Track-Before-Detect Labeled Multi-Bernoulli Smoothing for Multiple Extended Objects

Boqian Yu
Department of Informatics,
Technical University of Munich
Garching, Germany
boqian.yu@tum.de

Egon Ye
BMW Group
Unterschleissheim, Germany
egon.ye@bmw.de

Abstract—For the evaluation of autonomous driving systems, this paper provides a new approach of generating reference data for multiple extended object tracking. In our approach, we apply a forward-backward smoother for objects with star-convex shapes based on the Labeled Multi-Bernoulli (LMB) Random Finite Set (RFS) and recursive Gaussian processes. We further propose to combine a robust birth policy with a backward filter to solve the conflict between robustness and completeness of tracking. Thereby, cluster candidates are evaluated based on a quality measure to only initialize objects from more reliable clusters in the forward pass. Missing states will then be recovered by the backward filter through post-processing the unassociated data after the smoothing process. Simulations and real-world experiments demonstrate superior performance of the proposed method in both cardinality and individual state estimation compared to naive LMB filter and smoother for extended objects.

Index Terms—ground-truth generation, track-before-detect, multi-object tracking, random finite set, extended object modeling, Gaussian process, forward-backward smoothing

I. Introduction

Aiming to improve both safety and comfort, autonomous driving systems have gained increasing attention in many relevant research and industry fields. For validation and evaluation of autonomous driving systems, novel methodologies such as evaluation without ground-truth [1], [2] and automatic ground-truth generation [3], [4] have been investigated in the recent years to cope with the shortage of ground-truth data [5]. In this paper, a new method is proposed to generate more reliable reference data that can be used as ground truth: based on LiDAR data, we perform offline forward-backward Bayesian tracking for multiple extended objects, which benefits from simultaneous estimation of kinematics and object contours and information fusion of past, current, and future sensor data.

The objective of multi-object tracking is to jointly estimate the amount of existing tracks as well as their individual states from a sequence of noisy and cluttered recordings. Apart from temporary occlusion, the substantial difference between multi-object and single-object tracking is the varying amount of tracks due to object births and deaths [6], [7]. Algorithms for multi-object tracking can be classified into three paradigms according to the applied data association appraoch [8]: Joint Probabilistic Data Association methods [6], [9] assign weights to individual observations due to the association probabilities;

Multiple Hypothesis Tracking methods [10], [11] propagate association hypotheses in time; and the Random Finite Set (RFS) framework [7] involves modeling the multi-object states as a whole.

The fundamental methodology of the RFS framework is to model the collection of individual object states, including object births, deaths, occlusion and clutter, as finite set-valued random variables, and cast the data association problem in the classic Bayesian inference framework using Finite Set Statistics (FISST) developed in [7]. Tracking via propagating posterior densities using the RFS framework has gained popularity in recent years, of which the Probability Hypothesis Density filter [12], [13], the Multi-Target Multi-Bernoulli filter [7], [14], the Generalized Labeled Multi-Bernoulli (GLMB) filter [15] and its first moment approximation LMB filter [16] have been widely applied [17]–[19]. The formulae of the corresponding RFS-based smoothers have also been developed in the literature [20]–[23].

The previously-mentioned approaches model objects as single points without extents; however, one is often additionally interested in estimating the shape of objects. Objects that can occupy multiple resolution cells of the sensor are denoted as extended objects, for which trackers have been developed by integrating extended object models. In some simple cases, object contours can be modeled as basic shapes like circles [24], [25] or rectangles [26], [27]. The random matrix model is frequently applied to estimate elliptical shapes [28]–[30], while the random hypersurface model proposed in [31] is flexible to model various kinds of contours such as elliptical or star-convex [32], [33]. The recursive Gaussian process proposed in [34] can efficiently propagate shapes approximated by certain basis vectors [35], [36].

In most object tracking scenarios, measurement extraction involves some thresholding operation to prevent false tracks, which could in turn result in loss of information when the signal-to-noise ratio is low. A higher threshold usually leads to more robust estimations yet less complete tracks. The term *track-before-detect* refers to tracking approaches that process all sensor data without applying a threshold [37], and is typically used for object tracking on radar intensity images [38]. This term can be generalized to describe approaches based on LiDAR data [39], where the number of measurements

decreases as the object gets further from the sensor.

The contribution of this paper is as follows: first, we integrate the star-convex shape model into forward-backward LMB smoothing, enabling offline tracking for multiple extended objects to improve kinematics and shape estimations; second, we propose a backward filter process to generate track-before-detect estimations for objects approaching from remote positions. By applying robust birth and post-processing the unassociated data, our proposed method solves the conflict between robustness of extended object estimation and completeness of track. We evaluate the performance of the proposed method in both simulations and real-world experiments to demonstrate its capability of ground-truth generation.

II. PRELIMINARIES

This section provides a brief review of basic theories on tracking with LMB-RFS and extended object modeling, which provide the foundation of the method proposed in this paper.

A. Notations

In this paper, scalars are denoted with regular small letters (e.g., x), vectors with underlined small letters (e.g., $\underline{x}$), sets with regular capital letters (e.g., X, Z), and spaces with blackboard bold letters (e.g., $\mathbb{X}, \mathbb{L}$). Bold face letters (e.g., $\boldsymbol{x}, \boldsymbol{X}$) are used for labeled variables (see section II-B) to distinguish from unlabeled ones.

The multi-object exponential for a real-valued function h on a finite set X is denoted by $h^X \triangleq \prod_{\underline{x} \in X} h(\underline{x})$ with $h^\emptyset = 1$. The inner product is denoted by

$$\langle f, g \rangle \triangleq \int f(x) g(x) dx. \tag{1}$$

The inclusion function is given by

$$1_Y(X) \triangleq \begin{cases} 1, & \text{if } X \subseteq Y, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

If $X = \{x\}$, the notation $1_Y(x)$ is used for brevity.

B. Labeled Multi-Bernoulli RFS

A labeled RFS $\boldsymbol{X}$ defined on the state space $\mathbb{X}$ and discrete label space $\mathbb{L}$ has only distinct labels within each realization, i.e., let $\mathcal{L}(\boldsymbol{X}) = \{\ell : (\underline{x}, \ell) \in \boldsymbol{X}\}$ denote the set of labels in $\boldsymbol{X}$, the distinct label indicator

$$\Delta(\boldsymbol{X}) = \begin{cases} 1, & \text{if } |\mathcal{L}(\boldsymbol{X})| = |\boldsymbol{X}|, \\ 0, & \text{if } |\mathcal{L}(\boldsymbol{X})| \neq |\boldsymbol{X}| \end{cases} \tag{3}$$

always satisfies $\Delta(\boldsymbol{X}) = 1$.

A labeled multi-Bernoulli (LMB) RFS is a multi-Bernoulli RFS with distinct labels attached to non-empty Bernoulli components, defined by the parameter set

$$\pi = \left\{ \left(r^{(\ell)}, p^{(\ell)} \right) \right\}_{\ell \in \mathbb{L}}, \tag{4}$$

where $r^{(\ell)}$ is the existence probability and $p^{(\ell)}$ the probability density of the component with label $\ell \in \mathbb{L}$. The density of an LMB RFS is then given by [15]

$$\pi(\boldsymbol{X}) = \Delta(\boldsymbol{X}) w(\mathcal{L}(\boldsymbol{X})) p^{\boldsymbol{X}}, \tag{5}$$

where

$$w(L) = \prod_{i \in \mathbb{L}} \left(1 - r^{(i)} \right) \prod_{\ell \in L} \frac{1_{\mathbb{L}}(\ell) r^{(\ell)}}{1 - r^{(\ell)}}, \tag{6}$$

$$p(\underline{x}, \ell) = p^{(\ell)}(\underline{x}). \tag{7}$$

A generalized labeled multi-Bernoulli (GLMB) RFS is a labeled RFS distributed according to [15]:

$$\pi(\boldsymbol{X}) = \Delta(\boldsymbol{X}) \sum_{\xi \in \Xi} w^{(\xi)}(\mathcal{L}(\boldsymbol{X})) \left[p^{(\xi)} \right]^{\boldsymbol{X}}, \tag{8}$$

where Ξ is a discrete index set. The weights $w^{(\xi)}(L)$ and the spatial distributions $p^{(\xi)}$ satisfy the normalization conditions

$$\sum_{L \subseteq \mathbb{L}} \sum_{\xi \in \Xi} w^{(\xi)}(L) = 1, \quad \int_{\underline{x} \in \mathbb{X}} p^{(\xi)}(\underline{x}, \ell) d\underline{x} = 1. \tag{9}$$

It can be observed from (5) and (8) that an LMB is a special case of a GLMB (with one single component in Ξ).

C. Extended Object Modeling

1) Star-Convex: A shape $\mathcal{S} \subset \mathbb{R}^2$ is a star-convex, if there exists one reference point $\underline{x}_c$ such that for every $\underline{x} \in \mathcal{S}$, the line segment between $\underline{x}_c$ and $\underline{x}$ is fully contained in $\mathcal{S}$ [32]. With known reference point, the contour of a star-convex can be fully described by a radial function $\rho = f(\vartheta)$ in polar coordinates that maps angles to radii, i.e.,

$$\overline{\mathcal{S}} = \left\{ \underline{x}_c + \underline{p}(\vartheta) f(\vartheta), \ \vartheta \in (-\pi, \pi] \right\}, \tag{10}$$

where $\overline{\mathcal{S}}$ is the contour of $\mathcal{S}$ and $\underline{p}(\vartheta) \triangleq [\cos(\vartheta) \ \sin(\vartheta)]^\top$ the orientation vector for angle ϑ.

2) Gaussian Process: A Gaussian process (GP) [40]

$$f(u) \sim \mathcal{GP}\left(\mu(u), k\left(u, u'\right)\right) \tag{11}$$

is a stochastic process uniquely defined by its mean $\mu(u)$ and covariance (kernel) function $k\left(u, u'\right)$ of a multivariate Gaussian distribution

$$\begin{bmatrix} f(u_1) \\ \vdots \\ f(u_n) \end{bmatrix} \sim \mathcal{N}\left(\underline{\mu}, K\right), \text{ where } \underline{\mu} = \begin{bmatrix} \mu(u_1) \\ \vdots \\ \mu(u_n) \end{bmatrix},$$

$$\text{and } K = \begin{bmatrix} k(u_1, u_1) & \dots & k(u_1, u_n) \\ \vdots & \ddots & \vdots \\ k(u_n, u_1) & \dots & k(u_n, u_n) \end{bmatrix}. \tag{12}$$

Given measurements $\underline{z} = [z_1 \cdots z_N]^\top$ and their corresponding input $\underline{u} = [u_1 \cdots u_N]^\top$, a GP can be used to learn function values $\underline{f} = [f_1 \cdots f_{N_f}]^\top$ at certain points $\underline{u}^f = [u_1^f \cdots u_{N_f}^f]^\top$. Assuming the mean function $\underline{\mu} = \underline{0}$, the joint distribution of $\underline{z}$ and $\underline{f}$ is given by

$$\begin{bmatrix} \underline{z} \\ \underline{f} \end{bmatrix} \sim \mathcal{N}\left(\underline{0}, \begin{bmatrix} K(\underline{u}, \underline{u}) & K(\underline{u}, \underline{u}^f) \\ K(\underline{u}^f, \underline{u}) & K(\underline{u}^f, \underline{u}^f) \end{bmatrix}\right), \tag{13}$$

which yields

$$p(\underline{f}|\underline{z}) = \mathcal{N}\left(\underline{m}^f, P^f\right), \tag{14}$$

with

$$\begin{aligned} \underline{m}^f &= K\left(\underline{u}^f, \underline{u}\right) K^{-1}(\underline{u}, \underline{u}) \underline{z}, \\ P^f &= K\left(\underline{u}^f, \underline{u}^f\right) - K\left(\underline{u}^f, \underline{u}\right) K^{-1}(\underline{u}, \underline{u}) K(\underline{u}, \underline{u}^f). \end{aligned} \tag{15}$$

III. ROBUST LMB SMOOTHER FOR EXTENDED OBJECTS

A. Extended State Vector

Typically, while the orientation of road objects (vehicles, cyclists, pedestrians) can alter with time, the velocity and acceleration exist almost only in the direction of their orientation. To better model the motion of typical road objects, the *Constant Turn Rate and Acceleration* (CTRA) model with white noise [41] is used in this paper to describe the kinematics of objects, i.e., individual kinematic state can be defined by position, velocity, orientation, turn rate and acceleration as:

$$\underline{x}_k = \begin{bmatrix} x & y & v & \psi & \omega & a \end{bmatrix}^\top, \tag{16}$$

whose transition follows from

$$\underline{\dot{x}}_k = \begin{bmatrix} v\cos\psi & v\sin\psi & a & \omega & 0 & 0 \end{bmatrix}^\top + \underline{q} \tag{17}$$

with process noise $\underline{q} \sim \mathcal{N}(\underline{0}, \mathrm{diag}(0,0,0,0,\dot{\sigma}_\omega^2, \dot{\sigma}_a^2))$.

By modeling an object contour as star-convex, it is fully defined by its radial function that can be approximated by function values $\underline{\rho}^s = [\rho_1^s, \cdots, \rho_{N_s}^s]^\top$ at a certain set of angles $\underline{\vartheta}^s = [\vartheta_1^s, \cdots, \vartheta_{N_s}^s]^\top$. In this way, contours can be efficiently inferred over time by recursive Gaussian processes [34], [35]. The exponential sine square kernel with period 2π [42]

$$k(\vartheta, \vartheta') = \sigma_f^2 e^{-\frac{2\sin^2\left(\frac{|\vartheta - \vartheta'|}{2}\right)}{l^2}} \tag{18}$$

is applied for the Gaussian process to guarantee smoothness and closure of contour estimations while assuming no other prior knowledge about the shape to maximize the flexibility.

Combined together, an individual extended object can be fully described by its extended state vector $\underline{x} = \left[\underline{x}_k^\top \ (\underline{\rho}^s)^\top\right]^\top$.

B. Forward Pass

The forward computation of LMB smoother consists of two steps: the prediction step incorporates newborn objects and surviving objects from the last time step; the update step adjusts the predictions with the current measurement set.

1) Prediction: Suppose the posterior multi-object density on state space $\mathbb{X}$ and label space $\mathbb{L}$ is given by the parameter set $\pi = \left\{(r^{(\ell)}, p^{(\ell)})\right\}_{\ell \in \mathbb{L}}$ from the last time step, and the multi-object birth model on $\mathbb{X} \times \mathbb{B}$ is given by $\pi_B = \left\{\left(r_B^{(\ell)}, p_B^{(\ell)}\right)\right\}_{\ell \in \mathbb{B}}$. The predicted multi-object density is also an LMB RFS on the state space $\mathbb{X}$ and the label space $\mathbb{L}_+ = \mathbb{L} \cup \mathbb{B}$ defined by the parameter set [16]

$$\pi_+ = \left\{\left(r_{+,S}^{(\ell)}, p_{+,S}^{(\ell)}\right)\right\}_{\ell \in \mathbb{L}} \cup \left\{\left(r_B^{(\ell)}, p_B^{(\ell)}\right)\right\}_{\ell \in \mathbb{B}} \tag{19}$$

where

$$r_{+,S}^{(\ell)} = \eta_S(\ell) r^{(\ell)},$$
$$p_{+,S}(\underline{x}, \ell) = \frac{\langle p_{S,k|k-1}(\cdot, \ell) f(\underline{x}|\cdot, \ell), p(\cdot, \ell)\rangle}{\eta_S(\ell)}, \tag{20}$$
$$\eta_S(\ell) = \langle p_{S,k|k-1}(\cdot, \ell), p(\cdot, \ell)\rangle.$$

$p_{S,k|k-1}(x, \ell)$ is the state-dependent survival probability from $k-1$ to k, $\eta_S(\ell)$ the survival probability of a certain track ℓ, and $f(x|x', \ell)$ the single track transition density.
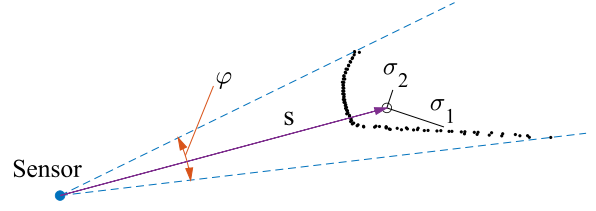

Fig. 1: Factors influencing the suitability of a cluster regarding birthing: measurement density (calculated by cardinality and azimuth angle range φ), mean distance s to the sensor, and shape ratio (represented by PCA variances σ_1^2, σ_2^2).

In the prediction step, a dynamic birth model π_B is realized by a consecutive process resembling the adaptive birth density in [16], i.e., clusters not assigned to any track are likely to correspond to new objects. Registration is applied between previously unassigned clusters from two different time steps, and new objects are initialized by the pairs yielding high shape similarity and valid kinematics.

Unlike tracking point objects, tracking extended objects requires a cluster of measurements with enough information to additionally yield robust initial contour estimation. This requirement is typically not fulfilled in sensor data for objects approaching from distant positions, leading to inaccurate initialization and negative influence on the subsequent inference processes. Thus, in this paper, we introduce a measure for the suitability of a cluster regarding birthing by three factors: its cardinality, its mean distance s to the sensor, and its shape ratio $V = \frac{\sigma_2}{\sigma_1}$, where $\sigma_1^2 \geq \sigma_2^2$ are the two variances from the principle component analysis (PCA) of cluster points. A *robust birth policy* then adjusts the existence probability based on the horizontal angular resolution D of the sensor and a constant parameter $\alpha > 0$, i.e.,

$$r_B^{\prime(\ell)} = r_B^{(\ell)} \cdot \left(t \cdot \frac{D \cdot |\mathcal{U}_{\theta(\ell)}(Z)|}{\varphi\left(\mathcal{U}_{\theta(\ell)}(Z)\right)} + (1-t) \cdot V\left(\mathcal{U}_{\theta(\ell)}(Z)\right)\right), \tag{21}$$

where $\mathcal{U}(Z)$ denotes a certain partition of the measurement set Z into clusters, $\theta : \mathbb{L} \to \{0, 1, \cdots, |\mathcal{U}(Z)|\}$ the track-to-cluster association map on label space $\mathbb{L}$, and $\varphi\left(\mathcal{U}_{\theta(\ell)}(Z)\right)$ the azimuth angle range of the cluster from the perspective of the sensor. $t = e^{-\alpha s(\mathcal{U}_{\theta(\ell)}(Z))}$ controls the weights of both terms based on the cluster's mean distance to the sensor, of which the first term rewards clusters with a high measurement density and the second term the clusters with measurements from different sides of an object. As we prefer possibly dense clusters with measurements over more dimensions of the object contour, the observation of a single side can still be sufficient if the cluster is close to the sensor (the ego-vehicle), e.g., to account for car-following maneuvers.

Higher quality of initialization can be achieved in such manner at the cost of losing some early states due to insignificant existence probability of certain candidates, which will be compensated later in the proposed backward filtering process.

2) Extended Object Likelihood: The LMB filter for extended objects differs from the classic one mainly in the partition-based extended object likelihood in the update step. Let $p_D(\underline{x}, \ell)$ be the state-dependent detection probability, κ the detection-independent clutter density, and $g_C(Z)$ the clutter distribution, the extended object likelihood is given by [30]

$$g(Z|\boldsymbol{X}) = g_C(Z) \sum_{i=1}^{|\boldsymbol{X}|+1} \sum_{\substack{\mathcal{U}(Z) \in \mathcal{P}_i(Z) \\ \theta \in \Theta(\mathcal{U}(Z))}} \left[\psi_{\mathcal{U}(Z)}(\cdot; \theta)\right]^{\boldsymbol{X}}, \quad (22)$$

where

$$\psi_{\mathcal{U}(Z)}(\underline{x}, \ell; \theta) = \begin{cases} \dfrac{p_D(\underline{x}, \ell)\tilde{g}\left(\mathcal{U}_{\theta(\ell)}(Z)|\underline{x}, \ell\right)}{[\kappa]^{\mathcal{U}_{\theta(\ell)}(Z)}}, & \theta(\ell) > 0, \\ 1 - p_D(\underline{x}, \ell), & \theta(\ell) = 0. \end{cases} \quad (23)$$

$\mathcal{P}_i$ is the set of partitions dividing the measurement set Z into exact i clusters, $\Theta(\mathcal{U})$ the space of possible track-to-cluster association maps under partition $\mathcal{U}$, and $\tilde{g}\left(\mathcal{U}_{\theta(\ell)}(Z)|\underline{x}, \ell\right)$ the likelihood for cluster $\mathcal{U}_{\theta(\ell)}(Z)$ given the single state $(\underline{x}, \ell)$.

3) Update: In contrast to GLMB, LMB is not closed under the update operation [15]. Thus, the predicted LMB from (19) is first interpreted as a GLMB with one single component to enable exact measurement updates. To constrain the exponentially increasing number of components, the posterior GLMB density is then approximated by an LMB density that matches its first moment, given by [30]

$$r^{(\ell)} = \sum_{L \subseteq \mathbb{L}_+} \sum_{i=1}^{|L|+1} \sum_{\substack{\mathcal{U}(Z) \in \mathcal{P}_i(Z) \\ \theta \in \Theta(\mathcal{U}(Z))}} 1_L(\ell) w^{(\theta)}_{\mathcal{U}(Z)}(L),$$

$$p^{(\ell)}(\underline{x}) = \frac{1}{r^{(\ell)}} \sum_{L \subseteq \mathbb{L}_+} \sum_{i=1}^{|L|+1} \sum_{\substack{\mathcal{U}(Z) \in \mathcal{P}_i(Z) \\ \theta \in \Theta(\mathcal{U}(Z))}} 1_L(\ell) w^{(\theta)}_{\mathcal{U}(Z)}(L) p^{(\theta)}(\underline{x}, \ell),$$

$$(24)$$

where

$$w^{(\theta)}_{\mathcal{U}(Z)}(L) \propto w_+(L) \left[\eta^{(\theta)}_{\mathcal{U}(Z)}\right]^L,$$

$$p^{(\theta)}(\underline{x}, \ell|\mathcal{U}(Z)) = \frac{p_+(\underline{x}, \ell)\psi_{\mathcal{U}(Z)}(\underline{x}, \ell; \theta)}{\eta^{(\theta)}_{\mathcal{U}(Z)}(\ell)}, \quad (25)$$

$$\eta^{(\theta)}_{\mathcal{U}(Z)}(\ell) = \left\langle p_+(\cdot, \ell), \psi_{\mathcal{U}(Z)}(\cdot, \ell; \theta)\right\rangle.$$

Computing the posterior density in (24) requires the set of all track-to-cluster associations, which is in general intractable due to its combinatorial growth in cardinality. In this paper, the *distance partitioning* algorithm [43] is used to generate only the most likely partitions, which is then pruned by thresholding the likelihood based on the predicted multi-object states (similar to *prediction partitioning* in [29]). A ranked assignment algorithm is also utilized to produce the most substantial track-to-cluster associations based on these partitions [30].

C. Backward Pass

The Bernoulli RFS is closed under the smoothing operation [23]. Moreover, let $\pi_{k|T} = \left\{(r^{(\ell)}_{k|T}, p^{(\ell)}_{k|T})\right\}_{\ell \in \mathbb{L}}$ denote the smoothed LMB density from time T to k where $T \geq k$ and assume no re-entry of track after death, the smoothed density at $k-1$ is also an LMB RFS $\pi_{k-1|T} = \left\{(r^{(\ell)}_{k-1|T}, p^{(\ell)}_{k-1|T})\right\}_{\ell \in \mathbb{L}}$ given by

$$r^{(\ell)}_{k-1|T} = 1 - \frac{\left(1 - r^{(\ell)}_{k-1|k-1}\right)\left(1 - r^{(\ell)}_{k|T}\right)}{\left(1 - r^{(\ell)}_{k|k-1}\right)},$$

$$p^{(\ell)}_{k-1|T}(\underline{x}) = p_{k-1|k-1}(\underline{x}, \ell) \times \left(\alpha_{S,k|T}(\underline{x}, \ell)\right.$$

$$\left. + \beta_{S,k|T}(\underline{x}, \ell) \left\langle f_{k|k-1}(\cdot|\underline{x}, \ell), \frac{p_{k|T}(\cdot, \ell)}{p_{k|k-1}(\cdot, \ell)}\right\rangle\right) / \gamma_{S,k|T}(\ell),$$

$$(26)$$

where

$$\alpha_{S,k|T}(\underline{x}) = \left(1 - p_{S,k|k-1}(\underline{x}, \ell)\right) \frac{\left(1 - r_{k|T}\right)}{\left(1 - r_{k|k-1}\right)},$$

$$\beta_{S,k|T}(\underline{x}) = p_{S,k|k-1}(\underline{x}, \ell) \frac{r_{k|T}}{r_{k|k-1}},$$

$$\gamma_{S,k|T}(\ell) = \left\langle p_{k-1|k-1}(\cdot, \ell), \alpha_{S,k|T}(\cdot, \ell) \right.$$

$$\left. + \beta_{S,k|T}(\cdot, \ell) \left\langle f_{k|k-1}(\cdot|\cdot, \ell), \frac{p_{k|T}(\cdot, \ell)}{p_{k|k-1}(\cdot, \ell)}\right\rangle\right\rangle.$$

$$(27)$$

As Gaussian distributions are assumed for object states, Gaussian mixture densities may be produced in the GLMB-to-LMB approximation from multiple parallel hypotheses in the forward update step [16]. In many real-world applications, however, hypotheses containing wrong cluster-to-object associations (if any, after prediction pruning discussed in section III-B3) will hardly survive due to insignificant likelihood that quickly approaches zero as the deviation from the predicted position increases. This is caused by the high frame rate of LiDAR and the smooth radial function of contour estimation. The mixture reduction algorithm based on Kullback-Leibler divergence [44] is used to guarantee that only the most significant component remains, and thus enables state smoothing in (26) with an extended Rauch-Tung-Striebel smoother [45]. Alternatively, a particle filter can be used as in [23], which requires much more computation yet is assumed to produce limited performance improvement in this case.

IV. TRACK-BEFORE-DETECT GENERATION

As the distance to the sensor grows larger, the number of measurements from the same object usually decreases, and states may be missing in the early time steps of the track for approaching objects due to insignificant existence probabilities of birth candidates from sparser clusters. We propose a backward LMB filter in this paper to generate track-before-detect estimations by further extracting states from unassociated data after forward-backward smoothing, as in the left part of Fig. 2. Note that instead of its original intention for tracking on radar intensity images, we use the term *track-before-detect* to denote the utilization of all recorded LiDAR data in the proposed offline tracking scheme to also produce state estimations before the first detection of an object.
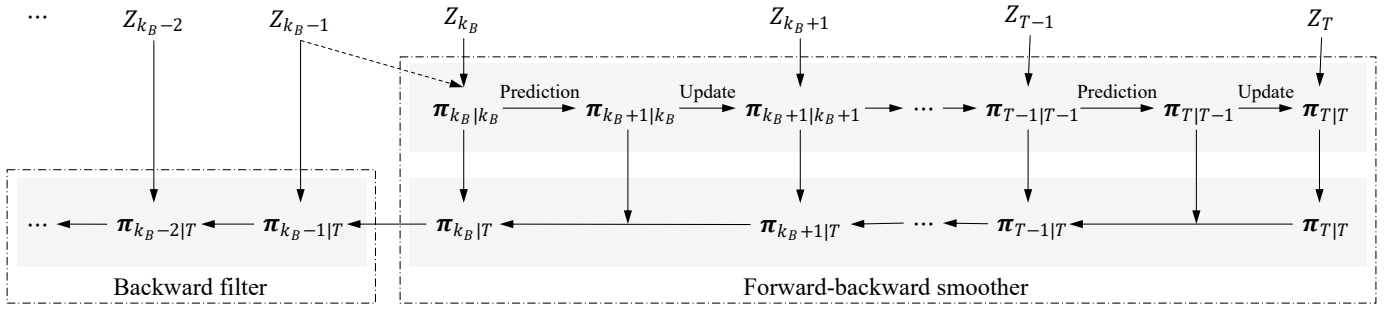
Fig. 2: The proposed method to generate track-before-detect estimations with a backward filter. k_B denotes the original birth time of a certain track in the forward pass, and the dashed arrow denotes the initial estimate of kinematics through cluster registration (as in section III-B1). The possibly missing states at $k_B - 1, k_B - 2, \cdots$ will be recovered by the backward filter.

After the smoothing operation, clusters associated to any extracted state are removed from the original dataset $Z = \{Z_k\}_{k=1,\cdots,T}$ to form a new one composed of measurements with no specific association, i.e., for each time step k with partition $\mathcal{U}_k$ and association map θ_k in estimation extraction,

$$Z'_k = Z_k \setminus \{\mathcal{U}_{\theta_k(\ell)}(Z_k) : \theta_k(\ell) > 0\}_{\ell \in \mathbb{L}}. \tag{28}$$

Such dataset $Z' = \{Z'_k\}_{k=1,\cdots,T}$ corresponds to measurements from untracked states or clutter. A modified LMB filter then starts from the last time step T and runs on Z' backward in time to search for missing states in each track. Instead of performing adaptive birth as in the forward pass (see section III-B1), the state of track ℓ is initialized automatically at its original birth time $k_B(\ell)$ in the forward pass with identical contour and reversed kinematics as from smoothing, i.e.,

$$\tilde{\pi}_B^{(\ell)} = \left(r_{k_B(\ell)|T}^{(\ell)}, \tilde{p}_{k_B(\ell)|T}^{(\ell)} \right), \tag{29}$$

where $\tilde{p}_{k_B(\ell)|T}^{(\ell)} \sim \mathcal{N}\left(A_R \underline{x}_{k_B(\ell)|T}^{(\ell)}, \Sigma_{k_B(\ell)|T}^{(\ell)} \right)$, $\underline{x}_{k_B(\ell)|T}^{(\ell)}$ and $\Sigma_{k_B(\ell)|T}^{(\ell)}$ denote the smoothed mean and covariance matrix of track ℓ at $k_B(\ell)$, and the matrix $A_R = \text{diag}(1, 1, -1, 1, -1, -1, I_{N_s})$ reverses the direction of velocity, turn rate and acceleration of the extended state vector. The state and covariance matrix of each track are then propagated backward by the filter in the same two-step manner as described in section III-B until death, which corresponds (or at least closer) to its true birth in the original dataset.

V. PERFORMANCE EVALUATION

In this section, the evaluation of the proposed method is performed in two simulations and one real-world experiment to verify its effectiveness in different aspects. In the first simulation, the proposed method is applied to a single maneuvering object to investigate its capability in simultaneously estimating kinematics and contour. In the second simulation, multiple objects perform typical traffic behaviors like turning, following, paralleling, etc. to further illustrate the method's performance in complicated situations and track-before-detect generation. In the real-world scenario, a vehicle track is extracted from LiDAR data using the proposed method and compared to the

corresponding Differential Global Positioning System (DGPS) recordings to demonstrate the practical capability.

In the simulations, objects are abstracted as rectangles of various geometries, and data points are generated as if coming from a one-layer LiDAR at $(0,0)$ with $360°$ field of view, horizontal angular resolution $D = \frac{1}{6}$ deg and sampling interval $\Delta T = 0.1s$. The surveillance region is the square $[-80, 80] \times [-80, 80]$ (units in meter), corresponding to the typical sensing range of a LiDAR, and the number of uniformly distributed clutter follows Poisson distribution with $\lambda = 15$. Process noises for acceleration and turn rate are assumed to be zero-mean Gaussian with $\dot{\sigma}_a = 3m/s^3, \dot{\sigma}_\omega = 20°/s^2$, which resemble the normal kinematics of road vehicles. Similarly, $\sigma_M = 5cm$ is set as measurement noise for each single LiDAR point. The survival and detection probabilities are set as $p_S = 0.99$, $p_D = 0.9$ for both forward and backward passes at all time steps.
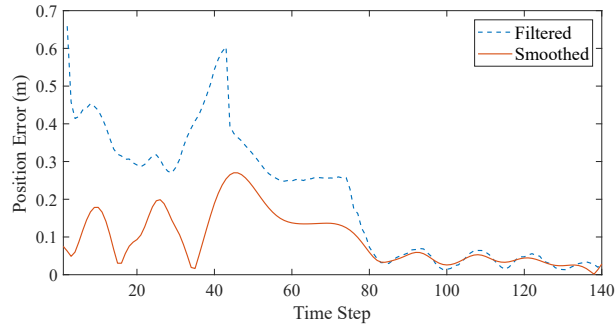
A. Single Object Simulation

In the single object scenario, one simulated object approaches the sensor along straight trajectory, performs a $90°$ left turn, and leaves the scene straightly. The estimated positions and velocities are compared to the true values. The Intersection-over-Union (IoU) ratio [26] is used as metric to evaluate contour estimation, i.e., let $\hat{A}$ and A_0 be the area of estimated and true contours, respectively,

$$\text{IoU} = \frac{\hat{A} \cap A_0}{\hat{A} \cup A_0} \in [0, 1] \tag{30}$$

where $\text{IoU} = 1$ denotes perfectly estimated contour.

The results for this scenario are depicted in Fig. 3. As can be noticed from Fig. 3c, the contour estimation in the forward pass has been largely improved after time step 80 due to the arrival of new information from the other side of the contour that just became visible to the sensor. By comparing Fig. 3a and 3c, fluctuations of IoU during forward pass can be partly explained as side effect of estimation errors in kinematics, and are mostly reduced in the final results.
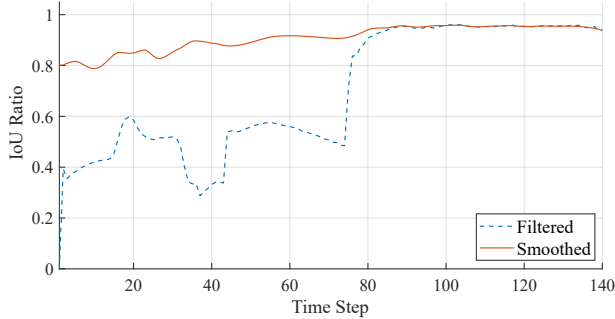
The improved contour is propagated backward in the backward pass and thus results in better estimation in the first 80
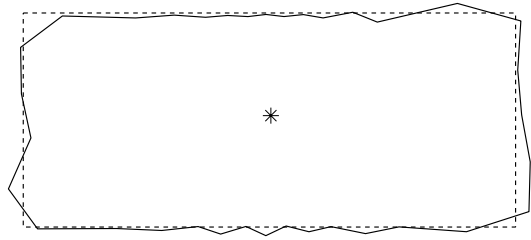
(a) Position error over time
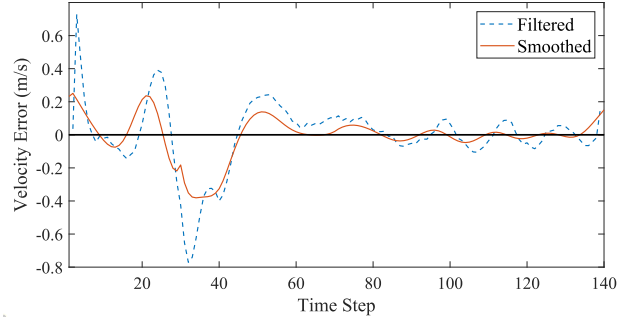
(b) Velocity error over time

(c) IoU ratio over time

(d) Contour estimation at the last time step. Dashed and solid lines represent the true and the estimated contours, respectively.

Fig. 3: Simulation results for single object scenario before and after smoothing

steps. As a consequence, the errors in position and velocity have been reduced as illustrated in Fig. 3a and 3b. It can also be noticed that the 1-step lag at the beginning of the forward pass caused by the adaptive birth has been corrected.

To test the flexibility of estimating free forms, we assume no prior knowledge of the object contour except for smoothness and closure, thus no advanced kernel function (e.g., symmetric, conservative kernels or explicit basis function [35]) is utilized to constrain the contour estimation with the Gaussian process, resulting in slightly irregular estimations as in Fig. 3d. Nonetheless, for specific applications where assumptions can be made for the geometry, the symmetry or the completeness of the contour, certain kernel functions can notably enhance the performance of the contour estimation [35], [36].

B. Multi-object Simulation

In the multi-object scenario, six objects with various geometries and birth/death intervals are present to simulate possible traffic behaviors at an intersection. The true and estimated trajectories are depicted in Fig. 4a. Using the IoU ratio and the Optimal SubPattern Assignment (OSPA) metric [46], the proposed track-before-detect method (Smoother with Robust Birth and Backward Filter, S-RB-BF) is compared against naive LMB smoother for extended objects and Smoother with Robust Birth (S-RB) in Table I and Fig. 4.

In the forward pass, the original birth times for objects are denoted as triangles, and some missing states can be observed as objects enter the scene (especially for the objects entering

TABLE I: Average IoU ratios over time

Objects	Smoother	S-RB	S-RB-BF
1	0.8387	0.8387	0.8445
2	0.8863	0.8863	0.8917
3	0.4837	0.8833	0.8939
4	0.8885	0.8885	0.8946
5	0.6744	0.6512	0.8675
6	0.7333	0.6424	0.7490

from bottom and right in Fig. 4a). After the backward pass, the proposed backward filter recovers these missing states and finally generates estimations which accurately follow the true tracks with a low cardinality error.

In Table I, the IoU ratios based on different methods are averaged over time for each object. Comparing the results of the naive smoother and S-RB, it can be seen that robust birth policy yields much better estimations for the object 3, yet worse results due to an initial cardinality error for objects 5 and 6 (entering from bottom and right in Fig. 4a). S-RB-BF then recovers almost all missing states and produces overall better results for all objects (comparing columns 2 and 4).

Comparing the OSPA errors of naive LMB smoother and S-RB in Fig. 4b, it can be observed that although the initial OSPA errors are low for the smoother without robust birth, some tracks are quickly lost due to a poor contour initialization (cf. Table I). In contrast, missing states in the early steps of S-RB cause high OSPA errors, which quickly decrease after all objects are tracked. The OSPA errors in the later steps

(a) True and estimated trajectories (only the center of mass is shown for the position estimates)

(b) OSPA errors of naive smoother, smoother with robust birth (S-RB), smoother with robust birth and backward filter (S-RB-BF)
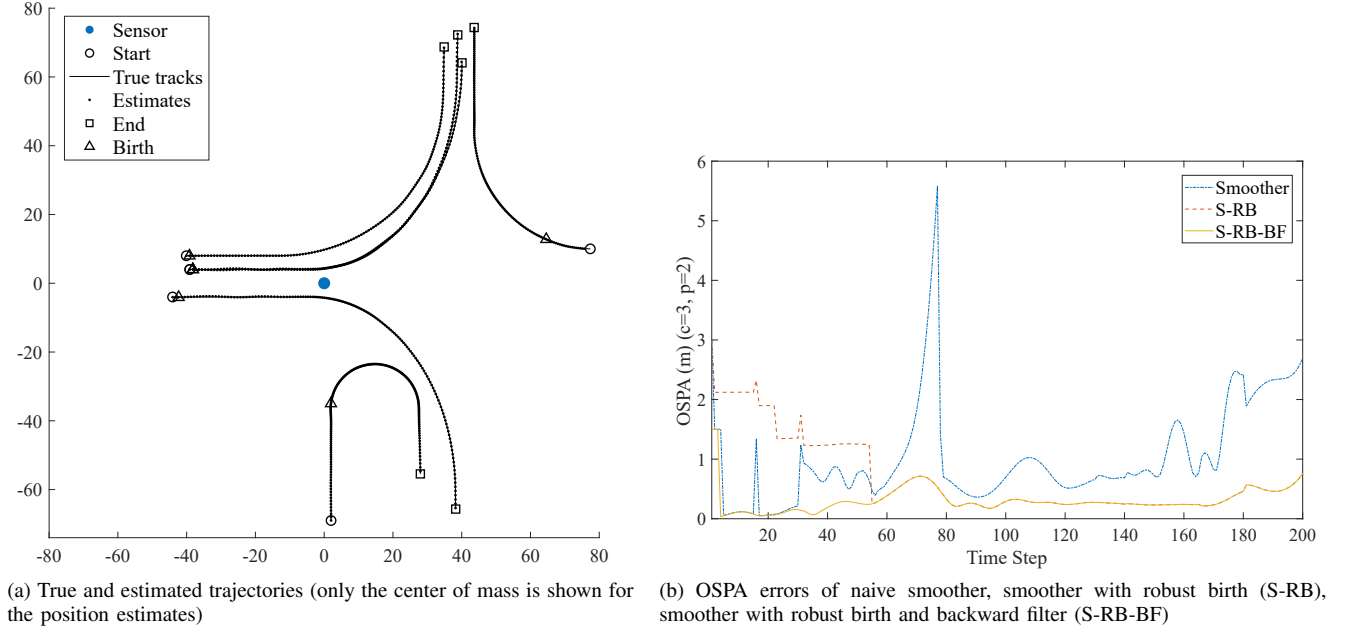
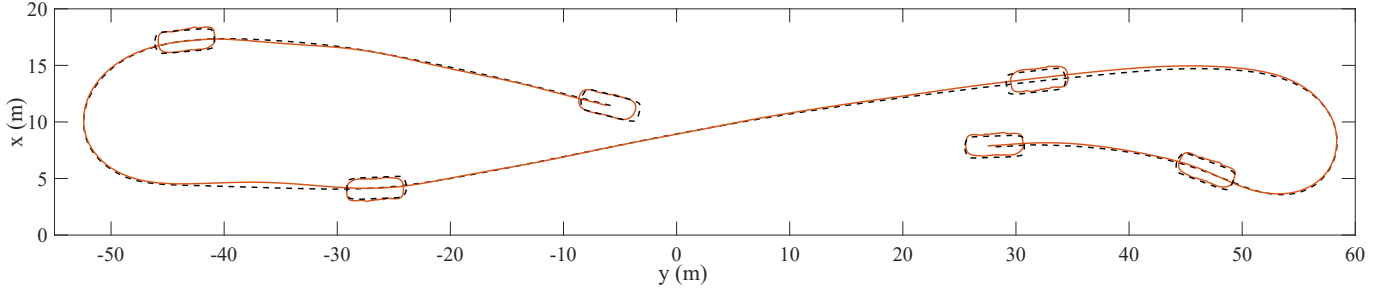Fig. 4: Test trajectories and simulation results for a multi-object scenario.

Fig. 5: Trajectory and contour estimations generated by the proposed method (solid lines), and DGPS-recorded trajectory with true contours (dashed lines). The corresponding sensor of the standing ego-vehicle is located at $(0,0)$.

are identical between S-RB and S-RB-BF, while S-RB-BF effectively recovers most missing states in the early steps and thus largely reduced the OSPA errors in this period.

C. Real-world Experiment

In this scenario, the proposed method is applied to a dataset collected by a Hesai Pandar LiDAR with 40 vertical channels, a horizontal angular resolution of $0.2°$, and the sampling interval $\Delta T = 0.1s$. All points are projected onto the X-Y plane and possible contour points (points visible to the sensor on the 2D plane) are extracted as described in [47]. For the validation, the motion data of the object vehicle is acquired using a RTK-supported DGPS system from OxTS. As the object to be tracked, we consider a BMW 740Li vehicle with a given contour shape to evaluate the estimations. As can be seen in Fig. 5, the generated trajectory is tightly bounded around the DGPS positions, and the contour estimation also yields a low deviation from the true contour, demonstrating the feasibility of the proposed method in practical cases.

VI. Conclusions

In this paper, the estimation of star-convex shapes using Gaussian processes is integrated in the LMB smoother to enable a simultaneous smoothing of the kinematics and the contour shapes for multiple extended objects. Robust birth policy is applied in the forward pass to guarantee enough information for the initialization of newborn objects, and thus improves both kinematics and contour estimations in the subsequent time steps. A novel approach is also proposed for generating track-before-detect estimations, which post-processes the unassociated data in a backward pass with a modified LMB filter to recover possibly missing states. Simulations and experiments have shown that our proposed method can yield more accurate estimations in both cardinality and individual states compared to the naive LMB smoother for extended object.

Our proposed approach allows developers to obtain reference data at low expense and high accuracy to fulfill the increasing demand in validating autonomous driving systems.

In the future, we would extend the contour model to 3D and investigate better metrics for evaluating cluster information to further improve robustness on processing real-world data.

REFERENCES

[1] L. Zhang, J. Lan, and X. R. Li, "Performance evaluation of multi-target tracking without knowing ground truth," in *Proc. of the 19th IEEE International Conference on Information Fusion*, 2016, pp. 185–192.

[2] C. Spampinato, S. Palazzo, and D. Giordano, "Evaluation of tracking algorithm performance without ground-truth data," in *Proc. of the 19th IEEE International Conference on Image Processing*, 2012, pp. 1345–1348.

[3] J. E. Stellet, L. Walkling, and J. M. Zöllner, "Post processing of laser scanner measurements for testing advanced driver assistance systems," in *Proc. of the 19th IEEE International Conference on Information Fusion*, 2016, pp. 1999–2006.

[4] U. Lages, M. Spencer, and R. Katz, "Automatic scenario generation based on laserscanner reference data and advanced offline processing," in *IEEE Intelligent Vehicles Symposium*, 2013, pp. 153–155.

[5] J. E. Stellet, M. R. Zofka, J. Schumacher, T. Schamm, F. Niewels, and J. M. Zöllner, "Testing of advanced driver assistance towards automated driving: A survey and taxonomy on existing approaches and open questions," in *Proc. of the 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 1455–1462.

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

[7] R. P. Mahler, *Statistical multisource-multitarget information fusion*. Norwood, MA, USA: Artech House, 2007.

[8] B. T. Vo, "Random finite sets in multi-object filtering," Ph.D. dissertation, University of Western Australia, Jun. 2008.

[9] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE journal of Oceanic Engineering*, vol. 8, no. 3, pp. 173–184, 1983.

[10] M. Mallick, S. Coraluppi, C. Carthel, V. Krishnamurthy, and B. Vo, "Multitarget tracking using multiple hypothesis tracking," in *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*. New York: Wiley Online Library, 2012, pp. 165–201.

[11] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, 2004.

[12] R. P. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[13] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[14] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 409–423, 2008.

[15] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, 2013.

[16] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, "The labeled multi-Bernoulli filter," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3246–3260, 2014.

[17] D. Meissner, S. Reuter, and K. Dietmayer, "Road user tracking at intersections using a multiple-model PHD filter," in *IEEE Intelligent Vehicles Symposium*, 2013, pp. 377–382.

[18] R. Hoseinnezhad, B.-N. Vo, and B.-T. Vo, "Visual tracking in background subtracted image sequences via multi-Bernoulli filtering," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 392–397, 2012.

[19] H. Deusch, S. Reuter, and K. Dietmayer, "The labeled multi-Bernoulli SLAM filter," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1561–1565, 2015.

[20] R. P. Mahler, B.-T. Vo, and B.-N. Vo, "Forward-backward probability hypothesis density smoothing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 1, pp. 707–728, 2012.

[21] M. Beard, B. T. Vo, and B.-N. Vo, "Generalised labelled multi-Bernoulli forward-backward smoothing," in *Proc. of the 19th International Conference on Information Fusion*. IEEE, 2016, pp. 688–694.

[22] S. Wong, B. T. Vo, and F. Papi, "Bernoulli forward-backward smoothing for track-before-detect," *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 727–731, 2014.

[23] B.-T. Vo, D. Clark, B.-N. Vo, and B. Ristic, "Bernoulli forward-backward smoothing for joint target detection and tracking," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4473–4477, 2011.

[24] M. Baum, V. Klumpp, and U. D. Hanebeck, "A novel Bayesian method for fitting a circle to noisy points," in *Proc. of the 13th International Conference on Information Fusion*. IEEE, 2010, pp. 1–6.

[25] N. Petrov, A. Gning, L. Mihaylova, and D. Angelova, "Box particle filtering for extended object tracking," in *Proc. of the 15th International Conference on Information Fusion*. IEEE, 2012, pp. 82–89.

[26] K. Granström, C. Lundquist, and U. Orguner, "Tracking rectangular and elliptical extended targets using laser measurements," in *Proc. of the 14th International Conference on Information Fusion*. IEEE, 2011, pp. 1–8.

[27] K. Granström, S. Reuter, D. Meissner, and A. Scheel, "A multiple model PHD approach to tracking of cars under an assumed rectangular shape," in *Proc. of the 17th International Conference on Information Fusion*. IEEE, 2014, pp. 1–8.

[28] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 1042–1059, 2008.

[29] K. Granstrom and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5657–5671, 2012.

[30] M. Beard, S. Reuter, K. Granström, B.-T. Vo, B.-N. Vo, and A. Scheel, "Multiple extended target tracking with labeled random finite sets," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1638–1653, 2015.

[31] M. Baum and U. D. Hanebeck, "Random hypersurface models for extended object tracking," in *IEEE International Symposium on Signal Processing and Information Technology*, 2009, pp. 178–183.

[32] ——, "Extended object tracking with random hypersurface models," *IEEE Transactions on Aerospace and Electronic systems*, vol. 50, no. 1, pp. 149–159, 2014.

[33] ——, "Shape tracking of extended objects and group targets with star-convex RHMs," in *Proc. of the 14th International Conference on Information Fusion*. IEEE, 2011, pp. 1–8.

[34] M. F. Huber, "Recursive Gaussian process: On-line regression and learning," *Pattern Recognition Letters*, vol. 45, pp. 85–91, 2014.

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

[36] T. Hirscher, A. Scheel, S. Reuter, and K. Dietmayer, "Multiple extended object tracking using Gaussian processes," in *Proc. of the 19th International Conference on Information Fusion*. IEEE, 2016, pp. 868–875.

[37] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter: Particle filters for tracking applications*. Norwood, MA, USA: Artech house, 2003.

[38] S. J. Davey, M. G. Rutten, and B. Cheung, "A comparison of detection performance for several track-before-detect algorithms," in *Proc. of the 11th International Conference on Information Fusion*. IEEE, 2008, pp. 1–8.

[39] T. Ogawa and G. Wanielik, "Track-before-detect approach on LiDAR signal processing for low SNR target detection," in *IEEE Intelligent Vehicles Symposium*, 2016, pp. 676–682.

[40] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. MIT press, 2006.

[41] R. Schubert, E. Richter, and G. Wanielik, "Comparison and evaluation of advanced motion models for vehicle tracking," in *Proc. of the 11th International Conference on Information Fusion*. IEEE, 2008, pp. 1–6.

[42] D. J. MacKay, "Introduction to Gaussian processes," in *Neural Networks and Machine Learning*, C. M. Bishop, Ed. Springer Verlag, 1998, pp. 133–165.

[43] K. Granström, C. Lundquist, and U. Orguner, "A Gaussian mixture PHD filter for extended target tracking," in *Proc. of the 13th International Conference on Information Fusion*. Edinburgh, UK: IEEE, Jul. 2010.

[44] A. R. Runnalls, "Kullback-Leibler approach to Gaussian mixture reduction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 989–999, 2007.

[45] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA journal*, vol. 3, no. 8, pp. 1445–1450, 1965.

[46] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE transactions on signal processing*, vol. 56, no. 8, pp. 3447–3457, 2008.

[47] S. Katz, A. Tal, and R. Basri, "Direct visibility of point sets," *ACM Transactions on Graphics (SIGGRAPH)*, vol. 26, no. 3, p. 24–es, 2007.