}_k - \mathbf{x}_{\text{BS}}\|))$ , where we assume arctan2 is used; We remind that the DOA is measured in the local frame of reference of the vehicle, so that the vehicle orientation must be accounted for:  $\theta_0^{\text{az}} = \pi + \arctan(y_k/x_k) - \alpha_k$ ,  $\theta_0^{\text{el}} = \arcsin((z_{\text{BS}} - z_k)/(\|\mathbf{x}_{\text{BS}} - \mathbf{v}_k\|))$ , since the DOA elevation measurement does not depend on the vehicle orientation

### B. Source is a reflecting surface

Each reflecting surface can be parameterized by a fixed virtual anchor (VA) location  $\mathbf{x}_{\text{VA}}$ , obtained by mirroring the BS with respect to the surface. Between a virtual anchor  $\mathbf{x}_{\text{VA}}$  and the user's position  $\mathbf{v}_k$ , the incidence point of the specular reflection on the reflecting surface is given by the point where the straight line between the VA and vehicle crosses the reflecting surface

$$\mathbf{x}_s = \mathbf{x}_{\text{VA}} + \frac{(\mathbf{f} - \mathbf{x}_{\text{VA}})^T \mathbf{u}}{(\mathbf{v}_k - \mathbf{x}_{\text{VA}})^T \mathbf{u}} (\mathbf{v}_k - \mathbf{x}_{\text{VA}}). \quad (42)$$

Here,  $\mathbf{u} = (\mathbf{x}_{\text{BS}} - \mathbf{x}_{\text{VA}}) / \|\mathbf{x}_{\text{BS}} - \mathbf{x}_{\text{VA}}\|$  and  $\mathbf{f} = (\mathbf{x}_{\text{BS}} + \mathbf{x}_{\text{VA}}) / 2$ . Note that this allows to find explicit expressions of  $\mathbf{x}_s$  that only depend on  $\mathbf{x}_{\text{VA}}$ ,  $\mathbf{x}_{\text{BS}}$ , and  $\mathbf{v}_k$  (not shown). Conversely, the location of a VA can be expressed as a function of the incidence point

$$\mathbf{x}_{\text{VA}} = \mathbf{v}_k + (\|\mathbf{v}_k - \mathbf{x}_s\| + \|\mathbf{x}_{\text{BS}} - \mathbf{x}_s\|) \frac{\mathbf{x}_s - \mathbf{v}_k}{\|\mathbf{x}_s - \mathbf{v}_k\|}. \quad (43)$$

Next, we state the relations between the channel parameters  $\tau$ ,  $\boldsymbol{\theta} = [\theta^{\text{el}}, \theta^{\text{az}}]^T$ , and  $\boldsymbol{\phi} = [\phi^{\text{el}}, \phi^{\text{az}}]^T$  and the system state:  $\tau = \|\mathbf{x}_{\text{VA}} - \mathbf{v}_k\|/c + B$ . This is equivalent to  $\tau = \|\mathbf{x}_{\text{BS}} - \mathbf{x}_s\|/c + \|\mathbf{x}_s - \mathbf{v}_k\|/c + B$ ;  $\phi^{\text{az}} = \arctan(y_s/x_s)$  and  $\phi^{\text{el}} = \arcsin((z_s - z_{\text{BS}})/(\|\mathbf{x}_s - \mathbf{x}_{\text{BS}}\|))$ ; and  $\theta^{\text{az}} = \arctan((y_{\text{VA}} - y_k)/(x_{\text{VA}} - x_k)) - \alpha_k$  and  $\theta^{\text{el}} = \arcsin((z_{\text{VA}} - z_k)/(\|\mathbf{x}_{\text{VA}} - \mathbf{v}_k\|))$ .

### C. Source is a small object

For small objects (SPs), the relations are largely a special case of the VAs. We here only note the differences, considering an SP with location  $\mathbf{x}_{\text{SP}}$ :  $\tau = \|\mathbf{x}_{\text{SP}} - \mathbf{x}_{\text{BS}}\|/c + \|\mathbf{x}_{\text{SP}} - \mathbf{v}_k\|/c + B$ ;  $\phi^{\text{az}} = \arctan(y_{\text{SP}}/x_{\text{SP}})$  and  $\phi^{\text{el}} = \arcsin((z_{\text{SP}} - z_{\text{BS}})/(\|\mathbf{x}_{\text{SP}} - \mathbf{x}_{\text{BS}}\|))$ ; and  $\theta^{\text{az}} = \arctan((y_{\text{SP}} - y_k)/(x_{\text{SP}} - x_k)) - \alpha_k$  and  $\theta^{\text{el}} = \arcsin((z_{\text{SP}} - z_k)/(\|\mathbf{x}_{\text{SP}} - \mathbf{v}_k\|))$ .

## APPENDIX C

### PROOF OF EXPECTED LIKELIHOOD (15)

We find  $f(\mathcal{X}|\mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) \propto \prod_{(\mathbf{x}, m) \in \mathcal{X}} D_{k|k-1}^i(\mathbf{x}, m)$  from (8) under a PPP prior approximation, while with  $\mathcal{Z}_k = \{\mathbf{z}_1, \dots, \mathbf{z}_{|\mathcal{Z}_k|}\}$ , the function  $f(\mathcal{Z}_k|\mathcal{X}, \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1})$  can be expressed as [38, eq. (13)]

$$f(\mathcal{Z}_k|\mathcal{X}, \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) \quad (44)$$

$$= \sum_{\mathcal{U} \uplus \mathcal{X}_1 \dots \uplus \mathcal{X}_{|\mathcal{Z}_k|} = \mathcal{X}} \prod_{(\mathbf{x}, m) \in \mathcal{U}} (1 - p_{\text{D}}^i(\mathbf{x}, m)) \prod_{r=1}^{|\mathcal{Z}_k|} \ell(\mathbf{z}_r|\mathcal{X}_r, \mathbf{s}_k^i),$$

where  $\uplus$  denotes the disjoint union and [38, eq. (14)]

$$\ell(\mathbf{z}_r|\mathcal{X}_r, \mathbf{s}_k^i) \quad (45)$$

$$= \begin{cases} p_{\text{D}}^i(\mathbf{x}, \mathbf{s}_k^i, m) f(\mathbf{z}_r|\mathbf{x}, \mathbf{s}_k^i, m) & \mathcal{X}_r = \{(\mathbf{x}, m)\}, \\ c(\mathbf{z}_r) & \mathcal{X}_r = \emptyset, \\ 0 & \text{else.} \end{cases}$$

Introducing a PPP density  $q(\mathcal{U}) \propto \prod_{(\mathbf{x}, m) \in \mathcal{U}} (1 - p_{\text{D}}^i(\mathbf{x}, m)) D_{k|k-1}^i(\mathbf{x}, m)$  and making use of the fact that

$\int \sum_{\mathcal{X} \uplus \mathcal{Y} = \mathcal{U}} f(\mathcal{X}) g(\mathcal{Y}) \delta \mathcal{U} = \int f(\mathcal{X}) \delta \mathcal{X} \int g(\mathcal{Y}) \delta \mathcal{Y}$ , we immediately find that

$$\begin{aligned} & f(\mathcal{Z}_k|\mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) \\ &= \int f(\mathcal{Z}_k|\mathcal{X}, \mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) f(\mathcal{X}|\mathbf{s}_{0:k}^i, \mathcal{Z}_{1:k-1}) \delta \mathcal{X} \\ &\propto \prod_{r=1}^{|\mathcal{Z}_k|} \int \ell(\mathbf{z}_r|\mathcal{X}_r, \mathbf{s}_k^i) D_{k|k-1}^i(\mathcal{X}_r) \delta \mathcal{X}_r \\ &= \prod_{\mathbf{z} \in \mathcal{Z}_k} (c(\mathbf{z}) + \sum_m \int p_{\text{D}}^i(\mathbf{x}, \mathbf{s}_k^i, m) f(\mathbf{z}|\mathbf{x}, \mathbf{s}_k^i, m) \\ &\quad \times D_{k|k-1}^i(\mathbf{x}, m) d\mathbf{x}), \end{aligned} \quad (46)$$

where the last step follows from substitution of (45) and the definition of set integrals (6).

## APPENDIX D

### IMPLEMENTATION OF BIRTH PROCESS

Here, we introduce the detailed implementation of the birth process (18). The mean  $\mathbf{x}_{b,k}^{i,j(\mathbf{z})}(m)$  and covariance  $\mathbf{P}_{b,k}^{i,j(\mathbf{z})}(m)$  corresponding to each measurement  $\mathbf{z}$  are inversely estimated by using sigma point principle of the CKF [40], details of which are described in Appendix D-A. For propagating the cubature points, the inverse of the nonlinear function in (4) is required, which in general is not defined (since a vehicle state gives rise to a noise-free measurement, but a noisy measurement may not correspond to a vehicle state). Thus, the cubature points are propagated using a simple optimization method, described in Appendix D-B. For simplification, all indices are dropped except for the source type  $m$ .

### A. Mean and Covariance Estimation

The mean  $\mathbf{x}(m)$  and covariance  $\mathbf{P}(m)$  are approximated by the following steps:

- 1) Factorize the covariance matrix of the measurement noise (i.e.,  $\boldsymbol{\Sigma}_{k,l}^{(n)}$  of (4))

$$\boldsymbol{\Sigma} = \mathbf{G} \mathbf{G}^T. \quad (47)$$

- 2) Evaluate the cubature point ( $c = 1, \dots, 2d_{\mathbf{z}}$ )

$$\mathbf{z}_c = \mathbf{G} \delta_{\mathbf{z},c} + \mathbf{z}, \quad (48)$$

where  $d_{\mathbf{z}} = \dim(\mathbf{z})$  and  $\delta_{\mathbf{z},c} = \sqrt{d_{\mathbf{z}}} [1]_{\mathbf{z},c}$ .  $[1]_{\mathbf{z},c}$  is defined as the  $c$ -th column vector of the matrix  $[\mathbf{I}_{d_{\mathbf{z}}, d_{\mathbf{z}}}, -\mathbf{I}_{d_{\mathbf{z}}, d_{\mathbf{z}}}] \in \mathbb{R}^{d_{\mathbf{z}} \times 2d_{\mathbf{z}}}$ , where  $\mathbf{I}_{d_{\mathbf{z}}, d_{\mathbf{z}}} \in \mathbb{R}^{d_{\mathbf{z}} \times d_{\mathbf{z}}}$  is the identity matrix.

- 3) Evaluate the propagated cubature point  $\mathbf{x}_c(m) \approx \mathbf{h}^{-1}(\mathbf{z}_c|\mathbf{s}, m)$  with the iterative maximum-likelihood estimation (explained further in Appendix D-B).
- 4) Evaluate birth mean  $\mathbf{x}(m) = \frac{1}{2d_{\mathbf{z}}} \sum_{c=1}^{2d_{\mathbf{z}}} \mathbf{x}_c(m)$  and covariance  $\mathbf{P}(m) = \frac{1}{2d_{\mathbf{z}}} \sum_{c=1}^{2d_{\mathbf{z}}} \mathbf{x}_c(m) \mathbf{x}_c(m)^T - \mathbf{x}(m) \mathbf{x}(m)^T$ .

### B. Simple Optimization Problem for Propagated Cubature Point

For estimating the propagated cubature point  $\mathbf{x}_c(m)$  of step 3) in Appendix D-A, we formulate an optimization problem as

$$\underset{\mathbf{x}}{\text{minimize}} (\mathbf{h}(\mathbf{x}, \mathbf{s}, m) - \mathbf{z}_c)^T \Sigma^{-1} (\mathbf{h}(\mathbf{x}, \mathbf{s}, m) - \mathbf{z}_c), \quad (49)$$

where  $\mathbf{h}(\mathbf{x}, \mathbf{s}, m)$  is the observation function for the source type  $m$  with the source location  $\mathbf{x}$  and vehicle state  $\mathbf{s}$ , and  $\mathbf{z}_c \in \mathbb{R}^{d_z}$  is the evaluated cubature point in (48). Note that the used function  $\mathbf{h}(\cdot)$  are determined with respect to the source type  $m$ , which were described in Appendix B. However, (49) does not admit a closed-form solution, and an optimal point is determined in a iterative manner. The optimum point at the iteration  $\epsilon + 1$  is designed as  $\bar{\mathbf{x}}_{\epsilon+1}(m) = (1 - \eta)\bar{\mathbf{x}}_{\epsilon}(m) + \eta\tilde{\mathbf{x}}_{\epsilon+1}(m)$ , where the design parameter  $\eta$  is set to 0.2, and the initial point  $\bar{\mathbf{x}}_0$  is obtained by geometric relations in Appendix B.  $\tilde{\mathbf{x}}_{\epsilon+1}$  is calculated as  $\tilde{\mathbf{x}}_{\epsilon+1} = \underset{\mathbf{x}}{\text{argmin}} \mathbf{A}(\mathbf{x}, m)^T \Sigma^{-1} \mathbf{A}(\mathbf{x}, m)$ , where  $\mathbf{A}(\mathbf{x}, m)$  is denoted by  $\mathbf{A}(\mathbf{x}, m) = \mathbf{H}(\bar{\mathbf{x}}_{\epsilon}, \mathbf{s}, m)(\mathbf{x} - \bar{\mathbf{x}}_{\epsilon}) + \mathbf{h}(\bar{\mathbf{x}}_{\epsilon}, \mathbf{s}, m) - \mathbf{z}_c$ , where  $\mathbf{H}(\bar{\mathbf{x}}_{\epsilon}, \mathbf{s}, m) \in \mathbb{R}^{5 \times 3}$  is a Jacobian matrix, and is calculated by the finite difference method [45]. The difference is set to  $10^{-3}$ . The iterative method is performed until the cost (49) increases, and then  $\mathbf{x}_c$  is determined.

### REFERENCES

- [1] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mm-wave positioning for vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, Dec. 2018.
- [2] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.
- [3] K. Witrals 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.
- [4] 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, Feb. 2019.
- [5] H. Wymeersch, N. Garcia, H. Kim, G. Seco-Granados, S. Kim, F. Wen, and M. Fröhle, "5G mmWave downlink vehicular positioning," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 206–212.
- [6] 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.
- [7] 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.
- [8] A. Yassin, Y. Nasser, A. Y. Al-Dubai, and M. Awad, "MOSAIC: Simultaneous localization and environment mapping using mmWave without a-priori knowledge," *IEEE Access*, vol. 6, pp. 68 932–68 947, Nov. 2018.
- [9] M. Aladsani, A. Alkhateeb, and G. C. Trichopoulos, "Leveraging mmWave imaging and communications for simultaneous localization and mapping," in *Proc. 2019 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brighton, UK, May 2019, pp. 4539–4543.
- [10] J. Palacios, P. Casari, and J. Widmer, "JADE: Zero-knowledge device localization and environment mapping for millimeter wave systems," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Atlanta, GA, USA, May 2017, pp. 1–9.
- [11] J. Palacios, G. Bielsa, P. Casari, and J. Widmer, "Communication-driven localization and mapping for millimeter wave networks," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Honolulu, HI, USA, Apr. 2018, pp. 2402–2410.
- [12] R. Mendrzik, H. Wymeersch, G. Bauch, and Z. Abu-Shaban, "Harnessing NLOS components for position and orientation estimation in 5G millimeter wave MIMO," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 93–107, 2018.
- [13] H. Kim, H. Wymeersch, N. Garcia, G. Seco-Granados, and S. Kim, "5G mmWave vehicular tracking," in *Proc. IEEE 52nd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2018, pp. 541–547.
- [14] R. Mahler, "Multitarget Bayes filtering via first-order multi target moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.
- [15] 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.
- [16] K. Y. Leung, F. Inostroza, and M. Adams, "An improved weighting strategy for Rao-Blackwellized Probability Hypothesis Density simultaneous localization and mapping," in *Proc. 13th Int. Conf. Control, Autom. and Inf. Sci. (ICCAIS)*, Kwangju, Korea, Oct. 2013, pp. 103–110.
- [17] G. Battistelli, L. Chisci, and A. Laurenzi, "Random set approach to distributed multivehicle SLAM," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2457 – 2464, 2017, 20th IFAC World Congress.
- [18] H. Deusch, "Random finite set-based localization and SLAM for highly automated vehicles," Ph.D. dissertation, Fakultät für Ingenieurwissenschaften, Informatik und Psychologie, Universität Ulm, 2015.
- [19] H. Deusch, S. Reuter, and K. Dietmayer, "The labeled multi-Bernoulli SLAM filter," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1561–1565, Oct. 2015.
- [20] M. Fatemi, K. Granström, L. Svensson, F. Ruiz, and L. Hammarstrand, "Poisson multi-Bernoulli mapping using Gibbs sampling," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2814–2827, Jun. 2017.
- [21] T. Li, J. M. Corchado, and S. Sun, "On generalized covariance intersection for distributed PHD filtering and a simple but better alternative," in *Proc. 20th Int. Conf. Inf. Fusion (FUSION)*, Xian, China, Jul. 2017, pp. 1–8.
- [22] T. Li, V. Elvira, H. Fan, and J. M. Corchado, "Local-diffusion-based distributed SMC-PHD filtering using sensors with limited sensing range," *IEEE Sensors J.*, vol. 19, no. 4, pp. 1580–1589, 2018.
- [23] L. Gao, G. Battistelli, and L. Chisci, "Multiobject fusion with minimum information loss," *arXiv preprint arXiv:1903.04239*, 2019.
- [24] S. Pasha, B.-N. Vo, H. Tuan, and W.-K. Ma, "A Gaussian mixture PHD filter for jump Markov system models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 3, pp. 919–936, Jul. 2009.
- [25] R. Mendrzik, H. Wymeersch, and G. Bauch, "Joint localization and mapping through millimeter wave MIMO in 5G systems-extended version," *arXiv preprint arXiv:1804.04417*, 2018.
- [26] F. Meyer, T. Kropfreiter, J. L. Williams, R. Lau, F. Hlawatsch, P. Braca, and M. Z. Win, "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, 2018.
- [27] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Autom. Control*, vol. 33, no. 8, pp. 780–783, Aug. 1988.
- [28] K. Granström and C. Lundquist, "On the use of multiple measurement models for extended target tracking," in *Proc. 16th Int. Conf. Inf. Fusion (FUSION)*, Istanbul, Turkey, Jul. 2013.
- [29] K. Granström, S. Reuter, D. Meissner, and A. Scheel, "A multiple model PHD approach to tracking of cars under an assumed rectangular shape," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, Salamanca, Spain, Jul. 2014.
- [30] R. Mahler, "On multitarget jump-Markov filters," in *Proc. 15th Int. Conf. Inf. Fusion (FUSION)*, Singapore, Jul. 2012, pp. 149–156.
- [31] S. Li, G. Battistelli, L. Chisci, W. Yi, B. Wang, and L. Kong, "Multi-sensor multi-object tracking with different fields-of-view using the lmb filter," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Cambridge, U.K., Jul. 2018, pp. 1201–1208.
- [32] G. Li, G. Battistelli, W. Yi, and L. Kong, "Distributed multi-sensor multi-view fusion based on generalized covariance intersection," *arXiv preprint arXiv:1903.06985*, 2019.
- [33] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents Series)*. MIT Press, 2005.
- [34] X. Rong-Li and V. Jilkov, "Survey of maneuvering target tracking: Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.
- [35] R. W. Heath, 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. 2016.

- [36] J. L. Williams, "Marginal multi-bernoulli filters: RFS derivation of MHT, JIPDA, and association-based MeMber," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 1664–1687, Jul. 2015.
- [37] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.
- [38] Á. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: Direct derivation and implementation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1883–1901, Aug. 2018.
- [39] B. Ristic, D. Clark, B. N. Vo, and B. T. Vo, "Adaptive target birth intensity for PHD and CPHD filters," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1656–1668, 2012.
- [40] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.
- [41] R. P. Mahler, "Optimal/robust distributed data fusion: a unified approach," in *Signal Processing, Sensor Fusion, and Target Recognition IX*, vol. 4052, 2000, pp. 128–139.
- [42] D. W. Bunn, "A Bayesian approach to the linear combination of forecasts," *Journal of the Operational Research Society*, vol. 26, no. 2, pp. 325–329, 1975.
- [43] 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 Dataport*, 2019. [Online]. Available: <http://dx.doi.org/10.21227/hyfa-fa96>
- [44] A. S. Rahmathullah, A. F. García Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Proc. 20th Int. Conf. Inf. Fusion (FUSION)*, Xian, China, Jul. 2017, pp. 1–8.
- [45] G. D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford University Press, 1985.



**Hyowon Kim** (S'17) received his B.S. degree in electronics and computer engineering from Hanyang University, Korea in 2015, where he is currently pursuing the Ph.D. degree in the department of electronics and computer engineering, Hanyang University, Korea. He was a Visiting Researcher with the Department of Signals and Systems, Chalmers University of Technology, Sweden, from 2019 to 2020. His research interests include wireless localization/navigation systems, multiobject tracking, and situational-aware communications.



**Karl Granström** (M'08) is a researcher at the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. He received the MSc degree in Applied Physics and Electrical Engineering in May 2008, and the PhD degree in Automatic Control in November 2012, both from Linköping University, Sweden. He previously held postdoctoral positions at the Department of Electrical and Computer Engineering at University of Connecticut, USA, from September 2014 to August 2015, and at the Department of Electrical

Engineering of Linköping University from December 2012 to August 2014. His research interests include multi-object tracking and mapping, especially for extended objects, and machine learning. He has received best paper awards at the Fusion 2011 and Fusion 2012 conferences. In 2018 the International Society of Information Fusion (ISIF) awarded him the ISIF Young Investigator Award for his contributions to extended object tracking and for his service to the research community. He has organised several well-attended workshops and tutorials on the topic Multiple Extended Object Tracking with applications to autonomous vehicles.



**Lin Gao** received the B.S., M.S. and Ph.D. degrees all in electronic and information engineering from University of Electronic Science and Technology of China, in 2011, 2014 and 2018, respectively. He is currently a Research Fellow in Department of Information Engineering from University of Florence.

His research interests include statistical signal processing, multi-target multi-sensor tracking, information fusion and multi-vehicle SLAM.



**Giorgio Battistelli** received the Laurea degree in electronic engineering and the Ph.D. degree in robotics from the University of Genoa, Genoa, Italy, in 2000 and 2004, respectively. From 2004 to 2006, he was a Research Associate with the Dipartimento di Informatica, Sistemistica e Telematica, University of Genoa. Since 2006 he has been with the University of Florence, Florence, Italy, where he is currently a Professor of automatic control with the Dipartimento di Ingegneria dell'Informazione. His current research interests include adaptive and learning

systems, real-time control reconfiguration, linear and nonlinear estimation, hybrid systems, sensor networks, and data fusion. He was a member of the editorial boards of the IFAC Journal Engineering Applications of Artificial Intelligence and of the IEEE Transactions on Neural Networks and Learning Systems. He is currently an Associate Editor of the IEEE Control Systems Letters and of the journals Sensors and Nonlinear Analysis: Hybrid Systems.



**Sunwoo Kim** (S'99-M'05-SM'17) received his B.S. degree from Hanyang University, Seoul, Korea in 1999, and his Ph.D. degree, in 2005, from the Department of Electrical and Computer Engineering, University of California, Santa Barbara. Since 2005, he has been working in the Department of Electronic Engineering at Hanyang University, Seoul, Korea, where he is currently a professor. He is also the director of the 5G/Unmanned Vehicle Research Center, funded by the Ministry of Science and ICT of Korea.

He was a visiting scholar to the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology from 2018 to 2019. He is an associate editor of IEEE Transactions on Vehicular Technology. He is a senior member of the IEEE. His research interests include wireless communication/positioning/localization, signal processing, vehicular networks, and location-aware communications.



**Henk Wymeersch** (S'01-M'05-SM'19) received the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium. He is currently a Professor of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Prof. Wymeersch served as Associate Editor for IEEE Communication

Letters (2009-2013), IEEE Transactions on Wireless Communications (since 2013), and IEEE Transactions on Communications (2016-2018). His current research interests include cooperative systems and intelligent transportation.