



Multiple Target Tracking With Uncertain Sensor State Applied To Autonomous Vehicle Data

Downloaded from: <https://research.chalmers.se>, 2025-06-18 04:19 UTC

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

Fröhle, M., Granström, K., Wymeersch, H. (2018). Multiple Target Tracking With Uncertain Sensor State Applied To Autonomous Vehicle Data. 2018 IEEE Statistical Signal Processing Workshop (SSP): 628-632. <http://dx.doi.org/10.1109/SSP.2018.8450842>

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

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

MULTIPLE TARGET TRACKING WITH UNCERTAIN SENSOR STATE APPLIED TO AUTONOMOUS VEHICLE DATA

Markus Fröhle, Karl Granström, Henk Wymeersch

Department of Electrical Engineering
Chalmers University of Technology, Gothenburg, Sweden
E-mail: {frohle, karl.granstrom, henkw}@chalmers.se

ABSTRACT

In a conventional multitarget tracking (MTT) scenario, the sensor position is assumed known. When the MTT sensor, e.g., an automotive radar, is mounted to a moving vehicle with uncertain state, it becomes necessary to relax this assumption and model the unknown sensor position explicitly. In this paper, we compare a recently proposed filter that models the unknown sensor state [1], to two versions of the track-oriented marginal MeMBer/Poisson (TOMB/P) filter: the first does not model the sensor state uncertainty; the second models it approximately by artificially increasing the measurement variance. The results, using real measurement data, show that in terms of tracking performance, the proposed filter can outperform TOMB/P without sensor state uncertainty, and is comparable to TOMB/P with increased variance.

Index Terms— multi-target tracking, localization, Kalman filter

1. INTRODUCTION

Automated and autonomous vehicles need to be provided with accurate position information, not only of the vehicle itself, but also of the environment they operate in. Location of static features, e.g., road segments and street lights, can be provided by a pre-recorded local dynamic map (LDM). The presence and state of dynamic features such as a pedestrian or a cyclist, need to be estimated using the vehicles' on-board sensors. Typically, the environment can be observed by an on-board radar or LIDAR, and the vehicle location can be determined by a global navigation satellite system (GNSS) receiver. Tracking the state of multiple dynamic features in the environment can be addressed by an MTT filter. Generally, MTT filters consider the sensor state to be known [2]. However, when the tracking sensor is mounted on a moving platform with uncertain state, the uncertain ego-position needs to be considered in the MTT filter, in order not to



Fig. 1: Measurement scenario, with a vehicle with uncertain sensor state tracking a moving pedestrian.

deteriorate tracking performance. An MTT filter which explicitly models the uncertain sensor state was proposed in [1], extending the TOMB/P filter from [3].

In this paper, we present results on the evaluation of [1] with real measurements collected by a Volvo XC90 with on-board Applanix POSLV GNSS receiver and stereo vision camera system from Autoliv, as shown in Fig. 1.

2. RELATION TO PRIOR WORK

Many MTT filters have been proposed to track dynamic features using radar-like sensors, when the sensor state is known [2]. For example, the multi-hypothesis tracker (MHT), whose complexity grows exponentially due to the unknown measurement-to-feature data association (DA), and the joint probability data association filter (JPDAF), which approximates the posterior using marginal association probabilities. More recent developments involve MTT filters based on random finite set (RFS) and finite-set statistics (FISST). See

This work is supported, in part, by the EU-H2020 project HIGHTS (High Precision Positioning for Cooperative ITS Applications) under grant no. MG-3.5a-2014-636537 and COPPLAR (campus shuttle cooperative perception and planning platform) project, funded under grant no. 2015-04849 from Vinnova.

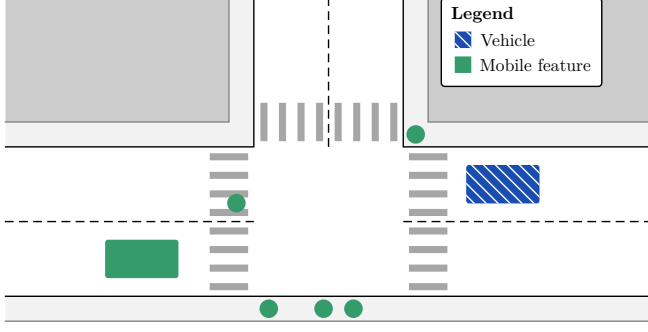


Fig. 2: A single vehicle equipped with GNSS and V2F sensor is driving in an urban environment.

[4] for more details about RFSs and FISST. For instance, the probability hypothesis density (PHD) filter propagates the first moment of the RFS density forward in time, and the TOMB/P filter approximates the posterior density by a Poisson multi-Bernoulli (PMB) multi-object density [3]. Mapping the environment when the sensor state is unknown, belongs to the topic of simultaneous localization and mapping (SLAM) [5]. It was also stated in terms of RFSs, e.g., in [6], where a particle filter to estimate the sensor trajectory is combined with a PHD filter to estimate the map. Differently in [7], the uncertain sensor state was considered to track at most one feature. Joint estimation of sensor and features state was addressed in [8], where an unknown sensor bias is estimated through MTT. A multi-sensor PHD filter for joint state estimation was proposed in [9, 10]. For the case, that the feature presence is known, the joint estimation problem can be solved via message passing on a factor graph as done in [11] centrally; and distributed in [12], where measurement-to-feature DA was additionally known. Methods to overcome those restrictions and relations between message passing and RFS based MTT filtering are discussed in [13].

3. PROBLEM STATEMENT AND SYSTEM MODELS

3.1. Problem statement

Consider the scenario illustrated in Fig. 2. The vehicle is equipped with a GNSS and vehicle-to-feature (V2F) measurement sensor. On the vehicle, a MTT filter is run, which should track the mobile features as accurately as possible considering the vehicles' own location uncertainty. We are therefore interested in the joint posterior distribution of the feature and vehicle states at every time step t .

3.2. Vehicle and feature dynamics

The time-varying vehicle state $\mathbf{x}_t \in \mathbb{R}^{N_x}$ is statistically modeled by the state transition probability density function (PDF) $p(\mathbf{x}_t|\mathbf{x}_{t-1})$, with linear state-space model

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{w}_t, \quad (1)$$

where \mathbf{A}_t denotes the state-transition matrix and $\mathbf{w}_t \sim \mathcal{N}(0, \mathbf{W}_t)$ with error covariance matrix \mathbf{W}_t . Since there is potentially more than one feature present, we model the features by an RFS \mathbf{F} with multi-object density $f(\mathbf{F})$. A single time-varying feature, e.g., a pedestrian, with state $\mathbf{f} \in \mathbb{R}^{N_f}$ survives in the next time step with survival probability p_S , and, if it survives, its state motion is statistically modeled by the state transition PDF $p(\mathbf{f}_t|\mathbf{f}_{t-1})$, with linear state-space model

$$\mathbf{f}_t = \mathbf{B}_t \mathbf{f}_{t-1} + \mathbf{v}_t, \quad (2)$$

where \mathbf{B}_t denotes the state-transition model and process noise $\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{V}_t)$ with error covariance matrix \mathbf{V}_t .

3.3. Measurement models

A vehicle collects measurements from two types of sensors. A GNSS sensor obtains measurements related to the absolute vehicle state \mathbf{x}_t , described by the likelihood function $p(\mathbf{z}_{G,t}|\mathbf{x}_t)$ with linear observation model

$$\mathbf{z}_{G,t} = \mathbf{H}_{G,t} \mathbf{x}_t + \mathbf{r}_t, \quad (3)$$

where $\mathbf{H}_{G,t}$ is the linear observation matrix and $\mathbf{r}_t \sim \mathcal{N}(0, \mathbf{R}_t)$ with error covariance matrix \mathbf{R}_t . The second type of sensor is capable of observing features, e.g., a camera, radar or LIDAR. This V2F sensor is susceptible to measurement noise, missed and false detections. A feature is detected (and thus produces a measurement \mathbf{z}_t) with detection probability $p_D(\mathbf{x}_t, \mathbf{f}_t)$ depending on the sensor field-of-view (FoV) and hence on the states \mathbf{x}_t and \mathbf{f}_t . The V2F measurement likelihood function $\ell(\mathbf{z}_t|\mathbf{x}_t, \mathbf{f}_t)$ is described by the linear observation model

$$\mathbf{z}_t = \mathbf{H}_{1,t} \mathbf{x}_t + \mathbf{H}_{2,t} \mathbf{f}_t + \mathbf{q}_t, \quad (4)$$

where $\mathbf{H}_{1,t}$ and $\mathbf{H}_{2,t}$ denote the observation matrices, and $\mathbf{q}_t \sim \mathcal{N}(0, \mathbf{Q}_t)$ with error covariance matrix \mathbf{Q}_t . In one scan, the V2F sensor obtains multiple V2F measurements from possibly multiple features present in the scene, as well as false alarm measurements due to clutter. The clutter is modeled by a Poisson point process (PPP), and all accumulated measurements, both feature detections and clutter, are modeled by the set \mathbf{Z}_t^F . We assume the standard point target measurement assumptions: each feature can generate at most one measurement per scan, and each measurement is the result of at most one feature [14]. Note that it is not known which feature is responsible for which measurement in \mathbf{Z}_t^F .

3.4. Reference methods

We consider an unknown sensor state with mean $\mu_{\mathbf{x}_t}$ and covariance $\Sigma_{\mathbf{x}_t}$, obtained from, e.g., a Kalman filter. MTT methods such as the TOMB/P filter ignore such uncertainty, though for a linear model, we can replace \mathbf{x}_t in (4) by its mean $\mu_{\mathbf{x}_t}$, and \mathbf{Q}_t is increased by $\mathbf{H}_1 \Sigma_{\mathbf{x}_t} \mathbf{H}_1^T$, which we denote variance inflation.

4. PROPOSED METHOD

The proposed MTT filter was derived in [1], and the reader is referred there for additional details. It belongs to the class of Bayesian filters calculated through altering prediction and measurement update steps, operating on the joint vehicle-feature density. Here, we summarize only the key ideas behind this filter. At time step t , the prior joint vehicle-feature density is given by

$$f(\mathbf{x}_{t-1}, \mathbf{F}_{t-1}) = p(\mathbf{x}_{t-1})f(\mathbf{F}_{t-1}), \quad (5)$$

where $p(\mathbf{x}_{t-1})$ denotes the vehicle state PDF described by a Gaussian distribution with mean vector and covariance matrix. The prior feature RFS density $f(\mathbf{F}_{t-1})$ is PMB consisting of detected and undetected features, i.e., features which are hypothesized to exist but have never been detected (c.f. [1] or [14, Def. I] for a general description on undetected features). Under the assumption that the vehicle and features move independently¹, we can predict the joint vehicle-feature density to the current time step t , denoted $f(\mathbf{x}_{t|t-1}, \mathbf{F}_{t|t-1})$. This is done via the Chapman-Kolmogorov equation applied separately on the vehicle state and on the feature RFS [4]. The prediction preserves the factorized form of the density (5). Updating the predicted joint vehicle-feature density by GNSS measurement $\mathbf{z}_{G,t}$ results in

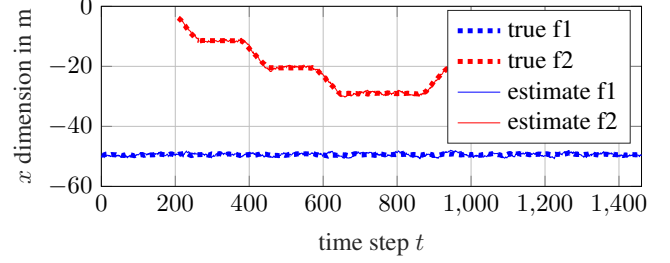
$$f(\mathbf{x}_t, \mathbf{F}_{t|t-1} | \mathbf{z}_{G,t}) = p(\mathbf{x}_t | \mathbf{z}_{G,t})f(\mathbf{F}_{t|t-1}), \quad (6)$$

where $p(\mathbf{x}_t | \mathbf{z}_{G,t})$ is computed through a standard Kalman update step influencing only the vehicle state [1, Sec. IV-B], [15]. Updating (6) (or the predicted joint vehicle-feature density in the absence of a GNSS measurement) by a V2F measurement set \mathbf{Z}_t^F is more involved, and includes approximations to ensure that the updated joint density is in the same form as the prior joint density (5). By applying suitable approximations, the complexity of the MTT filter will not increase over time. The details of the involved approximations are given in [1]. The updated joint density is then of form

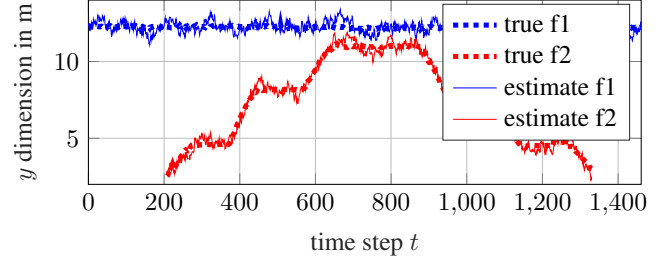
$$f(\mathbf{x}_t, \mathbf{F}_t | \mathbf{Z}_t^F) \propto \sum_{\mathbf{F}^u \uplus \mathbf{F}^d = \mathbf{F}_t} \hat{f}^u(\mathbf{F}^u) \times \sum_{A \in \mathcal{A}} w^A p^A(\mathbf{x}_t | \mathbf{Z}_t^F) \tilde{f}^{d,A}(\mathbf{F}^d | \mathbf{Z}_t^F), \quad (7)$$

where the operator \uplus denotes the disjoint set union, A is a DA hypothesis with weight $w^A : \sum_{A \in \mathcal{A}} w^A = 1$, belonging to the set of possible hypotheses \mathcal{A} , $\hat{f}^u(\mathbf{F}^u)$ is the undetected feature density, which depends only on RFS \mathbf{F}^u . For a given association A , $\tilde{f}^{d,A}(\mathbf{F}^d | \mathbf{Z}_t^F)$ is the approximate detected feature density, and $p^A(\mathbf{x} | \mathbf{Z}_t^F)$ is the vehicle state PDF. It can be seen that (7) is a Poisson multi-Bernoulli mixture (PMBM) density, which can be reduced to a PMB density of the form (5) using, e.g., the variational approximation

¹This assumption is valid either when the vehicle and features are spatially well separated and hence do not influence each other, or when the sampling time is short enough, such that any feature interactions are negligible.



(a) x dimension



(b) y dimension

Fig. 3: The estimated feature tracks using the proposed MTT filter are compared with the true filter tracks. Feature 1 is denoted f1, and similarly for feature 2.

presented in [16], or based on marginal DA probabilities [3] known as the TOMB/P step. Due to space constraints, we refer the reader to [1] for a more detailed explanation of (7).

5. NUMERICAL EXAMPLE

We investigate feature tracking performance of the proposed MTT filter and the TOMB/P filter (c.f. [3]) using real measurement data.

5.1. Setup

Measurement data was recorded using the COPPLAR project test vehicle, a Volvo XC90 equipped with different kinds of automotive sensors. The filters were then run offline. The GNSS sensor is a high-precision Applanix POSLV, and the onboard V2F sensor is a stereo vision camera from Autoliv. Measurements were considered as ground-truth and measurement noise was artificially added to the GNSS measurements with variance $\sigma_G^2 = 0.9216 \text{ m}^2$ corresponding to a differential GNSS (D-GNSS) receiver² and with variance $\sigma_{V2F}^2 = 0.42 \text{ m}^2$ for the V2F measurements, corresponding to the worst-case performance of the used V2F sensor. For the test recording, the vehicle was standing still and a pedestrian was walking in front of the vehicle resting at approximately 10, 20, and 30 m distance. The measurement scenario is outlined

²Note that this is the x/y position accuracy of a likewise GNSS receiver RT3000 from OXTS using D-GNSS according to [17] and [18].

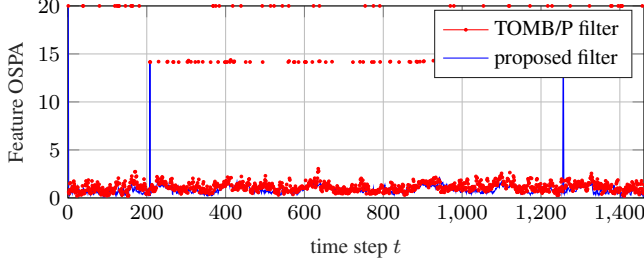


Fig. 4: Feature OSPA over time.

in Fig. 1. Additionally, a static feature was placed at around 50 m distance of the vehicle (not visible in Fig. 1). The total simulation time was $t = 1461$ time steps.

In the filters, the state of the vehicle (sensor) at time step t is described by $\mathbf{x} = [\mathbf{p}^\top, \mathbf{v}^\top]^\top$, where $\mathbf{p} \in \mathbb{R}^2$ denotes the position and $\mathbf{v} \in \mathbb{R}^2$ the velocity. Similarly, we describe the state of a feature \mathbf{f} by its two dimensional position and velocity. Vehicle and feature dynamics follow the linear constant velocity (CV) model described by (1) and (2) with

$$\mathbf{A}_t = \mathbf{B}_t = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} \otimes \mathbf{I}_2, \quad (8)$$

where the sampling time $T_s = 0.1$ s, the operator \otimes denotes the Kronecker product, and \mathbf{I}_2 is an identity matrix of dimension two. In (1), the error covariance matrix $\mathbf{W}_t = \mathbf{S}$ with

$$\mathbf{S} = s \begin{bmatrix} T_s^3/3 & T_s^2/2 \\ T_s^2/2 & T_s \end{bmatrix} \otimes \mathbf{I}_2 \quad (9)$$

with $s = 0.1 \text{ m}^2$, and in (2) the error covariance matrix $\mathbf{V}_t = \mathbf{S}$ with $s = 0.5 \text{ m}^2$. The GNSS measurement matrix in (3) is $\mathbf{H}_{G,t} = [1 \ 0] \otimes \mathbf{I}_2$, and the error covariance matrix is $\mathbf{R}_t = \sigma_G^2 \mathbf{I}_2$. The V2F measurement matrices in (4) are $\mathbf{H}_{1,t} \triangleq \mathbf{H}_{G,t}$ and $\mathbf{H}_{2,t} \triangleq -\mathbf{H}_{G,t}$, and the error covariance matrix is $\mathbf{Q}_t = \sigma_{V2F}^2 \mathbf{I}_2$. We set the initial undetected feature intensity to $D_-^u(\mathbf{f}) = 10\mathcal{N}(\mathbf{0}, \mathbf{P})$, where $\mathbf{P} = \text{diag}([100^2, 100^2, 1, 1]^\top)$, the feature birth intensity to $D^b(\mathbf{f}) = 0.05\mathcal{N}(\mathbf{0}, \mathbf{P})$, the average number of false alarms per scan to $\lambda_c = 10$, with uniform distribution on $(-200 \text{ m}, 200 \text{ m})^2$, the probability of survival to $p_s = 0.7$, and the probability of detection to $p_D(\mathbf{x}, \mathbf{f}) \triangleq p_D = 0.9$. Feature tracking performance is assessed in terms of optimal sub-pattern assignment (OSPA) metric with cut-off parameter $c = 20$ m and order $p = 2$ [19]. MATLAB is used as a simulation tool.

5.2. Results

In Fig. 3a and Fig. 3b, the true and estimated position of the two features are plotted over time for x and y dimensions, respectively. We observe that feature 1 (static) is present throughout the whole experiment. Feature 2 (pedestrian) enters the scene at time step $t = 208$ and leaves the scene at

time step $t = 1330$. Furthermore, the x position of feature 2 remains unchanged (pedestrian standing still) for some time at approximately $x \approx \{10 \text{ m}, 20 \text{ m}, 30 \text{ m}\}$. The proposed MTT filter correctly detects the presence of feature 1 and 2, as well as the disappearance of feature 2. Furthermore, the feature state estimates fluctuate around the true feature positions. These fluctuations come from the noisy measurements (GNSS and V2F), as well as from the noise in the motion models (feature and vehicle).

In Fig. 4, the feature OSPA is plotted over time for the proposed MTT filter as well as the TOMB/P filter, which does not model the present sensor state uncertainty. This filter uses the GNSS measurement as the true sensor location (c.f. (3)). We observe that the proposed MTT filter achieves a lower feature OSPA. There is a cardinality mismatch resulting in an increase of OSPA when feature 2 (pedestrian) enters and leaves the scene. In contrast, with the TOMB/P filter, the feature OSPA jumps between high and low values indicating the filter has problems detecting correctly the presence of the features. The feature OSPA of the proposed MTT filter averaged over the whole runtime is 0.87, and 2.61 for the TOMB/P filter.

Since the observation and motion models are all linear, we are able to incorporate any sensor state uncertainty in the TOMB/P filter by inflating the V2F measurement noise variance \mathbf{Q}_t by the GNSS measurement noise variance \mathbf{R}_t . In this way, a V2F measurement acquired by a sensor with uncertain state is interpreted as a V2F measurement with increased V2F noise variance and deterministic sensor position at the GNSS measurement location. For this case, the average feature OSPA reduces to 0.99. We can conclude that in a single vehicle (sensor) scenario incorporation of sensor state uncertainty, either in the way of the proposed MTT filter or by inflating the V2F measurement variance in the TOMB/P filter, enables to accurately track the features over time. The proposed MTT filter allows to explicitly distinguish between different sources of uncertainty and enables to track (and hence reduce) the uncertain sensor state. Note that further performance results including real measurement data from two moving vehicles are shown in [1].

6. CONCLUSIONS

A MTT filter that incorporates sensor state uncertainty was evaluated for a single sensor scenario. Two different kind of measurements, absolute GNSS measurements and relative V2F measurements were incorporated. The proposed MTT filter has low computational complexity due to its parametric implementation. Experiments with real measurement data showed the benefit of modeling sensor state uncertainty, quantified in terms of the OSPA distance. The proposed MTT filter can naturally incorporate different sources of state uncertainty, whereas in the TOMB/P filter this needs to be incorporated by inflating the measurement noise variance, both leading to comparable performance.

7. REFERENCES

- [1] Markus Fröhle, Christopher Lindberg, Karl Granström, and Henk Wymeersch, “Multisensor Poisson Multi-Bernoulli Filtering with Uncertain Sensor States,” *ArXiv e-prints*, Dec. 2017.
- [2] Yaakov Bar-Shalom, Peter K Willett, and Xin Tian, *Tracking and data fusion: A handbook of algorithms*, Storrs, CT: YBS Publishing, 2011.
- [3] Jason 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.
- [4] Ronald PS Mahler, *Statistical multisource-multitarget information fusion*, Artech House, Inc., 2007.
- [5] Hugh Durrant-Whyte and Tim Bailey, “Simultaneous localization and mapping: part I,” *IEEE robotics & automation magazine*, vol. 13, no. 2, pp. 99–110, 2006.
- [6] John Mullane, Ba-Ngu Vo, Martin D Adams, and Ba-Tuong Vo, “A random-finite-set approach to Bayesian SLAM,” *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 268–282, 2011.
- [7] Simon J Julier and Amadou Gning, “Bernoulli filtering on a moving platform,” in *Information Fusion (Fusion), 2015 18th International Conference on*. IEEE, 2015, pp. 1511–1518.
- [8] Branko Ristic, Daniel E Clark, and Neil Gordon, “Calibration of multi-target tracking algorithms using non-cooperative targets,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 390–398, 2013.
- [9] Murat Üney, Bernard Mulgrew, and Daniel E Clark, “A Cooperative Approach to Sensor Localisation in Distributed Fusion Networks,” *IEEE Trans. Signal Processing*, vol. 64, no. 5, pp. 1187–1199, 2016.
- [10] Murat Üney, Bernard Mulgrew, and Daniel E Clark, “Distributed localisation of sensors with partially overlapping field-of-views in fusion networks,” in *2016 19th International Conference on Information Fusion (FUSION)*, July 2016, pp. 1340–1347.
- [11] Markus Fröhle, Christopher Lindberg, and Henk Wymeersch, “Cooperative Localization of Vehicles without Inter-vehicle Measurements,” in *IEEE Wireless Communications and Networking Conference*, April 2018.
- [12] Florian Meyer, Ondrej Hlinka, Henk Wymeersch, Erwin Riegler, and Franz Hlawatsch, “Distributed localization and tracking of mobile networks including noncooperative objects,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 1, pp. 57–71, Mar. 2016.
- [13] Florian Meyer, Thomas Kropfreiter, Jason L Williams, Roslyn Lau, Franz Hlawatsch, Paolo Braca, and Moe Z Win, “Message Passing Algorithms for Scalable Multi-target Tracking,” *Proceedings of the IEEE*, vol. 106, no. 2, pp. 221–259, 2018.
- [14] Jason Williams and Roslyn Lau, “Approximate Evaluation of Marginal Association Probabilities with Belief Propagation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2942–2959, Oct. 2014.
- [15] Dan Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*, John Wiley & Sons, 2006.
- [16] Jason L Williams, “An Efficient, Variational Approximation of the Best Fitting Multi-Bernoulli Filter,” *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 258–273, 2015.
- [17] HIGHTS, Deliverable D5.2, “Specifications of implemented cooperative and fusion algorithms,” 2016.
- [18] Oxford Technical Solutions Ltd, “Datasheet RT3000 v2 GNSS/INS for high dynamic vehicle testing,” Tech. Rep., Oxford Technical Solutions Ltd, United Kingdom, 2018.
- [19] Dominic Schuhmacher, Ba-Tuong Vo, and Ba-Ngu Vo, “A consistent metric for performance evaluation of multi-object filters,” *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, 2008.