IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 15, 2020, accepted June 9, 2020, date of publication July 8, 2020, date of current version July 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3008007

Decentralized Poisson Multi-Bernoulli Filtering for Vehicle Tracking

MARKUS FRÖHLE[1], (Student Member, IEEE), KARL GRANSTRÖM[2,3], (Member, IEEE), AND HENK WYMEERSCH[3], (Senior Member, IEEE)

[1]Zenuity AB, 417 56 Gothenburg, Sweden
[2]Embark Trucks Inc., San Francisco, CA 94107, USA
[3]Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden

Corresponding author: Markus Fröhle (markus.frohle@zenuity.com)

This work was supported in part by the EU-H2020 Project HIGHTS (High Precision Positioning for Cooperative ITS Applications) under Grant MG-3.5a-2014-636537, and in part by the COPPLAR (Campus shuttle cooperative perception and planning platform) Project funded by Vinnova under Grant 2015-04849.

ABSTRACT A decentralized Poisson multi-Bernoulli filter is proposed to track multiple vehicles using multiple high-resolution sensors. Independent filters estimate the vehicles' presence, state, and shape using a Gaussian process extent model; a decentralized filter is realized through fusion of the filters posterior densities. An efficient implementation is achieved by parametric state representation, utilization of single hypothesis tracks, and fusion of vehicle information based on a fusion mapping. Numerical results demonstrate the performance.

INDEX TERMS Gaussian processes, multitarget tracking, posterior fusion, target extent.

I. INTRODUCTION

Multitarget tracking (MTT), i.e., tracking of independently moving targets, is important for surveillance and safety applications [1], [2]. Traditionally, it has been developed for surveillance of the sky using ground-to-air radar sensors. An MTT filter allows incorporating the peculiarities of those kind of sensors: false alarm measurements due to clutter; missed detections; unknown measurement-to-target correspondence; and target appearance and disappearance; which are all challenges that arise for radar-like sensors. In many typical MTT scenarios, the sensor resolution is low with respect to (w.r.t.) the target size, and a reasonable assumption is to model the targets as points having a kinematic state (e.g., position and velocity). An elegant way to track multiple targets is via the Poisson multi-Bernoulli mixture (PMBM) filter [3], [4], which preserves a PMBM form during prediction and update steps.

With the availability of high-resolution sensors, the point target assumption does not hold anymore, see, e.g., [5]. For instance, a high-resolution Lidar sensor can obtain in one scan multiple detections from a single vehicle [6] or cyclist [7]. This is because the sensor resolution is high w.r.t. the target (here, a vehicle) size. In such an application scenario, the vehicle extent needs to be modeled (and estimated) as well in the MTT filter, leading to an extended target tracking (ETT) filter. In ETT, extended targets (ETs) give rise to possibly multiple noisy detections, the vehicle extent (shape and size) is a priori unknown and may vary over time, and the objective is to estimate the ET's kinematic state as well as its extent [5]. A common model for the ET extent is based on Gaussian process (GP) modelling [8], [9]. Different models for the target extent (shape and size) exist in ETT, which may be classed according to complexity, ranging from assuming a specific geometric shape [6], [7], [10] with, e.g., unknown translation and rotation, to models that describe general shapes, see, e.g., [11]–[13]. Typically, more complex models provide a richer shape description. Two popular models are the random matrix approach [10], where the target shape is described by an ellipsoid; and a GP based approach [8], [9], where a star convex target shape is described by a GP. See [5] for an extensive overview of works on ETT.

For tracking multiple extended targets, several different filters have been developed: Probability Hypothesis Density (PHD) filter [14]–[17]; Cardinalized PHD (CPHD) filter [18]; δ-generalized labeled multi-Bernoulli (δ-GLMB) filter [19]; and PMBM filter [3], [4]. For applications involving tracking of vehicles see, e.g., [20]–[22]. The δ-GLMB filter and PMBM filter are so-called multi-object conjugate priors, meaning that if we start with the conjugate density form

The associate editor coordinating the review of this manuscript and approving it for publication was Qichun Zhang.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

(δ-GLMB or PMBM), then all subsequent predicted and updated densities will be of the same form. Based on the conjugate priors, computationally cheaper, approximate filters have been presented. The labeled multi-Bernoulli (LMB) filter is an approximation of the δ-GLMB filter, see [19]. The Poisson multi-Bernoulli (PMB) filter is an approximation of the PMBM filter, see [23], [24]. The PMBM conjugate prior [23] was originally developed for point targets; a PMBM conjugate prior for extended targets was presented in [3], [4]. In several simulation studies it has been shown that, compared to tracking filters built upon labeled random finite set (RFS), the PMBM conjugate prior has good performance for tracking the set of present target states, for both point targets [25]–[27] and extended targets [3], [4], [28]. The PMBM conjugate priors for point targets and extended targets have been shown to be versatile, and have been used with data from Lidars [29]–[32], radars [30], [31], and cameras [31], [33]. They have been successfully applied not only to tracking of moving targets, but also the mapping of stationary objects [34], as well as joint tracking and sensor localization [35]. Thus, it is well motivated to use the PMBM filter in this work; specifically, we work with the computationally cheaper PMB filter. Developing similar decentralized tracking algorithms for the δ-GLMB or LMB filter [19] is a topic for future work.

Neither MTT nor ETT filters are bound to single sensors. When multiple sensors are used, fusion can provide performance benefits. Fusion can be centralized (with a fusion center) or decentralized. The latter approach was proposed in [36], for fusing two multi-Bernoulli (MB) densities, though it did not account for a Poisson point process (PPP) component. The information exchanged can be in the form of measurements, so that fusion can be solved by performing multiple update steps (e.g., Kalman update or likewise) by augmenting the measurement model to incorporate all sensors, see, e.g., [37]. When communication capacity is limited, one can perform filtering already at the sensor and share only target track information, e.g., the parameters of a known posterior probability density function (PDF) family. Fusion should then ensure that (unknown) common information obtained by the independent tracking filters is not double-counted, e.g., the prior target density [38]. Depending on the type of posterior multiobject density in MTT/ETT filtering, several sub-optimal information fusion strategies have been developed based on, e.g, covariance intersection (CI) for Gaussian densities. This replaces the product form of Bayes' rule with the Kullback-Leibler average (KLA)[1]. CI (and consequently KLA) is a method to fuse information with unknown priors in a robust way in the sense that the fused posterior is conservative and never overconfident about the estimates and thus implicitly sub-optimal [38], [39]. Examples of CI/KLA include fusion of Bernoulli and independent identically distributed (IID) cluster processes posteriors [40], fusion for Bernoulli filters [41], fusion for PHD and CPHD filters [42]–[47], and fusion for LMB filters [36], [46], [48].

In this paper, we utilize and extend the shape model from [8], [9], and integrate it into an ETT filter, which allows tracking of multiple ETs. The resulting ETT filter's multiobject posterior is of the so-called PMB form [23]. Furthermore, we propose a novel fusion strategy that performs fusion separately for the PPP and MB parts of the PMB density. The implementation of the PMB filter yields a tracking filter with low computational cost[2] locally at each sensor, and globally low computational cost through the introduction of a fusion map based on the Kullback-Leibler divergence (KLD) between target tracks. Simulation results demonstrate the performance of the proposed independent ETT filter as well as of the decentralized ETT filtering approach. The main contributions of this paper are:

- We apply a state-of-the-art ETT filter [3], [4] and propose a novel distributed fusion strategy based on the KLA by fusion of the filters' posterior multiobject densities of PMB form; and
- We enable low complexity in distributed fusion through introduction of a fusion map based on KLD between target tracks.
- We extend, in Section VI-A2, the GP model from [8], [9] for the ET shape description to a multisensor scenario incorporating the sensors' state (position, orientation);

The remainder of this paper is organized as follows. Section II gives some background knowledge on RFSs, and Section III introduces the system model and the problem formulation. Section IV details the proposed ETT filter, Section V presents the decentralized posterior fusion approach using independent ETT filters. Simulation results are given in Section VI, and conclusions are drawn in Section VII.

II. BACKGROUND ON RANDOM FINITE SETS

Two types of RFSs relevant for this work deserve special attention: Bernoulli RFS, and a PPP. They can be extended to MB RFS, multi-Bernoulli mixture (MBM) RFS and combined in a PMBM RFS. These are described below, for more details on RFSs the reader is referred to [49].

A. COMMON RFS DENSITIES

1) BERNOULLI RFS

A Bernoulli RFS $\boldsymbol{X}$ has a multiobject density [49]

$$f(\boldsymbol{X}) = \begin{cases} 1-r, & \boldsymbol{X} = \emptyset, \\ rf(\boldsymbol{x}), & \boldsymbol{X} = \{\boldsymbol{x}\}, \\ 0, & |\boldsymbol{X}| \geq 2, \end{cases} \tag{1}$$

where $r \in [0, 1]$ denotes the probability that a target exists and if it exists $f(\boldsymbol{x})$ is its PDF.

[1] In some literature this is known as exponential mixture density (EMD).

[2] As measured by the average time it takes for one cycle of prediction and update.

2) MULTI-BERNOULLI RFS

An MB RFS is the disjoint union of independent Bernoulli RFSs indexed by i. It is fully parameterized by $\{r_i, f_i(\boldsymbol{x})\}_{i\in\mathbb{I}}$, where $\mathbb{I}$ is its index set. For $\boldsymbol{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$, the multiobject density can be written as

$$f(\boldsymbol{X}) = \sum_{\uplus_{i\in\mathbb{I}}\boldsymbol{X}_i=\boldsymbol{X}} \prod_{i\in\mathbb{I}} f_i(\boldsymbol{X}_i) \tag{2}$$

for $|\boldsymbol{X}| \leq |\mathbb{I}|$, and $f(\boldsymbol{X}) = 0$ otherwise. The notation, $\boldsymbol{X}_1 \uplus \boldsymbol{X}_2 = \boldsymbol{X}$ means $\boldsymbol{X}_1 \cup \boldsymbol{X}_2 = \boldsymbol{X}$ and $\boldsymbol{X}_1 \cap \boldsymbol{X}_2 = \emptyset$.

3) MULTI-BERNOULLI MIXTURE RFS

The multiobject density of an MBM is the normalized, weighted sum of multiobject densities of MBs, which can be stated as [3]

$$f(\boldsymbol{X}) = \sum_{j\in\mathbb{J}} w_j \sum_{\uplus_{i\in\mathbb{I}^j}\boldsymbol{X}_i=\boldsymbol{X}} \prod_{i\in\mathbb{I}^j} f_{j,i}(\boldsymbol{X}_i). \tag{3}$$

The MBM multiobject density is parametrized by $\{w_{j,i}, \{r_{j,i}, f_{j,i}(\boldsymbol{x})\}_{i\in\mathbb{I}^j}\}_{j\in\mathbb{J}}$, where w_j is the weight of MB j, and $\mathbb{J}$ is the index set of the MBs in the MBM. An MB is therefore a special case of an MBM with $|\mathbb{J}| = 1$.

4) PPP

A PPP is a type of RFS, where the cardinality follows a Poisson distribution and its elements are IID. It is parametrized by the intensity function $D(\boldsymbol{x}) = \lambda f(\boldsymbol{x})$, where $\lambda > 0$ is the Poisson rate and $f(\boldsymbol{x})$ is a PDF on the single element state $\boldsymbol{x}$. The multiobject density of a PPP is [3], [49]

$$f(\boldsymbol{X}) = e^{-\lambda} \prod_{i=1}^{n} \lambda f(\boldsymbol{x}_i). \tag{4}$$

5) PMBM

An PMBM RFS is the disjoint set union of an PPP and an MBM having multiobject density [3]

$$f(\boldsymbol{X}) = \sum_{\boldsymbol{X}^u\uplus\boldsymbol{X}^d=\boldsymbol{X}} f^u(\boldsymbol{X}^u) f^d(\boldsymbol{X}^d), \tag{5}$$

where $f^u(\cdot)$ has multiobject density (4), and $f^d(\boldsymbol{X}^d)$ has multiobject density (3). For $|\mathbb{J}| = 1$, (5) is called a PMB distribution.

B. RFS BAYESIAN FILTER

Similar to the random vector (RV) case, an RFS based filter can be described, conceptually at least, within the Bayesian filtering framework by performing a prediction step using the motion model [2, Ch.14]

$$f_+(\boldsymbol{X}) = \int f(\boldsymbol{X}|\boldsymbol{X}') f_-(\boldsymbol{X}') \delta\boldsymbol{X}', \tag{6}$$

where $f_-(\boldsymbol{X}')$ is the prior RFS density, $f(\boldsymbol{X}|\boldsymbol{X}')$ is the multiobject process model, and a Bayesian update step

$$f(\boldsymbol{X}|\boldsymbol{Z}) \propto \ell(\boldsymbol{Z}|\boldsymbol{X}) f_+(\boldsymbol{X}). \tag{7}$$

Here, $f_+(\boldsymbol{X})$ is the predicted RFS density, and $\ell(\boldsymbol{Z}|\boldsymbol{X})$ is the RFS measurement likelihood for measurement set $\boldsymbol{Z}$.

A typical way to estimate the set states from a Bernoulli process with RFS density $f(\boldsymbol{X})$ is by comparing the probability of existence r to an existence threshold r_{th}. For $r > r_{\text{th}}$, the target is said to exist and has PDF $f(\boldsymbol{x})$. Its state can then be estimated by the mean $\hat{\boldsymbol{x}} = \int \boldsymbol{x} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$. See, e.g., [26] for an elongated discussion on multiobject estimation.

III. PROBLEM FORMULATION AND SYSTEM MODEL

Here, we present first the problem formulation of this paper followed by the ET state and transition model and the ET set measurement likelihood function.

A. PROBLEM FORMULATION

We consider a scenario with N_{sens} sensing systems (each composed of sensors plus filter), each collecting measurements using its local sensors, to jointly surveil an environment $\mathcal{E}$ where vehicles pass (c.f. Fig. 1). Our goals are (i) to derive a low-complexity PMB-ETT filter for each sensing system $s = 1, 2, \ldots, N_{\text{sens}}$, computing in every time step k the posterior density $f_s(\boldsymbol{X}_k|\boldsymbol{Z}_{s,k:1})$ of the ETs, using only its own sensors with measurement set $\boldsymbol{Z}_{s,k:1}$; and (ii) to derive a decentralized method to combine posterior information of ETs obtained by N_{sens} independent ETT filters to obtain a global posterior density $\bar{f}_w(\boldsymbol{X}_k|\boldsymbol{Z}_{1,k:1}, \boldsymbol{Z}_{2,k:1}, \ldots, \boldsymbol{Z}_{N_{\text{sens}},k:1})$ (i.e., using only the *posterior* densities of each ETT filter).

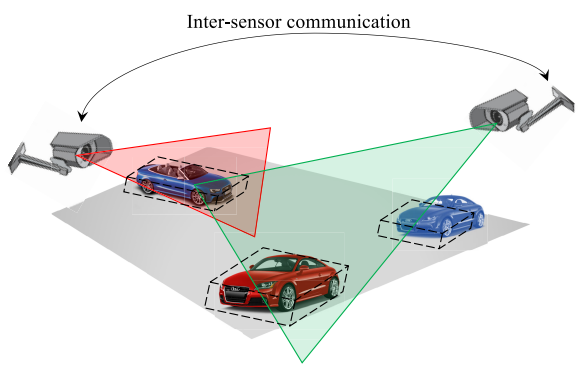


FIGURE 1. Scenario with three ETs observed by two sensors with partially overlapping FoVs.

B. ET STATE AND TRANSITION MODEL

A standard motion model for the ET is assumed, where ET motion follows IID Markov processes with single ET transition PDF $f_{k+1|k}(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)$, where $\boldsymbol{x}_k$ denotes the state at time k. Targets arrive according to a non-homogeneous PPP with intensity $D^b(\boldsymbol{x}_k)$, and depart according to IID Markov processes, where the survival probability in $\boldsymbol{x}_k$ is $p_S(\boldsymbol{x}_k)$. Additionally to this, $\boldsymbol{x}_k$ comprises an unknown Poisson rate γ_k describing the average number of measurements generated by the ET and the ET spatial state $\boldsymbol{y}_k$ (this will become clear in Sec. VI-A where a detailed description of the state and transition density is provided).

C. ET SET MEASUREMENT LIKELIHOOD FUNCTION

In one scan, a sensor receives a *set of measurements* $\boldsymbol{Z}$ consisting of target-generated measurements $\boldsymbol{z}$ and clutter, where the ETs are independently detected with state-dependent probability of detection $p_{\mathrm{D}}(\boldsymbol{x})$, which depends on the sensor FoV. Clutter is modeled by a PPP with intensity $\kappa(\boldsymbol{z}) = \lambda c(\boldsymbol{z})$ with mean λ and spatial distribution $c(\boldsymbol{z})$. Target-generated measurements are modeled by a PPP with intensity $\gamma(\boldsymbol{x})f(\boldsymbol{z}|\boldsymbol{x})$, where both the Poisson measurement rate $\gamma(\boldsymbol{x})$ and the single measurement likelihood $f(\boldsymbol{z}|\boldsymbol{x})$ are state-dependent. The measurement likelihood for ETs $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ and measurement set $\boldsymbol{Z}$ is

$$\begin{aligned}&\ell(\boldsymbol{Z}|\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\})\\ &= e^{-\lambda} \sum_{\boldsymbol{Z}_c \uplus \boldsymbol{Z}_1 \uplus \ldots \boldsymbol{Z}_n = \boldsymbol{Z}} [c(\cdot)]^{\boldsymbol{Z}_c} \prod_{i=1}^{n} \ell_{\boldsymbol{Z}_i}(\boldsymbol{x}_i),\end{aligned} \tag{8}$$

where $[c(\cdot)]^{\boldsymbol{Z}_c}$ is shorthand for $\prod_{\boldsymbol{z}\in\boldsymbol{Z}_c} c(\boldsymbol{z})$, $[c(\cdot)]^{\emptyset} = 1$ by definition, and [3]

$$\ell_{\boldsymbol{Z}}(\boldsymbol{x}) = \begin{cases} p_{\mathrm{D}}(\boldsymbol{x})e^{-\gamma(\boldsymbol{x})} \prod_{\boldsymbol{z}\in\boldsymbol{Z}} \gamma(\boldsymbol{x})f(\boldsymbol{z}|\boldsymbol{x}), & |\boldsymbol{Z}| > 0, \\ (1 - p_{\mathrm{D}}(\boldsymbol{x})) + p_{\mathrm{D}}(\boldsymbol{x})e^{-\gamma(\boldsymbol{x})}, & \boldsymbol{Z} = \emptyset. \end{cases} \tag{9}$$

Note that (i) equation (8) involves potentially multiple ETs, leading to a data association (DA) problem; (ii) a single ET can generate multiple measurements.

IV. INDEPENDENT PMB-ETT FILTER

In this section, we briefly describe the processing performed by each PMB-ETT (PMB-ETT) filter from [3]. We omit the time index k for brevity. The PMB model is a combination of a PPP describing the distribution of unknown targets, i.e., targets which are hypothesized to exist, but have not yet been detected; and an MB which describes targets that have been detected at least once. The target set can, therefore, be split into two disjoint subsets $\boldsymbol{X} = \boldsymbol{X}^u \uplus \boldsymbol{X}^d$ corresponding to the unknown target set $\boldsymbol{X}^u$ (with PPP intensity $D^u(\boldsymbol{x})$, modelled as a non-normalized mixture density with mixture components located in the region of interest denoted $\mathcal{E}$) and the detected target set $\boldsymbol{X}^d$ with density $f^d(\boldsymbol{X}^d)$ and index set $\mathbb{I}$. Hence, the PMB density is fully described by an MB component described by $\{r^i, f^i(\boldsymbol{x})\}_{i\in\mathbb{I}}$ and a PPP component $D^u(\boldsymbol{x})$.

1) PMB-ETT FILTER PREDICTION

The predicted density is a PMB density with parameters [3, Sec. IV], [4]

$$D^u_+ = D^b(\boldsymbol{x}) + \left\langle D^u, p_{\mathrm{S}} f_{k+1|k} \right\rangle, \tag{10}$$

$$r^i_+ = \left\langle f^i, p_{\mathrm{S}} \right\rangle r^i, \tag{11}$$

$$f^i_+ = \frac{\left\langle f^i, p_{\mathrm{S}} f_{k+1|k} \right\rangle}{\left\langle f^i, p_{\mathrm{S}} \right\rangle}, \tag{12}$$

where $\langle g, h\rangle = \int g(\boldsymbol{x})f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ denotes the inner product. The proof of the prediction step can be found in, e.g., [23].

2) PMB-ETT FILTER UPDATE

We introduce the set of valid DAs $\mathcal{A} = \mathcal{P}(\mathbb{M} \cup \mathbb{I})$, where $\mathbb{M}$ is the index set for $\boldsymbol{Z}$. Here, $A \in \mathcal{A}$ is a partition of $\mathbb{M} \cup \mathbb{I}$ into non-empty disjoint subsets $C \in A$ (called index cells), with the constraint that for each $C : |C \cap \mathbb{I}| \leq 1$ (i.e., measurements can only be associated with a single target). When $|C \cap \mathbb{I}| = 1$, let the entry in $C \cap \mathbb{I}$ be denoted by i_C and let $\boldsymbol{C}_C = \cup_{m\in C\cap\mathbb{M}} \boldsymbol{z}_m$ contain the associated measurements. Given the predicted prior PMB density with parameters (10), (11), (12), and a set of measurements $\boldsymbol{Z}$; the updated density is a PMBM density [3, Sec. IV], [4]

$$f(\boldsymbol{X}|\boldsymbol{Z}) = \sum_{\boldsymbol{X}^u \uplus \boldsymbol{X}^d = \boldsymbol{X}} f^u(\boldsymbol{X}^u) \sum_{A\in\mathcal{A}} w_A f^d_A(\boldsymbol{X}^d), \tag{13}$$

$$f^u(\boldsymbol{X}^u) = e^{-\langle D^u, 1\rangle} \prod_{\boldsymbol{x}\in\boldsymbol{X}^u} D^u(\boldsymbol{x}), \tag{14}$$

$$f^d_A(\boldsymbol{X}^d) = \sum_{\uplus_{C\in A} \boldsymbol{X}^C = \boldsymbol{X}^d} \prod_{C\in A} f_C(\boldsymbol{X}^C), \tag{15}$$

$$D^u(\boldsymbol{x}) = q_{\mathrm{D}}(\boldsymbol{x}) D^u_+(\boldsymbol{x}), \tag{16}$$

where $f_C(\boldsymbol{X}^C)$ is a Bernoulli density, with existence probability and spatial distribution provided in Appendix VII-A, together with the weights w_A of each DA hypothesis. Above, $q_{\mathrm{D}}(\boldsymbol{x})$ denotes the probability that the target $\boldsymbol{x}$ is not detected and is defined as

$$q_{\mathrm{D}}(\boldsymbol{x}) = 1 - p_{\mathrm{D}}(\boldsymbol{x}) + p_{\mathrm{D}}(\boldsymbol{x})e^{-\gamma(\boldsymbol{x})}. \tag{17}$$

To reduce computational complexity, we use standard methods and truncate the space of possible partitions by clustering measurements and consider DA w.r.t. different clusters [15], [16], [18]. Finally, the PMBM in (13) is converted to a PMB, which was detailed in [50] for point targets and in [24] for ETs.

V. DECENTRALIZED POSTERIOR FUSION

Here, we present the decentralized approach to robust fusion of posterior densities $f_s(\boldsymbol{X}|\boldsymbol{Z}_s)$ computed by independent PMB-ETT filters s with unknown prior densities. Fusion can be performed after every update step of the filters or based on a lower rate depending on the application and communication capabilities.

A. ROBUST POSTERIOR FUSION: PRINCIPLE

Robust posterior fusion can be achieved by minimizing the KLA between RFS densities $f(\cdot)$ and $f_s(\cdot)$ for $s = 1, \ldots, N_{\mathrm{sens}}$ with respect to $f(\cdot)$. The KLA is defined as [45]

$$\bar{f} = \arg\inf_{f} \sum_{s=1}^{N_{\mathrm{sens}}} w_s D(f\|f_s), \tag{18}$$

for any combinations of weights $w_s \in [0, 1] : \sum_s w_s = 1$. We have introduced $D(f\|f_s)$ as the KLD between RFS densities $f(\boldsymbol{X}|\boldsymbol{Z}_{1:N_{\mathrm{sens}}})$ and $f_s(\boldsymbol{X}|\boldsymbol{Z}_s)$ defined as [45]

$$D(f\|f_s) = \int f(\boldsymbol{X}|\boldsymbol{Z}_{1:N_{\mathrm{sens}}}) \log \frac{f(\boldsymbol{X}|\boldsymbol{Z}_{1:N_{\mathrm{sens}}})}{f_s(\boldsymbol{X}|\boldsymbol{Z}_s)} \delta\boldsymbol{X}. \tag{19}$$

The fused posterior (18) is robust in the sense that it is conservative and never overconfident w.r.t. the true target uncertainty [38]. Problem (18) was shown to have closed-form solution [45]

$$\bar{f}(\boldsymbol{X}|\boldsymbol{Z}_{1:N_{\text{sens}}}) = \frac{\prod_{s=1}^{N_{\text{sens}}} f_s(\boldsymbol{X}|\boldsymbol{Z}_s)^{w_s}}{\int \prod_{s'=1}^{N_{\text{sens}}} f_{s'}(\boldsymbol{X}|\boldsymbol{Z}_{s'})^{w_{s'}} \delta \boldsymbol{X}}. \tag{20}$$

Note that (20) is a generalization of the Uhlmann-Julier covariance intersection method (c.f. [38]) for posterior RFS densities with unknown priors [39]. The weights w_i can be chosen such that (20) is as peaky as possible [39]. An example of fusion of three Gaussian densities with equal weights is depicted in Fig. 2.

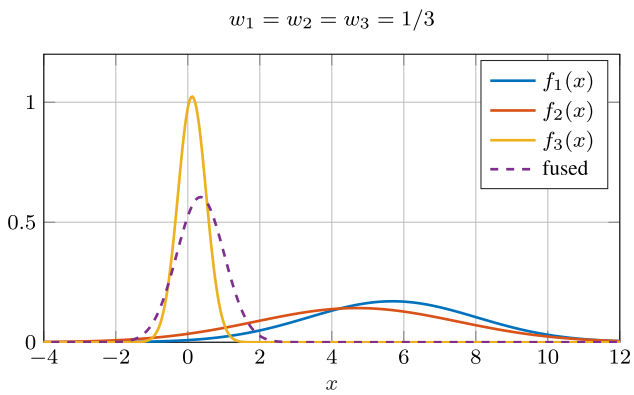


FIGURE 2. Example of posterior fusion with three Gaussian posteriors.

B. ROBUST PMB POSTERIOR FUSION

The posterior density of each PMB-ETT filter is a PMB, therefore, to be able to utilize the fused posterior density as prior for the next time step in each PMB-ETT filter, the fused posterior should be of the PMB form, i.e.,

$$\bar{f}(\boldsymbol{X}|\boldsymbol{Z}_{1:N_{\text{sens}}}) = \sum_{\boldsymbol{X}^u \uplus \boldsymbol{X}^d = \boldsymbol{X}} \bar{f}^u(\boldsymbol{X}^u|\boldsymbol{Z}_{1:N_{\text{sens}}}) \bar{f}^d(\boldsymbol{X}^d|\boldsymbol{Z}_{1:N_{\text{sens}}}) \tag{21}$$

comprised of an PPP $\bar{f}^u(\boldsymbol{X}^u|\boldsymbol{Z}_{1:N_{\text{sens}}})$ modeling the unknown targets and an MB $\bar{f}^d(\boldsymbol{X}^d|\boldsymbol{Z}_{1:N_{\text{sens}}})$ modeling detected targets. Note, for brevity we avoid the conditioning on the measurement set in the remainder of this section.

A challenge in the fusion step is the fact that sensors do not have the same FoVs (c.f. Fig. 1). In the process of fusion, any target that is in the FoV of one sensor and has been detected, but outside the FoV of another sensor, must be treated carefully. In this situation, prior to the fusion, for the first sensor the target corresponds to a detected target represented by a Bernoulli density, whereas for the second sensor the target corresponds to an unknown target, which is represented by the PPP. Because of this, in the closed-form KLA solution (20), we must fuse a Bernoulli with a part of the PPP intensity. From this, it follows that if we solve the KLA (18) separately for the PPP and the MB then we will obtain incorrect results. To enable a valid fusion, we propose the approach outlined in the following sub-sections.

Consider PMB densities with PPP intensity $\lambda_s(x)$ and N_s Bernoullis with parameters r_s^i and $f_s^i(x)$ indexed $i = 1, \ldots, N_s$. Fusion of the PMB densities is simplified if all have the same number of Bernoullis, however, in the general case one cannot assume this. Instead, we rely on a result from [49, Sec. 4.3.1]: any PPP with intensity $\lambda_s(x)$ can be divided into multiple independent PPPs with intensities $\lambda_s^j(x)$, where $\sum_j \lambda_s^j(x) = \lambda_s(x)$. Based on this result, for each sensor we divide $\lambda_s(x)$ into M_s parts such that $N_s + M_s = K$ for all sensors s, where parameter K is determined depending on the number of components of the MBs. A robust choice is $K = \sum_s N_s$, so that each Bernoulli component in one sensor can be assigned to any combination of PPP or Bernoullis in the other sensors. The FoV can be taken into account in order to reduce K[3].

If follows from the division into K parts that the PMB densities can be expressed as follows,

$$f_s(\boldsymbol{X}) = \sum_{\uplus_{i=1}^K \boldsymbol{X}^i = \boldsymbol{X}} \prod_{i=1}^K f_s^i(\boldsymbol{X}^i) \tag{22}$$

where $f_s^i(\boldsymbol{X}^i)$ is Bernoulli for $i \in \{1, \ldots, N_s\}$ and PPP with intensity $\lambda_s^i(x)$ for $i \in \{N_s + 1, \ldots, N_s + M_s\}$. Note that the sum in (22) is implicit and never has to be computed; it is sufficient to represent the parameters of the densities $f_s^i(\boldsymbol{X}^i)$.

We are seeking the fused density $\bar{f}(\boldsymbol{X})$ that minimizes the KLA (18). Assume that $\bar{f}(\boldsymbol{X})$ is of the format (22), i.e.,

$$\bar{f}(\boldsymbol{X}) = \sum_{\uplus_{i=1}^K \boldsymbol{X}^i = \boldsymbol{X}} \prod_{i=1}^K \bar{f}^i(\boldsymbol{X}^i). \tag{23}$$

In [50] it is shown that, under this assumption, the minimization problem (18) can be solved approximately by minimizing an upper bound,

$$\sum_{s=1}^{N_{\text{sens}}} w_s \left[\sum_{i=1}^K D\left(\bar{f}^i(\boldsymbol{X}) || f_s^{\pi_s(i)}(\boldsymbol{X}) \right) \right] \tag{24}$$

where $\pi_s \in \Pi_{(1:K)}$ for all s, and $\Pi_{(1:K)}$ is the set of all permutations of the integers $\{1, \ldots, K\}$. For $\pi \in \Pi$ and $i \in \{1, \ldots, K\}$ we have $\pi(i) \in \{1, \ldots, K\}$ and $\pi(i) \neq \pi(i')$ for $i \neq i'$. The fusion results will depend on how the permutations π_s are chosen, which we discuss in Section V-D.

Based on (24), we compute the K components of the fused density $\bar{f}(\boldsymbol{X})$ as

$$\bar{f}^i(\boldsymbol{X}) = \frac{\prod_{s=1}^{N_{\text{sens}}} \left(f_s^{\pi_s(i)}(\boldsymbol{X}) \right)^{w_s}}{\int \prod_{s=1}^{N_{\text{sens}}} \left(f_s^{\pi_s(i)}(\boldsymbol{X}) \right)^{w_s} \delta \boldsymbol{X}}. \tag{25}$$

[3] In particular, we partition the deployment region into $2^{N_{\text{sens}}}$ sub-regions, each determined by a subset of sensors that have each partition in the FoV. For each sub-region $l \in \{1, \ldots, 2^{N_{\text{sens}}}\}$, sensor s has $N_{s,l}$ Bernoulli components (with $N_s = \sum_l N_{s,l}$), we determine $K_l = \max N_{s,l}$ and $K = \sum_l K_l$.

In summary, our proposed approach to finding the (approximately) optimal fused PMB density consists of the following steps:

1) Depending on how the FoVs overlap, we select how to divide the PPP intensities such that the PMB densities all have K parts, and we find permutations π_s such that the fused densities $f_s^{\pi_s(i)}(\boldsymbol{X})$ have a high degree of similarity in the sense of the KL-divergence. Note that this entails determining K, $\lambda_s^j(x)$, and π_s.
2) Fuse the matched PPPs and Bernoullis, see details below in Section V-C.
3) After the fusion of the K parts of $\bar{f}(\boldsymbol{X})$, we add all the PPP intensities such that a single PPP is obtained.
4) Lastly, we recycle any Bernoullis in $\bar{f}(\boldsymbol{X})$ that have very low probability of existence.

C. FUSION OF BERNOULLIS AND PPPs

In this subsection, we provide expressions for the fusion of Bernoullis, fusion of PPPs, as well as the fusion of both Bernoullis and PPPs. Lastly, expressions for fusion of Gaussian densities are provided.

1) FUSION OF PPPs

Let $\mathbb{I}$ be an index set for PPP densities $f_i(\boldsymbol{X})$ with intensities $\lambda_i(x) = \mu_i f_i(x)$. Fusion of densities $f_i(\boldsymbol{X})$, $i \in \mathbb{I}$ with fusion weights w_i, $\sum_{i\in\mathbb{I}} w_i = 1$, yields a PPP density with intensity [40]

$$\bar{\lambda}(x) = \prod_{i\in\mathbb{I}} (\lambda_i(x))^{w_i} = \bar{\mu}\bar{f}(x), \tag{26a}$$

$$\bar{\mu} = C \prod_{i\in\mathbb{I}} \mu_i^{w_i}, \tag{26b}$$

$$\bar{f}(x) = \frac{\prod_{i\in\mathbb{I}} (f_i(x))^{w_i}}{C}, \tag{26c}$$

$$C = \int \prod_{i\in\mathbb{I}} (f_i(x))^{w_i}\, dx. \tag{26d}$$

2) FUSION OF BERNOULLI RFSs

Consider Bernoulli densities $f_i(\boldsymbol{X})$ with probability of existence r_i and state PDF $f_i(x)$. Fusion of densities $f_i(\boldsymbol{X})$, $i \in \mathbb{I}$, $|\mathbb{I}| \geq 1$, with fusion weights w_i, $\sum_{i\in\mathbb{I}} w_i = 1$, yields a Bernoulli density with parameters

$$\bar{r} = \frac{C \prod_{i\in\mathbb{I}} r_i^{w_i}}{\prod_{i\in\mathbb{I}}(1-r_i)^{w_i} + C \prod_{i\in\mathbb{I}} r_i^{w_i}}, \tag{27a}$$

$$\bar{f}(x) = \frac{\prod_{i\in\mathbb{I}} (f_i(x))^{w_i}}{C}, \tag{27b}$$

$$C = \int \prod_{i\in\mathbb{I}} (f_i(x))^{w_i}\, dx. \tag{27c}$$

3) FUSION OF BERNOULLI RFSs AND PPPs

Let $\mathbb{I}$, $|\mathbb{I}| \geq 1$, be an index set for PPP densities $f_i(\boldsymbol{X})$ with intensities $\lambda_i(x) = \mu_i f_i(x)$, and let $\mathbb{J}$, $|\mathbb{J}| \geq 1$, be an index set for Bernoulli densities $f_j(\boldsymbol{X})$ with probabilities of existence r_j and state densities $f_j(x)$. Fusion of PPPs and Bernoullis with fusion weights w_i and w_j, $\sum_{i\in\mathbb{I}} w_i + \sum_{j\in\mathbb{J}} w_j = 1$, yields (without approximation) a Bernoulli density with parameters

$$\bar{r} = \frac{C \prod_{j\in\mathbb{J}} r_j^{w_j} \prod_{i\in\mathbb{I}} \mu_i^{w_i}}{\prod_{j\in\mathbb{J}}(1-r_j)^{w_j} + C \prod_{j\in\mathbb{J}} r_j^{w_j} \prod_{i\in\mathbb{I}} \mu_i^{w_i}}, \tag{28a}$$

$$\bar{f}(x) = \frac{\prod_{i\in\mathbb{I}} (f_i(x))^{w_i} \prod_{j\in\mathbb{J}} \left(f_j(x)\right)^{w_j}}{C}, \tag{28b}$$

$$C = \int \prod_{i\in\mathbb{I}} (f_i(x))^{w_i} \prod_{j\in\mathbb{J}} \left(f_j(x)\right)^{w_j}\, dx. \tag{28c}$$

The fusion in (28) holds for any PPP intensities $\lambda_i(x)$, however, given that a Bernoulli RFS represents zero or one object, the fusion results will be more accurate for $\mu_i < 1$.

4) FUSION OF GAUSSIAN DENSITIES

When the Bernoulli state densities, and/or the PPP intensities are Gaussian, the fused densities $\bar{f}(\cdot)$ and the normalizing constants C can be computed exactly, see, e.g., [45, Eqn. 36] and [51]. Let $f_i(x) = \mathcal{N}(x; m_i, P_i)$ for $i \in \mathbb{I}$. Fusion of the densities, with weights w_i, $\sum_{i\in\mathbb{I}} w_i = 1$, yields a Gaussian density,

$$\bar{f}(x) = \frac{\prod_{i\in\mathbb{I}} (f_i(x))^{w_i}}{C} = \mathcal{N}\left(x; \bar{m}, \bar{P}\right), \tag{29a}$$

where

$$\bar{m} = \bar{P}\left(\sum_{i\in\mathbb{I}} w_i \left(P_i\right)^{-1} m_i\right), \tag{29b}$$

$$\bar{P} = \left(\sum_{i\in\mathbb{I}} w_i \left(P_i\right)^{-1}\right)^{-1}, \tag{29c}$$

$$\begin{aligned} C &= \int \prod_{i\in\mathbb{I}} (f_i(x))^{w_i}\, dx \\ &= \frac{\det(2\pi\bar{P})^{1/2}}{\prod_{i\in\mathbb{I}} \det(2\pi P_i)^{w_i/2}} \\ &\quad \times \exp\left[\frac{1}{2}\left(\bar{m}^T \bar{P}^{-1}\bar{m} - \sum_{i\in\mathbb{I}} w_i m_i^T P_i^{-1} m_i\right)\right]. \end{aligned} \tag{29d, 29e}$$

D. COMPLEXITY REDUCTION

From (24), we see that we need to fuse components of the PMBs over all permutations $\pi_s \in \Pi_{(1:K)}$. The KLD is the lowest when the RFS densities are identical. In our application, this means when the single-target posterior densities represent the same target. To reduce computational complexity towards finding the fused posterior density minimizing the KLA, we propose to use the optimal permutation, denoted the *best possible fusion map*. It is defined and found as follows. Let there be two RFSs densities with index sets $\mathbb{I}_1$ and $\mathbb{I}_s$ respectively. We define the best fusion map $\pi_s^* \in \Pi_{(1:K)}$ as

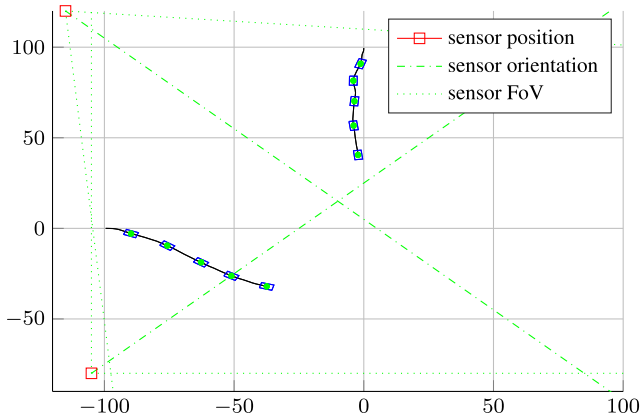


FIGURE 3. Simulation scenario with two ETs observed by two sensors with overlapping FoVs. The ETs states are plotted every 20 time steps (position is visualized by a bold green dot and contour by solid blue lines).

the solution of the optimal assignment problem [52]

$$\begin{aligned} \underset{\mathbf{a}}{\text{minimize}} \quad & \sum_{n=1}^{|\mathbb{I}_1|} \sum_{m=1}^{|\mathbb{I}_s|} a_{n,m} C_{n,m} \\ \text{subject to} \quad & \sum_{m=1}^{|\mathbb{I}_s|} a_{n,m} = 1, \qquad \forall n, \\ & \sum_{n=1}^{|\mathbb{I}_1|} a_{n,m} = 1, \qquad \forall m, \\ & a_{n,m} \in \{0, 1\}, \end{aligned} \tag{30}$$

where $C_{n,m}$ denotes the cost for assigning (mapping) component n in $f_1(\boldsymbol{X})$ with component m in $f_s(\boldsymbol{X})$. To solve (30), we can use, e.g., the Munkres algorithm [53], [54]. For $N_{\text{sens}} > 2$ the best fusion map is found by sequentially solving (30) for $s = 2, \ldots, N_{\text{sens}}$.

We define the cost metric in terms of the KLD between the PDFs of the components in $f_1(\boldsymbol{X})$ and $f_s(\boldsymbol{X})$. For a component $n \in \mathbb{I}_1$ with PDF $f_{1,n}(\boldsymbol{y})$, and similarly for $m \in \mathbb{I}_s$, the cost is defined

$$C_{n,m} = \frac{1}{2}\left[D(f_{1,n}\|f_{s,m}) + D(f_{s,m}\|f_{1,n})\right], \tag{31}$$

which admits a closed-form expression for Gaussian PDFs. Using the procedure above, we find the best fusion map and use it to solve (24).

VI. NUMERICAL RESULTS

We present first the target extent model used for the simulations. This is followed by the simulation setup, the used performance metrics, and a discussion of the obtained results using the independent PMB-ETT filter with and without posterior fusion. As was mentioned in Section I, the only comparable work is [36], which considered fusion of only the MB components, but not the PPP part. Hence, it can be seen as a special case of our proposed filter.

A. SINGLE ET MODEL

The specific target extent model is based on [8], here extended to include the sensor state (including position and orientation) in the measurement model. In the following, we describe the ET state and measurement model, and how this leads to extended Kalman filter (EKF) prediction and update equations that are utilized in the PMB-ETT filter.

1) ET STATE AND MOTION MODEL

The augmented state $\boldsymbol{x}_k = [\gamma_k, \boldsymbol{y}_k^{\mathsf{T}}]^{\mathsf{T}}$ of a single ET at time k comprises an unknown Poisson rate γ_k of number of measurements generated by the ET and the ET spatial state $\boldsymbol{y}_k$. The state has prior[4] PDF

$$f(\boldsymbol{x}_k) = \mathcal{G}(\gamma_k; \alpha_k, \beta_k)\mathcal{N}(\boldsymbol{y}_k; \hat{\boldsymbol{y}}_k, \boldsymbol{P}_k), \tag{32}$$

where the gamma distribution with parameters α_k and β_k is a conjugate prior for the rate γ_k, and the Gaussian distribution with mean $\hat{\boldsymbol{y}}_k$ and covariance matrix $\boldsymbol{P}_k$ describes the a priori knowledge regarding the ET spatial state $\boldsymbol{y}_k$. The ET spatial state (which includes the ET center, orientation, and extent) and motion model are based on [8] and detailed in Appendix B. Since the measurement rate is independent of the sensor state, the decentralized fusion of Sec. V is only applied to the spatial state $\boldsymbol{y}_k$, while each sensor maintains a local density of the rate γ_k (i.e., the average number of measurements obtained from target with specific sensor) of each target.

2) ET MEASUREMENT MODEL

In this section, the time index k will be omitted for the sake of brevity. A sensor s located at $\boldsymbol{p}_s \in \mathbb{R}^2$ with orientation α_s observes the ET contour in its local coordinate frame. We distinguish between three coordinate frames: a quantity in the *sensor coordinate frame* is indicated by superscript S, in the *ET coordinate frame* by superscript L, and in the *global coordinate frame* by superscript G. A measurement $\boldsymbol{z}^S \in \mathbb{R}^2$ is thus

$$\boldsymbol{z}^S = h(\boldsymbol{y}) + \boldsymbol{w}, \tag{33}$$

where $\boldsymbol{w} \sim \mathcal{N}(0, R)$ with measurement noise covariance R and

$$h(\boldsymbol{y}) = \boldsymbol{y}_c^S + \boldsymbol{e}(\theta^S)^{\mathsf{T}} f(\theta^L), \tag{34}$$

where $\theta^S = \angle(\boldsymbol{z}^S - \boldsymbol{y}_c^S)$, $\boldsymbol{e}(\theta^S)$ is a unit vector in direction θ^S, and $\theta^L = \theta^S - \psi^S$ in which $\boldsymbol{y}_c^S$ and ψ^S are the target location and orientation in the sensor frame of reference[5].

[4]Note that in the augmented state PDF (32) the measurement rate γ_k and the ET state (including its extent) $\boldsymbol{y}_k$ are considered independent, which is a common assumption in ETT (see e.g. [55]). Due to the sensor-to-target geometry (e.g., for a Lidar its angular resolution and the sensor-target distance), the estimated measurement rate of the ET is implicitly sensor dependent, i.e., a different sensor configuration yields a different estimate for γ_k. The explicit modeling of the dependence of γ_k on the sensor state and $\boldsymbol{y}_k$ is not considered here.

[5]Note that here the unknown angle θ^S is replaced by a point estimate, which is a simple but inaccurate approach and can be seen as a greedy association model [5]. It is also the approach taken by [8].

Here, $f(\theta^L) \geq 0$ is the extent of the target along local angle θ^L. We now express the observation in the global coordinate frame. It can be shown that $\boldsymbol{z}^S \sim \mathcal{N}(\tilde{h}(\boldsymbol{y}), \tilde{R})$ with

$$\tilde{h}(\boldsymbol{y}) = R(\alpha_s)(\boldsymbol{y}_c - \boldsymbol{p}_s) + \boldsymbol{e}^{\mathsf{T}}(\theta^S)H^f(\theta^L)\boldsymbol{y}^f, \tag{35}$$

$$\tilde{R} = R + k^f \boldsymbol{e}(\theta^S)\boldsymbol{e}^{\mathsf{T}}(\theta^S), \tag{36}$$

where $R(\alpha_s)$ is a rotation matrix, $H^f(\theta^L)$, and k^f are defined in Appendix C.

3) ET EXTENDED KALMAN FILTER EQUATIONS

Prediction and update equations of an EKF filter using the ET motion and measurement models are given in Appendix D. The prediction and update equations for the measurement rate and the predicted likelihood are also stated. All these are utilized in the prediction and update step of the PMB-ETT filter (c.f. Section IV).

B. SETUP

If not stated otherwise, there are two ETs present in the scene and their visibility and number of measurements produced per scan depend on the sensor FoV and its configuration. We use a rectangular target of length 5 m and width 3 m to model ETs representing vehicles. Furthermore, we use Lidar type sensors with the following simplified sensor models. Sensor 1 is located at $\boldsymbol{p}_{S_1} = [-115, 120]^{\mathsf{T}}$, with orientation $\alpha_{S_1} = -45°$, opening angle of 80°, angular resolution of 0.15°, and maximum range of 300 m. We generate a measurement when a ray from the sensor hits an ET and add noise with covariance matrix $R_{S_1} = 0.5\boldsymbol{I}_2$. Sensor 2 is located at $\boldsymbol{p}_{S_2} = [-105, -80]^{\mathsf{T}}$, with orientation $\alpha_{S_2} = 45°$, opening angle of 90°, angular resolution of 0.15°, maximum range of 300 m, and $R_{S_2} = 0.02\boldsymbol{I}_2$. Each sensor produces clutter measurements with rate $\lambda = 2$ and uniform spatial distribution $c(\boldsymbol{z}) = \mathcal{U}[-200, 200]^2$.

In the simulation, each independent PMB-ETT filter has only access to measurements from one sensor (denoted indep. filter). The fusion filters are independent PMB-ETT filters (denoted fusion filter), but perform posterior fusion according to Sec. V. We approximate the PMBM posterior by a PMB which consists of the MB in the MBM that has the highest weight. If not stated otherwise, posterior fusion is applied in every time step. We set the weights $w_s = 1/N_{\text{sens}}$, resulting in a more conservative estimate than achievable. The fused posterior is then used as the prior for the next filter iteration. For comparison, an PMB-ETT filter that incorporates measurements from all sensors is used (denoted centralized filter). There, measurements from each sensor are incorporated separately by applying multiple sequential PMB-ETT filter update steps. The timing of the updates for the different filters is visualized in Fig. 4: the independent filters perform a measurement and an update each sampling period T (see further). The centralized filter takes both these measurements and performs a single update each period T. Finally, the posterior fusion takes the posteriors from both sensors and fuses them each period T.

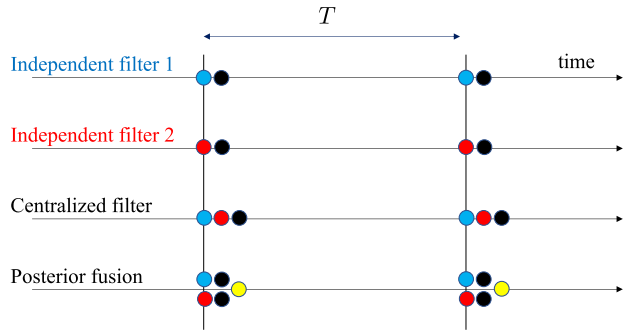


FIGURE 4. Timing of updates: blue/red circles are measurements, black circles are posterior updates, and yellow circles represent a posterior fusion.

For all filter variants, spatially close measurements are clustered into measurement cells using the DBSCAN clustering algorithm [56], where we set the maximum radius for the neighborhood to 4 m and the minimum number of points for a core point to 4. The simulation scenario is outlined in Fig. 3. The hyperparameters of the GP (see Appendix B) are $l^2 = \pi/8$, $\sigma_f^2 = 2$, $\sigma_r^2 = 2$, and 20 support points are used to track the target extent. Note that the dimension of the ET state is 26 (xy-position, orientation, xy-velocity, angular velocity, target extent support points). The target motion model and its model in the filter are identical with sampling time $T = 0.5$ s,

$$\bar{F}_k = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \otimes \boldsymbol{I}_3, \tag{37}$$

$$\bar{W}_k = \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix} \otimes \text{diag}([0.01, 0.01, 0.001]), \tag{38}$$

$\beta = 0.001$ in F_f (c.f. (51)), and the forgetting factor is set to $\frac{1}{\eta} = \frac{1}{1.11}$. In the filters, the birth intensity has rate $\lambda^b = \frac{1}{10}$ and for the spatial distribution we use a single Gaussian centered at location $x = [0, 100]^{\mathsf{T}}$ with covariance matrix $P = 30\boldsymbol{I}_2$. The probability of ET survival is $p_{\text{S}} = 0.999$, the probability of detection is $p_{\text{D}} = 0.99$. Hence, the filter is subject to both missed detections and false alarms.

C. PERFORMANCE METRICS

Multiple target tracking performance is measured by three errors: estimation error for localized targets, number of missed targets, and number of false targets, see, e.g., [57, Sec. 13.6]. The generalized optimal subpattern assignment (GOSPA) metric [58], [59] measures all three errors, hence we use it for performance evaluation[6]. Performance of the estimated number of targets as well as their center location is thus assessed as follows. Let sets $\hat{\boldsymbol{X}} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ and $\boldsymbol{Y} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$ be finite subsets of $\mathbb{R}^N$, where without loss

[6]In tracking literature, the OSPA metric [60] is often used. However, recent work [61] has shown that the OSPA metric is susceptible to "spooky action at a distance", which is undesirable. Hence, we do not use the OSPA metric in this paper.

of generality $n \leq m$. Then [58], [59]

$$d_{\mathrm{GOSPA}}^{(c,\alpha,p)} = \left(\min_{I_n \in \mathcal{F}_n(\{1,\ldots,n\})} \sum_{i=1}^{n} d^{(c)}(\boldsymbol{x}_i, \boldsymbol{y}_{I_n(i)})^p + \frac{c^p}{\alpha}(m-n) \right)^{\frac{1}{p}}, \tag{39}$$

where we set the power parameter $p = 2$, cut-off distance $c = 20$, $\alpha = 2$, and $d^{(c)}(\boldsymbol{x}, \boldsymbol{y}) = \min(\|\boldsymbol{x} - \boldsymbol{y}\|_2, c)$, i.e., the minimum of the Euclidean distance and value c. To obtain $\hat{\boldsymbol{X}}$, we estimate the detected ETs from the (fused) posterior through comparison of the probability of existence of each Bernoulli component against the threshold $r_{\mathrm{th}} = 0.5$. The set $\boldsymbol{Y}$ contains the true ETs. Note that we only use the mean position of the target center (c.f. Sec. II-B).

The performance of the target extent estimation is assessed with the intersection over union (IOU) of the true target shape (in the global coordinate frame) and the estimated shape. Let A_k be the true ET area in xy-dimension at time step k, and $\hat{A}_k$ its estimate. Then, the IOU is defined as, see, e.g., [8],

$$\mathrm{IOU}(A_k, \hat{A}_k) = \frac{\mathrm{area}(A_k \cap \hat{A}_k)}{\mathrm{area}(A_k \cup \hat{A}_k)}. \tag{40}$$

Note that the IOU is, by definition, always between zero for non-overlapping target shapes and one when they fully overlap. Thus a well-performing ETT filter will yield a high IOU for every ET. For both GOSPA and IOU average performance results were obtained by averaging over 50 Monte-Carlo runs.

D. DISCUSSION OF RESULTS

Here, we first discuss the ETT filter performance with decentralized posterior fusion in terms of GOSPA and IOU performance metrics. This is followed by performing posterior fusion at a lower rate. After that, we investigate the case when more than two ETT filters are used for posterior fusion.

1) FILTER PERFORMANCE WITH POSTERIOR FUSION

In Fig. 5, the true and estimated target state is plotted for ET 2 using the different filter variants. We observe that independent filter 1 (Fig. 5a) estimates the target state with a clear position error (in the positive y-direction). This filter uses measurements from sensor 1, where most of the measurements provide information only from the ET's top edge due to the horizontal movement of the target and the sensor pose. In the measurement model (33), occlusions caused by the ET itself are not modeled. Due to the simple data association that is used, the received measurements can then be explained by an ET whose target contour is in the proximity of the measurements and the target center is placed north of it. In contrast, independent filter 2 and the centralized filter estimate the target center closer to the true position (Fig. 5b and Fig. 5c). The former filter overestimates the target size in the direction where no measurement is provided, whereas the latter filter utilizes measurements from both sensors and can, therefore, estimate the target size accurately. The fusion filter utilizes information from both sensors through posterior fusion.

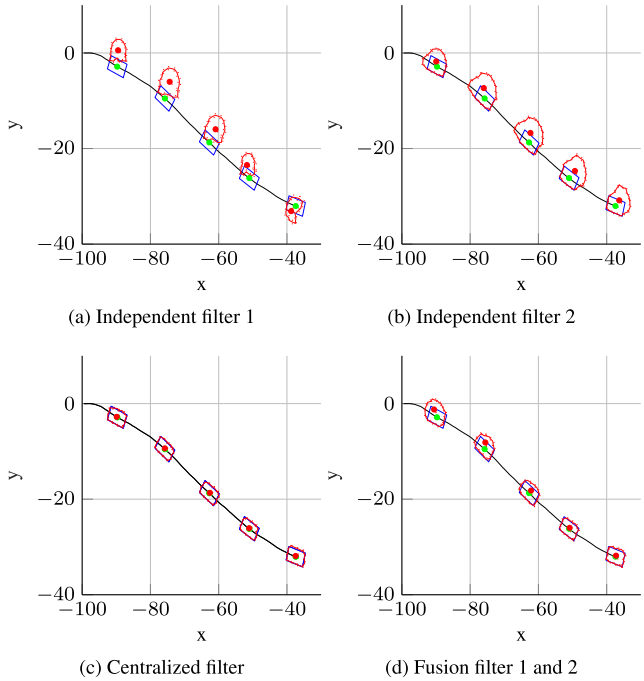


FIGURE 5. Estimated shape of one ET for different PMB-ETT filters. The true ET center and extent is plotted (green dot, blue solid line), as well as the estimated ones (red dot, red solid line for mean extent, red dotted line for one standard deviation). The standard deviation of the mean is not shown for clarity.

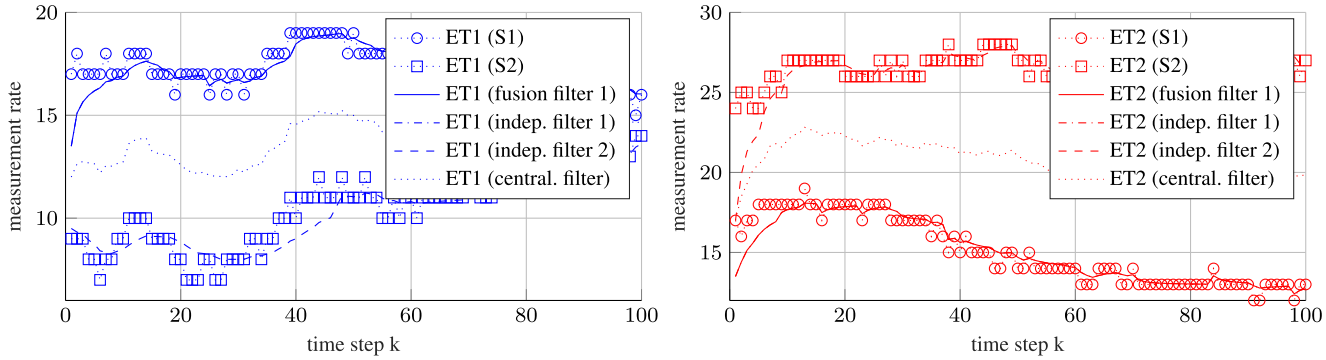
In Fig. 6, the true and estimated measurement rate γ is plotted over time. We can observe that the true measurement rate of the ETs is varying over time. Although only a simple process model for the measurement rate is used in the filters, they can correctly track the rate. This is true for all filter variants except for the centralized filter. This filter performs two sequential filter update steps using measurements from different sensors. Since the number of measurements for an ET are different for each sensor, it follows that after centralized fusion the filter estimates the measurement rate of the ET as the average of the two.

In Fig. 7, the average GOSPA is plotted over time for the simulation scenario illustrated in Fig. 3. The centralized filter has the best performance followed by the fusion filter. Independent filter 2 has superior performance compared to the independent filter 1. Note that independent filter 1 is provided with measurements from sensor 1, which has a higher measurement noise compared to sensor 2.

In Table 1, the average IOU is stated for the different ETs and filter variants. We observe that the IOU for ET 2 is low with independent filter 1 due to the misplaced target center, and with independent filter 2 due to the overestimation of the target size. The fusion and the centralized filter show comparable performance.

2) LOW RATE POSTERIOR FUSION

In a real system, it may not be feasible to perform posterior fusion after every filter update step. This can occur when the computers on which the filters run are geographically



(a) ET 1 (b) ET 2

FIGURE 6. The true and estimated measurement rate for each ET and sensor/filter is plotted over time.

TABLE 1. Average IOU.

Filter	ET 1	ET 2
Independent 1	0.65	0.25
Independent 2	0.71	0.43
Fusion	0.72	0.75
Centralized	0.70	0.85

TABLE 2. Average GOSPA.

Filter	$N = 1$	$N = 15$	$N = 30$	$N = 50$
Fusion 1	1.12	1.65	2.20	2.76
Fusion 2	1.12	1.80	2.03	2.10
Independent 1	3.20	3.20	3.20	3.20
Independent 2	2.17	2.17	2.17	2.17
Centralized	0.89	0.89	0.89	0.89

separated and need to communicate over the wireless channel. Therefore, it is worthwhile to investigate the filter performance when posterior fusion is performed only every N time steps.

In Table 2, the average GOSPA is stated for different values of N. We see that with increasing N the performance of the fusion filters deteriorates, since the information transfer (through fusion) between the filters over time is too low. Fusion filter 1 has worse performance compared to fusion filter 2 for $N \geq 15$, since it is equipped with the low-quality sensor 1.

3) POSTERIOR FUSION WITH MORE FILTERS

In Sec. V, we proposed a procedure to fuse PMB posteriors when there are more than two independent PMB-ETT filters. We implement this sequentially, where first posteriors from two filters are fused. The outcome is then used to fuse with a not yet fused posterior from one of the remaining filters. This process is repeated until all filter posteriors have been incorporated. We placed four sensors at $\boldsymbol{p}_{S_1} = [-150, -80]^\mathsf{T}$, $\boldsymbol{p}_{S_2} = [-150, -50]^\mathsf{T}$, $\boldsymbol{p}_{S_3} = [-150, -20]^\mathsf{T}$, and $\boldsymbol{p}_{S_4} = [-150, 10]^\mathsf{T}$ all with overlapping sensor FoVs towards the ETs. The remaining sensor parameters are the same as for sensor 1 used in the previous simulations. In Table 3, the average GOSPA, as well as the average IOU per ET are stated for

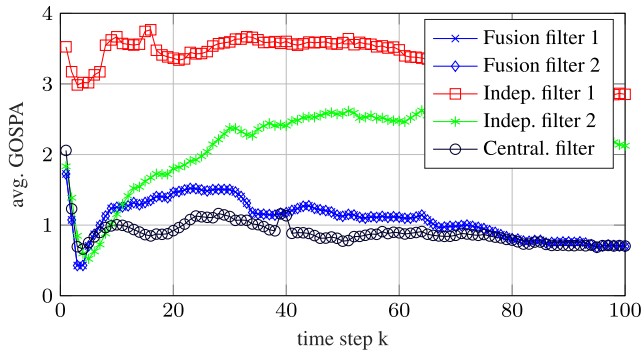


FIGURE 7. The average GOSPA value is plotted over time. Posterior fusion is performed in every time step.

TABLE 3. Multiple posterior fusion.

No. Posteriors Fused	GOSPA	IOU (ET 1)	IOU (ET 2)
(no fusion)	1.22	0.61	0.62
2	0.90	0.66	0.67
4	0.86	0.68	0.66

a single independent PMB-ETT filter (no fusion), two filters with posterior fusion performed after every filter update, and four filters with posterior fusion. With posterior fusion, performance increases, visible by a decrease of the GOSPA value. Also, the average IOU increases for all ETs with posterior fusion. Furthermore, fusion of two filter posteriors and four filter posteriors show similar performance in this scenario. Note, however, that for the IOU of ET 2, two sensors perform slightly better than four sensors. We believe this can be ascribed to the randomness of the simulations.

VII. CONCLUSIONS

We proposed a low-complexity decentralized ETT filter that is capable of estimating the presence, state, and shape of ETs accurately. It operates by combining the multiobject densities of PMB form computed by independent ETT filters. Fusion is performed in the minimum KLA sense yielding a fused posterior which is conservative but never overconfident about the estimated states. A low-complexity implementation is

highlighted, which permits the use of an optimal fusion mapping between pairs of sensors. The fusion map was identified as the solution of a linear optimal assignment problem based on a cost matrix comprised of the symmetric KLD between target state estimates.

In the simulation results, we observed how the independent PMB-ETT filters together with the GP target extent model are capable of estimating the state and shape of the present ETs. Furthermore, we observed that fusion of the filters posterior permits a holistic view of the surveillance region spanned by all sensors combined. This resulted in a reduced state estimation error quantified by the GOSPA distance metric, and for the ET shape estimation by an increased area overlap quantified by the IOU value.

Additional work is needed to evaluate the performance when the number of targets in the FoV is varying and FoVs are partially overlapping.

APPENDIX

A. ET-PMB PARAMETERS

Following [3, Sec. IV], we find that

$$w_A = \frac{\prod_{C\in A}\mathcal{L}_C}{\sum_{A\in\mathcal{A}}\prod_{C\in A}\mathcal{L}_C}, \tag{41}$$

$$\mathcal{L}_C = \begin{cases} \kappa^{\boldsymbol{C}_C} + \left\langle D_+^u, \ell_{\boldsymbol{C}_C}\right\rangle, & \text{if } C\cap\mathbb{I}=\emptyset, |\boldsymbol{C}_C|=1,\\ \left\langle D_+^u, \ell_{\boldsymbol{C}_C}\right\rangle, & \text{if } C\cap\mathbb{I}=\emptyset, |\boldsymbol{C}_C|>1,\\ 1-r_+^{i_C}+r_+^{i_C}\left\langle f_+^{i_C}, q_{\mathrm{D}}\right\rangle, & \text{if } C\cap\mathbb{I}\neq\emptyset, \boldsymbol{C}_C=\emptyset,\\ r_+^{i_C}\left\langle f_+^{i_C}, \ell_{\boldsymbol{C}_C}\right\rangle, & \text{if } C\cap\mathbb{I}\neq\emptyset, \boldsymbol{C}_C\neq\emptyset. \end{cases} \tag{42}$$

The density $f_C(\boldsymbol{X}^C)$ in (15) is a Bernoulli density with parameters

$$r_C = \begin{cases} \dfrac{\left\langle D_+^u, \ell_{\boldsymbol{C}_C}\right\rangle}{\kappa^{\boldsymbol{C}_C} + \left\langle D_+^u, \ell_{\boldsymbol{C}_C}\right\rangle}, & \text{if } C\cap\mathbb{I}=\emptyset, \quad |\boldsymbol{C}_C|=1,\\ 1, & \text{if } C\cap\mathbb{I}=\emptyset, \quad |\boldsymbol{C}_C|>1,\\ \dfrac{r_+^{i_C}\left\langle f_+^{i_c}, q_{\mathrm{D}}\right\rangle}{1-r_+^{i_C}+r_+^{i_C}\left\langle f_+^{i_C}, q_{\mathrm{D}}\right\rangle}, & \text{if } C\cap\mathbb{I}\neq\emptyset, \quad \boldsymbol{C}_C=\emptyset,\\ 1, & \text{if } C\cap\mathbb{I}\neq\emptyset, \quad \boldsymbol{C}_C\neq\emptyset, \end{cases} \tag{43}$$

$$f_C(\boldsymbol{x}) = \begin{cases} \dfrac{\ell_{\boldsymbol{C}_C} D_+^u(\boldsymbol{x})}{\left\langle D_+^u, \ell_{\boldsymbol{C}_C}\right\rangle}, & \text{if } C\cap\mathbb{I}=\emptyset,\\ \dfrac{q_D(\boldsymbol{x}) f_+^{i_C}(\boldsymbol{x})}{\left\langle f^{i_C}, q_{\mathrm{D}}\right\rangle}, & \text{if } C\cap\mathbb{I}\neq\emptyset, \quad \boldsymbol{C}_C=\emptyset,\\ \dfrac{\ell_{\boldsymbol{C}_C}(\boldsymbol{x}) f_+^{i_C}(\boldsymbol{x})}{\left\langle f_+^{i_C}, \ell_{\boldsymbol{C}_C}\right\rangle}, & \text{if } C\cap\mathbb{I}\neq\emptyset, \quad \boldsymbol{C}_C\neq\emptyset. \end{cases} \tag{44}$$

B. ET SPATIAL STATE AND MOTION MODEL

1) SPATIAL STATE

The ET spatial state including target extent is given by

$$\boldsymbol{y} = [(\bar{\boldsymbol{y}})^\mathsf{T}, (\boldsymbol{y}^{\boldsymbol{f}})^\mathsf{T}]^\mathsf{T}, \tag{45}$$

where

$$\bar{\boldsymbol{y}} = [(\boldsymbol{y}_c)^\mathsf{T}, \psi, \boldsymbol{y}^{*\mathsf{T}}]^\mathsf{T} \tag{46}$$

comprising the ET center $\boldsymbol{y}_c$, the ET orientation ψ, and any additional quantities (e.g, velocity) in $\boldsymbol{y}^*$. The variable $\boldsymbol{y}^{\boldsymbol{f}}$ models the target extent, following [8]: Let u denote the local angle w.r.t. the ET orientation and y_i^f denotes the unknown target extent along input (angle) u_i, for a fixed and finite set of N angles. Then

$$\boldsymbol{y}^{\boldsymbol{f}} = [y^{f_1}, \ldots, y^{f_N}]^\mathsf{T}. \tag{47}$$

This vector is modeled as a zero-mean GP [62], [63]

$$\boldsymbol{y}^{\boldsymbol{f}} \sim \mathcal{GP}(\boldsymbol{0}, K(\boldsymbol{u}, \boldsymbol{u})) \tag{48}$$

with covariance matrix $[K(\boldsymbol{u}, \boldsymbol{u})]_{i,j} = k(u_i, u_j)$, in which $k(\cdot,\cdot)$ is a periodic kernel function. The input of the GP is $\boldsymbol{u} = [u_1, \ldots, u_N]^\mathsf{T}$, and the output is $\boldsymbol{y}^{\boldsymbol{f}}$. We utilize the periodic kernel function proposed in [8]

$$k(u, u') = \sigma_f^2 \exp\left(-\frac{2\sin^2\left(\frac{|u-u'|}{2}\right)}{l^2}\right) + \sigma_r^2, \tag{49}$$

where σ_f, l, and σ_r are the (known) model hyper-parameters. This function is 2π periodic, i.e., $k(u+2\pi, u') = k(u, u')$, and, thanks to σ_r, star convex object shapes of different sizes can be described. See [8], [9] for further details and different choices for the kernel function to describe the extent of an ET with the help of a GP.

2) MOTION MODEL

Extending Sec. III-B, the ET follows the linear dynamic model

$$\bar{\boldsymbol{y}}_{k+1} = \bar{F}_k \bar{\boldsymbol{y}}_k + \bar{\boldsymbol{w}}_k, \tag{50}$$

where $\bar{F}_k$ denotes the state transition matrix, and $\bar{\boldsymbol{w}}_k \sim \mathcal{N}(0, \bar{W}_k)$ with process noise covariance $\bar{W}_k$. The motion model of the ET contour is [8]

$$\boldsymbol{y}_{k+1}^{\boldsymbol{f}} = F_k^{\boldsymbol{f}} \boldsymbol{y}_k^{\boldsymbol{f}} + \boldsymbol{w}_k^{\boldsymbol{f}}, \tag{51}$$

where $F^{\boldsymbol{f}} = e^{-\beta T}\boldsymbol{I}_N$ with $\boldsymbol{I}_N$ denoting the identity matrix of dimension N, and $\boldsymbol{w}_k^{\boldsymbol{f}} \sim \mathcal{N}(0, W^{\boldsymbol{f}})$ with $W^{\boldsymbol{f}} = (1-e^{-2\beta T})K(\boldsymbol{u}^{\boldsymbol{f}}, \boldsymbol{u}^{\boldsymbol{f}})$. Here, $\beta \geq 0$ denotes the forgetting factor allowing to accommodate targets with extents that change slowly, and T is the sampling time.

According to (32), the measurement rate of the ET is assumed independent of the ETs' spatial state. To allow the measurement rate to change over time an exponential forgetting factor $1/\eta$ is used in the motion model, where the predicted rate is given by the motion of the gamma distribution parameters with [3], [55]

$$\alpha_{k|k-1} = \alpha_{k-1}/\eta, \tag{52}$$

$$\beta_{k|k-1} = \beta_{k-1}/\eta. \tag{53}$$

C. PROOF OF (49)–(50)

Due to the GP model, we follow [8] and express $f(\theta^L)$ as

$$f(\theta^L) = H^f(\theta^L)\boldsymbol{y}^f + e^f, \tag{54}$$
$$H^f(\theta^L) = K(\theta^L, \boldsymbol{u}^f)[K(\boldsymbol{u}^f, \boldsymbol{u}^f)]^{-1}, \tag{55}$$

in which $e^f \sim \mathcal{N}(0, k^f)$ with

$$k^f = k(\theta^L, \theta^L) - H^f(\theta^L)K(\theta^L, \boldsymbol{u}^f)^\mathsf{T}. \tag{56}$$

Separating signal and noise contribution, and expressing the local states in the global frame of reference through

$$\boldsymbol{y}_c^S = R(\alpha_s)(\boldsymbol{y}_c - \boldsymbol{p}_s) \tag{57}$$
$$\psi^S = \psi - \alpha_s, \tag{58}$$

where $R(\alpha)$ denotes the rotation matrix

$$R(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}, \tag{59}$$

we find that with the GP contour model measurement $\boldsymbol{z}^S$ in (33) has mean and covariance given by (35)–(36).

D. ET PREDICTION AND UPDATE STEPS

Here, we first describe the EKF prediction and update steps for the ET's spatial state. This is followed by the update step of the ET measurement rate, and lastly the predicted likelihood utilized in the update step of the PMB-ETT filter for the ET state model of Section III.

1) EKF PREDICTION AND UPDATE EQUATIONS

With the linear ET motion model the standard EKF prediction step with initial state $\boldsymbol{y}_0 \sim \mathcal{N}(\hat{\boldsymbol{y}}_0, P_0)$ is [37]

$$\hat{\boldsymbol{y}}_{k|k-1} = F_k \hat{\boldsymbol{y}}_{k-1}, \tag{60}$$
$$P_{k|k-1} = F_k P_{k-1} F_k^\mathsf{T} + W_k, \tag{61}$$

where $F_k = \text{blkdiag}(\bar{F}_k, F_k^f)$, and $W_k = \text{blkdiag}(\bar{W}_k, W_k^f)$.

We now extend the EKF update steps derived in [8] to incorporate the (known) sensor state (position $\boldsymbol{p}_s$ and orientation α_s). The standard EKF measurement update equations for a detection $\boldsymbol{z}_k$ are [37]

$$H_k = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{y}_k}\tilde{h}(\boldsymbol{y})|_{\boldsymbol{y}=\hat{\boldsymbol{y}}_{k|k-1}}, \tag{62}$$
$$S_k = H_k P_{k|k-1} H_k^\mathsf{T} + R_k, \tag{63}$$
$$K_k = P_{k|k-1} H_k^\mathsf{T} S_k^{-1}, \tag{64}$$
$$\hat{\boldsymbol{y}}_k = \hat{\boldsymbol{y}}_{k|k-1} + K_k(\boldsymbol{z}_k - \tilde{h}(\hat{\boldsymbol{y}}_{k|k-1})), \tag{65}$$
$$P_k = P_{k|k-1} - K_k H_k P_{k|k-1}, \tag{66}$$

where $\tilde{h}(\cdot)$ was defined in (35) and (36). To linearize the measurement function, we need to compute [8]

$$H_k = \left[\frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\boldsymbol{y}_c}, \frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\psi}, \frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\boldsymbol{y}^*}, \frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\boldsymbol{y}^f}\right], \tag{67}$$

where in our case $\frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\boldsymbol{y}^*} = 0$. We get

$$\frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\boldsymbol{y}^f} = \boldsymbol{e}(\theta^S)H^f(\theta^L), \tag{68}$$
$$\frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\psi} = \boldsymbol{e}(\theta^S)\frac{\mathrm{d}}{\mathrm{d}\psi}H^f(\theta^L)\boldsymbol{y}^f, \tag{69}$$
$$\frac{\mathrm{d}}{\mathrm{d}\psi}H^f(\theta^L) = -\left.\frac{\partial}{\partial u}H^f(u)\right|_{u=\theta^L}, \tag{70}$$

where [8]

$$\frac{\mathrm{d}H^f(u)}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}u}K(u, \boldsymbol{u}^f)[K(\boldsymbol{u}^f, \boldsymbol{u}^f)]^{-1}, \tag{71}$$
$$\frac{\mathrm{d}K(u, \boldsymbol{u}^f)}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}u}[k(u, u_1^f), \ldots, k(u, u_N^f)], \tag{72}$$
$$\frac{\mathrm{d}k(u, u_i^f)}{\mathrm{d}u} = -\frac{1}{l^2}\sin(u - u_i^f)k(u, u_i^f). \tag{73}$$

Further,

$$\frac{\mathrm{d}\tilde{h}(\boldsymbol{y})}{\mathrm{d}\boldsymbol{y}_c} = R(\alpha_s) + \left.\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{u})\right|_{\boldsymbol{u}=\boldsymbol{y}_c}(H^f(\theta^L)\boldsymbol{y}^f)^\mathsf{T} + \boldsymbol{e}(\boldsymbol{y}_c) \times \left(\left(\left(\left.\frac{\mathrm{d}}{\mathrm{d}u}H^f(u)\right|_{u=\theta^L}\right)^\mathsf{T}\left.\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{w}}\theta^L(\boldsymbol{w})\right|_{\boldsymbol{w}=\boldsymbol{y}_c}\right)^\mathsf{T}\boldsymbol{y}^f\right)^\mathsf{T}, \tag{74}$$

where [8]

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{u}) = \frac{(\boldsymbol{z}^S - \boldsymbol{u})(\boldsymbol{z}^S - \boldsymbol{u})^\mathsf{T}}{\|\boldsymbol{z}^S - \boldsymbol{u}\|^3} - \frac{1}{\|\boldsymbol{z}^S - \boldsymbol{u}\|}\boldsymbol{I}_2, \tag{75}$$
$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{w}}\theta^L(\boldsymbol{w}) = \frac{1}{\|\boldsymbol{z}^S - \boldsymbol{y}_c^S\|^2} \times \left[\boldsymbol{z}^{SY} - \boldsymbol{y}_c^{SY}, -(\boldsymbol{z}^{SX} - \boldsymbol{y}_c^{SX})\right]R(\alpha_s). \tag{76}$$

Here, the superscript $\boldsymbol{b}^X$ and $\boldsymbol{b}^Y$ correspond to the first and the second dimension of the vector $\boldsymbol{b}$. Furthermore, we wrote $\boldsymbol{e}(\boldsymbol{y}_c)$ for $\boldsymbol{e}(\theta^S)$ to indicate the state dependency.

To update the ET spatial state by a set of detections $\boldsymbol{W} = \{\boldsymbol{z}_{k,l}\}_{l=1}^{n_k}$, we augment the measurement vector

$$\boldsymbol{z}_k = [\boldsymbol{z}_{k,1}^\mathsf{T}, \ldots, \boldsymbol{z}_{k,n_k}^\mathsf{T}]^\mathsf{T}, \tag{77}$$

where

$$R_k = \text{diag}(R_{k,1}, \ldots, R_{k,n_k}), \tag{78}$$
$$\tilde{h}_k(\boldsymbol{y}_k) = [\tilde{h}_{k,1}(\boldsymbol{y}_k)^\mathsf{T}, \ldots, \tilde{h}_{k,n_k}(\boldsymbol{y}_k)^\mathsf{T}]^\mathsf{T}. \tag{79}$$

2) ET MEASUREMENT RATE

The predicted ET measurement rate $\gamma_{k|k-1}$ has parameters $\alpha_{k|k-1}$ and $\beta_{k|k-1}$ (c.f. (32), (53)), which are updated for a set of detections $\mathbf{W}$ by [3]

$$\alpha_k = \alpha_{k|k-1} + |\boldsymbol{W}|, \tag{80}$$
$$\beta_k = \beta_{k|k-1} + 1. \tag{81}$$

3) PREDICTED LIKELIHOOD FOR ET-PMB FILTER

The predicted likelihood, used in the update step of the PMB-ETT filter (c.f. Sec. IV-2 and Appendix A), for a set of detections $\boldsymbol{W}$ for a single ET is

$$\ell_{\boldsymbol{W}} = \frac{\Gamma(\alpha_k)\beta_{k|k-1}^{\alpha_{k|k-1}}}{\Gamma(\alpha_{k|k-1})\beta_k^{\alpha_k}} \prod_{l=1}^{|\boldsymbol{W}|} \mathcal{N}(z_{k,l} - \tilde{h}(\hat{\mathbf{y}}_{k|k-1,l}), S_{k,l}), \tag{82}$$

where $\Gamma(\cdot)$ denotes the gamma function.

REFERENCES

[1] Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, vol. 19. London, U.K.: YBS, 1995.
[2] R. P. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.
[3] K. Granström, M. Fatemi, and L. Svensson, "Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 208–225, Feb. 2020, doi: 10.1109/TAES.2019.2920220.
[4] K. Granström, M. Fatemi, and L. Svensson, "Gamma Gaussian inverse-Wishart Poisson multi-Bernoulli filter for extended target tracking," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2016, pp. 893–900.
[5] K. Granström, M. Baum, and S. Reuter, "Extended object tracking: Introduction, overview, and applications," *J. Adv. Inf. Fusion*, vol. 12, no. 2, pp. 139–174, Dec. 2017.
[6] 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. Int. Conf. Inf. Fusion*, Salamanca, Spain, Jul. 2014, pp. 1–8.
[7] K. Granström and C. Lundquist, "On the use of multiple measurement models for extended target tracking," in *Proc. Int. Conf. Inf. Fusion*, Istanbul, Turkey, Jul. 2013, pp. 1534–1541.
[8] N. Wahlström and E. Özkan, "Extended target tracking using Gaussian processes," *IEEE Trans. Signal Process.*, vol. 63, no. 16, pp. 4165–4178, Aug. 2015.
[9] T. Hirscher, A. Scheel, S. Reuter, and K. Dietmayer, "Multiple extended object tracking using Gaussian processes," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, 2016, pp. 868–875.
[10] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 1042–1059, Jul. 2008.
[11] M. Baum and U. D. Hanebeck, "Extended object tracking with random hypersurface models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 149–159, Jan. 2014.
[12] W. Aftab, A. De Freitas, M. Arvaneh, and L. Mihaylova, "A Gaussian process approach for extended object tracking with random shapes and for dealing with intractable likelihoods," in *Proc. 22nd Int. Conf. Digit. Signal Process. (DSP)*, Aug. 2017, pp. 1–5.
[13] E. Özkan, N. Wahlström, and S. J. Godsill, "Rao-Blackwellised particle filter for star-convex extended target tracking models," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, 2016, pp. 1193–1199.
[14] R. Mahler, "PHD filters for nonstandard targets, I: Extended targets," in *Proc. 12th Int. Conf. Inf. Fusion*, Seattle, WA, USA, Jul. 2009, pp. 915–921.
[15] K. Granström, C. Lundquist, and O. Orguner, "Extended target tracking using a Gaussian-mixture PHD filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3268–3286, Oct. 2012.
[16] K. Granström and U. Orguner, "A phd filter for tracking multiple extended targets using random matrices," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5657–5671, Nov. 2012.
[17] A. Swain and D. Clark, "The PHD filter for extended target tracking with estimable extent shape parameters of varying size," in *Proc. Int. Conf. Inf. Fusion*, Singapore, Jul. 2012, pp. 1111–1118.
[18] C. Lundquist, K. Granström, and U. Orguner, "An extended target CPHD filter and a gamma Gaussian inverse Wishart implementation," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 472–483, Jun. 2013.
[19] M. Beard, S. Reuter, K. Granström, B.-T. Vo, B.-N. Vo, and A. Scheel, "Multiple extended target tracking with labeled random finite sets," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1638–1653, Apr. 2016.
[20] A. Scheel, S. Reuter, and K. Dietmayer, "Using separable likelihoods for laser-based vehicle tracking with a labeled multi-Bernoulli filter," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, 2016, pp. 1200–1207.
[21] A. Scheel, S. Reuter, and K. Dietmayer, "Vehicle tracking using extended object methods: An approach for fusing radar and laser," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 231–238.
[22] M. Michaelis, P. Berthold, D. Meissner, and H.-J. Wuensche, "Heterogeneous multi-sensor fusion for extended objects in automotive scenarios using Gaussian processes and a GMPHD-filter," in *Proc. Sensor Data Fusion, Trends, Solutions, Appl. (SDF)*, Oct. 2017, pp. 1–6.
[23] J. L. Williams, "Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA, and association-based member," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 1664–1687, Jul. 2015.
[24] Y. Xia, K. Granström, L. Svensson, and M. Fatemi, "Extended target Poisson multi-Bernoulli filter," 2018, *arXiv:1801.01353*. [Online]. Available: http://arxiv.org/abs/1801.01353
[25] Y. Xia, K. Granstrcöm, L. Svensson, and A. F. Garcia-Fernandez, "Performance evaluation of multi-Bernoulli conjugate priors for multi-target filtering," in *Proc. 20th Int. Conf. Inf. Fusion (Fusion)*, Xi'an, China, Jul. 2017.
[26] 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.
[27] Y. Xia, K. Granström, L. Svensson, and A. F. Garcia-Fernandez, "An implementation of the Poisson multi-Bernoulli mixture trajectory filter via dual decomposition," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Cambridge, U.K., Jul. 2018, pp. 1–8.
[28] Y. Xia, K. Granström, L. Svensson, A. F. G. Fernández, and J. L. Williams, "Extended target Poisson multi-Bernoulli mixture trackers based on sets of trajectories," in *Proc. Int. Conf. Inf. Fusion*, Ottawa, ON, Canada, Jul. 2019, pp. 1–8.
[29] K. Granström, S. Reuter, M. Fatemi, and L. Svensson, "Pedestrian tracking using velodyne data—Stochastic optimization for extended object tracking," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Redondo Beach, CA, USA, Jun. 2017, pp. 39–46.
[30] L. Cament, M. Adams, J. Correa, and C. Perez, "The δ-generalized multi-Bernoulli Poisson filter in a multi-sensor application," in *Proc. Int. Conf. Control, Autom. Inf. Sci. (ICCAIS)*, Oct. 2017, pp. 32–37.
[31] L. Cament, M. Adams, and J. Correa, "A multi-sensor, gibbs sampled, implementation of the multi-Bernoulli Poisson filter," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Cambridge, U.K., Jul. 2018, pp. 2580–2587.
[32] K. Granström, L. Svensson, S. Reuter, Y. Xia, and M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods," *IEEE Trans. Intell. Vehicles*, vol. 3, no. 1, pp. 30–45, Mar. 2018.
[33] S. Scheidegger, J. Benjaminsson, E. Rosenberg, A. Krishnan, and K. Granström, "Mono-camera 3D multi-object tracking using deep learning detections and PMBM filtering," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Changshu, Suzhou, China, Jun. 2018, pp. 433–440.
[34] M. Fatemi, K. Granström, L. Svensson, F. J. R. Ruiz, and L. Hammarstrand, "Poisson multi-Bernoulli mapping using Gibbs sampling," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2814–2827, Jun. 2017.
[35] M. Fröhle, C. Lindberg, K. Granström, and H. Wymeersch, "Multisensor Poisson multi-Bernoulli filter for joint Target–Sensor state tracking," *IEEE Trans. Intell. Vehicles*, vol. 4, no. 4, pp. 609–621, Dec. 2019, doi: 10.1109/TIV.2019.2938093.
[36] B. Wang, W. Yi, R. Hoseinnezhad, S. Li, L. Kong, and X. Yang, "Distributed fusion with multi-Bernoulli filter based on generalized covariance intersection," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 242–255, Jan. 2017.
[37] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, NJ, USA: Wiley, 2006.
[38] J. K. Uhlmann, "General data fusion for estimates with unknown cross covariances," in *Proc. 5th Signal Process., Sensor Fusion, Target Recognit.*, vol. 2755, Jun. 1996, pp. 536–548.
[39] R. P. S. Mahler, "Optimal/robust distributed data fusion: A unified approach," in *Proc. 15th Signal Process., Sensor Fusion, Target Recognit.*, vol. 4052, Aug. 2000, pp. 128–139.
[40] D. Clark, S. Julier, R. Mahler, and B. Ristic, "Robust multi-object sensor fusion with unknown correlations," in *Proc. Sensor Signal Process. Defence (SSPD)*. Edison, NJ, USA: IET, 2010, p. 1–14.
[41] M. B. Guldogan, "Consensus Bernoulli filter for distributed detection and tracking using multi-static Doppler shifts," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 672–676, Jun. 2014.
[42] M. Üney, S. Julier, D. Clark, and B. Ristic, "Monte Carlo realisation of a distributed multi-object fusion algorithm," in *Proc. Sensor Signal Processing for Defence (SSPD)*. Edison, NJ, USA: IET, 2010, pp. 1–13.

[43] M. Üney, D. E. Clark, and S. J. Julier, "Distributed fusion of PHD filters via exponential mixture densities," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 521–531, Jun. 2013.

[44] 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)*, Jul. 2017, pp. 1–8.

[45] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 508–520, Jun. 2013.

[46] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and B.-N. Vo, "Average Kullback-Leibler divergence for random finite sets," in *Proc. 18th Int. Conf. Inf. Fusion (Fusion)*, 2015, pp. 1359–1366.

[47] C. Fantacci, B.-N. Vo, B.-T. Vo, G. Battistelli, and L. Chisci, "Consensus labeled random finite set filtering for distributed multi-object tracking," 2015, *arXiv:1501.01579*. [Online]. Available: http://arxiv.org/abs/1501.01579

[48] S. Li, W. Yi, R. Hoseinnezhad, G. Battistelli, B. Wang, and L. Kong, "Robust distributed fusion with labeled random finite sets," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 278–293, Jan. 2018.

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

[50] J. L. Williams, "An efficient, variational approximation of the best fitting multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 258–273, Jan. 2015.

[51] P. Bromiley, "Products and convolutions of Gaussian probability density functions," *Tina-Vision Memo*, vol. 3, no. 4, p. 1, 2003.

[52] Y. Bar-Shalom, *Multitarget-Multisensor Tracking: Advanced Applications*. Norwood, MA, USA: Artech House, 1990.

[53] J. Munkres, "Algorithms for the assignment and transportation problems," *J. Soc. Ind. Appl. Math.*, vol. 5, no. 1, pp. 32–38, Mar. 1957.

[54] F. Bourgeois and J.-C. Lassalle, "An extension of the Munkres algorithm for the assignment problem to rectangular matrices," *Commun. ACM*, vol. 14, no. 12, pp. 802–804, Dec. 1971.

[55] K. Granström and U. Orguner, "Estimation and maintenance of measurement rates for multiple extended target tracking," in *Proc. 15th Int. Conf. Inf. Fusion (FUSION)*, 2012, pp. 2170–2176.

[56] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. Kdd*, 1996, vol. 96, no. 34, pp. 226–231.

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

[58] Á. F. García-Fernández, A. S. Rahmathullah, and L. Svensson, "A metric on the space of finite sets of trajectories for evaluation of multi-target tracking algorithms," 2016, *arXiv:1605.01177*. [Online]. Available: http://arxiv.org/abs/1605.01177

[59] A. S. Rahmathullah, A. F. Garcia-Fernandez, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Proc. 20th Int. Conf. Inf. Fusion (Fusion)*, Jul. 2017, pp. 1–8.

[60] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

[61] A. Garcia-Fernandez and L. Svensson, "Spooky effect in optimal OSPA estimation and how GOSPA solves it," in *Proc. Int. Conf. Inf. Fusion (FUSION)*, Ottawa, ON, Canada, Jul. 2019, pp. 1–8.

[62] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[63] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006. [Online]. Available: https://www.springer.com/gp/book/9780387310732

MARKUS FRÖHLE (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in telematics from the Graz University of Technology, Graz, Austria, in 2009 and 2012, respectively, and the Ph.D. degree in signals and systems from the Chalmers University of Technology, Gothenburg, Sweden, in 2018. From 2012 to 2013, he was with the Signal Processing and Speech Communication Laboratory, Graz University of Technology. From 2013 to 2018, he was with the Department of Electrical Engineering, Chalmers University of Technology. He joined Zenuity AB, in 2019. His current research interests include localization and tracking.

KARL GRANSTRÖM (Member, IEEE) received the M.Sc. degree in applied physics and electrical engineering and the Ph.D. degree in automatic control from Linköping University, Sweden, in May 2008 and November 2012, respectively. He is currently a Perception Specialist at Embark Trucks Inc., San Francisco, CA, USA. Before joining Embark, he worked at the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden, first as a Postdoctoral Researcher and then as a Researcher. He held postdoctoral positions at the Department of Electrical and Computer Engineering, University of Connecticut, USA, from September 2014 to August 2015, and at the Department of Electrical Engineering, Linköping University, from December 2012 to August 2014. His research interests include estimation theory, multiple model estimation, sensor fusion, and target tracking, especially for extended targets. He received paper awards at the Fusion 2011 and Fusion 2012 conferences. He has organized several workshops and tutorials on the topic of Multiple Extended Target Tracking and Sensor Fusion. At Fusion 2018, the International Society for Information Fusion (ISIF) awarded him the ISIF Young Investigator Award for his contributions to extended target tracking research and his services to the research community.

HENK WYMEERSCH (Senior Member, IEEE) received the Ph.D. degree in electrical engineering/applied sciences from Ghent University, Belgium, in 2005. He is currently a Professor of communication systems at the Department of Electrical Engineering, Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a Postdoctoral Researcher with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, from 2005 to 2009. His current research interests include cooperative systems and intelligent transportation. He served as an Associate Editor for the IEEE Communication Letters, from 2009 to 2013, the IEEE Transactions on Wireless Communications, since 2013, and the IEEE Transactions on Communications, from 2016 to 2018.