Probabilistic Vehicle Tracking with Sparse Radar Detection Measurements

PHILIPP BERTHOLD
MARTIN MICHAELIS
THORSTEN LUETTEL
DANIEL MEISSNER
HANS-JOACHIM WUENSCHE

Most automotive perception systems leverage radar sensors for their long-range measuring capability and weather robustness at economic costs. A downside is the rather low spatial resolution. It complicates the estimation of pose and size of an extended object. High-resolution sensors facilitate techniques like shape recognition based on a single measurement. But even these sensors only provide sparse measurements at larger distances, which makes instantaneous object detection highly ambiguous. We propose an approach that incorporates the current state estimate to probabilistically identify the true origin of a detection and thereby decreases its association ambiguity. It uses all given measurement data, including the radial speed. This improves the information gain for mass-market sensors with a high measurement uncertainty. We first perform a parametrization of the object using a set of components. They describe the characteristics of a detection in dependency of the current state estimate and various physical relations. Their superposition resembles the spatial detection likelihood of the entire object. Subsequently, we perform a computationally efficient state update that exploits the probabilistic association of the detection to the components. All steps take about 20 μs of computing time. In this article, we demonstrate this technique in an application that tracks vehicles with radar detections. Besides providing details on the algorithm and a formal description of the components, we also illustrate the probabilistic association with examples. Finally, we discuss the performance in real-world tracking scenarios and outline interfaces to multi-hypotheses and multi-sensor fusion algorithms. This paper is accompanied by an exemplary MATLAB implementation and a demonstration video.

Manuscript received February 17, 2022; revised September 27, 2022; released for publication March 20, 2023.

P. Berthold, M. Michaelis, T. Luettel, and H.-J. Wuensche are with the Institute for Autonomous Systems Technology (TAS), University of the Bundeswehr Munich, 85577 Neubiberg, Germany (e-mail: philipp.berthold@unibw.de; martin.michaelis@unibw.de; thorsten.luettel@unibw.de; joe.wuensche@unibw.de).

D. Meissner is with Objects, Obstacles and AI based low level Perception, BMW Group, 80788 Munich, Germany (e-mail: daniel.da.meissner@bmw.de).

1557-6418/22/$17.00

I. INTRODUCTION

Advanced automotive perception systems have to meet high expectations in terms of cost-effectiveness, performance, and robustness. The fusion of different sensor types accommodates these requirements by exploiting the cumulative strengths. Monocular cameras are widely used to identify objects as they provide semantic information. However, they do not provide range measurements. This impairs the immediate estimation of the position and the extent of objects and often requires the incorporation of model knowledge. Stereo cameras are typically limited to short ranges [11]. On the other hand, LiDAR sensors mainly capture high-resolution spatial information on an object's contour, which encapsulates the pose and the extent of an object. Radar sensors also provide full spatial information, but with a lower resolution. However, their major advantages are their ability to directly measure radial speed and their resistance to tough weather conditions due to their lower frequency range.

The nature of *extended objects* states that multiple detections might be caused by arbitrary parts of the extent of the object. High-resolution sensors provide such a large quantity of detections that the contour of objects can be spotted in a single measurement [8]. The thereby captured object instances can be directly filtered to their corresponding tracks [28]. However, these approaches are not feasible if only sparse measurement data are available, resulting in few or no resolved detections per object. This issue is not necessarily limited to mass-market sensors; also high-performance sensors only provide sparse measurement data at respective distances. At this point, a contour (or structure) extraction from a single measurement is no longer possible. To sustain the tracking, the detections need to be directly filtered to their tracks. The arising challenge is the correct determination of the origin of each detection without any structural information from the current measurement data. Especially in the case of a radar sensor, the association problem is tough: The lateral measurement noise is substantial due to its measurement principle [25] and depends on the complexity of the surrounding.

We propose a filtering approach that tackles this association problem. First, it splits the object in components with individual, physically deduced detection characteristics. Second, it incorporates the current object state estimate to model the current statistical appearance of these components. We apply this approach to radar sensors for vehicle tracking in this work. The utilized radar sensors provide a set of points, which is called *scatter* data. Each point represents a so-called *detection*, which represents a maximum of local reflectivity and is given by measurements of position, radial speed, and amplitude. Our approach not only uses the position measurement but also exploits the radial speed measurement. This shifts the association problem to a space of higher dimension and improves its resolution.

A. Related Work

Adaptive cruise control has been one of the first popular automotive radar applications. The radar detects vehicles in direction of travel and returns their distance and speed. For this application, it is sufficient to obtain a single measurement for an object. Modern advanced driver assistance systems like the lane change assistant require pose and extent information. Filter techniques, which use radar measurements to estimate both pose and extent of objects, can be grouped into four categories. In 2017, Granström *et al.* [14] defined three categories: generic spatial models, set of points on a rigid body (SPRB) models, and physical models. In the past few years, however, machine learning approaches have also been adopted in this area and represent a popular fourth category.

Spatial models define the extent by a shape or function. Popular approaches incorporate ellipse-based models [24], probabilistic density fields given by Poisson processes [12], [13], radius functions [14], [38], or any kind of functional shape descriptions by Gaussian processes [27], [35]. As these approaches can be expressed in closed functions, they show high runtime performance. Additionally, they do not require explicit model knowledge and are therefore suitable for a wide spectrum of objects. However, the lack of model knowledge impairs the extent estimation accuracy and, consequently, also the estimation of the pose and the kinematics. The impact of the observation perspective and model-specific features like micro-Doppler cannot be exploited.

SPRB models use discretized spatial model descriptions instead of continuous ones. According to SPRB, the object can be modeled by a set of discrete scattering points. The location of these points can be estimated online [16], or by incorporating some model knowledge. In this manner, Bühren *et al.* [9] place the points on typical reflection sources of vehicles like wheels and corners. They also consider some visibility constraints. Hammarstrand *et al.* [17] propose an adequate SPRB filtering approach. A major downside of SPRB is the missing ability to model continuous, extended parts of the object.

Physical models are powerful and accurate in describing the object and predicting its expected measurements. They are often composed of an object model and a measurement model. The accuracy of the object model varies from geometric shapes to 3D computer models. The measurement model is an inference of physical considerations. Ray tracing methods [23] incorporate any desired level of model knowledge and achieve high reproduction accuracy. The prevalent downside is a substantial runtime overhead, which often renders them futile in real-time multi-object tracking applications. The poor runtime performance is not only due to the demanding modeling computation but also due to the tracking itself that often requires particle filters.

A both new and by now very popular approach to model radar detections is machine learning. It correlates the state of objects to their obtained detection characteristics in annotated training data. These approaches vary from variational Gaussian mixtures (VGM) [19], [21], [30], [37] to deep neural networks [10], [36]. The latter was facilitated by the recent progress in 4D high-resolution imaging radars that provide a large number of detections per target in a single measurement frame [41]. Machine learning approaches allow accurate measurement reproductions and circumvent expensive manual statistical studies on the sensor model. The sensor-specific measurement characteristics are learned from the measurement data. Their overall performance depends on the spectrum of the scenarios in the training data. If the training data does not contain more complex scenarios like different kinds of occlusion, then the outcome is undefined. Additionally, the network needs to learn new training data to adapt to new sensors or object types; it cannot be parametrized easily. However, thanks to recent advances in GPU development, their runtime performance allows real-time usage.

B. Previous Research

With the exception of some machine learning approaches, most of these models do not really match our observed data. This seems to be mainly due to oversimplification or incomplete modeling of the objects, i.e., vehicles. As a result, our aim is the development of a model that is physically derived to ensure generalizability. Its abstraction is chosen at a level that allows for its real-time usage in tracking applications, but without sacrificing performance potential.

Our work has began with a radar measurement analysis. In [2], we performed measurement campaigns to record the reflection characteristics of vehicles. These campaigns cover a spectrum of relative poses between the radar sensor and the target vehicle. An algorithm, which sorts, accumulates, and statistically re-weighs the measurement data, extracts a detection probability map in target coordinates for any desired relative observation pose. These results reveal a high impact of the observation angle. Unsurprisingly, the outer parts of the vehicle, which are oriented perpendicular toward the radar sensor, cause the most significant portion of the object's reflectivity. Moreover, the corners are highly reflective as a part of the round curve is always perfectly orthogonal toward the sensor. Next to the vehicle sides and corners, the wheels are also significant reflection sources. Wheels that are facing toward the radar reflect well due to the wheel rim. But the measurement analysis reveals that the opposite wheels are also often spotted in the radar measurement data. The low mounting height of series radars often causes a line of sight between the sensor and the opposite wheels. Opposite-wheel visibility is also given by underbody reflections, i.e., depending on the elevation angle of incidence, the beam is reflected by the

ground surface and the vehicle underbody. This effect causes a slight spatial detection probability for the complete underbody extent of the vehicle. On the contrary, inner parts of the vehicle are rarely visible. Varnish and windows heavily attenuate the beam amplitude. Roof structures are visible if the vertical field of view of the sensor is sufficiently large.

In [3], we complemented this work with a radar measurement model. Its primary aim is the preferably accurate probabilistic prediction of measurement data for any given target state. The measurement model is a physical one. It achieves a generic measurement reproduction and an inherent incorporation of effects like mutual occlusion. Typically, the latter is hard to accomplish when dealing with statistical or oversimplified models. Our model separates the generation process of detection measurements in abstraction layers like physical wave distribution, signal reception, and peak detection. We utilize this measurement model to evaluate models against real-world data, but its runtime performance hinders an immediate usage in multi-hypotheses tracking applications.

In [4], we enhanced both the measurement analysis and the measurement model with the Doppler-derived radial speed measurement. The radial speed measurement provides valuable information as it is directly measured and subject to only low measurement noise. Besides, it plays a crucial role in determining the angle of the detections. The radial speed measurement can be predicted for any point of the rigid object body as long as the relative kinematics of the object and the radar sensor are known. Parts like legs or wheels that move relative to the rigid body span a range of potential Doppler measurements. The radial speed measurements of moving parts of a moving object are known as micro-Doppler measurements and are subject to ongoing research [18], [32]. Current approaches [20] explicitly detect micro-Doppler measurements of vehicles in imaging radar data, extract the wheels by exploiting the Doppler spectrum [39], and use their position to track their pose.

C. Our Contribution

While our previous work has primarily elaborated a preferably precise and well-founded but computationally expensive physical model, this article presents its abstraction that can be utilized in real-time tracking applications. As far as possible, its functional structure is derived from physical and technical interrelations. According to our findings, the division of an object into different components yields a good modularizability and allows for individual measurement characteristics. The proposed abstracted model uses this mechanism and therefore resembles primarily the SPRB approaches. The main differences are that our model also supports and utilizes spatially extended, continuous components. Besides, we exploit the kinematic measurement of a

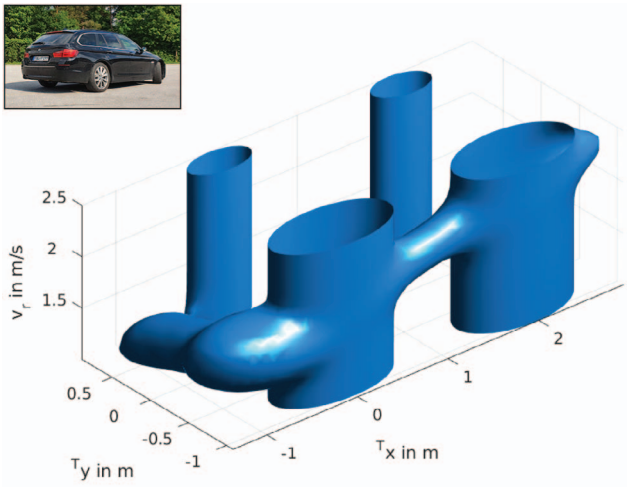

Fig. 1. The expected spatial detection likelihood of a moving vehicle in a left turn. The measurement space consists of the two-dimensional Euclidean position in target coordinates $^{T}(x, y)$ and the Doppler-deduced radial speed measurement v_r shown in the z-axis. To illustrate the three-dimensional detection likelihood, this plot shows the isosurface of an examplary detection likelihood. The vertically extended tubes are caused by the micro-Doppler effect of the wheels. The azimuth view angle corresponds to the observation angle of the radar sensor. Due to the left turn, the front of the vehicle moves away faster than the rear.

detection, e.g., the radial speed, to improve the origin search and to perform a direct kinematic state update. Figure 1 outlines the modeled joint measurement space and illustrates the expected multidimensional detection likelihood of a moving vehicle. The state of the components is linked to the state vector using type-dependent definitions. Concerning the components themselves, we consider not only angular visibility regions (similar as proposed by [9]) but also the reflectivity, the kinematic Doppler properties, and physical effects like scattering. This addition of features, though, requires more modeling effort. The proposed spatial measurement function described in this article can be used independently from the proposed tracking approach in any Bayes-based filters. It also outputs the expected number of detections for any given object state.

Learning-based approaches have the essential advantage that learning the reflectivity of an object does not require expert knowledge. However, learning methods usually learn the complete stack of measurement generation and cannot split different components, e.g., sensor model from object model. The proposed approach aims for parameterability and exchangeability of all relevant modules. New object types can be supported by partial adjustments of the object model on the basis of a datasheet, for example. There is no need for gathering and annotating new training data and relearning, especially if only partial properties have to be adjusted. Physical effects like scattering are mathematically described and therefore generically utilizable, and it is possible to apply assumptions like symmetry for a subset of the object model. Another drawback is the computing effort that comes, i.e., with the high dimen-

sionality of the VGM. Scheel *et al.* [30] still manage, though, to incorporate the radial speed profile of an object. The origin association and object tracking is commonly performed using probabilistic multi-hypothesis tracking (PMHT) [33], particle filters or labeled multi-Bernoulli filter (LMB) [30]. In case of PMHT applications, the computation of the association probabilities [7, eq. (19)] is similar to our approach. The subsequent tracking, however, optimizes the association cost using the expectation-maximization algorithm. We use a probabilistic Kalman-based approach [5]: The usage of explicit mathematical functions allows for a very fast execution time of our approach. The complete state update takes about 20 µs, rendering it suitable for multi-hypotheses tracking applications. Besides, our approach inherently supports and utilizes extended structures of the object model like vehicle sides. This avoids their approximation with a large number of Gaussian mixture components and reduces both computing effort and bias effects.

Another key functionality of the proposed filter is its treatment of sparse measurement data. In contrast to applications requiring imaging or high-resolution data, where, e.g., wheels can be spotted in a single measurement, this filter is suitable for radars receiving about two detections per target and measurement epoch on average. It performs a probabilistic association and uses the state estimates to infer the origin of the detections.

We provide an implementation of the proposed filtering approach for vehicle tracking and finally a discussion of its tracking performance based on real-data examinations. We also provide a MATLAB code that implements the proposed object modeling and state update.[1]

D. Structure

The goal of the presented approach is to update the state $\mathbf{x}$ using radar detection measurements $\mathbf{y}$ at timestamp k:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) \propto p(\mathbf{y}_k|\mathbf{x}_k) \cdot p(\mathbf{x}_k|\mathbf{y}_{1:k-1}). \tag{1}$$

To solve the origin search problem, we model the target object as a complex of spatially distributed scattering sources (*components*). They show different statistical properties in terms of detection rates and kinematic measurements. We exploit this heterogeneity to obtain statistical inference regarding the possible origin of the detection. The division of the object into these components $j \in J$ resembles Gaussian mixtures, allowing the marginal measurement likelihood to be specified in the following format, where $\mathrm{o}^{(j)}$ is the mixture weight:

$$p(\mathbf{y}_k|\mathbf{x}_k) \propto \sum_{j \in J} \mathrm{o}^{(j)}(\mathbf{x}) \cdot \mathcal{N}(\cdot). \tag{2}$$

[1] Available at https://github.com/UniBwTAS/sparse_radar_tracking.

Table I
The Notation and Some Variables of This Article

Symbol	Description
$\mathbf{x}$	(target) state vector
$\mathbf{y}$	single detection measurement
$\mathbf{x}^\star$	predicted state (prior)
$\mathbf{y}^\star$	predicted measurement (given $\mathbf{x}^\star$)
$\hat{\mathbf{x}}$	updated state (posterior)
x, y	scalar Euclidean coordinates
(x, y)	Euclidean position vector (2D)
${}^F(\cdot)$	reference coordinate frame $F \in$ { world frame W, ego vehicle frame E, sensor frame S, target object frame T}
$(\cdot)^{(K)}$	component identifier K

Each component thereby abstracts technical principles that strongly depend on the object type and its current state.

This article is structured bottom-up: Section II specifies the notation and the utilized variables. Section III states our sensor model. Section IV denotes the object modeling concept and its implementation for a vehicle. The fundamental technical properties for each component are stated. Section V carries this on to a spatial measurement function that resembles the marginal measurement likelihood. It is given for each component; the superposition of all likelihoods describes the spatial measurement function for the complete object. This measurement function can be used in Bayes filters. In Section VI, we utilize the predictive measurement likelihood $p(\mathbf{y}_k|\mathbf{y}_{1:k-1})$ based on the state prediction to determine the origin for a given detection measurement $\mathbf{y}$ in a probabilistic manner. Section VII proposes our Bayesian filtering approach. It explains the state update of a single state hypothesis. The performance of the algorithm is then examined with real-world tracking scenarios in Section VIII. We briefly outline interfaces to multi-object multi-sensor tracking frameworks in Section IX and finally discuss the filter in Section X. This article is concluded in Section XI and provides an outlook for future work in Section XII.

II. NOTATION, VARIABLES, AND COORDINATE FRAMES

Table I briefly outlines frequently used variables and the notation of this article. The therein referenced coordinate frames are required for coordinate transformations of the detection measurements and illustrated in Fig. 2. A detection measurement $\mathbf{y}$ is initially obtained in polar sensor coordinates and consists of

- range measurement r with measurement noise σ_r,
- angle measurement α with measurement noise σ_α,
- Doppler measurement $\dot{r}$ with measurement noise $\sigma_{\dot{r}}$,
- amplitude or radar cross section measurement a.

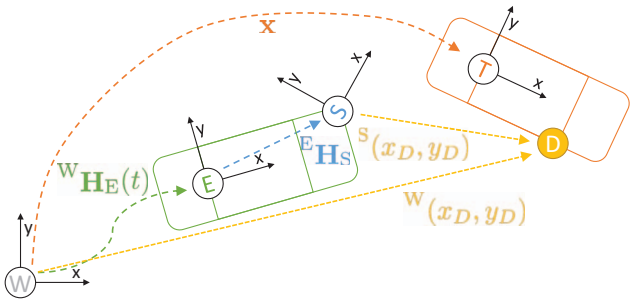

Fig. 2. The coordinate systems. The odometry provides a time-variant transform $^{W}\mathbf{H}_{E}(t)$ from the world coordinates W to the ego vehicle coordinates E. The static mounting location $^{E}\mathbf{H}_{S}$ of the treated sensor establishes the sensor coordinate system S. These transforms project the detection measurement D given in sensor coordinates $^{S}(x_D, y_D)$ into world coordinates $^{W}(x_D, y_D)$. The state estimate $\mathbf{x}$ provides the reference point of the target T in world coordinates.

The position of the detection in Cartesian sensor coordinates $^{S}(x_D, y_D)$ is given by

$$^{S}(x_D, y_D) := {}^{S}\begin{bmatrix} x_D \\ y_D \end{bmatrix} = {}^{S}\begin{bmatrix} r \cdot \cos\alpha \\ r \cdot \sin\alpha \end{bmatrix}, \tag{3}$$

and the Cartesian measurement noise matrix in sensor coordinates $^{S}\mathbf{R}_{\mathbf{xy}}$ is

$$^{S}\mathbf{R}_{\mathbf{xy}} = \mathrm{R}(\alpha) \cdot \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & (2r \cdot \tan(\sigma_\alpha/2))^2 \end{bmatrix} \cdot \mathrm{R}(\alpha)^{\mathsf{T}}, \tag{4}$$

where $\mathrm{R}(\cdot)$ is the two-dimensional rotation matrix. These conversions are subject to bias effects when used in state filters. Dedicated compensation techniques are provided by Bordonaro *et al.* [6], but these effects are negligible compared to real-world measurement phenomena the filter has to deal with in this application. A transformation matrix $^{E}\mathbf{H}_{S}$ from sensor coordinates S to ego coordinates E, which reflects the mounting position, and a time-variant egomotion transformation matrix $^{W}\mathbf{H}_{E}(t)$ from ego to world coordinates convert these parameters into world coordinates. The position of the detection in world coordinates $^{W}(x_D, y_D)$ is given by

$$^{W}\begin{bmatrix} x_D \\ y_D \end{bmatrix} = {}^{W}\mathbf{H}_{E}(t) \cdot {}^{E}\mathbf{H}_{S} \cdot {}^{S}\begin{bmatrix} x_D \\ y_D \end{bmatrix}. \tag{5}$$

Therefore, its measurement noise in world coordinates can be derived as

$$^{W}\mathbf{R}_{\mathbf{xy}} = \mathrm{R}\left({}^{W}\varphi_{E}(t) + {}^{E}\varphi_{S}\right) \cdot {}^{S}\mathbf{R}_{\mathbf{xy}} \cdot \mathrm{R}\left({}^{W}\varphi_{E}(t) + {}^{E}\varphi_{S}\right)^{\mathsf{T}}, \tag{6}$$

with the mounting yaw $^{E}\varphi_{S}$ being deduced from $^{E}\mathbf{H}_{S}$ and the heading of the ego vehicle $^{W}\varphi_{E}(t)$ provided by $^{W}\mathbf{H}_{E}(t)$.

Further variables in this article, like the state vectors, are explained when they are introduced.

III. SENSOR MODEL

This section briefly denotes some sensor-specific parameters and their derivations. Given a specific target reflector, the radar sensor will measure a detection with a particular detection probability. It manifests in the *detection rate* o, which describes at which rate a detection is invoked by a specific reflectivity at a certain distance r. As this rate depends both on the sensor and the individually measured reflectivity, we factor out the sensor-specific part: the *reference rate* $\mathrm{o}_{\mathrm{R}}(r)$. We address not only the resolution ability of the sensor but also firmware-sided tuning. In fact, the signal strength of a radar echo decreases with the fourth power of the distance, but the firmware often neutralizes this effect by applying adapted trigger and noise thresholds in the constant false alarm rate algorithms [29]. This way, the detection rate of an object keeps almost constant over the distance, until the maximum measurement range of the sensor is reached. At this point, the detection rate drops rapidly. However, the exact effect should be determined by measurement analysis for each sensor model. We model the reference rate $\mathrm{o}_{\mathrm{R}}(r)$ of our sensor with

$$\mathrm{o}_{\mathrm{R}}(r) = a \cdot \mathrm{erf}\left((r_{\max} - r)/d\right), \tag{7}$$

using the error function $\mathrm{erf}(\cdot)$

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} \mathrm{d}t, \tag{8}$$

and the parameters effective maximum range $r_{\max}$ (e.g., 40 m), decay magnitude d (e.g., 10 m) and amplitude a, which depends on the reference reflectivity. The reference reflectivity can be chosen arbitrarily here; we have selected the corner of a car. If the amplitude a is determined by measurement analysis, then the reference rate already incorporates the ratio for false negative and true positive detections. The false positive detection rate causes clutter measurements and needs to be determined independently.

The detection measurement is spread around the true position of the target according to its measurement noise characteristics. They are mostly given in sensor coordinates to match the physical measurement process and are described by the scalar uncertainties introduced in Section II. Some sensors provide these uncertainties for each measured detection themselves. This allows for the incorporation of certain ambiguities in the sensor-internal preprocessing that depend on the environment [4].

Another important parameter is the bandwidth of the sensor. It affects the resolution capability. Besides the number of resolved detections, it also determines a kind of longitudinal "penetration" depth: The larger the range resolution is, the more reflections of a larger depth of the object (e.g., caused by vertical tapers) are received inseparably. The peak detection of the sensor then signals a longitudinal measurement, which is smeared over the penetration depth, and therefore needs to be considered as an additive term of the range measurement. The actual incorporation of this term is firmware-specific and depends on the structure of the target.

These parameters describe the radar sensor and aim for its exchangeability without the need to touch the object model. However, if a sensor is utilized that changes the relative detection rate ratio between the individual components of the object, an adjustment of the object model might be required, since this ratio is primarily exploited in the presented approach. We have not observed such a behavior in comparable sensor series, though. Obtaining the extrinsic and intrinsic sensor parameters by dedicated measurement campaigns and ground-truth-assisted analysis is often not trivial despite automation, as it requires manual effect decomposition and can take several days.

IV. OBJECT MODEL AND A VEHICLE IMPLEMENTATION

The abstraction of the physical model (i.e., [3], [4]) has to meet diverse requirements. Its usage in a multi-hypotheses tracking framework must meet real-time constraints, although an oversimplification of the precise physical model impairs the tracking performance. A major criterion is the correct reproduction of the object contour, since any deviation causes a position bias of the object estimate. The measurement data show a strong impact of the viewing angle, which thus needs to be modeled accordingly. Additionally, the data does not only show a dependency on occlusions by other objects but also on self-occlusions. For instance, wheels that are facing the sensor might shadow opposite wheels.

Our approach is to model an object by the superposition of a certain number of *components*: the object is split into a set of separate, individually, and formally described parts. This method gives the opportunity to use different measurement models, visibility constraints, and detection likelihoods for each component. The number of components should preferably be small to achieve fast computation times, but sufficiently large to allow a precise representation of the object. In the example of our vehicle model, suitable component classes are wheels, corners, sides, and the body. A component class can have multiple instances. Physical effects like the micro-Doppler can be specifically implemented for each class and can be exploited to tightly associate a detection to a component. The visibility constraint of a component class can also depend on other components in advanced models. This enables the modeling of self-occlusions or, in multi-object tracking applications, occlusions by other objects.

The following describes a set of components that jointly define the measurement characteristic of a vehicle. This set and its configuration have been obtained by recording and analyzing the reflection characteristics of various vehicles, among them a compact class vehicle and a sport utility vehicle in particular as edge cases [2]. Besides, short-range and far-range sensors have been utilized. According to our findings, the set of components comprising wheels, corners, sides, and the body is a good compromise between precision and complexity regarding the utilized radar sensors. This set results in 4 component classes and 13 component instances. Each component class is defined by a set of attributes $A^{(\cdot)} = \{$ position ${}^{\mathrm{T}}(x_C, y_C)^{(\cdot)}(\mathbf{x})$, position uncertainty $\mathrm{Cov}\left({}^{\mathrm{T}}(x_C, y_C)^{(\cdot)}\right)$, detection rate $\mathrm{o}^{(\cdot)}(\mathbf{x})$, radial speed model $v_r^{(\cdot)}(\mathbf{x})\}$. The *position* ${}^{\mathrm{T}}(x_C, y_C)^{(\cdot)}(\mathbf{x})$ denotes the position of the component in the target frame T. The *position uncertainty* $\mathrm{Cov}\left({}^{\mathrm{T}}(x_C, y_C)^{(\cdot)}\right)$ describes the uncertainty of this position and can also be used to model a slight extent with an additive noise term. The detection rate $\mathrm{o}^{(\cdot)}(\mathbf{x})$ describes the expected number of detections this component invokes. It depends mainly on the reflection characteristics of the component and its visible angular extent. The visible angular extent is usually estimated using the target state and the pose of the radar sensor toward the object. The radial speed model $v_r^{(\cdot)}(\mathbf{x})$ denotes the measurement model of the radial speed measurements for the given component. We deduce the mathematical correlation of these attributes to the state vector using the physical relations found in [3], and parametrize those accordingly to match the observed data. In this work, we use the plain extent state vector $\mathbf{x}_{\mathrm{ext}} = [l, w]^{\intercal}$ to estimate the length l and the width w of the object to preserve a low computing expense. As a result, all remaining required information, like the wheel positions, is statistically derived from both variables. Alternatively, any desired parameter can also be included in the state vector. Apart from that, the following component descriptions are based on a constant turnrate and velocity (CTRV) state model, which describes the position (x, y) and heading φ of the object along with its kinematical properties translational speed v and yaw rate ω. The state vector is finally given as

$$\mathbf{x}_{\mathrm{kin}} = [x, y, \varphi, v, \omega]^{\intercal}, \tag{9}$$

$$\mathbf{x} = [\mathbf{x}_{\mathrm{kin}}^{\intercal}, \mathbf{x}_{\mathrm{ext}}^{\intercal}]^{\intercal}. \tag{10}$$

The kinematic transition matrix yields [31], [34]:

$$\mathbf{x}_{\mathrm{kin},k+1} = \begin{bmatrix} x_k + v_k/\omega_k \cdot (+\sin(\omega_k \Delta t + \varphi_k) - \sin(\varphi_k)) \\ y_k + v_k/\omega_k \cdot (-\cos(\omega_k \Delta t + \varphi_k) + \cos(\varphi_k)) \\ \varphi_k + \omega_k \Delta t \\ v_k \\ \omega_k \end{bmatrix}, \tag{11}$$

and predicts the state epoch $k+1$ from epoch k by integrating the sample time Δt. When the yaw rate ω is close to zero, the transition matrix should be simplified to avoid numeric issues:

$$\mathbf{x}_{\mathrm{kin},k+1} = \begin{bmatrix} x_k + v_k \cdot \cos(\varphi_k)\Delta t \\ y_k + v_k \cdot \sin(\varphi_k)\Delta t \\ \omega_k \Delta t + \varphi_k \\ v_k \\ \omega_k \end{bmatrix}. \tag{12}$$

The transition matrix of the extent state vector $\mathbf{x}_{\mathrm{ext}}$ for rigid objects is the unit matrix $\mathbf{1}_{2\times 2}$.

In the following, the attributes for the four component classes—corners, wheels, sides, and body—are declared.

A. Vehicle Corners

The corners of a vehicle benefit from a good reflection effect. As discussed in Section I-B, the curvature of a visible corner exposes a spot that is perfectly perpendicular to the line of sight of the radar sensor. This spot reflects electromagnetic waves back to the radar sensor with only minimal deflection and thereby obtains an excellent visibility in the measurement data. Consequently, we model the corner as a point target. The exact position of the reflective spot depends on the shape of the vehicle and is empirically derived from the extent state vector. We model the uncertainty of the position with an additive Gaussian noise term. The resulting description is noted in Table II and graphically represented in Fig. 3.

The detection rate depends on several factors, where each models a specific influence on the detection rate. The corner detection rate $\mathrm{o}^{(C)}$ depends here on three factors. The first one is the reference rate, $\mathrm{o}_{\mathrm{R}}(r)$, as discussed in Section III. The second factor, the *component base rate* o_{C}, now puts the reflectivity of a component into relation to the reference reflectivity. The product of both factors thus resembles the detection rate of a specific component. In the case of the corners, $\mathrm{o}_{\mathrm{C}}^{(C)}$ is consequently 1, and the amplitude a of the reference rate has been adjusted to our findings. The third factor $\mathrm{o}_{\mathrm{V}}^{(C)}$ is

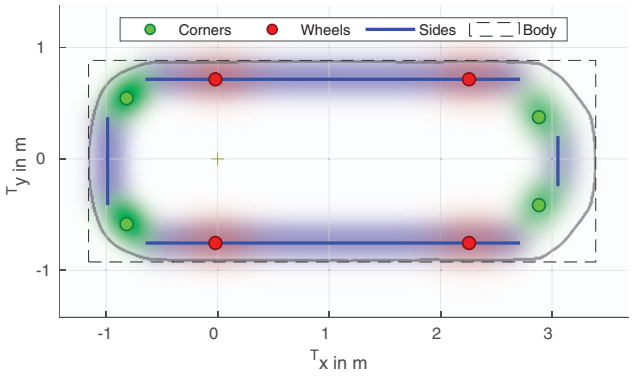

Fig. 3. The location of the components of the reflection model. The dots and lines denote the location of the components in the target frame; the gradients in the background the uncertainty of their location. The contour of a vehicle is overlayed for illustration purposes (gray).

a simple visibility constraint that checks if the corner is visible:

$$\mathrm{o}_{\mathrm{V}}^{(C)}(\mathbf{x}) = \begin{cases} 1 & \text{if adjacent vehicle sides are visible,} \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

which implies that a corner is considered visible if both adjacent sides of the vehicle are visible. The resulting detection rate $\mathrm{o}^{(C)}(\mathbf{x})$ of a corner is the product of all factors:

$$\mathrm{o}^{(C)}(\mathbf{x}) = \mathrm{o}_{\mathrm{R}}(r) \cdot \mathrm{o}_{\mathrm{C}}^{(C)}(\mathbf{x}) \cdot \mathrm{o}_{\mathrm{V}}^{(C)}(\mathbf{x}). \tag{14}$$

Table II
The Component Parametrization for Our Vehicle Model

Components	**Position** ${}^{\mathrm{T}}(x_C, y_C)^{(\cdot)}(\mathbf{x}_{\mathrm{ext}})$	**Position uncertainty** $\mathrm{Cov}\left({}^{\mathrm{T}}(x_C, y_C)^{(\cdot)}\right)$	**Detection rate** $\mathrm{o}^{(\cdot)}(\mathbf{x})$	**Doppler model**
Corner front {left, right}	$\begin{pmatrix} 0.65 \cdot l \\ \pm 0.25 \cdot w \end{pmatrix}$	$\mathrm{R}(\mp 45°) \cdot \begin{pmatrix} (0.15\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \\ (0\,\mathrm{m})^2 & (0.05\,\mathrm{m})^2 \end{pmatrix} \cdot \mathrm{R}(\mp 45°)^{\mathsf{T}}$	Reference Rate × Base Rate (1) × Visibility Constraint	CTRV
Corner rear {left, right}	$\begin{pmatrix} -0.2 \cdot l \\ \pm 0.35 \cdot w \end{pmatrix}$	$\mathrm{R}(\pm 45°) \cdot \begin{pmatrix} (0.15\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \\ (0\,\mathrm{m})^2 & (0.05\,\mathrm{m})^2 \end{pmatrix} \cdot \mathrm{R}(\pm 45°)^{\mathsf{T}}$	Reference Rate × Base Rate (0.66) × Visibility Constraint	none
Wheel front {left, right}	$\begin{pmatrix} 0.5 \cdot l \\ \pm 0.5 \cdot w \mp 0.15\,\mathrm{m} \end{pmatrix}$	$\begin{pmatrix} (0.2\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \\ (0\,\mathrm{m})^2 & (0.1\,\mathrm{m})^2 \end{pmatrix}$		
Wheel rear {left, right}	$\begin{pmatrix} 0\,\mathrm{m} \\ \pm 0.5 \cdot w \mp 0.15\,\mathrm{m} \end{pmatrix}$	$\begin{pmatrix} (0.2\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \\ (0\,\mathrm{m})^2 & (0.1\,\mathrm{m})^2 \end{pmatrix}$		
Side {left, right}	$\begin{pmatrix} -0.15 \cdot l \to 0.6 \cdot l \\ \pm 0.5 \cdot w \mp 0.15\,\mathrm{m} \end{pmatrix}$	$\begin{pmatrix} (0\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \\ (0\,\mathrm{m})^2 & (0.05\,\mathrm{m})^2 \end{pmatrix}$	Reference Rate × Base Rate (0.29/1°) × Angular Width × Visibility Constraint × Scattering	CTRV
Side front	$\begin{pmatrix} 0.67 \cdot l \\ -0.125 \cdot w \to 0.125 \cdot w \end{pmatrix}$	$\begin{pmatrix} (0.05\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \\ (0\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \end{pmatrix}$		
Side rear	$\begin{pmatrix} -0.2 \cdot l \\ -0.15 \cdot w \to 0.15 \cdot w \end{pmatrix}$	$\begin{pmatrix} (0.05\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \\ (0\,\mathrm{m})^2 & (0\,\mathrm{m})^2 \end{pmatrix}$		
Body	-	-	Reference Rate × Base Rate (0.11)	CTRV

The coordinates are given in the target frame T. The predicted state estimate $\mathbf{x}^\star$ provides the length l and the width w of the vehicle. $\mathrm{R}(\cdot)$ is the two-dimensional rotation matrix.

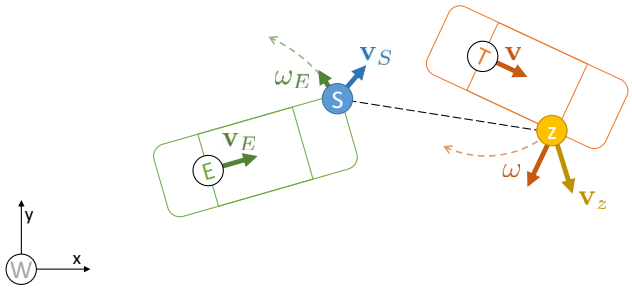

Fig. 4. The computation of the expected radial speed measurement. The velocity vector over ground of the radar sensor is the sum of the translational speed of the ego vehicle v_E and the rotary movement of the sensor caused by the yaw rate ω_E. The velocity vector of the requested point z on the target vehicle is similarly computed using v and ω. The radial speed measurement v_r of this point is the projection of the difference of both velocity vectors along the line of sight (dashed line).

The expected radial speed measurement is computed using the CTRV model [22]. Figure 4 accompanies the following calculation. Firstly, the velocity vector of the radar sensor over ground $\mathbf{v}_S$ is computed:

$$\mathbf{v}_S = \begin{pmatrix} \cos\varphi_E \cdot v_E - \omega_E \cdot (y_S - y_E) \\ \sin\varphi_E \cdot v_E + \omega_E \cdot (x_S - x_E) \end{pmatrix}. \tag{15}$$

This requires the longitudinal speed of the ego vehicle v_E, its heading in world coordinates φ_E, its yaw rate $\omega_E := \dot{\varphi}_E$, its pivot point (x_E, y_E), and the position of the sensor (x_S, y_S) in world coordinates. Secondly, the velocity vector $\mathbf{v}_z$ of the requested point $\mathbf{z}$, which lies on the target vehicle, is determined analogously:

$$\mathbf{v}_z(\mathbf{z}) = \begin{pmatrix} \cos\varphi \cdot v - \omega \cdot (y_\mathbf{z} - y) \\ \sin\varphi \cdot v + \omega \cdot (x_\mathbf{z} - x) \end{pmatrix}, \tag{16}$$

where the speed v, heading φ, yaw rate ω, and position of the target (x, y) are obtained from the target state estimate, while $(x_\mathbf{z}, y_\mathbf{z})$ are the world coordinates of the requested point. Thirdly, the orientation to the detection originating from the sensor Θ is determined:

$$\Theta(\mathbf{z}) = \operatorname{atan2}(x_\mathbf{z} - x_S, y_\mathbf{z} - y_S). \tag{17}$$

And fourthly, the difference of both velocity vectors is rotated to the radar frame:

$$v_r(\mathbf{z}) = \big(\cos\Theta(\mathbf{z}),\ \sin\Theta(\mathbf{z})\big) \cdot (\mathbf{v}_z(\mathbf{z}) - \mathbf{v}_S). \tag{18}$$

v_r returns the longitudinal velocity component or rather the radial speed. These equations are outlined in detail in [4, Section II.C].

B. Vehicle Wheels

The wheels of a vehicle are good reflectors, especially due to the metal rim and the suspension. We model the wheels as point targets, as their extents are also rather small. The rotating wheels cause radial speed measurements that do not match the body of the vehicle and cause the micro-Doppler effect. Thus, the radial speed measurement cannot be used for the kinematic state estimate of the vehicle. However, the radial speed measurement can be exploited to associate a nearby strong detection to a wheel: If the radial speed measurement mismatches the expected radial speed measurement of the body, then the wheel gains a high association probability. The component base rate $\mathrm{o}_\mathrm{C}^{(W)}$ is 0.66 according to our measurement analysis. The visibility constraint $\mathrm{o}_\mathrm{V}^{(W)}$, however, is more complex to model. For example, if the sensor is mounted at a typical low height, then the opposite wheels are in line of sight to it. Although the visible area of an opposite wheel can be computed, necessary parameters like the underbody height of the target are still unknown, and the estimation of it can be challenging. As a result, we reduce this problem to an empiric constant that corresponds to the average detection rate of opposite wheels according to our measurement analysis:

$$\mathrm{o}_\mathrm{V}^{(W)}(\mathbf{x}) = \begin{cases} 1 & \text{if the corresponding side is visible,} \\ 0.3 & \text{otherwise.} \end{cases} \tag{19}$$

The detection rate $\mathrm{o}^{(W)}$ is again the product of all factors:

$$\mathrm{o}^{(W)}(\mathbf{x}) = \mathrm{o}_\mathrm{R}(r) \cdot o_\mathrm{C}^{(W)}(\mathbf{x}) \cdot \mathrm{o}_\mathrm{V}^{(W)}(\mathbf{x}). \tag{20}$$

C. Vehicle Sides

The sides of a vehicle resemble the largest part of the shape of the vehicle. Therefore, they are a significant source of detections. Their extent no longer justifies a point target approximation. Especially, in static scenarios, where a specific part of a side has the highest reflectivity and causes nonuniformly distributed detections along the side, a bias occurs. It shifts the center, or mean, of the side toward that part. Instead, each point of the side has to be regarded as a potentially independent detection source. As a result, we consider each point of the line as a subcomponent of the side of the car, with each point having an independent measurement function to obtain the expected position and radial speed measurements. The detection rate $\mathrm{o}^{(S)}$ for the complete side consists of multiple factors. Figure 5 outlines the calculation of these factors. The first ones are again the reference rate $\mathrm{o}_\mathrm{R}(r)$ and the component base rate $\mathrm{o}_\mathrm{C}^{(S)}$. The latter must now be referenced to a certain angular width like 1° to consider the actual observed width. Multiplying this reference angular width with the actual observed angular width, ψ, then gives the final component base rate. The observed angular width, ψ, depends on the distance, the absolute length, and the orientation of the side. It is computed using the edge points of the side $\mathbf{A} = (x_A, y_A)$ and $\mathbf{B} = (x_B, y_B)$, which are defined in a counter-clockwise order around the center of the vehicle:

$$\psi = \left|\angle(\overrightarrow{\mathbf{SA}}, \overrightarrow{\mathbf{SB}})\right|. \tag{21}$$

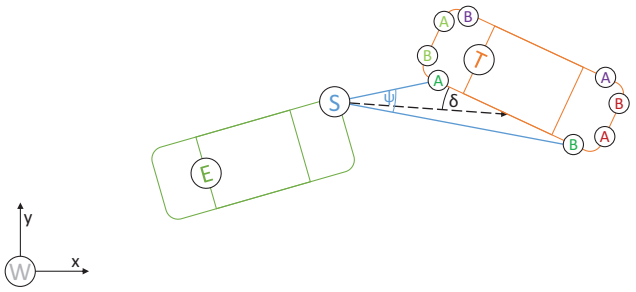

Fig. 5. The computation of the detection occurrence likelihood for the vehicle sides. The rear and right side of the target vehicle T are considered as visible because the sensor S is "right" of the respective vector $\overrightarrow{AB}$. The observed angular width ψ and the angle of incidence δ are drawn for the right side.

When measuring $\mathrm{o}_{\mathrm{C}}^{(S)}$, the vehicle side needs to precisely face the sensor. According to our evaluation, this factor is $0.29/1^{\circ}$.

The next factor is the visibility constraint $\mathrm{q_v}^{(S)}$, which checks if the side is oriented to the radar sensor. This is mathematically performed by computing the non-normalized signed distance $d_S^{(S)}$ of the radar sensor $\mathbf{S} = (x_S, y_S)$ toward the side:

$$d_S^{(S)} = (x_S - x_A) \cdot (y_B - y_A) - (y_S - y_A) \cdot (x_B - x_A)\,. \tag{22}$$

The boolean visibility constraint $\mathrm{o}_{\mathrm{V}}^{(S)}$ is then given by a sign check:

$$\mathrm{o}_{\mathrm{V}}^{(S)}(\mathbf{x}) = \begin{cases} 1 & \text{if } d_S^{(S)} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

The last factor represents the *scattering effect*. The reflections of the radar waves are scattered depending on the angle of incidence δ:

$$\delta = \operatorname{atan2}(y_B - y_A, x_B - x_A) - \angle(\overrightarrow{\mathbf{SM}}, \overrightarrow{\mathbf{AB}}), \tag{24}$$

where the point $\mathbf{M}$ is the midpoint of the side and used as approximative reference:

$$\mathbf{M} = 1/2(\overrightarrow{\mathbf{A}} + \overrightarrow{\mathbf{B}}). \tag{25}$$

This approximation is required as each point of the line has a different angle of incidence. It is fully sufficient for vehicles that are not in the immediate vicinity. Later, in the state update, $\overrightarrow{\mathbf{SM}}$ can also be replaced by the actual orientation of the radar toward the detection.

The steeper the angle between the vehicle side and the radar, the more signal power is scattered and the less signal power is reflected back to the sensor. This damping factor $\mathrm{o}_{\mathrm{S}}^{(S)}$ is provided by a model proposed in [15] and [3]:

$$\mathrm{o}_{\mathrm{S}}^{(S)}(\mathbf{x}) = \sin(\delta)^2. \tag{26}$$

The detection rate finally results in

$$\mathrm{o}^{(S)}(\mathbf{x}) = \mathrm{o_R}(r) \cdot o_{\mathrm{C}}^{(S)}(\mathbf{x}) \cdot \frac{\psi}{1^{\circ}} \cdot \mathrm{o}_{\mathrm{V}}^{(S)}(\mathbf{x}) \cdot \mathrm{o}_{\mathrm{S}}^{(S)}(\mathbf{x}). \tag{27}$$

D. Vehicle Body

Finally, a portion of detections is caused by arbitrary, model-specific parts of the vehicle body. It is not possible to perform a position update of the estimate as the origin of the detection is unknown. However, a kinematic update is conceivable for the current position of the detection. We derived the component base rate $\mathrm{o}^{(B)}$ of the vehicle body from the measurements (i.e., 0.11):

$$\mathrm{o}^{(B)}(\mathbf{x}) = \mathrm{o_R}(r) \cdot o_{\mathrm{C}}^{(B)}(\mathbf{x}). \tag{28}$$

At this point, all relevant components of a vehicle have been abstracted to a set of generic functions.

V. EXPECTED SPATIAL DETECTION LIKELIHOOD

This section depicts the computation of the expected spatial detection likelihood based on the generic component descriptions. The expected spatial detection likelihood serves as a measurement function for an arbitrary object state $\mathbf{x}$ and indicates the expected number of detections for any point in the measurement space $\mathbf{z}$. Thereby, it also takes into account the expected state uncertainty $\mathbf{P}$. The spatial detection likelihood can also be interpreted as the detection rate or frequency for a given point, or as a probabilistic detection density. The spatial sum of the detection likelihood corresponds to the expected number of detections the complete object presumably invokes.

The following equations are given for a single time step. Hence, the corresponding indices are omitted for the sake of simplicity. As a prerequisite, the algorithm requires the component locations $(x_C, y_C)^{(\cdot)}$ to be transformed from target coordinates T to world coordinates W. The mean transformation is given by

$${}^{\mathrm{W}}\begin{bmatrix} x_C \\ y_C \end{bmatrix}^{(\cdot)} = \begin{bmatrix} x^\star \\ y^\star \end{bmatrix} + \mathrm{R}(\varphi^\star) \cdot {}^{\mathrm{T}}\begin{bmatrix} x_C \\ y_C \end{bmatrix}^{(\cdot)}, \tag{29}$$

and its uncertainty transformation by

$$\mathrm{Cov}\left({}^{\mathrm{W}}(x_C, y_C)^{(\cdot)}\right) = \mathrm{R}(\varphi^\star) \cdot \mathrm{Cov}\left({}^{\mathrm{T}}(x_C, y_C)^{(\cdot)}\right) \cdot \mathrm{R}(\varphi^\star)^{\mathsf{T}}. \tag{30}$$

In the following, we perform all computations in world coordinates and omit the coordinate frame index W and time indices to simplify the formal representation.

The spatial detection likelihood indicates the chance of obtaining a detection for any desired point in the measurement space. The likelihood is computed for each component and depends on the object state, the position of the component, its extent model, and also on the uncertainties of both the state and the measurement.

A. Vehicle Corners

The expected measurement vector of a corner consists of a position and radial speed measurement:

$$\mathbf{y}^{\star(C)} = \begin{bmatrix} x_C^{(C)}(\mathbf{x}^\star) \\ y_C^{(C)}(\mathbf{x}^\star) \\ v_r\left((x_C, y_C)^{(C)}(\mathbf{x}^\star)\right) \end{bmatrix}. \tag{31}$$

The measurement matrix $\mathbf{C}^{(C)}$ for a corner can be determined using linearization:

$$\mathbf{C}^{(C)} = \left.\frac{\partial \begin{bmatrix} x_C^{(C)}(\mathbf{x}) \\ y_C^{(C)}(\mathbf{x}) \\ v_r\left((x_C, y_C)^{(C)}(\mathbf{x})\right) \end{bmatrix}}{\partial \mathbf{x}}\right|_{\mathbf{x}^\star}. \tag{32}$$

Subsequently, the innovation covariance matrix $\mathbf{S}^{(C)}$ yields

$$\mathbf{S}^{(C)} = \mathbf{C}^{(C)}\mathbf{P}^\star\mathbf{C}^{(C)\mathsf{T}} + \begin{bmatrix} \mathrm{Cov}\left((x_C, y_C)^{(C)}\right) + \mathbf{R}_{\mathbf{xy}} & 0 \\ 0 & \sigma_{\dot{r}}^2 \end{bmatrix}, \tag{33}$$

which treats the location uncertainty $\mathrm{Cov}((x_C, y_C)^{(C)})$ of the corner as an additive measurement uncertainty. At this point, the spatial detection likelihood caused by a corner $\gamma^{(C)}(\mathbf{z})$ can be computed using the Gaussian distribution $\mathcal{N}(\cdot)$:

$$\begin{aligned} \gamma^{(C)}(\mathbf{z}) &= \mathrm{o}^{(C)}(\mathbf{x}^\star) \cdot \mathcal{N}\left(\mathrm{x}{=}\mathbf{z},\, \mu{=}\mathbf{y}^{\star(C)},\, \sigma^2{=}\mathbf{S}^{(C)}\right) \\ &= \frac{\mathrm{o}^{(C)}(\mathbf{x}^\star)}{\sqrt{(2\pi)^3 \det\left(\mathbf{S}^{(C)}\right)}} \cdot \\ &\exp\left(-\frac{1}{2}\left(\mathbf{z} - \mathbf{y}^{\star(C)}\right)^\mathsf{T} \mathbf{S}^{(C)^{-1}} \left(\mathbf{z} - \mathbf{y}^{\star(C)}\right)\right). \end{aligned} \tag{34}$$

B. Vehicle Wheels

The spatial detection likelihood of the wheels is similar to the corners, but the micro-Doppler effect prevents the usage of the radial speed measurement. The expected measurement $\mathbf{y}^{\star(W)}$ is given by

$$\mathbf{y}^{\star(W)} = \begin{bmatrix} x_C^{(W)}(\mathbf{x}^\star) \\ y_C^{(W)}(\mathbf{x}^\star) \end{bmatrix}, \tag{35}$$

and the measurement matrix $\mathbf{C}^{(W)}$ is given by

$$\mathbf{C}^{(W)} = \left.\frac{\partial \begin{bmatrix} x_C^{(W)}(\mathbf{x}) \\ y_C^{(W)}(\mathbf{x}) \end{bmatrix}}{\partial \mathbf{x}}\right|_{\mathbf{x}^\star}. \tag{36}$$

The innovation covariance matrix $\mathbf{S}^{(W)}$ is

$$\mathbf{S}^{(W)} = \mathbf{C}^{(W)}\mathbf{P}^\star\mathbf{C}^{(W)\mathsf{T}} + \mathrm{Cov}\left((x_C, y_C)^{(W)}\right) + \mathbf{R}_{\mathbf{xy}}. \tag{37}$$

Therefore, the spatial detection likelihood can be described by

$$\gamma^{(W)}(\mathbf{z}) = \mathrm{o}^{(W)}(\mathbf{x}^\star) \cdot \mathcal{N}\left(\mathrm{x}{=}\mathbf{z},\, \mu{=}\mathbf{y}^{\star(W)},\, \sigma^2{=}\mathbf{S}^{(W)}\right). \tag{38}$$

C. Vehicle Sides

As discussed, the length of the vehicle sides demands a more sophisticated handling than the approximation as a point target. Instead, we model a vehicle side as a line. Each point of the line can be the possible source of a detection. As a result, the uncertainties of the state and the measurement reveal a subordinated, continuous association ambiguity for a given detection, as there is a span of possible point sources for a given detection. We aim for a continuous approach [5, Section II.B] to solve the association ambiguity: splitting the line into segments would result in more runtime efforts and only attenuate the bias effect that is evoked by discretized sampling points. To begin, we consider a point $\mathbf{s}^{(S)}(u) \in S, u \in [0; 1]$. Its parametrization can be formally represented as

$$\mathbf{s}^{(S)}(u) = \begin{bmatrix} x_A(\mathbf{x}^\star) + u \cdot (x_B(\mathbf{x}^\star) - x_A(\mathbf{x}^\star)) \\ y_A(\mathbf{x}^\star) + u \cdot (y_B(\mathbf{x}^\star) - y_A(\mathbf{x}^\star)) \end{bmatrix}, \tag{39}$$

by utilizing both end points of the side $(x_A(\mathbf{x}^\star), y_A(\mathbf{x}^\star))$ and $(x_B(\mathbf{x}^\star), y_B(\mathbf{x}^\star))$ in world coordinates W. This allows for the denotation of the expected measurement vector as

$$\mathbf{y}^{\star(s)}(u) = \begin{bmatrix} \mathbf{s}^{(S)}(u) \\ v_r\left(\mathbf{s}^{(S)}(u)\right) \end{bmatrix}, \tag{40}$$

the measurement matrix as

$$\mathbf{C}^{(s)}(u) = \left.\frac{\partial \begin{bmatrix} \mathbf{s}^{(S)}(u) \\ v_r(\mathbf{s}^{(S)}(u)) \end{bmatrix}}{\partial \mathbf{x}}\right|_{\mathbf{x}^\star}, \tag{41}$$

the innovation covariance matrix as

$$\begin{aligned} \mathbf{S}^{(s)}(u) &= \mathbf{C}^{(s)}(u)\mathbf{P}^\star\mathbf{C}^{(s)\mathsf{T}}(u) \\ &+ \begin{bmatrix} \mathrm{Cov}\left((x_C, y_C)^{(S)}\right) + \mathbf{R}_{\mathbf{xy}} & 0 \\ 0 & \sigma_{\dot{r}} \end{bmatrix}, \end{aligned} \tag{42}$$

and finally, the spatial detection likelihood as

$$\gamma^{(s)}(u, \mathbf{z}) = \mathrm{o}^{(S)}(\mathbf{x}^\star) \cdot \mathcal{N}\left(\mathrm{x}{=}\mathbf{z},\, \mu{=}\mathbf{y}^{\star(s)}(u),\, \sigma^2{=}\mathbf{S}^{(s)}\right). \tag{43}$$

The spatial detection likelihood invoked by the complete vehicle side can be computed by summing up the detection likelihoods of all the points:

$$\gamma^{(S)}(\mathbf{z}) = \int_0^1 \gamma^{(s)}(u, \mathbf{z})\ \mathrm{d}u. \tag{44}$$

This integral is known as *stick model* in the literature. Some simplifications of the stick model and the following integrals yield short closed functions. As the computation of the expected radial speed measurement is rather complex, we approximate it by a linear function. This approximation is only used for the association steps

and uses both end points as sampling points. The measurement matrix $\mathbf{C}^{(s)}(u)$ is different for all points, mainly because of the altering impact of the uncertainty of the yaw of the state estimate. Due to its symmetric characteristic, we consider a static innovation covariance matrix, which is either sampled for the center of the line or for the nearest point of the line from the detection.

D. Vehicle Body

The last component class represents arbitrary detections on the body of the vehicle. As the actual source is unknown, a position update is not feasible. However, the expected radial speed measurement can be computed for any point $\mathbf{z}$. This allows for purely kinematic association hypotheses and state updates. The latter is viable by filtering the radial speed measurement at the position of the measured detection. The measurement vector for the body component consequently consists only of the radial speed measurement $\dot{r}$:

$$\mathbf{y}^{(B)} = \dot{r}, \tag{45}$$

while the corresponding expected measurement is given by

$$\mathbf{y}^{\star(B)} = v_r\left(\mathbf{z}\right). \tag{46}$$

By considering the partial derivative, the body measurement matrix $\mathbf{C}^{(B)}$ yields

$$\mathbf{C}^{(B)} = \left.\frac{\partial v_r(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}^\star}, \tag{47}$$

and the innovation covariance matrix $\mathbf{S}^{(B)}$ yields

$$\mathbf{S}^{(B)} = \mathbf{C}^{(B)}\mathbf{P}^\star\mathbf{C}^{(B)\mathsf{T}} + \sigma_{\dot{r}}^2. \tag{48}$$

With these equations, we can compute the detection likelihood $\gamma^{(B)}$:

$$\gamma^{(B)}(\mathbf{z}) = \mathrm{o}^{(B)}(\mathbf{x}^\star) \cdot \mathcal{N}\left(\mathrm{x}{=}\mathbf{z},\, \mu{=}\mathbf{y}^{\star(B)},\, \sigma^2{=}\mathbf{S}^{(B)}\right), \tag{49}$$

which indicates the likelihood that the body component is the source of the detection.

E. Clutter (Optional)

Similar to probabilistic data association filter (PDAF) applications, a clutter hypothesis can be added to the association problem. Clutter is caused by false positive detection measurements and part of the sensor model (Section III). The clutter likelihood $\gamma^{(0)}$ can be modeled with a Poisson distribution. Its parameters usually depend on the distance and the environmental complexity and are firmware-specific. They can be determined with an appropriate measurement analysis. Clutter is not part of the object model itself. However, for single-object tracking applications, it can be interpreted as an additional virtual component that does not invoke a state innovation.

Figure 6 shows the spatial detection likelihoods for different observation angles. Besides, the superposition of all expected spatial detection likelihoods $\gamma^{(J)}(\mathbf{z}) = \sum_{j\in J}\gamma^{(j)}(\mathbf{z})$ is illustrated and compared with the measurement data. As the measurement data can only be visualized for a span of observation angles, the superposed spatial likelihoods are sampled and averaged over this span to match the data visualization. Note that Fig. 6 only shows the position components (x, y) of the measurement space vector $\mathbf{z} = (x, y, \dot{r} = 0)$. A dynamic scenario is shown in Fig. 1 and rendered in the full measurement space.

At this point, our proposed (spatial) measurement function for an extended object is available. It can be used in a Bayes filter like a particle filter to judge state hypotheses and thereby to estimate the state of an object.

VI. DETECTION-TO-COMPONENT ASSOCIATION PROBABILITY

This section outlines the computation of the association probability $\beta^{(j)}$ of a single detection measurement $\mathbf{y}$ toward any component j. This is done by normalizing their detection likelihoods $\gamma(\cdot)$ for the given detection $\mathbf{y}$:

$$\beta^{(j)} = \frac{\gamma^{(j)}(\mathbf{y})}{\sum_{k\in J}\gamma^{(k)}(\mathbf{y})}. \tag{50}$$

Figure 7 illustrates the prior detection likelihoods and the association hypotheses for an example target state estimate with a realistic state and measurement uncertainty. The target is positioned at ${}^{\mathrm{W}}(0\,\mathrm{m}, 0\,\mathrm{m})$ and parked at a heading of 30°. The radar sensor is positioned at ${}^{\mathrm{W}}(0\,\mathrm{m}, -10\,\mathrm{m})$. At this distance, the (Euclidean) lateral measurement noise of the radar sensor is significantly higher than the longitudinal one. First, the priors are computed. The visibility constraints predict visibility for the rear and right vehicle sides, the rear right corner, and all wheels. The a priori detection likelihood of the right vehicle side is higher than the detection likelihood of the rear side because the angle of incidence causes a significantly higher scattering effect at the rear side. The sum of all a priori detection likelihoods is approximately two, i.e., the measurement model expects two detections to be obtained in this scenario. The prior spatial detection likelihood takes all uncertainties into account and predicts the occurrence of detections in the measurement space. It is visualized in the background (gradients). As it is not possible to print the three-dimensional detection likelihood, the gradients are rendered for the sectional plane given by $\mathbf{z} = (x^\star, y^\star, \dot{r}^\star = \dot{r})$, i.e., the plane in the z-axis of Fig. 1 that corresponds to the actually measured radial speed. This foreknowledge about the detection measurement at this point is limited to this illustrative purpose. According to the illustrated spatial detection likelihood, there is a high probability that these are located on the

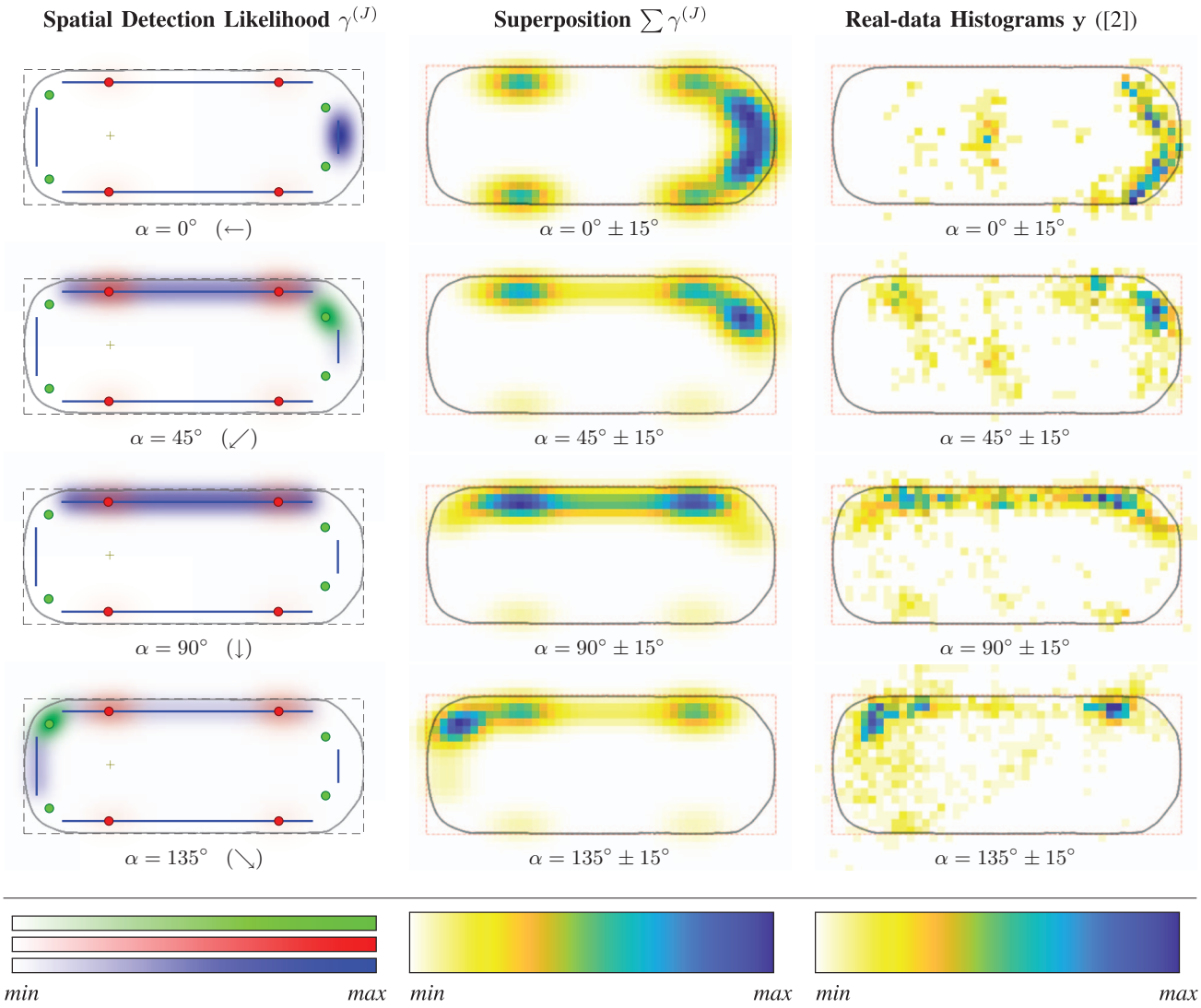

Fig. 6. Comparison of the approximated model and the measurement data: The left column shows the spatial prior detection likelihood for various observation angles α. The middle column shows their superposition for a certain range of observation angles, for which the measurement data on the right column have been respectively recorded. The resolutions are adjusted. The colors of the spatial detection likelihood plots correspond to Fig. 3. The data histograms are collected over several minutes.

rear right corner or on the right vehicle side. The likelihood that a detection is measured at the back side or the rear right wheel is lower.

As the illustrated detection measurement (blue point) is received, the association probabilities can be computed. The detection results in a high association probability of the right side, a moderate probability of the back side, but a low probability of the rear right corner (primarily due to the lower longitudinal measurement uncertainty). The shown orange arrows represent the association hypotheses. In case of the sides, they point to their mean origin point $\bar{\mathbf{s}}^{(S)}$. This point is the average of all points on the side, but weighted by their individual association likelihood:

$$\bar{u} = \left(\int_0^1 u \cdot \gamma^{(s)}(u, \mathbf{y}) \, \mathrm{d}u \right) \Big/ \left(\int_0^1 \gamma^{(s)}(u, \mathbf{y}) \, \mathrm{d}u \right), \quad (51)$$

$$\bar{\mathbf{s}}^{(S)} = \begin{bmatrix} x_A(\mathbf{x}^\star) + \bar{u} \cdot (x_B(\mathbf{x}^\star) - x_A(\mathbf{x}^\star)) \\ y_A(\mathbf{x}^\star) + \bar{u} \cdot (y_B(\mathbf{x}^\star) - y_A(\mathbf{x}^\star)) \end{bmatrix}. \quad (52)$$

The obtained mean origin point $\bar{\mathbf{s}}^{(S)}$ can be used to recompute the measurement matrix $\mathbf{C}^{(s)}(u)$ in a recursive approach. As the lateral innovation uncertainty is higher than the longitudinal uncertainty, the mean origin points of the sides are mainly laterally shifted from the detection measurement.

Figure 8 introduces dynamics in the scenario: The target drives in a curve to the left. The radial speed measurement is set in a manner that it matches the expected radial speed measurement of the rear side. As the rear slightly moves toward the sensor in a left curve, the radial speed measurement is negative. Given the low uncertainty of the radial speed measurement, the detection is now associated with the rear side with high significance. Additionally, the hypothesis that the detection originates from the wheel emerges as its

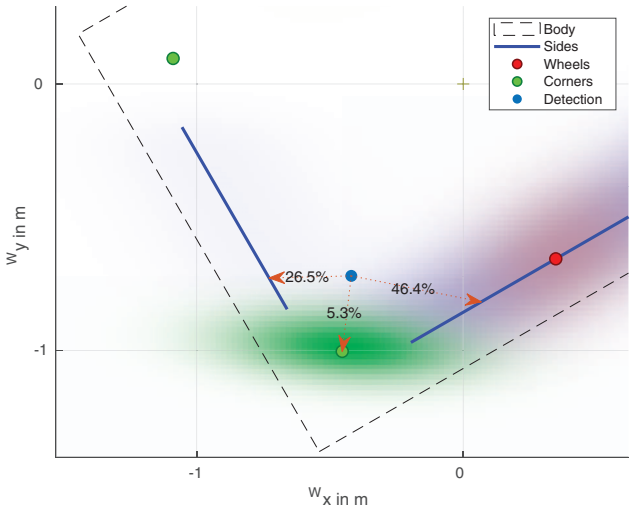

Fig. 7. Association probabilities of a static scenario. The orange arrows denote the association probabilities. The probabilities for the body (18.0%) and clutter (3.6%) are not shown.

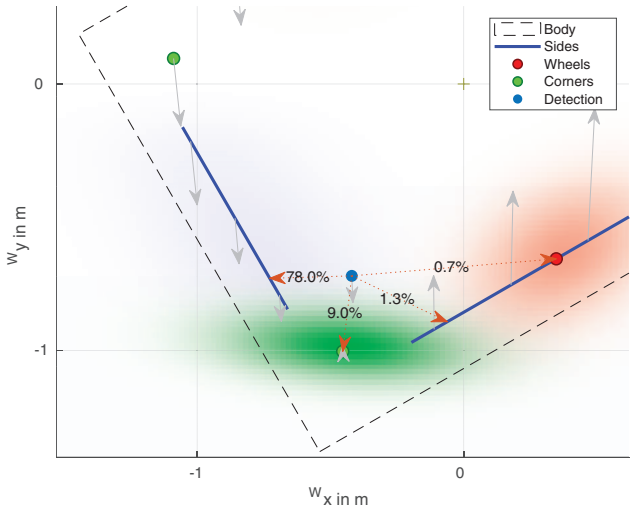

Fig. 8. Association probabilities of a dynamic scenario. The probability for the body is ∼0% and for the clutter is 10.9%. The gray arrows denote the radial speed measurement of the detection and the expected radial speed measurements of all components ($1\,\mathrm{m} \mathrel{\widehat{=}} 1\,\mathrm{ms}^{-1}$).

association ignores the mismatching radial speed differences due to the micro-Doppler effect. The estimated mean origin of the right side is shifted to the left to better match the radial speed measurement, but it still loses any significant association likelihood.

The association probabilities $\beta^{(j)}$ can now be used to obtain a probabilistic indication about the origin of a detection.

VII. STATE UPDATE

This section describes the filter principle and its actual implementation to update the state and uncertainty of a single-object state hypothesis $\mathbf{x}$.

A. Principle

Each component is a possible source of a detection. Especially when considering the uncertainty of both measurement and state, a given detection could originate from multiple components. A particle filter that matches the complete spatial detection frequency or a multiple hypothesis tracking (MHT)-adapted approach that tracks the associations of the detections to the components over time is not feasible in an application where a multi-hypotheses tracking is run upstream. A simple maximum a posteriori estimate, or hard association, though, does not establish a robust tracking due to the high ambiguity of the association problem. We aim for a soft association approach, which represents a suitable compromise according to our findings. The association ambiguities are resolved probabilistically and are still encased in a Gaussian state formulation. Association uncertainties are thereby incorporated in the state uncertainty. The utilized association and tracking algorithm has been developed previously [5] as preparation for this work. It shares its basic principles with the PDAF [1] and is briefly stated in the following.

Every component j has a state-dependent association likelihood $\gamma^{(j)}$. It denotes the presumption that component j has caused a given detection measurement. The absolute association probability, $\beta^{(j)}$, is determined by computing the association likelihoods for all components J and by normalizing them, as done in equation (50). At this point, clutter measurements are not yet considered. As each component description correlates the object state with the component, it can also provide a state update $\hat{\mathbf{x}}^{(j)}$ of the predicted target state $\mathbf{x}^\star$. This update is conditioned on the assumption that the detection is actually caused by the component j:

$$\hat{\mathbf{x}}^{(j)} = \mathbf{x}^\star + \mathbf{K}^{(j)} \cdot \left(\mathbf{y} - \mathbf{y}^{\star(j)}\right), \tag{53}$$

where $\mathbf{K}^{(j)}$ is the Kalman gain of component j, $\mathbf{y}$ is the detection measurement, and $\mathbf{y}^{\star(j)}$ is the expected measurement if component j is assumed to be the origin of the detection. In a last step, the conditional state updates $\hat{\mathbf{x}}^{(j)}$ are fused according to their association probabilities $\beta^{(j)}$ to obtain the updated state $\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = \sum_{j \in J} \beta^{(j)} \cdot \hat{\mathbf{x}}^{(j)}. \tag{54}$$

The updated state uncertainty $\hat{\mathbf{P}}$ is calculated similarly by

$$\hat{\mathbf{P}} = \sum_{j \in J} \beta^{(j)} \left((\mathbf{1} - \mathbf{K}^{(j)}\mathbf{C}^{(j)})\mathbf{P}^\star + \underline{(\hat{\mathbf{x}}^{(j)} - \hat{\mathbf{x}})(\hat{\mathbf{x}}^{(j)} - \hat{\mathbf{x}})^\intercal} \right) \tag{55}$$

and depends on the predicted state uncertainty $\mathbf{P}^\star$ and measurement matrices $\mathbf{C}^{(j)}$ for all components $j \in J$. The term marked with a dashed underline represents the uncertainty of the association, the so-called *spread*

of means. This term carries the information how certain the association search is. It increases according to the cardinality, the likelihood, and the impact of alternative association hypotheses.

This probabilistic association requires a far lower computing expense than multi-hypotheses trackers, which resolve the combinatorial association problem over time. The downside of probabilistic associations is that even wrong associations are filtered in with a certain weight. In this application, significant association probabilities are only invoked by components that are of comparable likelihood to have caused the detection. Such components are mostly close together, and their state updates are similar as they are part of a rigid body. This is a major difference to PDAF applications where multiple, independent object tracks are updated with a single measurement.

B. Implementation

In the following, a Kalman filter update of the target state, considering a single-object hypothesis, is performed.

The update of the target state is composed of update steps for each single component. According to the principle of the probabilistic origin association, each individual component update step is performed in the assumption that the respective component is the origin of the given detection, regardless of its actual association probability. For every component $j \in J$, the Kalman gain $\mathbf{K}^{(j)}$ is computed as

$$\mathbf{K}^{(j)} = \mathbf{P}^{\star}\mathbf{C}^{(j)\mathsf{T}}\mathbf{S}^{(j)-1} \tag{56}$$

and utilized for the component-wise state updates $\hat{\mathbf{x}}^{(j)}$ according to equation (53), and subsequently for the fused posterior state estimate $\hat{\mathbf{x}}$ according to equation (54) and the posterior state uncertainty $\hat{\mathbf{P}}$ according to equation (55). While the state updates $\hat{\mathbf{x}}^{(j)}$ can be directly obtained for the other components, the vehicle sides require a more elaborate treatment. Their native posterior state estimates $\hat{\mathbf{x}}^{(S)}$ yield

$$\begin{aligned}\hat{\mathbf{x}}^{(S)} &= \int_0^1 \beta^{(s)}(u) \cdot \hat{\mathbf{x}}^{(s)}(u)\ \mathrm{d}u \\ &= \int_0^1 \beta^{(s)}(u) \cdot \left(\mathbf{x}^{\star} + \mathbf{K}^{(s)}(u) \cdot \left(\mathbf{y} - \mathbf{y}^{\star(s)}(u)\right)\right)\ \mathrm{d}u.\end{aligned} \tag{57}$$

The recursive approximation of the measurement matrix $\mathbf{C}^{(s)}(u)$ for a static replacement $\mathbf{C}^{(s)}(\bar{\mathbf{s}}^{(S)})$, previously discussed in Section VI, simplifies the posterior state estimate to

$$\hat{\mathbf{x}}^{(S)} = \mathbf{x}^{\star} + \mathbf{K}^{(s)}(\bar{u}) \cdot \left(\mathbf{y} - \underline{\int_0^1 \beta^{(s)}(u) \cdot \mathbf{y}^{\star(s)}(u)\ \mathrm{d}u}\right), \tag{58}$$

where the dashed underlined term represents the mean of the origins. This mean has already been computed in equation (51) and yields

$$\hat{\mathbf{x}}^{(S)} = \mathbf{x}^{\star} + \mathbf{K}^{(s)}(\bar{u}) \cdot \left(\mathbf{y} - \mathbf{y}^{\star(s)}(\bar{u})\right). \tag{59}$$

This approximation implies that the Kalman gain of the expected mean origin is applied to the nearby, less likely origins in a symmetric manner. Therefore, the native posterior state uncertainty of a vehicle side $\hat{\mathbf{P}}^{(S)}$ is given by

$$\begin{aligned}\hat{\mathbf{P}}^{(S)} = \int_0^1 \beta(u)\Bigg(&\left(\mathbf{1} - \mathbf{K}^{(s)}(u)\mathbf{C}^{(s)}(u)\right)\mathbf{P}^{\star} \\ &+ \left(\hat{\mathbf{x}}^{(s)}(u) - \hat{\mathbf{x}}^{(S)}\right)\left(\hat{\mathbf{x}}^{(s)}(u) - \hat{\mathbf{x}}^{(S)}\right)^{\mathsf{T}}\Bigg)\ \mathrm{d}u\end{aligned} \tag{60}$$

and can be simplified with the same approximation $\mathbf{C}^{(s)}(u) \approx \mathbf{C}^{(s)}(\bar{u})$ to obtain a closed equation, albeit too long to be printed here. Details on the analytic solution can be found in the supplied MATLAB code. Similar to the discrete association problem, this term incorporates the uncertainty of the association search into the resulting innovation uncertainty. In this way, high state or measurement uncertainties increase the possible association range and are—in contrast to pure greedy decision approaches—probabilistically resolved.

Another implementation issue concerns the computation of the expected radial speed measurement. In the association step, it is computed for the mean position of the respective component (x_C, y_C). Its advantage is the improved search for the origin of a detection by comparing it with the precise radial speed measurement. However, the actual origin of the detection can be located anywhere on the extent of the component; it is spatially distributed according to its position mean and uncertainty parametrization in Table II. Depending on the size of the respective extent, this discrepancy might cause a pseudo-systematic bias in the state update, especially when perceiving the component repeatedly from a similar angle. An alternative is the usage of the measured position of the detection (x_D, y_D): It is instead subject to (zero-mean) measurement noise. The choice depends on the magnitude of the estimated uncertainties, the measurement noise and the extent of the components. We have performed an ablation study and gained the result that both variants perform almost identical, primarily because the Gaussian extents are minor relative to the Doppler gradient [4]. We utilize the second option for (and only for) the state update to achieve higher generality with respect to extent sizes. This requires the reprocessing of equations (31) and (40) to use $v_r\big((x_D, y_D)\big)$, and subsequently the reprocessing of their respective measurement and innovation covariance matrices described by equations (32), (33), (41), and (42).

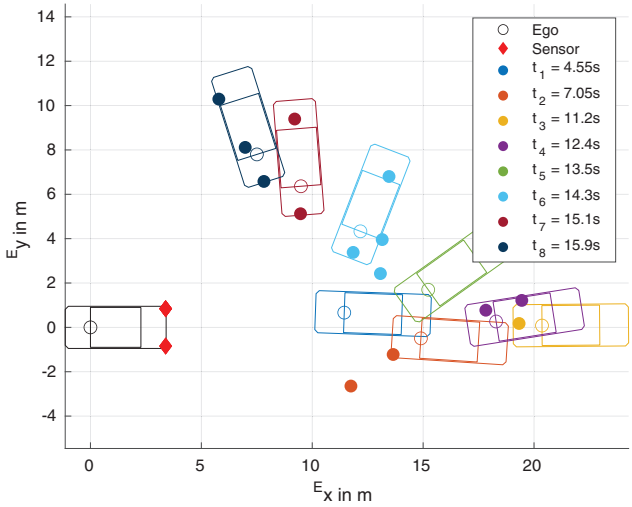

Fig. 9. Country road trailing: Exemplary measurement time steps and their respective true target positions *in the ego frame*. The blank circles denote the reference points of the ego vehicle (black) and the target vehicle (colored according to the timestamp). The filled circles denote the respective detection measurements. No detections were obtained in t_1 and t_5.

VIII. SINGLE-OBJECT TRACKING PERFORMANCE

This section focuses on single-object tracking and illustrates three different tracking scenarios. The measurement data has been recorded with low-resolution radar sensors, which are mounted on the corners of the ego vehicle. We utilize a RTK-GNSS/IMU-based ground truth with centimeter-level accuracy for both the gating of the radar detections (within a radius of 4 m to the center of the target) and the evaluation of the tracking performance. Figure 9 shows the measured detections and the ground-truth-provided true object state of some representative measurement frames of the first curve of the first scenario. This illustration reveals the challenge that the tracking algorithm has to tackle. On the one hand, the number of detections is low, the detections are generated at unknown positions and are subject to significant measurement noise. On the other hand, the dynamic variables of the object state can change abruptly. Wrong associations would directly impair the tracking robustness. The association algorithm primarily exploits the statistical detection characteristic, provided by the object model, and the radial speed measurements to solve the association problem. These difficulties should be considered when assessing the resulting tracking performance.

All scenarios use the same parametrization. The process noise of the CTRV model has been obtained by the inspection of a larger dataset and regards slight model inconsistencies concerning unpaved roads, slopes, and varying driving styles. We parameterize it by

$$\sigma^2_{\text{kin}} = \text{diag}\big([(4.5\,\text{cm})^2, (4.5\,\text{cm})^2, \tag{61}$$

$$(1.1^\circ)^2, (0.67\,\text{ms}^{-1})^2, (6.3^\circ\text{s}^{-1})^2]\big). \tag{62}$$

There is no process noise modeled for the extent state model. The initial CTRV position is roughly set to the first encountered measurement. All kinematic means are zero. All CTRV state parameters are initialized as extremely uncertain. The initialization of the extent state, though, depends on the application. In general, the low number of detections obtained in the usual observation time, as in urban scenarios, does not permit a very precise extent estimation. In such applications, the extent state should be initialized with an average extent state (like $\mathbf{x}^{\mathsf{T}}_{\text{ext},0} = [4.85\,\text{m}, 1.85\,\text{m}]$) and with a low uncertainty. Longer observation times, as given in the presented scenarios, render extent estimation feasible. The initial extent is set to $\mathbf{x}^{\mathsf{T}}_{\text{ext},0} = [4.7\,\text{m}, 1.75\,\text{m}]$. The length of the utilized target vehicles exceeds this by up to ∼0.6 m: This initial mismatch additionally challenges the component association search. The initial extent uncertainty is set to $\sigma^2_{\text{l},0} = 0.1\ \text{m}^2$ and $\sigma^2_{\text{w},0} = 0.015\ \text{m}^2$. Especially the variety of the width among typical vehicles is obviously bounded by regulations [21].

To deal with clutter in the single-object tracking (without a track management that handles clutter itself), we utilize the clutter hypothesis from Section V-E. We set $\gamma^{(0)} = 0.01$. This implies that roughly 1% of all detection measurements in immediate proximity of the target are clutter. This value is conservatively modeled without dependency to distance or signal strength, as such factors are already regarded in the clutter suppression of the sensor firmware. The clutter hypothesis has a certain "association" probability in equation (50) depending on the matching of a detection measurement to the real components. The clutter is then ignored in the subsequent state updates [equations (54) and (55)]. Consequently, the clutter detection is resolved probabilistically. A detection measurement that does not match any component at all (after consideration of all uncertainties), will not invoke a state update.

A. Country Road Trailing

In this scenario, the ego vehicle follows the target (a mid-class sedan) in a winding round trip. Clutter measurements are obtained from vegetation on the road side. The short-range mass-market, 77 GHz radar sensors utilized in this scenario have a substantial lateral measurement noise. A sensor is mounted at each corner of the vehicle, although only the front two sensors perceive the target. Each sensor provides measurements at a rate of roughly 20 Hz.

Figure 10 outlines the path of the ego vehicle (both by estimate and ground truth) in world coordinates, while Fig. 11 illustrates the accumulation of detection measurements and position estimates over time in the target frame (based on the ground truth). The lateral measurement noise also manifests itself in a lateral position estimate error. Moreover, Fig. 12 provides the yaw estimate over time, while Figs. 13 and 14 outline the estimation errors of the dynamic states. Their estimation

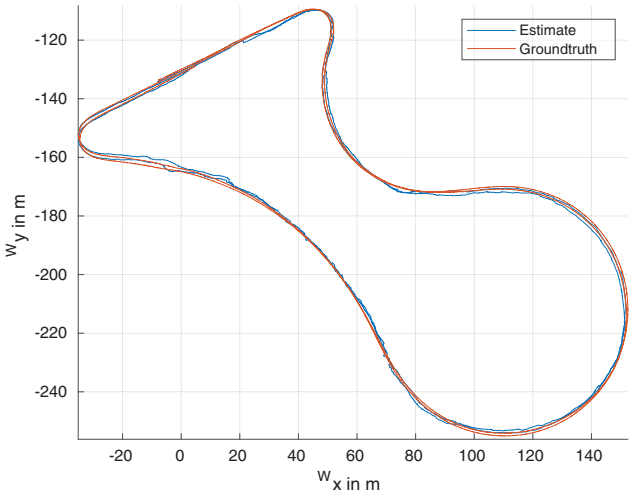

Fig. 10. Country road trailing: Position estimate in world coordinates.

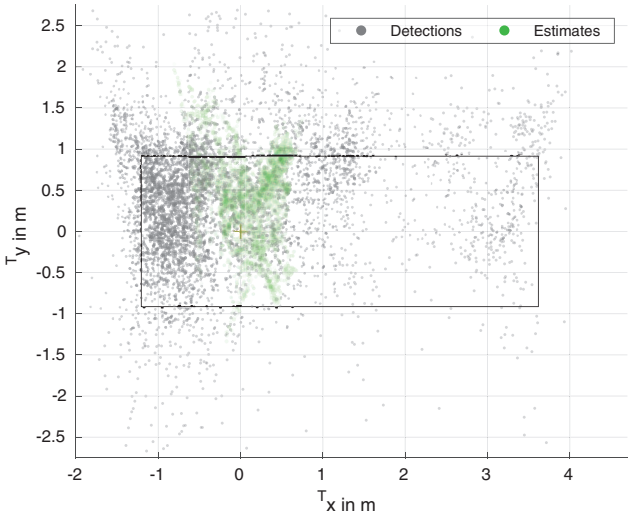

Fig. 11. Country road trailing: Position estimates in target coordinates (accumulation over whole dataset). The rectangle resembles the true extent of the vehicle.

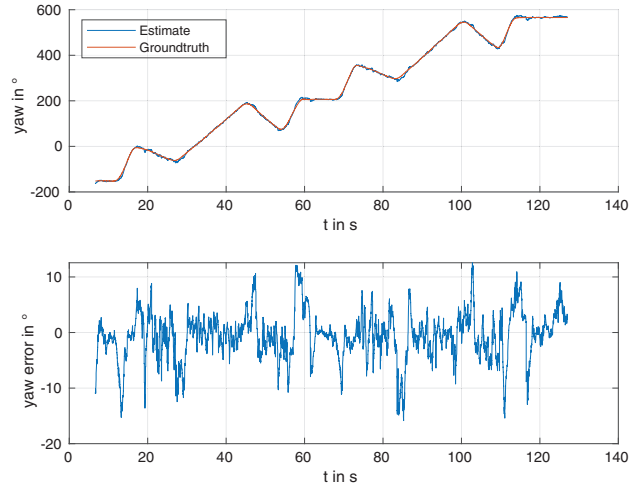

Fig. 12. Country road trailing: Yaw estimate.

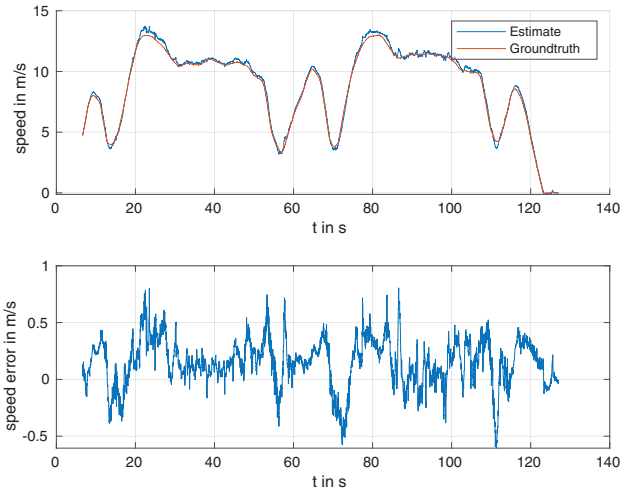

Fig. 13. Country road trailing: Speed estimate.

is subject to higher noise as they are the highest-order states of the CTRV model. Figure 15 shows the length estimation. After the first curve and progressing kinematic estimation, it is steadily improving. An error of roughly 5% remains. The repository referred in Section I-C contains a video that illustrates the association technique based on the first curve of this scenario.

B. Circling

In this scenario, the target is a long-wheelbase luxury sedan. Its extent exceeds the dimensions of the vehicles used as reference in the modeling. Besides, the radar sensors of the ego vehicle are slightly more recent and provide more detections but also more clutter measurements than the ones used in the first scenario. In addition, two additional sensors are mounted near the centers of the left and right vehicle side. The target vehicle drives circles around the parked ego vehicle, and is thus perceived by all sensors in rotating manner. However, only one side of the target vehicle is observed. Al-

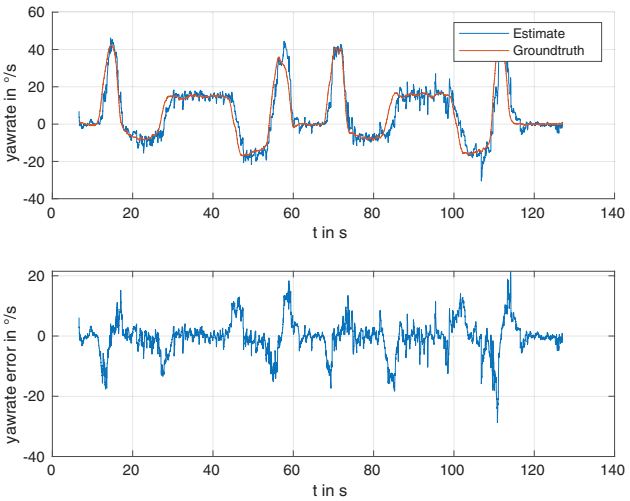

Fig. 14. Country road trailing: Yaw rate estimate.

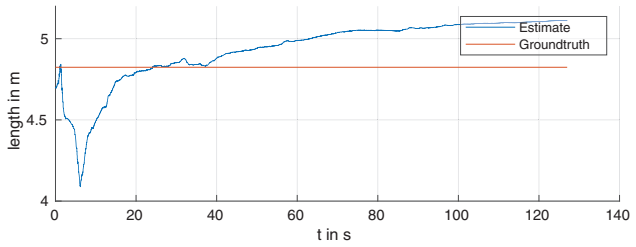

Fig. 15. Country road trailing: Length estimate.

though the lateral measurement noise is lower than in the previous example, the lateral association problem arises for the entire vehicle side. The longitudinal movement of the car is mainly inferred by detections at the ends of the side because they restrict the possible longitudinal position. In this scenario, the tracking algorithm benefits from the different component characteristics for the association and subsequent position estimation. Figures 16 and 17 show the position estimates, while Figs. 18–20 outline the estimation of the yaw, the speed and the yaw rate over time. Figure 21 shows the length estimation. It has settled from the 20th second. This scenario shows an interesting effect. Although the sensors only observe the left side of the target, a width estimation is feasible if the initial extent uncertainty is chosen accordingly: The filter inherently exploits the visibility of the opposite wheels to directly infer the width state. Figure 22 shows its estimation over time, although the resulting accuracy is attributable to the long observation time. This specific figure has been obtained by utilizing an initial width uncertainty of $\sigma^2_{w,0} = 0.2\ \mathrm{m}^2$. This parametrization, though, is far higher than the statistical variety of typical vehicles and needlessly reduces the robustness of the filter especially in challenging situations.

C. Urban Trailing

The ego vehicle follows the target vehicle again in this scenario. The target vehicle and the sensors are identical to the those utilized in the second scenario. Nearby metallic containers and buildings on this narrow track cause mirrored (ghost) detections and signal overexposures. They lead to biased detections and the loss of detections from the target vehicle. The estimation performance decreases, especially concerning the yaw due to the biased detections, but the tracking stays robust. Figures 23–27 show the respective tracking performance. Figure 28 illustrates the length estimate. Again, after the kinematic quantities have roughly been estimated, it is able to resolve the initial extent error. It is steadily improving as the vehicle is mostly observed from behind.

Table III depicts a root-mean-square error (RMSE) comparison of all scenarios. Changes of the parametrization of the CTRV process noise in the range of ±20% (standard deviation) have not shown a worse degradation than 6% of the RMSE of any state variable; some state variables also show better accuracy. The position of features like the wheels is yet purely statistically derived, and its modeling error and the extent are mod-

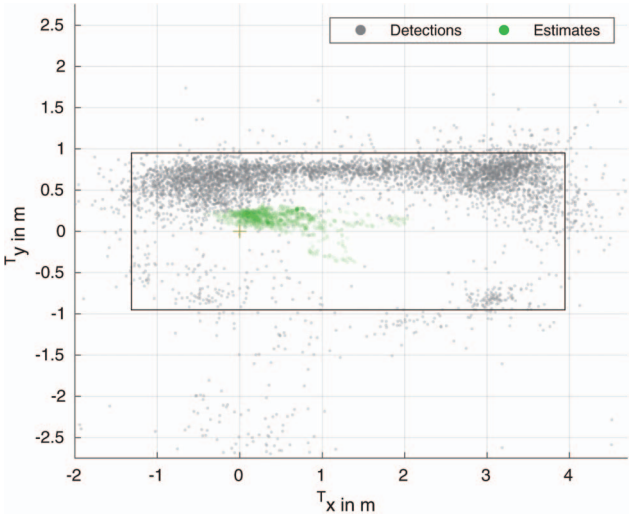

Fig. 17. Circling around the ego: Position estimates in target coordinates.

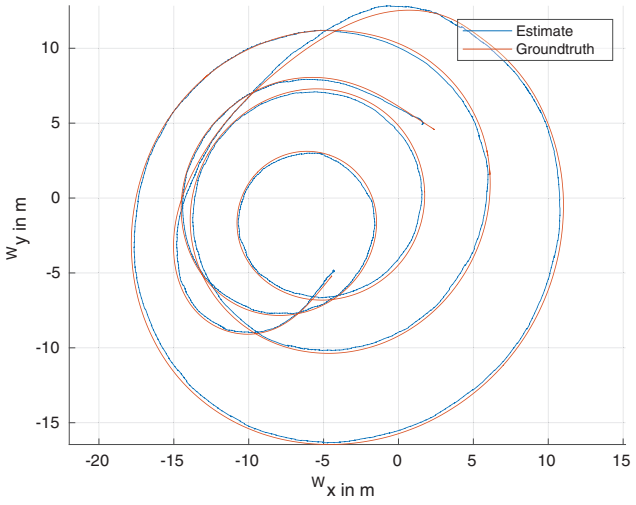

Fig. 16. Circling around the ego: Position estimate in world coordinates.

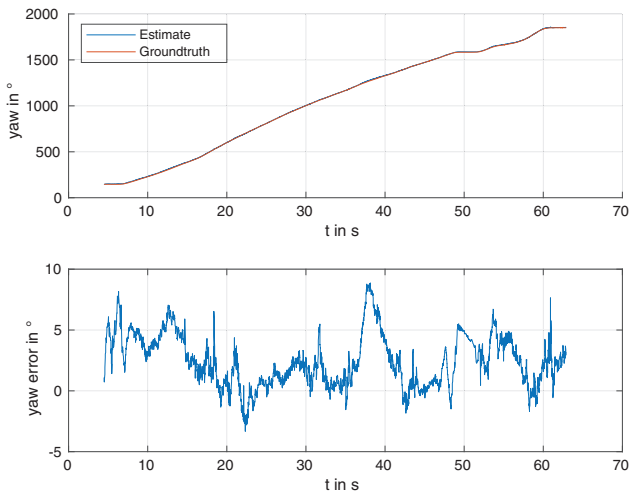

Fig. 18. Circling around the ego: Yaw estimate.

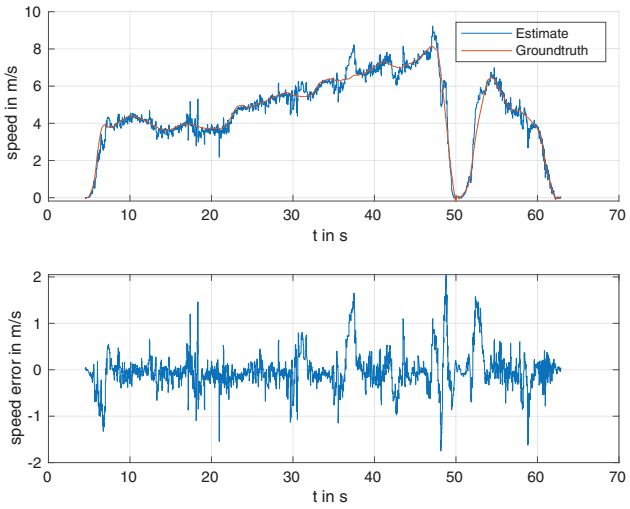

Fig. 19. Circling around the ego: Speed estimate.

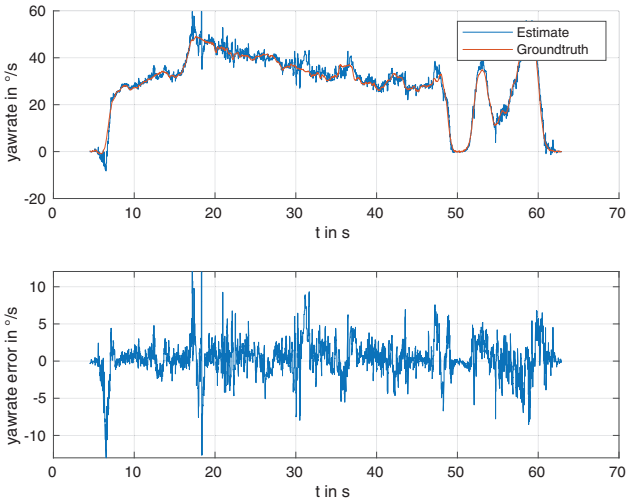

Fig. 20. Circling around the ego: Yawrate estimate.

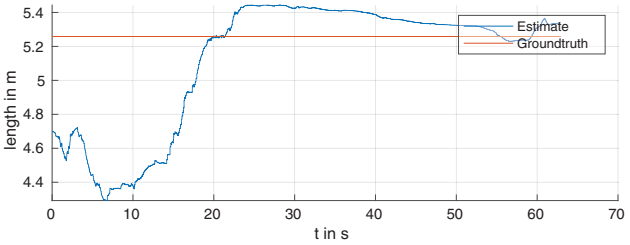

Fig. 21. Circling around the ego: Length estimate.

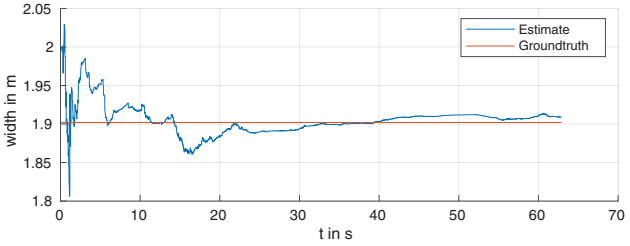

Fig. 22. Circling around the ego: Width estimate.

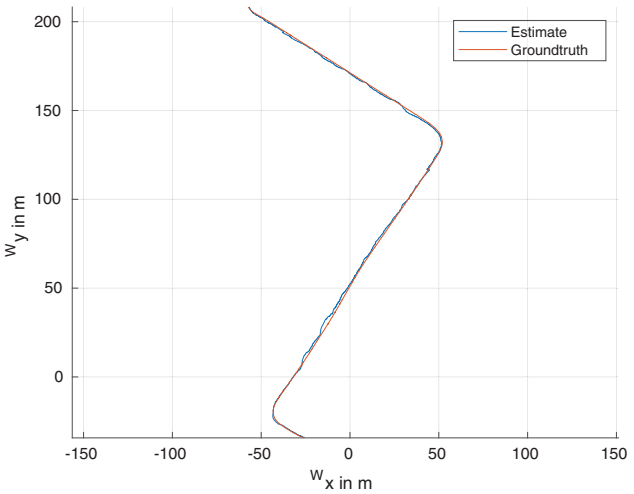

Fig. 23. Urban trailing: Position estimate in world coordinates.

Table III
Tracking Performance

RMSE	${}^{\mathrm{T}}x$	${}^{\mathrm{T}}y$	${}^{\mathrm{W}}\varphi$	${}^{\mathrm{W}}v$	${}^{\mathrm{W}}\omega$
1) Country road	0.34 m	0.66 m	4.3°	0.25 ms^{-1}	5.4°s^{-1}
2) Circling	0.60 m	0.18 m	3.2°	0.39 ms^{-1}	2.3°s^{-1}
3) Urban trailing	0.30 m	0.69 m	5.1°	0.15 ms^{-1}	4.3°s^{-1}

The Position Error References the CTRV Pivot Point and is Given in Target Coordinates T, While the Other Errors are Given in World Coordinates W.

eled with a Gaussian noise term. A possible improvement is to correlate this noise term with the estimated extent size. However, the induced change is negligible considering typical vehicles. The incorporation of the position of such features in the state vector and their explicit estimation support the extent estimation, as the individual modeling error can be corrected over time. However, their precise estimation requires an observation time that exceeds typical urban scenarios (consider-

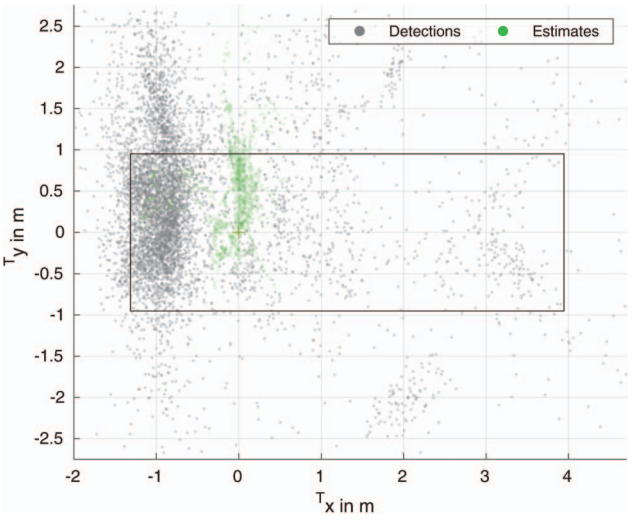

Fig. 24. Urban trailing: Position estimate in target coordinates.

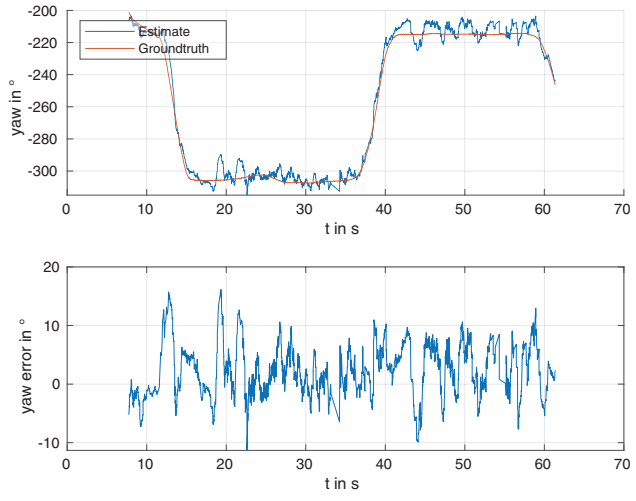

Fig. 25. Urban trailing: Yaw estimate.

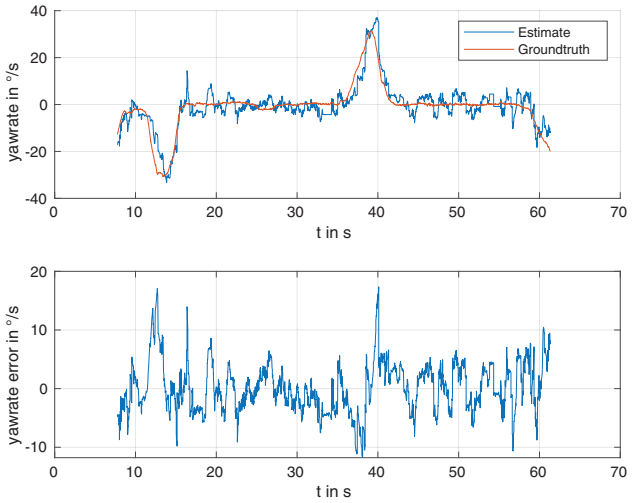

Fig. 27. Urban trailing: Yaw rate estimate.

ing sparse measurements). Next to groundtruth-assisted scenarios, our dataset also contains urban scenarios with numerous vehicles. As no ground truth is available for those vehicles, only a qualitative evaluation of the robustness using LiDAR scans could be performed there.

IX. TRACKING FRAMEWORK INTERFACES

This section outlines interfaces which integrate the proposed filter into larger tracking frameworks.

A. Interfaces to Low-Level Fusion Algorithms

The abstraction of an object to its physical components offers a convenient opportunity to fuse heterogeneous sensor data. Camera sensors and their processing chains often utilize semantic segmentation to detect features of objects. Following the example of vehicles, these are wheels, lights, license plates, and corners. Furthermore, LiDAR sensors detect license plates particularly well due to their reflectivity. If mounted closer to the ground, they also obtain point measurements from opposite wheels.

Those features are either already directly observable with the radar sensor (e.g., wheels and corners) or can be added as additional components. The description of additional components with the parameter set A correlates them to the state vector. The abstraction of objects into components is thus a suitable interface for a low-level or feature-level fusion.

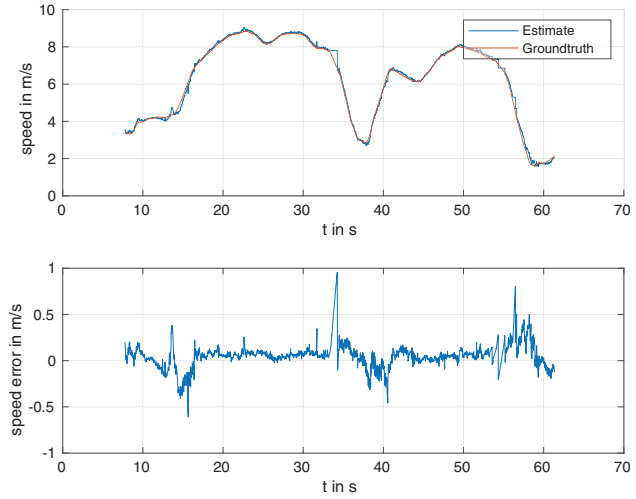

Fig. 26. Urban trailing: Speed estimate.

B. Interfaces to Multi-Object Trackers

We use this radar tracker in interaction with a multihypotheses track management in a C++/ROS-based real-time tracking application. Although the tracking shows robustness against local optima like turned vehicles or wrong wheel associations, a multihypotheses tracker speeds up the correction. The track management usually requires some additional interfaces to the underlying trackers in addition to the actual state updates of the hypotheses. Due to the size of this article, an actual implementation of a multi-target tracking and the interaction of objects cannot be covered here.

1) Track Initialization: When a track for a new object is created, either because it enters the range of visibility or it leaves an occluded area, the first obtained measurement is usually a single detection y. The center of the new object is normally set to the position of the detection as

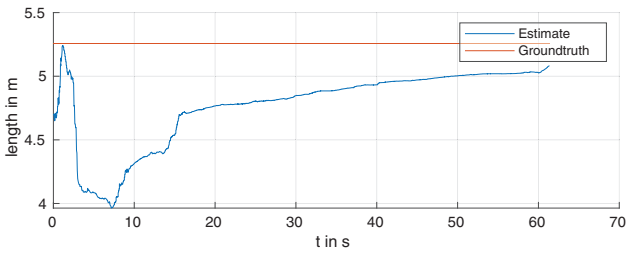

Fig. 28. Urban trailing: Length estimate.

its orientation, and therefore its side facing the sensor, is unknown. Instead, we exploit the radial speed measurement to obtain a first rough velocity vector $\mathbf{v}_W$ in world coordinates. Subsequently, we derive the orientation of the object φ (by assuming that it moves forward), determine the facing side, and place this side on the obtained detection. The velocity vector $\mathbf{v}_W$ is the sum of the velocity vector of the radar sensor $\mathbf{v}_S$ (see equation (15)) and the rotated radial speed measurement $\dot{r}$, which is transformed to world coordinates:

$$\mathbf{v}_W = \mathbf{v}_S + \mathrm{R}\left(\Theta(\mathbf{y})\right) \cdot [\dot{r}, 0]^\intercal, \tag{63}$$

and results in the initial yaw φ and speed v estimate:

$$\varphi = \operatorname{atan2}(\mathbf{v}_W), \tag{64}$$

$$v = \|\mathbf{v}_W\|. \tag{65}$$

The orientation to the detection originating from the sensor $\Theta(\mathbf{y})$ is provided by equation (17). The radar-facing side of the object is now determined and placed in the position of the detection. If no additional information on the lateral position is available, then the center of the side can be simply placed on the detection position. Although only one component of the velocity vector of the target can be measured, this approach reduces the initialization time significantly. If further detections are measured, then a complete velocity vector can be obtained [22].

We also use the radial speed measurement to determine if an unassociated detection belongs to a dynamic object. If the longitudinal component of $\mathbf{v}_S$, regarded in the sensor frame S, plus the radial speed measurement $\dot{r}$ is above a noise-dependent threshold, a new object hypothesis is created.

2) Expected Number of Detections: Track existence checks require the expected number of detections N that an object presumably generates. This number corresponds to the sum of the detection occurrence likelihoods of all components:

$$N = \sum_{j \in J} \mathrm{o}^{(j)}. \tag{66}$$

3) Detection-to-Track Association Likelihood: The crucial problem multi-object trackers deal with is the association of a detection to multiple plausible objects. This process demands the association likelihoods of one detection to all of these objects. The association likelihood for a complete object γ and a given detection measurement $\mathbf{y}$ is the sum of the association likelihoods of all components for this detection:

$$\gamma(\mathbf{y}) = \sum_{j \in J} \gamma^{(j)}(\mathbf{y}). \tag{67}$$

The radial speed measurement appears here as a valuable support for the detection-to-track association.

X. DISCUSSION

The focus of this work is the both precise and fast filtering of sparse radar detections. The problem is split into two parts: an accurate modeling and prediction of measurement data, and the respective state update using this representation. The usage of a set of components to model the measurements is conformable and predicts all relevant measurement effects. The noise of the measurement data does not justify any further particularities in our case. The state update is performed with a probabilistic association approach, which shows robust results and demands far less computing resources than combinatorial approaches. The computing time is a crucial factor as the upstream track management itself usually utilizes multi-hypotheses approaches. A single Intel i7-4790k (2014) core performs a typical complete state update in MATLAB within 2 ms, and in automatically generated C within 20 µs.

However, a difficulty results from the *base point error*. Besides the measurement matrices, also the visibility computations rely on the current state estimate. Estimates always differ from the true state. Assuming an example vehicle that is observed from the front, slight changes in the yaw estimate predict either a good visibility of the left or the right corner of the vehicle (similar to the actual physical characteristic). A workaround is to exploit multiple samplings and evaluations of the measurement functions in relation to the uncertainty of the estimate. However, this effect plays a minor role and workarounds are not necessarily required according to our findings: If a component is wrongly assumed to have just become visible, then its estimated angular extent is still small and invokes a weak, thus insignificant detection likelihood.

A major benefit is the inherent dealing with the *error in variables*. The filter considers all the uncertainties of both the state and measurements, as well as all possible sources of origin. This is in contrast to models, which assume that the regression variables can be determined in an exact manner. For example, most radial functions determine and apply the difference between measurement and contour solely in radial direction and ignore any measurement errors in the tangential direction. Dedicated precautions [26], [40] have to be applied to circumvent this issue.

To use this approach for different objects, component descriptions need to be modeled and parametrized statistically. Machine learning approaches, which derive the object characteristics from measurement data, can be deployed more quickly and are easier to implement. On the other hand, our proposed algorithm offers a low run time. Its parameterability allows the exchange of sensors without extensively recording new data. As the algorithm is based on physical relations, the model is scalable for any desired precision. The behaviour of the algorithm is deterministic and predictable. It can be modularized and extended with interfaces for a low-level fusion.

XI. CONCLUSION

This work deals with the filtering of sparse measurement data. When techniques like feature recognition in measurement data are no longer conceivable, the measurements need to be associated with their estimated origins that are derived from the state vector. We split the object into a set of different components with different measurement characteristics to compute the association more robustly. Association uncertainties are resolved probabilistically to achieve a low computing expense. This approach shows remarkable results in the prediction of measurements according to real-data comparisons. It also places a low demand on computing resources. A tracking evaluation proves the possibility of robust tracking with a low number of measurements. This low number of detections would usually not allow an accurate object state estimation with a single measurement epoch or a even few of them—the estimation is achieved by the filtering over time. The proposed algorithm is developed to be utilized in interaction with a multihypotheses track management and a heterogeneous low-level sensor fusion.

XII. FUTURE WORK

This article outlines the usage of the probabilistic component association with sparse radar measurement data and applies it for vehicle tracking. Further works can focus on other objects like pedestrians, cyclists, and trucks. After an evaluation of the usage of camera and LiDAR sensors, the algorithm can be extended with heterogeneous sensor data fusion. With the availability of radar sensors with an adequate elevation measurement performance, the component descriptions can be extended to 3D models.

The extension of the PDAF adaption to a joint probabilistic data association filter (JPDAF) variant might further improve the performance when obtaining dense measurement data [17].

A statistical study on the structure of objects and/or their measurement characteristics based on a large sample with spatial and kinematic reference data can improve the generality of the component descriptions.

Our current work, though, continues the proposed approach to support (and exploit) mutual occlusion.

ACKNOWLEDGMENT

The authors would like to thank Dominik Kellner, Andreas Rauch, Dominik Bauch, Marco Baumgartl, and colleagues at the BMW group for their support and constructive collaboration.

REFERENCES

[1] Y. Bar-Shalom, F. Daum, and J. Huang
"The probabilistic data association filter,"
IEEE Control Syst. Mag., vol. 29, no. 6, pp. 82–100, Dec. 2009.

[2] P. Berthold, M. Michaelis, T. Luettel, D. Meissner, and H.-J. Wuensche
"Radar reflection characteristics of vehicles for contour and feature estimation,"
in *Proc. Sensor Data Fusion: Trends, Solutions, Appl.*, 2017, pp. 1–6.

[3] P. Berthold, M. Michaelis, T. Luettel, D. Meissner, and H.-J. Wuensche
"An abstracted radar measurement model for extended object tracking,"
in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2018, pp. 770–76.

[4] P. Berthold, M. Michaelis, T. Luettel, D. Meissner, and H.-J. Wuensche
"A radar measurement model for extended object tracking in dynamic scenarios,"
in *Proc. IEEE Intell. Veh. Symp.*, 2019, pp. 770–776.

[5] P. Berthold, M. Michaelis, T. Luettel, D. Meissner, and H.-J. Wuensche
"A continuous probabilistic origin association filter for extended object tracking,"
in *Proc. IEEE Int. Conf. Multisensor Fusion and Integration*, 2020, pp. 323–329.

[6] S. Bordonaro, P. Willett, and Y. Bar-Shalom
"Consistent linear tracker with converted range, bearing, and range rate measurements,"
IEEE Trans. Aerosp. Electron. Syst., vol. 53, no. 6, pp. 3135–3149, Dec. 2017.

[7] S. Bordonaro, P. Willett, Y. Bar-Shalom, M. Baum, and T. Luginbuhl
"Extended object tracking with exploitation of range rate measurements,"
ISIF J. Adv. Inf. Fusion, vol. 12, pp. 228–242, Dec. 2017.

[8] P. Burger, B. Naujoks, and H.-J. Wuensche
"Fast dual-decomposition based mesh-graph clustering for point clouds,"
in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2018, pp. 1129–1135.

[9] M. Bühren and B. Yang
"Automotive radar target list simulation based on reflection center representation of objects,"
in *Proc. Int. Workshop Intell. Transp.*, 2006, pp. 161–166.

[10] J. Ebert, T. Gumpp, S. Münzner, A. Matskevych, A. P. Condurache, and C. Gläser
"Deep radar sensor models for accurate and robust object tracking,"
in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2020, pp. 1–6.

[11] B. Forkel and H.-J. Wuensche
"LiDAR-SGM: Semi-global matching on LiDAR point clouds and their cost-based fusion into stereo matching,"
in *Proc. IEEE Int. Conf. Robot. Automat.*, 2023, pp. 1–7.

[12] K. Gilholm, S. Godsill, S. Maskell, and D. Salmond
"Poisson models for extended target and group tracking,"
Proc. SPIE, vol. 5913, pp. 230–241, Sep. 2005.

[13] K. Gilholm and D. Salmond
"Spatial distribution model for tracking extended objects,"
IEE Proc. Radar, Sonar Navigation, vol. 152, no. 5, pp. 364–371, Oct. 2005.

[14] K. Granström, M. Baum, and S. Reuter
"Extended object tracking: Introduction, overview and applications,"
ISIF J. Adv. Inf. Fusion, vol. 12, no. 2, pp. 139–174, Dec. 2017.

[15] D. Gubelli, O. A. Krasnov, and O. Yarovyi
"Ray-tracing simulator for radar signals propagation in radar networks,"
in *Proc. 10th Eur. Radar Conf.*, 2013, pp. 73–76.

[16] L. Hammarstrand, M. Lundgren, and L. Svensson
"Adaptive radar sensor model for tracking structured extended objects,"
IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 3, pp. 1975–1995, Jul. 2012.

[17] L. Hammarstrand, L. Svensson, F. Sandblom, and J. Sorstedt
"Extended object tracking using a radar resolution model,"
IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 3, pp. 2371–2386, Jul. 2012.

[18] P. Held, D. Steinhauser, A. Kamann, A. Koch, T. Brandmeier, and U. T. Schwarz
"Micro-Doppler extraction of bicycle pedaling movements using automotive radar,"
in *Proc. IEEE Intell. Veh. Symp.*, 2019, pp. 744–749.

[19] J. Honer and H. Kaulbersch
"Bayesian extended target tracking with automotive radar using learned spatial distribution models,"
in *Proc. IEEE Int. Conf. Multisensor Fusion Integration*, 2020, pp. 316–322.

[20] A. Kamann, D. Steinhauser, F. Gruson, T. Brandmeier, and U. T. Schwarz
"Extended object tracking using spatially resolved micro-Doppler signatures,"
IEEE Trans. Intell. Veh., vol. 6, no. 3, pp. 440–449, Sep. 2021.

[21] H. Kaulbersch, J. Honer, and M. Baum
"EM-based extended target tracking with automotive radar using learned spatial distribution models,"
in *Proc. Int. Conf. Inf. Fusion*, 2019, pp. 1–8.

[22] D. Kellner, M. Barjenbruch, K. Dietmayer, J. Klappstein, and J. Dickmann
"Instantaneous lateral velocity estimation of a vehicle using Doppler radar,"
in *Proc. Int. Conf. Inf. Fusion*, 2013, pp. 877–884.

[23] C. Knill, A. Scheel, and K. Dietmayer
"A direct scattering model for tracking vehicles with high-resolution radars,"
in *Proc. IEEE Intell. Veh. Symp.*, 2016, pp. 298–303.

[24] 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.

[25] A. I. Leonov and K. I. Fomichev
Monopulse Radar. Norwood, MA: Artech House, 1986.

[26] M. Michaelis, P. Berthold, T. Luettel, D. Meissner, and H.-J. Wuensche
"Extended object tracking with an improved measurement-to-contour association,"
in *Proc. Int. Conf. Inf. Fusion*, 2020, pp. 1–6.

[27] 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.*, 2017, pp. 1–6.

[28] B. Naujoks, P. Burger, and H.-J. Wuensche
"The greedy dirichlet process filter—An online clustering multi-target tracker,"
in *Proc. IEEE Glob. Conf. Signal Inf. Process.*, 2018, pp. 1233–1237.

[29] H. Rohling
"Radar CFAR thresholding in clutter and multiple target situations,"
IEEE Trans. Aerosp. Electron. Syst., AES, vol. 19, no. 4, pp. 608–621, Aug. 1983.

[30] A. Scheel and K. Dietmayer
"Tracking multiple vehicles using a variational radar model,"
IEEE Trans. Intell. Transp. Syst., vol. 20, no. 10, pp. 3721–3736, Oct. 2019.

[31] R. Schubert, E. Richter, and G. Wanielik
"Comparison and evaluation of advanced motion models for vehicle tracking,"
in *Proc. Int. Conf. Inf. Fusion*, 2008, pp. 1–6.

[32] D. Steinhauser, P. Held, A. Kamann, A. Koch, and T. Brandmeier
"Micro-Doppler extraction of pedestrian limbs for high resolution automotive radar,"
in *Proc. IEEE Intell. Veh. Symp.*, 2019, pp. 764–769.

[33] R. L. Streit and T. E. Luginbuhl
"Probabilistic multi-hypothesis tracking," Naval Undersea Warfare Center Division, Newport, RI, USA, Tech. Rep. ADA298501, Feb. 1995.

[34] D. Svensson
"Derivation of the discrete-time constant turn rate and acceleration motion model,"
in *Proc. Sensor Data Fusion: Trends, Solutions, Appl.*, 2019, pp. 1–5.

[35] N. Wahlström and E. Özkan
"Extended target tracking using Gaussian processes,"
IEEE Trans. Signal Process., vol. 63, no. 16, pp. 4165–4178, Aug. 2015.

[36] T. A. Wheeler, M. Holder, H. Winner, and M. Kochenderfer
"Deep stochastic radar models,"
in *Proc. IEEE Intell. Veh. Symp.*, 2017, pp. 47–53.

[37] Y. Xia, P. Wang, K. Berntorp, L. Svensson, K. Granström, H. Mansour, P. Boufounos, and P. V. Orlik
"Learning-based extended object tracking using hierarchical truncation measurement model with automotive radar,"
IEEE J. Sel. Topics Signal Process., vol. 15, no. 4, pp. 1013–1029, Jun. 2021.

[38] D. Zhang and G. Lu
"Study and evaluation of different fourier methods for image retrieval,"
Image Vis. Comput., vol. 23, no. 1, pp. 33–49, Nov. 2005.

[39] L. Zhang, H. Jiang, Y. Zhou, Y. Zhai, D. Zhou, and Y. Bai
"Wheeled vehicle detection under the asymmetric micro-Doppler interference,"
IEEE Sensors J., vol. 22, no. 18, pp. 18085–18092, Sep. 2022.

[40] S. Zhou, D. Pati, T. Wang, Y. Yang, and R. J. Carroll
"Gaussian processes with errors in variables: Theory and computation,"
arXiv: Statistics Theory, 2019.

[41] Y. Zhou, L. Liu, H. Zhao, M. López-Benítez, L. Yu, and Y. Yue
"Towards deep radar perception for autonomous driving: Datasets, methods, and challenges,"
Sensors, vol. 22, no. 11, Art. no. 4208, 2022.

Philipp Berthold studied electrical engineering at the Technical University of Munich (TUM), München, Germany. He has actively developed autonomous systems and localization techniques. Since 2016, he has been a Research Assistant at the Institute for Autonomous Systems Technology at the University of the Bundeswehr Munich. His current research involves heterogeneous multisensor data fusion for automated driving with focus on radar perception.

Martin Michaelis studied mathematics at the University of Bonn, Bonn, Germany. Afterwards, he worked for two years as a Research Assistantt at the Fraunhofer Institute FKIE, Wachtberg, Germany. Since 2016, he has been with the University of the Bundeswehr Munich. His main research interests are multiobject tracking, multisensor fusion, automotive radar perception, and extended target tracking.

Thorsten Luettel studied electrical engineering with a focus on mechatronics at Leibniz Universität Hannover, Hannover, Germany, where he worked primarily on bipedal walking robots. Since 2006, he has been conducting research in the field of autonomous driving at the Institute for Autonomous Systems Technology at the University of the Bundeswehr Munich, Neubiberg, Germany. His main areas of interest are sensor data fusion and system integration. As a research associate, he has led several externally funded projects in recent years as well as led "Team MuCAR" to success in international competitions. Since 2020, he also leads the research area "Networking and Autonomy" in the dtec.bw-funded project MORE (Munich Mobility Research Campus).

Daniel Meissner was born in 1984 in Ansbach, Germany. He received a Diploma degree (equivalent to M.Sc. degree) in electrical engineering from the Technical University of Munich (TUM), München, Germany, in 2009, and the Dr.-Ing. degree (equivalent to Ph.D.) in 2015 from Ulm University, Ulm, Germany. In 2014, he joined the BMW Group as a developer for driver assistance systems. In 2021, he became the Chief Engineer for object and freespace detection.

Hans-Joachim Wuensche got his Ph.D. from the University of the Bundeswehr Munich, Neubiberg, Germany, in 1987 with Ernst D. Dickmanns, where he codeveloped the 4D-approach to computer vision. After many years in management, he returned to the same university to lead the Institute for Autonomous Systems Technology in 2004. His research interests include autonomous robots, especially on- and off-road vehicles exploring and navigating unknown terrain.