Scalable Detection and Tracking of Geometric Extended Objects

Florian Meyer [ORCID], *Member, IEEE*, and Jason L. Williams [ORCID], *Senior Member, IEEE*

Abstract—**Multiobject tracking provides situational awareness that enables new applications for modern convenience, public safety, and homeland security. This paper presents a factor graph formulation and a particle-based sum-product algorithm (SPA) for scalable detection and tracking of extended objects. The proposed method dynamically introduces states of newly detected objects, efficiently performs probabilistic multiple-measurement to object association, and jointly infers the geometric shapes of objects. Scalable extended object tracking (EOT) is enabled by modeling association uncertainty by measurement-oriented association variables and newly detected objects by a Poisson birth process. Contrary to conventional EOT methods, a fully particle-based approach makes it possible to describe different geometric object shapes. The proposed method can reliably detect, localize, and track a large number of closely-spaced extended objects without gating and clustering of measurements. We demonstrate significant performance advantages of our approach compared to the recently introduced Poisson multi-Bernoulli mixture filter. In particular, we consider a simulated scenarios with up to twenty closely-spaced objects and a real autonomous driving application where measurements are captured by a lidar sensor.**

Index Terms—**Extended object tracking, data association, factor graphs, sum-product algorithm, random matrix theory.**

I. INTRODUCTION

EMERGING sensing technologies and innovative signal processing methods will lead to new capabilities for autonomous operation in a wide range of problems such as applied ocean sciences and ground navigation. A technology that enables new capabilities in this context is EOT. EOT methods make it possible to detect and track an unknown number of objects with unknown shapes by using modern sensors such as high-resolution radar, sonar, and lidar [1].

Due to the high resolution of these sensors, the point object assumption used in conventional multiobject tracking algorithms [2]–[4] is no longer valid. Objects have an unknown

shape that must be inferred together with their kinematic state. In addition, data association is particularly challenging due to the overwhelmingly large number of possible association events [5]. Even if only a single extended object is considered, the number of possible measurement-to-object association events scales combinatorially in the number of measurements. In the more realistic case of multiple extended objects, this "combinatorial explosion" of possible events is further exacerbated.

A. State of the Art

An important aspect of EOT is the modeling and inference of object shapes. The shape of an object determines how the measurements that originate from the object are spatially distributed around its center. An important and widely used model [5]–[9] based on random matrix theory has been introduced in [10]. This model considers elliptically shaped objects. Here, unknown reflection points of measurements on extended objects are modeled by a random matrix that acts as the covariance matrix in the measurement model. This random extent state is estimated sequentially together with the kinematic state. The approach proposed in [10] introduces a closed-form update step for linear dynamics. It exploits that the Gaussian inverse Wishart distribution is a conjugate prior for linear Gaussian measurement models.

A major drawback of the original random matrix-based tracking method is that the driving noise variance in the state transition function of the kinematic state must be proportional to the extent of the object [10]. Furthermore, the orientation of the random matrix is constant. These limitations have been addressed by improved and refined random matrix-based extent models introduced in [11]–[13] and particle-based methods [14]. A significant number of additional application-dependent object extent models have been proposed recently (see [1] and references therein). The state-of-the-art algorithm for tracking an unknown number of objects with elliptical shape is the Poisson multi-Bernoulli mixture (PMBM) filter [5]. The PMBM filter is based on a model [15] where objects that have produced a measurement are described by a multi-Bernoulli probability distribution and objects that exist but have not yet produced a measurement yet are described by a Poisson distribution. This PMBM model is a conjugate prior with respect to the prediction and update steps of the extended object tracking problem [5].

Important methods for the tracking of point objects include joint probabilistic data association (JPDA) [2], multi-hypothesis tracking (MHT) [16]–[18], and approaches based on random

finite sets (RFS) [3], [15], [19]–[21]. Many of these point object tracking methods have been adapted for extended object tracking. In [22], a probabilistic data association algorithm for the tracking of a single object that can produce more than one measurement has been introduced. A multiple-measurement JPDA filter for the tracking of multiple objects has been independently developed in [23] and [24]. The method in [6] combines a random matrix extent model with multisensor JPDA for the tracking of multiple extended objects. The computational complexity of these methods scales combinatorially in the number of objects and the number of measurements. They rely on gating, a suboptimal preprocessing technique that excludes unlikely data association events. Thus, these methods are only suitable in scenarios where objects produce few measurements and no more than four objects come into close proximity at any time [25].

Another widely used approach to limit computational complexity for data association with extended objects is to perform clustering of spatially close measurements. Here, based on the assumption that all measurements in a cluster correspond to the same object, the majority of possible data association events is discarded in a suboptimal preprocessing step. Based on this approach, tracking methods for objects that produce multiple measurements have been developed in the MHT [26] and the JPDA [27], [28] frameworks. Practical implementations of the update step of RFS-based methods for EOT [5], [7]–[9] including the one of the PMBM filter [5] also rely on gating and clustering and suffer from a reduced performance if objects are in close proximity. An alternative approach is to perform data association with extended objects using sampling methods [29], [30]. Here, in scenarios with objects that are in proximity, a better estimation performance can be achieved compared to methods based on clustering. These methods based on sampling, however, also discard the majority of possible data association events, and thus eventually fail if the number of relevant events becomes overwhelmingly large [31].

An innovative approach to high-dimensional estimation problems is the framework probabilistic graphical models [32]. In particular, the SPA [32]–[34] can provide scalable solutions to high-dimensional estimation problems. The SPA performs local operations through "messages" passed along the edges of a factor graph. Many traditional sequential estimation methods such as the Kalman filter, the particle filter, and JPDA can be interpreted as instances of SPA. In addition, SPA has led to a variety of new estimation methods in a wide range of applications [35]–[37]. For probabilistic data association (PDA) with point objects, an SPA-based method referred to as the sum-product algorithm for data association (SPADA) is obtained by executing the SPA on a bipartite factor graph and simplifying the resulting SPA messages [4], [31], [38]. The computational complexity of this algorithm scales as the product of the number of measurements and the number of objects. Sequential estimation methods that embed the SPADA to reduce computational complexity have been introduced for multiobject tracking (MOT) [15], [39]–[43], indoor localization [44]–[47], as well as simultaneous localization and tracking [48]–[50].

SPA methods for probabilistic data association based on a bipartite factor graph have recently been investigated for the tracking of extended objects [25], [51], [52]. Here, the number of measurements produced by each object is modeled by an arbitrary truncated probability mass function (PMF). An SPA for data association with extended objects with a computational complexity that scales as the product of the number of measurements and the number of objects has been introduced [25], [51]. This method can track multiple objects that potentially generate a large number of measurements but was found unable to reliably determine the probability of existence of objects detected for the first time. An alternative method for scalable data association with extended objects has been developed independently and has been presented in [53], [54]. These methods however either assume that the number of objects are known, or that a certain prior information of new object locations is available. They are thus unsuitable for scenarios with spontaneous object birth.

B. Contributions and Notations

In this paper, we introduce a Bayesian particle-based SPA for scalable detection and tracking of an unknown number of extended objects. Object birth and the number of measurements generated by an object are described by a Poisson PMF. The proposed method is derived on a factor graph that depicts the statistical model of the extended object tracking problem and is based on a representation of data association uncertainty by means of measurement-oriented association variables. Contrary to the factor graph used in [51], it makes it possible to reliably determine the existence of objects by means of the SPA. In our statistical model, the object extent is modeled as a basic geometric shape that is represented by a positive semidefinite extent state. Contrary to conventional EOT algorithms with a random-matrix model, the proposed fully particle-based method is not limited to elliptical object shapes. A fully particle-based approach also makes it possible to consider a uniform distribution of measurements on the object extent. Furthermore, our method has a computational complexity that scales only quadratically in the number of objects and the number of measurements. It can thus accurately detect, localize, and track multiple closely-spaced extended objects that generate a large number of measurements. The contributions of this paper are as follows.

- We derive a new factor graph for the problem of detection and tracking of multiple extended objects and introduce the corresponding message passing equations.
- We establish a particle-based scalable message passing algorithm for the detection and tracking of an unknown number of extended objects.
- We demonstrate performance advantages in a challenging simulated scenario and by using real lidar measurements acquired by an autonomous vehicle.

This paper advances over the preliminary account of our method provided in the conference publication [55] by (i) simplifying the representation of data association uncertainty, (ii) including object extents in the statistical model, (iii) presenting a detailed derivation of the factor graph, (iv) establishing a particle-based implementation of the proposed method, (v) discussing scaling properties, (vi) demonstrating performance advantages compared to the recently proposed
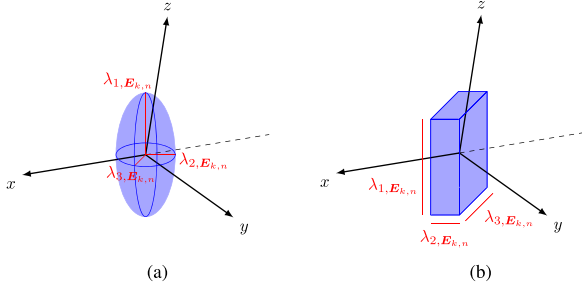
Fig. 1. Object shapes represented by extent state $\mathbf{E}_{k,n}$ in a 3-D scenario. An ellipsoid (a) or cube (b) is defined by the eigenvalues $\lambda_{1,\mathbf{E}_{k,n}} > \lambda_{2,\mathbf{E}_{k,n}} > \lambda_{3,\mathbf{E}_{k,n}} \geq 0$ of the 3×3 symmetric, positive-semidefinite matrix $\mathbf{E}_{k,n}$. The local reference frame and thus the orientation of the object is defined by the eigenvectors $\boldsymbol{\lambda}_{1,\mathbf{E}_{k,n}}$, $\boldsymbol{\lambda}_{2,\mathbf{E}_{k,n}}$, and $\boldsymbol{\lambda}_{3,\mathbf{E}_{k,n}}$ of $\mathbf{E}_{k,n}$.

Poisson multi-Bernoulli mixture filter [5] using synthetic and real measurements.

The probabilistic model established in this paper is closely related to the probabilistic model employed by the PMBM filter (cf. [4]). Our focus is on obtaining a factor graph formulation for which the SPA scales well, and on exploiting the accuracy and flexibility of a particle-based implementation.

Notation: Random variables are displayed in sans serif, upright fonts, while their realizations are in serif, italic fonts. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For example, a random variable and its realization are denoted by x and x; a random vector and its realization by $\mathbf{x}$ and $\boldsymbol{x}$; and a random matrix and its realization are denoted by $\mathbf{X}$ and $\mathbf{X}$. Furthermore, $\boldsymbol{x}^{\mathrm{T}}$ denotes the transpose of vector $\boldsymbol{x}$; $\propto$ indicates equality up to a normalization factor; $f(\boldsymbol{x})$ denotes the probability density function (PDF) of random vector $\mathbf{x}$. $\mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the Gaussian PDF (of random vector $\mathbf{x}$) with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $\mathcal{U}(\boldsymbol{y}; \mathcal{S})$ denotes the uniform PDF (of random vector $\mathbf{y}$) with support $\mathcal{S}$, and $\mathcal{W}(\mathbf{X}; q, \mathbf{Q})$ denotes the Wishart distribution (of random matrix $\mathbf{X}$) with degrees of freedom q and mean $q\,\mathbf{Q}$. The determinant of matrix $\mathbf{Q}$ is denoted $|\mathbf{Q}|$. Finally, $\mathrm{bdiag}(\mathbf{M}_1, \dots, \mathbf{M}_I)$ denotes the block diagonal matrix that consists of submatrices $\mathbf{M}_1, \dots, \mathbf{M}_I$.

II. SYSTEM MODEL

At time n, object k is described by a kinematic state and an extent state. The kinematic state $\mathbf{x}_{k,n} \triangleq [\mathbf{p}_{k,n}^{\mathrm{T}} \ \mathbf{m}_{k,n}^{\mathrm{T}}]^{\mathrm{T}}$ consists of the object's position $\mathbf{p}_{k,n} \in \mathbb{R}^{d_p}$ and possibly further parameters $\mathbf{m}_{k,n} \in \mathbb{R}^{d_m}$ such as velocity or turn rate. The extent state $\mathbf{E}_{k,n} \in \mathbb{R}^{d_p \times d_p}$ is a symmetric, positive semidefinite random matrix that can either model an ellipsoid or a cube. Example realizations of $\mathbf{E}_{k,n}$ and corresponding object shapes are shown in Fig. 1. Formally, we also introduce the vector notation $\mathbf{e}_{k,n}$ of extent state $\mathbf{E}_{k,n}$, which is the concatenation of diagonal and unique off-diagonal elements of $\mathbf{E}_{k,n}$, e.g., in a 2-D tracking scenario the vector notation of extent matrix $\mathbf{E}_{k,n} = [[\mathsf{e}_{k,n}^{(11)} \ \mathsf{e}_{k,n}^{(21)}]^{\mathrm{T}} \ [\mathsf{e}_{k,n}^{(12)} \ \mathsf{e}_{k,n}^{(22)}]^{\mathrm{T}}]$ is given by $\mathbf{e}_{k,n} = [\mathsf{e}_{k,n}^{(11)} \ \mathsf{e}_{k,n}^{(21)} \ \mathsf{e}_{k,n}^{(22)}]^{\mathrm{T}}$. Note that the support of $\mathbf{e}_{k,n}$ corresponds to all positive-semidefinite matrices in $\mathbb{R}^{d_p \times d_p}$. In what follows,

we will use extent matrix $\mathbf{E}_{k,n}$ and its vector notation $\mathbf{e}_{k,n}$ interchangeably.

As in [4], [40], [50], we account for the time-varying and unknown number of extended objects by introducing potential objects (POs) $k \in \{1, \dots, K_n\}$. The number of POs K_n is the maximum possible number of actual objects that have produced a measurement so far [4]. (Note that K_n increases with time n.) The existence/non-existence of PO k is modeled by the binary existence variable $\mathsf{r}_{k,n} \in \{0, 1\}$ in the sense that PO k represents an actual object if and only if $r_{k,n} = 1$. Augmented PO states are denoted as $\mathbf{y}_{k,n} = [\mathbf{x}_{k,n}^{\mathrm{T}} \ \mathbf{e}_{k,n}^{\mathrm{T}} \ \mathsf{r}_{k,n}]^{\mathrm{T}}$.

Formally, PO k is also considered if it is non-existent, i.e., if $r_{k,n} = 0$. The states $\mathbf{x}_{k,n}$ of non-existent POs are obviously irrelevant and have no impact on the estimation solution. Therefore, all hybrid continuous/discrete PDFs defined on augmented PO states, $f(\boldsymbol{y}_{k,n}) = f(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n}, r_{k,n})$, are of the form $f(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n}, 0) = f_{k,n} f_{\mathrm{d}}(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n})$, where $f_{\mathrm{d}}(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n})$ is an arbitrary "dummy PDF" and $f_{k,n} \in [0, 1]$ is the probability that the PO does not exist[1]. For every PO state at previous times, there is one "legacy" PO state $\underline{\mathbf{y}}_{k,n}$ at current time n. The number of legacy objects and the joint legacy PO state at time n are introduced as $\underline{K}_n = K_{n-1}$ and $\underline{\mathbf{y}}_n \triangleq [\underline{\mathbf{y}}_{1,n}^T \cdots \underline{\mathbf{y}}_{\underline{K}_n,n}^T]^{\mathrm{T}}$, respectively.

A. State-Transition Model

The state-transition pdf for legacy PO state $\underline{\mathbf{y}}_n$ factorizes as

$$f(\underline{\boldsymbol{y}}_n | \boldsymbol{y}_{n-1}) = \prod_{k=1}^{\underline{K}_n} f(\underline{\boldsymbol{y}}_{k,n} | \boldsymbol{y}_{k,n-1}) \tag{1}$$

where the single-object augmented state-transition pdf $f(\underline{\boldsymbol{y}}_{k,n} | \boldsymbol{y}_{k,n-1}) = f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n} | \boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1})$ is given as follows. If PO k does not exist at time $n-1$, i.e., $r_{k,n-1} = 0$, then it does not exist at time n either, i.e., $\underline{r}_{k,n} = 0$, and thus its state pdf is $f_{\mathrm{d}}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n})$. This means that

$$f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n} | \boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1} = 0)$$
$$= \begin{cases} f_{\mathrm{d}}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}), & \underline{r}_{k,n} = 0 \\ 0, & \underline{r}_{k,n} = 1. \end{cases} \tag{2}$$

On the other hand, if PO k exists at time $n-1$, i.e., $r_{k,n-1} = 1$, then the probability that it still exists at time n, i.e., $\underline{r}_{k,n} = 1$, is given by the survival probability p_{s}, and if it still exists at time n, its state $\underline{\mathbf{x}}_{k,n}$ is distributed according to a single-object state-transition pdf $f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n} | \boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1})$. Thus,

$$f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n} | \boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1} = 1)$$
$$= \begin{cases} (1 - p_{\mathrm{s}}) f_{\mathrm{d}}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}), & \underline{r}_{k,n} = 0 \\ p_{\mathrm{s}} f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n} | \boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}), & \underline{r}_{k,n} = 1. \end{cases} \tag{3}$$

---

[1]Note that representing objects states and existence variables by this type of PDFs is analogous to a multi-Bernoulli formulation in the RFS framework [4], [5], [15].

It is assumed that at time $n = 0$, the prior distributions $f(\boldsymbol{y}_{k,0})$ are statistically independent across POs k. If no prior information is available, we have $K_0 = 0$.

A single-object state-transition pdf $f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n} | \boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1})$ that was found useful in many extended object tracking scenarios is given by [12]

$$f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n} | \boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}) = \mathcal{N}\left(\underline{\boldsymbol{x}}_{k,n}; f(\boldsymbol{x}_{k,n-1}), \boldsymbol{\Sigma}_{k,n}\right)$$
$$\times \mathcal{W}\left(\underline{\mathbf{E}}_{k,n}; q_{k,n}, \frac{\mathbf{V}(\boldsymbol{m}_{k,n-1}) \mathbf{E}_{k,n-1} \mathbf{V}(\boldsymbol{m}_{k,n-1})^{\mathrm{T}}}{q_{k,n}}\right) \quad (4)$$

where $f(\boldsymbol{x}_{k-1,n})$ is the state transition function of the kinematic state $\mathbf{x}_{k-1,n}$, $\boldsymbol{\Sigma}_{k,n}$ is the kinematic driving noise covariance matrix, and $\mathbf{V}(\boldsymbol{m}_{k,n})$ is a rotation matrix. The degrees of freedom of the Wishart distribution determine the uncertainty of object extent prediction. A small $q_{k,n}$ implies a large state-transition uncertainty.

B. Measurement Model

At time n, a sensor produces measurements $\mathbf{z}_{l,n} \in \mathbb{R}^{d_{\mathbf{z}}}$, $l \in \{1, \dots, \mathsf{M}_n\}$. The joint measurement vector is denoted as $\mathbf{z}_n \triangleq [\mathbf{z}_{1,n}^{\mathrm{T}} \cdots \mathbf{z}_{\mathsf{M}_n,n}^{\mathrm{T}}]^{\mathrm{T}}$. (Note that the total number of measurements M_n is random.) Each measurement is either generated by a single object or is a false alarm.

Let $\mathbf{z}_{l,n}$ be a measurement generated by the kth PO[2]. In particular, let $\mathbf{v}_{k,n}^{(l)}$ be the relative position (with respect to $\mathbf{p}_{k,n}$) of the reflection point that generated $\mathbf{z}_{l,n}$. The considered general nonlinear measurement model is now given by

$$\mathbf{z}_{l,n} = \boldsymbol{d}\left(\mathbf{p}_{k,n} + \mathbf{v}_{k,n}^{(l)}\right) + \mathbf{u}_{l,n} \quad (5)$$

where $\boldsymbol{d}(\cdot)$ is an arbitrary nonlinear function and $\mathbf{u}_{l,n} \sim \mathcal{N}(\boldsymbol{u}_{l,n}; \mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{u}_{l,n}})$ is the measurement noise. (Note that the subspace $\mathbf{m}_{k,n}$ remains unobserved.) This measurement model directly determines the conditional PDF $f(\mathbf{z}_{l,n} | \boldsymbol{x}_{k,n}, \boldsymbol{v}_{k,n}^{(l)})$.

We consider three geometric object shape models. The random extent state $\mathbf{E}_{k,n} = \mathbf{E}_{k,n}$ either defines (i) the covariance matrix of a Gaussian PDF, i.e., $\mathbf{v}_{k,n}^{(l)} \sim \mathcal{N}(\boldsymbol{v}_{k,n}^{(l)}; \mathbf{0}, \mathbf{E}_{k,n}^2)$; (ii) the elliptical support $\mathcal{S}_{\mathrm{e}}(\mathbf{E}_{k,n})$ of a uniform PDF, i.e., $\mathbf{v}_{k,n}^{(l)} \sim \mathcal{U}(\boldsymbol{v}_{k,n}^{(l)}; \mathcal{S}_{\mathrm{e}}(\mathbf{E}_{k,n}))$; or (iii) the cubical support $\mathcal{S}_{\mathrm{c}}(\mathbf{E}_{k,n})$ of a uniform PDF, i.e., $\mathbf{v}_{k,n}^{(l)} \sim \mathcal{U}(\boldsymbol{v}_{k,n}^{(l)}; \mathcal{S}_{\mathrm{c}}(\mathbf{E}_{k,n}))$[3]. The considered object shape model directly determines the conditional PDF $f(\boldsymbol{v}_{k,n}^{(l)} | \boldsymbol{e}_{k,n})$. The PDF of measurement vector $\mathbf{z}_{l,n}$ conditioned

on $\mathbf{x}_{k,n}$ and $\mathbf{e}_{k,n}$ can now be obtained as

$$f(\boldsymbol{z}_{l,n} | \boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n})$$
$$= \int f\left(\boldsymbol{z}_{l,n} | \boldsymbol{x}_{k,n}, \boldsymbol{v}_{k,n}^{(l)}\right) f\left(\boldsymbol{v}_{k,n}^{(l)} | \boldsymbol{e}_{k,n}\right) \mathrm{d}\boldsymbol{v}_{k,n}^{(l)}. \quad (6)$$

An important special case of this model is the additive-Gaussian case, i.e.,

$$\mathbf{z}_{l,n} = \mathbf{p}_{k,n} + \mathbf{v}_{k,n}^{(l)} + \mathbf{u}_{l,n}. \quad (7)$$

For extent model (i), in this case, there exist closed-form expressions for (6). In particular, the likelihood function $f(\boldsymbol{z}_{l,n} | \boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n})$ can be expressed as $\mathcal{N}(\boldsymbol{z}_{l,n}; \boldsymbol{p}_{k,n}, \mathbf{E}_{k,n}^2 + \boldsymbol{\Sigma}_{\boldsymbol{u}_{l,n}})$. Similarly, for models (ii) and (iii), there exists a simple approximation discussed in [56, Section 2], that also results in a closed-form of (6).

If PO k exists ($\mathsf{r}_{k,n} = 1$) it generates a random number of object-originated measurements $\mathbf{z}_{l,n}$, which are distributed according to the conditional PDF $f(\boldsymbol{z}_{l,n} | \boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n})$ discussed above. The number of measurements it generates is Poisson distributed with mean $\mu_{\mathrm{m}}(\mathbf{x}_{k,n}, \mathbf{e}_{k,n})$. For example, this Poisson distribution can be determined by a spatial measurement density ρ related to the sensor resolution and by the volume or surface area of the object extent, i.e., $\mu_{\mathrm{m}}(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}) = \rho |\mathbf{E}_{k,n}|$. False alarm measurements are modeled by a Poisson point process with mean μ_{fa} and strictly positive PDF $f_{\mathrm{fa}}(\boldsymbol{z}_{l,n})$.

C. New POs

Newly detected objects, i.e., actual objects that generated a measurement for the first time, are modeled by a Poisson point process with mean μ_{n} and PDF $f_{\mathrm{n}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n})$. Following [4], [5], [15], newly detected objects are represented by new PO states $\overline{\mathbf{y}}_{k,n}$, $k \in \{1, \dots, \mathsf{M}_n\}$. Each new PO state corresponds to a measurement $\mathbf{z}_{k,n}$; $\overline{\mathsf{r}}_{k,n} = 1$ means that measurement $\mathbf{z}_{k,n}$ was generated by a newly detected object. Since newly detected objects can produce more than one measurement, we define a mapping from measurements to new POs by the following rule[4]. At time n, if multiple measurements $l_1, \dots, l_L$ with $L \leq \mathsf{M}_n$ are generated by the same newly detected object, we have $\overline{\mathsf{r}}_{k_{\min},n} = 1$ for $k_{\min} = \min(l_1, \dots, l_L)$ and $\overline{\mathsf{r}}_{k,n} = 0$ for all $k \in \{l_1, \dots, l_L\} \backslash \{k_{\min}\}$. As will be further discussed in Section II-D, with this mapping every association event related to newly detected objects can be represented by exactly one configuration of new existence variables $\overline{\mathsf{r}}_{k,n}$, $k \in \{1, \dots, \mathsf{M}_n\}$. We also introduce $\overline{\mathbf{y}}_n \triangleq [\overline{\mathbf{y}}_{1,n}^{\mathrm{T}} \cdots \overline{\mathbf{y}}_{\mathsf{M}_n,n}^{\mathrm{T}}]^{\mathrm{T}}$ representing the joint vector of all new PO states. Note that at time n, the total number of POs is given by $\mathsf{K}_n = \underline{K}_n + \mathsf{M}_n$.

Since new POs are introduced as new measurements are incorporated, the number of PO states would grow indefinitely. Thus, for the development of a feasible method, a suboptimal pruning step is employed. This pruning step removes unlikely POs and will be further discussed in Section III-A.

---

[2] The fact that the PO with index k generates a measurement implies $r_{k,n} = 1$.

[3] The cubical support $\mathcal{S}_{\mathrm{c}}(\mathbf{E}_{k,n})$ and the elliptical support $\mathcal{S}_{\mathrm{e}}(\mathbf{E}_{k,n})$ are defined by the eigenvalues and eigenvectors of $\mathbf{E}_{k,n}$ as shown in Fig 1. For example, let us consider a 3-D scenario and let $\lambda_{1,\mathbf{E}_{k,n}} > \lambda_{2,\mathbf{E}_{k,n}} > \lambda_{3,\mathbf{E}_{k,n}} \geq 0$ be the eigenvalues of the 3×3 symmetric, positive-semidefinite matrix $\mathbf{E}_{k,n}$. Length, width, and height of the cube $\mathcal{S}_{\mathrm{c}}(\mathbf{E}_{k,n})$ are given by $\lambda_{1,\mathbf{E}_{k,n}}, \lambda_{2,\mathbf{E}_{k,n}}$, and $\lambda_{3,\mathbf{E}_{k,n}}$, respectively. Similarly, the lengths of the three principal semi-axes of the ellipsoid $\mathcal{S}_{\mathrm{e}}(\mathbf{E}_{k,n})$, are given by $\lambda_{1,\mathbf{E}_{k,n}}, \lambda_{2,\mathbf{E}_{k,n}}$, and $\lambda_{3,\mathbf{E}_{k,n}}$. The orientation of $\mathcal{S}_{\mathrm{c}}(\mathbf{E}_{k,n})$ and $\mathcal{S}_{\mathrm{e}}(\mathbf{E}_{k,n})$ is determined by the eigenvectors $\boldsymbol{\lambda}_{1,\mathbf{E}_{k,n}}, \boldsymbol{\lambda}_{2,\mathbf{E}_{k,n}}$, and $\boldsymbol{\lambda}_{3,\mathbf{E}_{k,n}}$. The eigenvalues $\lambda_{1,\mathbf{E}_{k,n}}, \lambda_{2,\mathbf{E}_{k,n}}$, and $\lambda_{3,\mathbf{E}_{k,n}}$ are assumed distinct. This assumption is always valid for real sensor data and objects of practical interest.

[4] A detailed derivation and discussion is provided in the supplementary material [56].

D. Data Association Uncertainty

Tracking of multiple extended objects is complicated by the data association uncertainty: it is unknown which measurement $z_{l,n}$ originated from which object k. To reduce computational complexity, following [4], [38], [40] we use *measurement-oriented association variables*

$$\mathsf{b}_{l,n} \triangleq \begin{cases} k \in \{1,\ldots,\underline{K}_n + l\}, & \text{if measurement } l \text{ is generated} \\ & \text{by PO } k \\ 0, & \text{if measurement } l \text{ is not generated by a PO} \end{cases}$$

and define the measurement-oriented association vector as $\mathbf{b}_n = [\mathsf{b}_{1,n} \cdots \mathsf{b}_{M_n,n}]^{\mathrm{T}}$. This representation of data association makes it possible to develop scalable SPAs for object detection and tracking. Note that the maximum value in the support of $\mathsf{b}_{l,n}$ is $\underline{K}_n + l$. (As discussed above, a measurement with index l cannot be associated to a new PO with index $l' > l$, i.e., the event in which measurements l and l' are generated by the same newly detected object is represented through the new PO l.) This is a direct result of the mapping introduced in Section II-C. In what follows, we write $\mathsf{b}_{l,n} \neq k$ short for $\mathsf{b}_{l,n} \in \{0, 1, \ldots, \underline{K}_n + l\}\backslash\{k\}$.

For a better understanding of new POs and measurement-oriented association vectors, we consider simple examples for fixed $M_n = 3$ and $\underline{K}_n = 0$. The event where all three measurements are generated by the same newly detected object, is represented by $\overline{r}_{1,n} = 1$, $\overline{r}_{2,n} = 0$, $\overline{r}_{3,n} = 0$, and $\mathbf{b}_n = [b_{1,n}\, b_{2,n}\, b_{3,n}]^{\mathrm{T}} = [1\,1\,1]^{\mathrm{T}}$. Furthermore, the event where all three measurements are generated by three different newly detected objects, is represented by $\overline{r}_{1,n} = 1$, $\overline{r}_{2,n} = 1$, $\overline{r}_{3,n} = 1$, and $\mathbf{b}_n = [1\,2\,3]^{\mathrm{T}}$. Finally, the event where measurements $m \in \{2, 3\}$ are generated by the same newly detected object and measurement $m = 1$ is a false alarm, is represented by $\overline{r}_{1,n} = 0$, $\overline{r}_{2,n} = 1$, $\overline{r}_{3,n} = 0$, and $\mathbf{b}_n = [0\,2\,2]^{\mathrm{T}}$. Note how every event related to newly detected objects is represented by exactly one configuration of new existence variables $\overline{r}_{n,k}$, $k \in \{1, 2, 3\}$ and association vector $\mathbf{b}_n$.

III. PROBLEM FORMULATION AND FACTOR GRAPH

In this section, we formulate the detection and estimation problem of interest and present the joint posterior pdf and factor graph underlying the proposed EOT method.

A. Object Detection and State Estimation

We consider the problem of detecting legacy and new POs $k \in \{1, \ldots, \underline{K}_n + M_n\}$ based on the binary existence variables $r_{k,n}$, as well as estimation of the object states $\mathbf{x}_{k,n}$ and extents $\mathbf{e}_{k,n}$ from the observed measurement history vector $z_{1:n} = [z_1^{\mathrm{T}} \cdots z_n^{\mathrm{T}}]^{\mathrm{T}}$. Our Bayesian approach aims to calculate the marginal posterior existence probabilities $p(r_{k,n} = 1|z_{1:n})$ and the marginal posterior pdfs $f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}|r_{k,n} = 1, z_{1:n})$. We perform detection by comparing $p(r_{k,n} = 1|z_{1:n})$ with a threshold P_{th}, i.e., if $p(r_{k,n} = 1|z_{1:n}) > P_{\text{th}}$ [57, Ch. 2], PO k is considered to exist. Estimates of $\mathbf{x}_{k,n}$ and $\mathbf{e}_{k,n}$ are then obtained for the detected objects k by the minimum mean-square error (MMSE) estimator [57, Ch. 4]. In particular, an MMSE estimate

of the kinematic state is obtained as

$$\hat{\mathbf{x}}_{k,n}^{\mathrm{MMSE}} \triangleq \int \mathbf{x}_{k,n} \int f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}|r_{k,n} = 1, z_{1:n}) \mathrm{d}\mathbf{e}_{k,n} \mathrm{d}\mathbf{x}_{k,n}. \quad (8)$$

Similarly, an MMSE estimate $\hat{e}_{k,n}^{\mathrm{MMSE}}$ of the extent state is obtained by replacing in (8) $\mathbf{x}_{k,n}$ with $\mathbf{e}_{k,n}$.

We can compute the marginal posterior existence probability $p(r_{k,n} = 1|z_{1:n})$ needed for object detection as discussed above, from the marginal posterior pdf, $f(\mathbf{y}_{k,n}|z_{1:n}) = f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}, r_{k,n}|z_{1:n})$, as

$$p(r_{k,n} = 1|z_{1:n})$$
$$= \int\int f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}, r_{k,n} = 1|z_{1:n}) \mathrm{d}\mathbf{x}_{k,n} \mathrm{d}\mathbf{e}_{k,n}. \quad (9)$$

Note that $p(r_{k,n} = 1|z_{1:n})$ is also needed for the suboptimal pruning step discussed in Section V-D.

Similarly, we can calculate the marginal posterior pdf $f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}|r_{k,n} = 1, z_{1:n})$ underlying MMSE state estimation (see (8)) from $f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}, r_{k,n}|z_{1:n})$ according to

$$f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}|r_{k,n} = 1, z_{1:n}) = \frac{f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}, r_{k,n} = 1|z_{1:n})}{p(r_{k,n} = 1|z_{1:n})}. \quad (10)$$

For this reason, the fundamental task is to compute the pdf $f(\mathbf{y}_{k,n}|z_{1:n}) = f(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}, r_{k,n}|z_{1:n})$. This pdf is a marginal density of $f(\mathbf{y}_{0:n}, \mathbf{b}_{1:n}|z_{1:n})$, which involves all the augmented states and measurement-oriented association variables up to the current time n.

The main problem to be solved is to find a computationally feasible recursive (sequential) calculation of marginal posterior PDFs $f(\mathbf{y}_{k,n}|z_{1:n})$. By performing message passing by means of the SPA rules [33] on the factor graph that represents our statistical model discussed in Section II, approximations ("beliefs") $\tilde{f}(\mathbf{y}_{k,n})$ of this marginal posterior pdfs can be obtained in an efficient way for all legacy and new POs.

B. The Factor Graph

We now assume that the measurements $z_{1:n}$ are observed and thus fixed. By using common assumptions [1]–[5], the joint posterior PDF of $\mathbf{y}_{0:n}$ and $\mathbf{b}_{1:n}$, conditioned on $z_{1:n}$ is given by[5]

$$f(\mathbf{y}_{0:n}, \mathbf{b}_{1:n}|z_{1:n})$$
$$\propto \left(\prod_{\ell=1}^{K_0} f(\mathbf{y}_{\ell,0}) \right) \prod_{n'=1}^{n} \left(\prod_{k=1}^{M_{n'}} \overline{q}(\overline{\mathbf{y}}_{k,n'})\, g_k\left(\overline{\mathbf{y}}_{k,n'}, b_{k,n'}; z_{k,n'}\right) \right.$$
$$\times \prod_{l=k+1}^{M_{n'}} h_{\underline{K}_{n'}+k}\left(\overline{\mathbf{y}}_{k,n'}, b_{l,n'}; z_{l,n'}\right) \bigg)$$
$$\times \prod_{k'=1}^{\underline{K}_{n'}} \underline{q}\left(\underline{\mathbf{y}}_{k',n'}|\mathbf{y}_{k',n'-1}\right) \prod_{l'=1}^{M_{n'}} h_{k'}\left(\underline{\mathbf{y}}_{k',n'}, b_{l',n'}; z_{l',n'}\right) \quad (11)$$

where we introduced the functions $h_k(\mathbf{y}_{k,n}, b_{l,n}; z_{l,n})$, $g_k(\overline{\mathbf{y}}_{k,n}, b_{k,n}; z_{k,n})$, $\underline{q}(\underline{\mathbf{y}}_{k,n}|\mathbf{y}_{k,n-1})$, and $\overline{q}(\overline{\mathbf{y}}_{k,n})$ that will be discussed next.

---

[5]A detailed derivation of this joint posterior PDF is provided in [56, Section 1].

The *pseudo likelihood functions* $h_k(\boldsymbol{y}_{k,n}, b_{l,n}; \boldsymbol{z}_{l,n}) = h_k(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n}, r_{k,n}, b_{l,n}; \boldsymbol{z}_{l,n})$ and $g_k(\overline{\boldsymbol{y}}_{k,n}, b_{k,n}; \boldsymbol{z}_{k,n}) = g_k(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}, \overline{r}_{k,n}, b_{k,n}; \boldsymbol{z}_{k,n})$ are given by

$$h_k\left(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n}, 1, b_{l,n}; \boldsymbol{z}_{l,n}\right)$$

$$\triangleq \begin{cases} \frac{\mu_{\mathrm{m}}(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n}) f(\boldsymbol{z}_{l,n}|\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n})}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_{l,n})}, & b_{l,n} = k \\ 1, & b_{l,n} \neq k \end{cases} \quad (12)$$

and $h_k(\boldsymbol{x}_{k,n}, \boldsymbol{e}_{k,n}, 0, b_{l,n}; \boldsymbol{z}_{l,n}) \triangleq 1 - \delta(b_{l,n} - k)$ as well as

$$g_k\left(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}, 1, b_{k,n}; \boldsymbol{z}_{k,n}\right)$$

$$\triangleq \begin{cases} \frac{\mu_{\mathrm{m}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}) f(\boldsymbol{z}_{k,n}|\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n})}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_{k,n})}, & b_{k,n} = \underline{K}_n + k \\ 0, & b_{k,n} \neq \underline{K}_n + k \end{cases} \quad (13)$$

and $g_k(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}, 0, b_{k,n}; \boldsymbol{z}_{k,n}) \triangleq 1 - \delta(b_{k,n} - (\underline{K}_n + k))$. Note that the second line in (13) is zero because, as discussed in Section II-D, the new PO with index k exists ($\overline{r}_{k,n} = 1$) if and only if it produced measurement k.

Furthermore, the *factors containing prior distributions* for new POs $\overline{q}(\overline{\boldsymbol{y}}_{k,n}) = \overline{q}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}, \overline{r}_{k,n})$, $k \in \{1, \ldots, M_n\}$ are given by

$$\overline{q}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}, \overline{r}_{k,n})$$

$$\triangleq \begin{cases} \mu_{\mathrm{n}} f_{\mathrm{n}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}) \frac{e^{-\mu_{\mathrm{m}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n})}}{1 - e^{-\mu_{\mathrm{m}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n})}}, & \overline{r}_{k,n} = 1 \\ f_{\mathrm{d}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}), & \overline{r}_{k,n} = 0 \end{cases}$$

and the *pseudo transition functions* (cf. (2) and (3)) for legacy POs $\underline{q}(\underline{\boldsymbol{y}}_{k,n}|\boldsymbol{y}_{k,n-1}) \triangleq \underline{q}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n}|\boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1})$ are given by $\underline{q}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n} = 1|\boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1}) = e^{-\mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n})} f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n} = 1|\boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1})$ and $\underline{q}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n} = 0|\boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1}) = f(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n}, \underline{r}_{k,n} = 0|\boldsymbol{x}_{k,n-1}, \boldsymbol{e}_{k,n-1}, r_{k,n-1})$.

A detailed derivation of this factor graph is provided in the supplementary material [56]. The factor graph representing factorization (11) is shown in Fig. 2. An interesting observation is that this factor graph has the same structure as a conventional multi-scan tracking problem [4], [58] with M scans. (Here, every measurement is considered an individual scan.) In what follows, we consider a single time step and remove the time index n for the sake of readability.

IV. THE PROPOSED SUM-PRODUCT ALGORITHM

Since our factor graph in Fig. 2 has cycles, we have to decide on a specific order of message computation [33]. We choose this order according to the following rules: (i) messages are not sent backward in time[6] [4], [40]; and (ii) at each time step messages are computed and processed in parallel. With these rules, the generic message passing equations of the SPA [33] yield the following operations at each time step. The corresponding messages are shown in Fig. 2.

[6]This is equivalent to a joint filtering approach with the assumption that the PO states are conditionally independent given the past measurements.
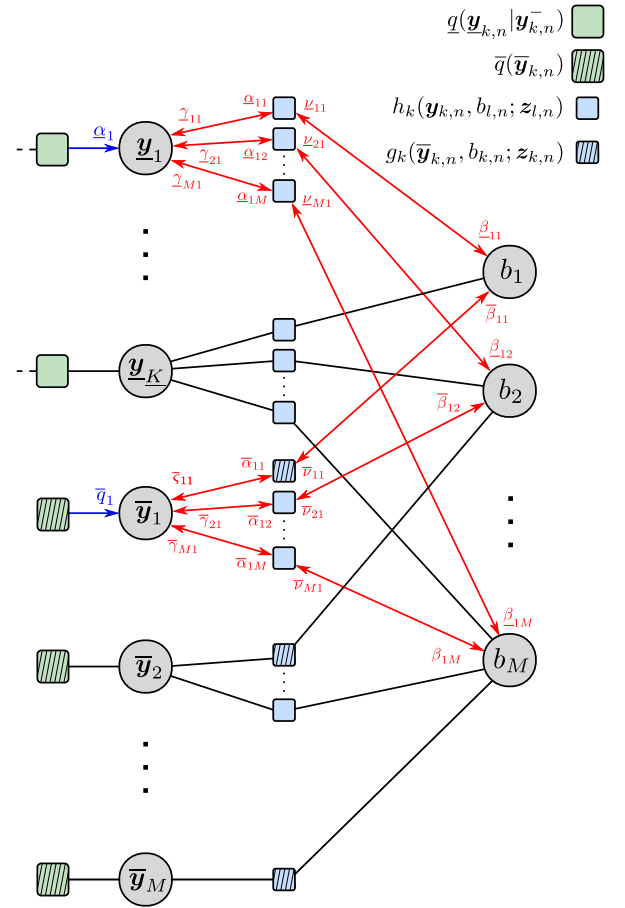

Fig. 2. Factor graphs for EOT corresponding to the factorization (11). Factor nodes and variable nodes are depicted as boxes and circles, respectively. Messages in blue are calculated only once and messages in red are calculated multiple times due to iterative message passing. The time index n is omitted. The following short notations are used for the messages: $\underline{\alpha}_k \triangleq \underline{\alpha}(\underline{\boldsymbol{y}}_k)$, $\overline{q}_k \triangleq \overline{q}(\overline{\boldsymbol{y}}_k)$, $\underline{\gamma}_{lk} \triangleq \underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{y}}_k)$, $\overline{\gamma}_{lk} \triangleq \overline{\gamma}_{lk}^{(p)}(\overline{\boldsymbol{y}}_k)$, $\overline{\varsigma}_{kk} \triangleq \overline{\varsigma}_{kk}^{(p)}(\overline{\boldsymbol{y}}_k)$, $\underline{\beta}_{kl} \triangleq \underline{\beta}_{kl}^{(p)}(b_l)$, $\overline{\beta}_{kl} \triangleq \overline{\beta}_{kl}^{(p)}(b_l)$, $\underline{\nu}_{lk} \triangleq \underline{\nu}_{lk}^{(p)}(b_l)$, $\overline{\nu}_{lk} \triangleq \overline{\nu}_{lk}^{(p)}(b_l)$, $\overline{\alpha}_{kl} = \overline{\alpha}_{kl}^{(p)}(\overline{\boldsymbol{y}}_k)$, and $\underline{\alpha}_{kl} = \underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{y}}_k)$.

A. Prediction Step

First, a *prediction* step is performed for all legacy POs $k \in \underline{K}$. Based on SPA rule [33, Eq. (6)], we obtain

$$\underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k) = \sum_{r_k^- \in \{0,1\}} \int \int \underline{q}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k | \boldsymbol{x}_k^-, \boldsymbol{e}_k^-, r_k^-)$$
$$\times \tilde{f}(\boldsymbol{x}_k^-, \boldsymbol{e}_k^-, r_k^-) \, \mathrm{d}\boldsymbol{x}_k^- \, \mathrm{d}\boldsymbol{e}_k^- \quad (14)$$

where $\tilde{f}(\boldsymbol{x}_k^-, \boldsymbol{e}_k^-, r_k^-)$ is the belief that was calculated at the previous time step. Recall that the integration $\int \mathrm{d}\boldsymbol{e}_k^-$ in (14) is performed over the support of $\boldsymbol{e}_k^-$, which corresponds to all positive-semidefinite matrices $\mathbf{E}_k^-$.

Next, we use the expression for $\underline{q}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k | \boldsymbol{x}_k^-, \boldsymbol{e}_k^-, r_k^-)$ as introduced in Section III-B and in turn (2) and (3) for $f(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k | \boldsymbol{x}_k^-, \boldsymbol{e}_k^-, r_k^-)$. In this way, we obtain the following

expressions for (14)

$$\underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1) = p_{\mathrm{s}}\, e^{-\mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_{k,n}, \underline{\boldsymbol{e}}_{k,n})} \int\!\!\int f(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k | \boldsymbol{x}_k^-, \boldsymbol{e}_k^-)$$

$$\times\, \tilde{f}(\boldsymbol{x}_k^-, \boldsymbol{e}_k^-, 1)\, \mathrm{d}\boldsymbol{x}_k^-\, \mathrm{d}\boldsymbol{e}_k^- \tag{15}$$

and $\underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 0) = \alpha_k^{\mathrm{n}} f_{\mathrm{d}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k)$ with

$$\alpha_k^{\mathrm{n}} \triangleq \tilde{f}_k^- + (1 - p_{\mathrm{s}}) \int\!\!\int \tilde{f}(\boldsymbol{x}_k^-, \boldsymbol{e}_k^-, 1)\, \mathrm{d}\boldsymbol{x}_k^-\, \mathrm{d}\boldsymbol{e}_k^-$$

$$= \tilde{f}_k^- + (1 - p_{\mathrm{s}})\big(1 - \tilde{f}_k^-\big). \tag{16}$$

We note that $\tilde{f}_k^- = \int\!\int \tilde{f}(\boldsymbol{x}_k^-, \boldsymbol{e}_k^-, 0)\, \mathrm{d}\boldsymbol{x}_k^-\, \mathrm{d}\boldsymbol{e}_k^-$ approximates the probability of non-existence of legacy PO k at the previous time step. After the prediction step, the iterative message passing is performed. For future reference, we also introduce $\underline{\alpha}_k \triangleq \int\!\int \underline{\alpha}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1)\, \mathrm{d}\underline{\boldsymbol{x}}_k\, \mathrm{d}\underline{\boldsymbol{e}}_k + \alpha_k^{\mathrm{n}}$.

B. Iterative Message Passing

At iteration $p \in \{1, \ldots, P\}$, the following operations are computed for all legacy and new POs.

1) Measurement Evaluation: The messages $\beta_{kl}^{(p)}(b_l)$, $k \in \{1, \ldots, \underline{K}\}$, $l \in \{1, \ldots, M\}$ sent from factor nodes $h_k(\underline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l) = h_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k, b_l; \boldsymbol{z}_l)$ to variable nodes b_l can be calculated as discussed next. First, by using the SPA rule [33, Eq. (6)], we obtain

$$\underline{B}_{kl}^{(p)}(b_l) = \sum_{\underline{r}_k \in \{0,1\}} \int\!\!\int h_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k, b_l; \boldsymbol{z}_l)$$

$$\times\, \alpha_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)\, \mathrm{d}\underline{\boldsymbol{x}}_k\, \mathrm{d}\underline{\boldsymbol{e}}_k. \tag{17}$$

Note that for $p = 1$, we set $\underline{\alpha}_{kl}^{(1)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k) \triangleq \underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)$, and for $p > 1$ we use $\underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)$ (represented by $\underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1)$ and $\underline{\alpha}_{kl}^{\mathrm{n}(p)}$) calculated in Section IV-B4. By using the expressions for $h_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k, b_l; \boldsymbol{z}_l)$ introduced in Section III-B, (17) can be further simplified, i.e.,

$$\underline{B}_{kl}^{(p)}(b_l = k) = \frac{1}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_l)} \int\!\!\int \mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k) f(\boldsymbol{z}_l | \underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k)$$

$$\times\, \alpha_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1)\, \mathrm{d}\underline{\boldsymbol{x}}_k\, \mathrm{d}\underline{\boldsymbol{e}}_k \tag{18}$$

and $\underline{B}_{kl}^{(p)}(b_l \neq k) = \alpha_{kl}^{(p)}$ with $\alpha_{kl}^{(p)} \triangleq \int\!\int \underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1)\, \mathrm{d}\underline{\boldsymbol{x}}_k\, \mathrm{d}\underline{\boldsymbol{e}}_k + \alpha_{kl}^{\mathrm{n}(p)}$. After multiplying these two expressions by $1/\alpha_{kl}^{(p)}$, the message $\underline{\beta}_{kl}^{(p)}(b_l)$ is given by[7]

$$\underline{\beta}_{kl}^{(p)}(b_l = k) = \frac{1}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_l) \alpha_{kl}^{(p)}} \int\!\!\int \mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k) f(\boldsymbol{z}_l | \underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k)$$

$$\times\, \alpha_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1)\, \mathrm{d}\underline{\boldsymbol{x}}_k\, \mathrm{d}\underline{\boldsymbol{e}}_k \tag{19}$$

and $\underline{\beta}_{kl}^{(p)}(b_l \neq k) = 1$. This final normalization step makes it possible to perform data association and measurement update

---

[7]Multiplying SPA messages by a constant factor does not alter the resulting approximate marginal posterior PDFs [33].

discussed in the next two sections more efficiently. For future reference, we introduce the short notation $\underline{\beta}_{kl}^{(p)} \triangleq \underline{\beta}_{kl}^{(p)}(b_l = k)$ for all $k \in \{1, \ldots, \underline{K}\}$, $l \in \{1, \ldots, M\}$.

The messages $\overline{\beta}_{kl}^{(p)}(b_l)$, $k \in \{1, \ldots, M - 1\}$, $l \in \{k + 1, \ldots, M\}$ sent from factor nodes $h_k(\overline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l)$ to variable nodes b_l can be obtained similarly. In particular, by replacing $\underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1)$ and $\underline{\alpha}_{kl}^{(p)}$ in (19) by $\overline{\alpha}_{kl}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k = 1)$ and $\overline{\alpha}_{kl}^{(p)}$, respectively, we obtain $\overline{\beta}_{kl}^{(p)}(b_l = \underline{K} + k)$. Furthermore, we have $\overline{\beta}_{kl}^{(p)}(b_l \neq \underline{K} + k) = 1$. Note that for $p = 1$ we set $\overline{\alpha}_{kl}^{(1)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k) \triangleq q(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k)$ and for $p > 1$ we again calculate $\overline{\alpha}_{kl}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k)$ as discussed in Section IV-B4.

Finally, the messages $\overline{\beta}_{kk}^{(p)}(b_k)$, $k \in \{1, \ldots, M\}$ sent from factor nodes $g_k(\overline{\boldsymbol{y}}_k, b_k; \boldsymbol{z}_k) = g_k(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k, b_k; \boldsymbol{z}_k)$ to variable nodes b_k are calculated by also replacing $h_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k, b_l; \boldsymbol{z}_l)$ in (17) by $g_k(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k, b_k; \boldsymbol{z}_k)$ and performing similar simplification steps. In particular, we get $\overline{\alpha}_{kk}^{\mathrm{n}(p)} \triangleq \int\!\int \overline{\alpha}_{kk}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k = 0)\, \mathrm{d}\overline{\boldsymbol{x}}_k\, \mathrm{d}\overline{\boldsymbol{e}}_k$ for $b_k \neq \underline{K} + k$ and a result equal to (18) (with l replaced by k and $\underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)$ replaced by $\overline{\alpha}_{kk}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k)$) for $b_k \neq k$. After multiplying both expressions by $1/\overline{\alpha}_{kk}^{\mathrm{n}(p)}$, the message $\overline{\beta}_{kk}^{(p)}(b_k)$ finally reads

$$\overline{\beta}_{kk}^{(p)}(b_k = \underline{K} + k) = \frac{1}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_k) \overline{\alpha}_{kk}^{\mathrm{n}(p)}} \int\!\!\int f(\boldsymbol{z}_k | \overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k)$$

$$\times\, \mu_{\mathrm{m}}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k) \overline{\alpha}_{kk}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k = 1)\, \mathrm{d}\overline{\boldsymbol{x}}_k\, \mathrm{d}\overline{\boldsymbol{e}}_k \tag{20}$$

and $\overline{\beta}_{kk}^{(p)}(b_k \neq \underline{K} + k) = 1$. For future reference, we again introduce the short notation $\overline{\beta}_{kl}^{(p)} \triangleq \overline{\beta}_{kl}^{(p)}(b_l = \underline{K} + k)$ for all $k \in \{1, \ldots, M\}$, $l \in \{k, \ldots, M\}$.

2) Data Association: The messages $\underline{\nu}_{lk}^{(p)}(b_l)$ sent from variable nodes b_l, $l \in \{1, \ldots, M\}$ to factor nodes $h_k(\boldsymbol{y}_k, b_l; \boldsymbol{z}_l)$, $k \in \{1, \ldots, \underline{K}\}$, can be expressed as [33, Eq. (5)]

$$\underline{\nu}_{lk}^{(p)}(b_l) = \left(\prod_{\substack{k'=1 \\ k' \neq k}}^{K} \underline{\beta}_{k'l}^{(p)}(b_l) \right) \prod_{\ell=1}^{l} \overline{\beta}_{\ell l}^{(p)}(b_l). \tag{21}$$

By using (19) and (20) in (21), we obtain $\underline{\nu}_{lk}^{(p)}(b_l = 0) = \underline{\nu}_{lk}^{(p)}(b_l = k) = 1$, $\underline{\nu}_{lk}^{(p)}(b_l = k') = \underline{\beta}_{k'l}^{(p)}$, $k' \in \{1, \ldots, \underline{K}\} \backslash \{k\}$, and $\underline{\nu}_{lk}^{(p)}(b_l = \underline{K} + \ell) = \overline{\beta}_{\ell l}^{(p)}$, $\ell \in \{1, \ldots, l\}$.

A similar expression is obtained for the messages $\overline{\nu}_{lk}^{(p)}(b_l)$ sent from variable nodes b_l, $l \in \{1, \ldots, M\}$ to factor nodes $h_{\underline{K}+k}(\boldsymbol{y}_k, b_l; \boldsymbol{z}_l)$, $k \in \{1, \ldots, l-1\}$ and factor node $g_k(\underline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_k)$, $k = l$, i.e.,

$$\overline{\nu}_{lk}^{(p)}(b_l) = \left(\prod_{\ell=1}^{K} \underline{\beta}_{\ell l}^{(p)}(b_l) \right) \prod_{\substack{k'=1 \\ k' \neq k}}^{l} \overline{\beta}_{k'l}^{(p)}(b_l). \tag{22}$$

By again using (19) and (20) in (22), we obtain $\overline{\nu}_{lk}^{(p)}(b_l = 0) = \overline{\nu}_{lk}^{(p)}(b_l = \underline{K} + k) = 1$, $\overline{\nu}_{lk}^{(p)}(b_l = k') = \underline{\beta}_{k'l}^{(p)}$, $k' \in \{1, \ldots, \underline{K}\}$, and $\overline{\nu}_{lk}^{(p)}(b_l = \underline{K} + \ell) = \overline{\beta}_{\ell l}^{(p)}$, $\ell \in \{1, \ldots, l\} \backslash \{k\}$.

3) Measurement Update: Next, the messages sent from factor nodes $h(\underline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l)$, $k \in \{1, \ldots, \underline{K}\}$, $l \in \{1, \ldots, M\}$ to variable nodes $\underline{\boldsymbol{y}}_k$ are calculated as [33, Eq. (6)]

$$\underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{y}}_k) = \sum_{b_l=0}^{K+l} h(\underline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l)\, \underline{\nu}_{lk}^{(p)}(b_l). \tag{23}$$

By using again the expression for $h(\underline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l)$ introduced in Section III-B, these messages $\underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{y}}_k) \triangleq \underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)$ can be further simplified as

$$\underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 0) = \sum_{\substack{b_l=0 \\ b_l \neq k}}^{K+l} \underline{\nu}_{lk}^{(p)}(b_l)$$

$$\underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1) = \frac{\mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k) f(\boldsymbol{z}_l | \underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k)}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_l)}\, \underline{\nu}_{lk}^{(p)}(b_l = k)$$

$$+ \sum_{\substack{b_l=0 \\ b_l \neq k}}^{K+l} \underline{\nu}_{lk}^{(p)}(b_l). \tag{24}$$

By using the simplification of (21) discussed in the previous Section IV-B2, we can rewrite (24) as

$$\underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 0) = \underline{\xi}_{kl}^{(p)}$$

$$\underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1) = \frac{\mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k) f(\boldsymbol{z}_l | \underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k)}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_l)} + \underline{\xi}_{kl}^{(p)}$$
$$\tag{25}$$

where we introduced the sum

$$\underline{\xi}_{kl}^{(p)} \triangleq \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \underline{\beta}_{k'l}^{(p)} + \sum_{\ell=1}^{l} \overline{\beta}_{\ell l}^{(p)} + 1. \tag{26}$$

A similar result can be obtained for the messages $\overline{\gamma}_{lk}^{(p)}(\overline{\boldsymbol{y}}_k)$ sent from factor nodes $h(\overline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l)$, $k \in \{1, \ldots, M-1\}$, $l \in \{k+1, \ldots, M\}$ to variable nodes $\overline{\boldsymbol{y}}_k$. In particular, these messages are obtained by replacing in (23) the functions $h_k(\underline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l)$ and $\underline{\nu}_{lk}^{(p)}(b_l)$ by $h_k(\overline{\boldsymbol{y}}_k, b_l; \boldsymbol{z}_l)$ and $\overline{\nu}_{lk}^{(p)}(b_l)$ as well as by performing the same simplification step discussed above.

By performing similar steps, we can then also obtain the messages $\varsigma_{kk}^{(p)}(\overline{\boldsymbol{y}}_k) \triangleq \varsigma_{kk}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k)$, $k \in \{1, \ldots, M\}$ sent from $g_k(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k, b_k; \boldsymbol{z}_k)$ to new PO state $\overline{\boldsymbol{y}}_k$ as

$$\varsigma_{kk}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k = 0) = \overline{\xi}_{kk}^{(p)}$$

$$\varsigma_{kk}^{(p)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k = 1) = \frac{\mu_{\mathrm{m}}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k) f(\boldsymbol{z}_k | \overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k)}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_k)}. \tag{27}$$

The calculation of $\overline{\gamma}_{lk}^{(p)}(\overline{\boldsymbol{y}}_k)$, $k \in \{1, \ldots, M-1\}$, $l \in \{k+1, \ldots, M\}$ and $\varsigma_{kk}^{(p)}(\overline{\boldsymbol{y}}_k)$, $k \in \{1, \ldots, M\}$ is based on the sum

$$\overline{\xi}_{kl}^{(p)} \triangleq \sum_{k'=1}^{K} \underline{\beta}_{k'l}^{(p)} + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{l} \overline{\beta}_{\ell l}^{(p)} + 1.$$

4) Extrinsic Information: Finally, updated messages used in the next message passing iteration $p+1$ for legacy POs $k \in \{1, \ldots, \underline{K}\}$ are obtained as [33, Eq. (5)]

$$\underline{\alpha}_{kl}^{(p+1)}(\underline{\boldsymbol{y}}_k) = \underline{\alpha}(\underline{\boldsymbol{y}}_k) \prod_{\substack{l'=1 \\ l' \neq l}}^{M} \underline{\gamma}_{l'k}^{(p)}(\underline{\boldsymbol{y}}_k). \tag{28}$$

Note that for non-existent objects, i.e., $\underline{r}_k = 0$, the messages $\underline{\alpha}_{kl}^{(p+1)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k) \triangleq \underline{\alpha}_{kl}^{(p+1)}(\underline{\boldsymbol{y}}_k)$ have the form $\underline{\alpha}_{kl}^{(p+1)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 0) = \underline{\alpha}_{kl}^{\mathrm{n}(p+1)} f_{\mathrm{d}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k)$.

Similarly for new POs $k \in \{1, \ldots, M\}$, we obtain

$$\overline{\alpha}_{kl}^{(p+1)}(\overline{\boldsymbol{y}}_k) = \overline{q}(\overline{\boldsymbol{y}}_k)\, \varsigma_{kk}^{(p)}(\overline{\boldsymbol{y}}_k) \prod_{\substack{l'=k+1 \\ l' \neq l}}^{M} \overline{\gamma}_{l'k}^{(p)}(\overline{\boldsymbol{y}}_k) \tag{29}$$

for $l \in \{k+1, \ldots, M\}$ and $\overline{\alpha}_{kl}^{(p+1)}(\overline{\boldsymbol{y}}_k) = \overline{q}(\overline{\boldsymbol{y}}_k) \prod_{l'=k+1}^{M} \overline{\gamma}_{l'k}^{(p)}(\overline{\boldsymbol{y}}_k)$ for $l = k$. Again, for $\overline{r}_k = 0$, the messages $\overline{\alpha}_{kl}^{(p+1)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k) \triangleq \overline{\alpha}_{kl}^{(p+1)}(\overline{\boldsymbol{y}}_k)$ have the form $\overline{\alpha}_{kl}^{(p+1)}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k = 0) = \overline{\alpha}_{kl}^{\mathrm{n}(p+1)} f_{\mathrm{d}}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k)$.

C. Belief Calculation

After the last iteration $p = P$, the belief $\tilde{f}(\underline{\boldsymbol{y}}_k) \triangleq \tilde{f}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)$ of legacy PO state $k \in \{1, \ldots, \underline{K}\}$ can be calculated as the normalized product of all incoming messages [33], i.e.,

$$\tilde{f}(\underline{\boldsymbol{y}}_k) = \underline{C}_k\, \underline{\alpha}(\underline{\boldsymbol{y}}_k) \prod_{l=1}^{M} \underline{\gamma}_{lk}^{(P)}(\underline{\boldsymbol{y}}_k) \tag{30}$$

where the normalization constant (cf. (16) and (25)) reads

$$\underline{C}_k = \left(\int \underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1) \right.$$
$$\left. \prod_{l=1}^{M} \underline{\gamma}_{lk}^{(P)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1) \mathrm{d}\underline{\boldsymbol{x}}_k\, \mathrm{d}\underline{\boldsymbol{e}}_k + \underline{\alpha}_k^{\mathrm{n}} \prod_{l=1}^{M} \underline{\xi}_{kl}^{(P)} \right)^{-1}. \tag{31}$$

Similarly, the belief $b(\overline{\boldsymbol{y}}_k) \triangleq b(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k, \overline{r}_k)$ of augmented new PO state $k = \{1, \ldots, M\}$, is given by

$$\tilde{f}(\overline{\boldsymbol{y}}_k) = \overline{C}_k\, \overline{q}(\overline{\boldsymbol{y}}_k)\, \varsigma_{kk}^{(P)}(\overline{\boldsymbol{y}}_k) \prod_{l=k+1}^{M} \overline{\gamma}_{lk}^{(P)}(\overline{\boldsymbol{y}}_k). \tag{32}$$

Here, $\overline{C}_k$ is again the normalization constant that guarantees that (32) is a valid probability distribution.

Note that a message passing order where messages are calculated sequentially and for each measurement individually is discussed in [55]. As demonstrated in [55, Section V], parallel processing leads to improved performance compared to sequential processing.

D. Computational Complexity and Scalability

In the prediction step, (18) has to be performed $\underline{K}$ times. Thus, its computational complexity scales as $\mathcal{O}(\underline{K})$. The computational complexity related to each message passing iteration $p \in \{1, \ldots, P\}$ is discussed next. In the measurement evaluation

step, for legacy PO $k \in \{1, \ldots, \underline{K}\}$, a total of M messages $\underline{\beta}_{kl}^{(p)}(b_l)$ have to be calculated. Similarly, for every new PO $k \in \{1, \ldots, M\}$, a total of $l \in \{k, \ldots, M\}$ messages $\overline{\beta}_{kl}^{(p)}(b_l)$ has to be obtained. Thus, the total number of messages is $\underline{K}M + 1/2M^2$. The computational complexity related to calculating each individual message is constant in $\underline{K}$ and M. Also in the data association, measurement update, and extrinsic information steps, a total of $\underline{K}M + 1/2\,M^2$ messages have to be calculated at each of the three steps. The computational complexity related to the calculation of the individual messages is again constant in $\underline{K}$ and M. For the data association and measurement update steps, this constant computational complexity is obtained by pre-computing the sums $\sum_{k=1}^{\underline{K}} \underline{\beta}_{kl}^{(p)}(k) + \sum_{\ell=1}^{l} \overline{\beta}_{\ell l}^{(p)}(\underline{K}+\ell) + 1$ for each $l \in \{1, \ldots, M\}$. Similarly, for the extrinsic information step, this constant computational complexity is obtained by precomputing the products $\prod_{l=1}^{M} \underline{\gamma}_{lk}^{(p)}(\underline{\boldsymbol{y}}_k), k \in \{1, \ldots, \underline{K}\}$ and $\prod_{l=k+1}^{M} \overline{\gamma}_{lk}^{(p)}(\overline{\boldsymbol{y}}_k), k \in \{1, \ldots, M\}$.

For P fixed, it thus has been verified that the overall computational complexity only scales as $\mathcal{O}(\underline{K}M + M^2)$. (Here we have omitted $1/2$ since the $\mathcal{O}(\cdot)$ notation does not track constants.) Recalling that $K = \underline{K} + M$ is the total number of legacy and new POs, we can equivalently write $\mathcal{O}(KM)$. Conservatively, we consider this to be quadratic in the number of measurements and objects, since existing objects are also expected to contribute measurements. We observed that increasing the number of message passing iterations beyond $P = 3$, does not significantly improve performance in typical EOT scenarios. Note that the computational complexity can be further reduced by preclustering of measurements (see, e.g., [7, Section IV]) and censoring of messages (see, e.g., [55, Section IV]). Preclustering combines the M measurements to a smaller number $M' < M$ of joint "hyper measurements" and replaces the single measurement ratios in (12) and (13) by the corresponding product of ratios. Censoring aims to omit messages related to new POs that are unlikely to represent an actual object.

V. PARTICLE-BASED IMPLEMENTATION

For general state evolution and measurement models, the integrals in (15)–(18) as well as the message products in (28)–(32) typically cannot be evaluated in closed form. Therefore, we next present an approximate particle-based implementation of these operations that can be seen as a generalization of the particle-based implementation introduced in [40] to EOT. Pseudocode for the presented particle-based implementation is provided in [56, Section 3]. Each belief $\tilde{f}(\boldsymbol{y}_k) \triangleq \tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, r_k)$ is represented by a set of particles and corresponding weights $\{(\boldsymbol{x}_k^{(j)}, \boldsymbol{e}_k^{(j)}, w_k^{(j)})\}_{j=1}^{J}$. More specifically, $\tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, 1)$ is represented by $\{(\boldsymbol{x}_k^{(j)}, \boldsymbol{e}_k^{(j)}, w_k^{(j)})\}_{j=1}^{J}$, and $\tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, 0)$ is given implicitly by the normalization property of $\tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, r_k)$, i.e., $\tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, 0) = 1 - \iint \tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, 1) \mathrm{d}\boldsymbol{x}_k \mathrm{d}\boldsymbol{e}_k$. Contrary to conventional particle filtering [59], [60] and as in [40], the particle weights $w_k^{(j)}, j \in \{1, \ldots, J\}$ do not sum to one; instead,

$$p_k^{\mathrm{e}} \triangleq \sum_{j=1}^{J} w_k^{(j)} \approx \iint \tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, 1) \mathrm{d}\boldsymbol{x}_k \mathrm{d}\boldsymbol{e}_k. \quad (33)$$

Note that since $\iint \tilde{f}(\boldsymbol{x}_k, \boldsymbol{e}_k, 1) \mathrm{d}\boldsymbol{x}_k \mathrm{d}\boldsymbol{e}_k$ approximates the posterior probability of existence for the object, it follows that the sum of weights p_k^{e} is approximately equal to the posterior probability of existence.

A. Prediction

The particle operations discussed in this section are performed for all legacy POs $k \in \{1, \ldots, \underline{K}\}$ in parallel. For each legacy PO k, J particles and weights $\{(\boldsymbol{x}_k^{-(j)}, \boldsymbol{e}_k^{-(j)}, w_k^{-(j)})\}_{j=1}^{J}$ representing the previous belief $\tilde{f}(\boldsymbol{x}_k^-, \boldsymbol{e}_k^-, r_k^-)$ were calculated at the previous time $n-1$ as described further below. Weighted particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_k^{(1,j)})\}_{j=1}^{J}$ representing the message $\underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, 1)$ in (15) are now obtained as follows. First, for each particle $(\boldsymbol{x}_k^{-(j)}, \boldsymbol{e}_k^{-(j)})$, $j \in \{1, \ldots, J\}$, one particle $(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)})$ is drawn from $f(\boldsymbol{x}_k, \boldsymbol{e}_k | \boldsymbol{x}_k^{-(j)}, \boldsymbol{e}_k^{-(j)})$. Next, corresponding weights $w_k^{(1,j)}, j \in \{1, \ldots, J\}$ are obtained as

$$\underline{w}_k^{(1,j)} = p_{\mathrm{s}} \, e^{-\mu_{\mathrm{m}}\left(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}\right)} \, w_k^{-(j)}, j \in \{1, \ldots, J\}. \quad (34)$$

Note that the proposal distribution [59], [60] underlying (34) is $f(\boldsymbol{x}_k, \boldsymbol{e}_k | \boldsymbol{x}_k^{-(j)}, \boldsymbol{e}_k^{-(j)})$ for $j \in \{1, \ldots, J\}$. A particle-based approximation $\underline{\tilde{\alpha}}_k^{\mathrm{n}}$ of $\underline{\alpha}_k^{\mathrm{n}}$ in (16) is now obtained as

$$\underline{\tilde{\alpha}}_k^{\mathrm{n}} = \left(1 - p_k^{-\mathrm{e}}\right) + (1 - p_{\mathrm{s}}) \, p_k^{-\mathrm{e}} \quad (35)$$

where $p_k^{-\mathrm{e}} \triangleq \sum_{j=1}^{J} w_k^{-(j)}$. Finally, a particle approximation $\underline{\tilde{\alpha}}_k$ of $\underline{\alpha}_k$ introduced in Section IV-A is given by

$$\underline{\tilde{\alpha}}_k = \sum_{j=1}^{J} \underline{w}_k^{(1,j)} + \underline{\tilde{\alpha}}_k^{\mathrm{n}}.$$

B. Measurement Evaluation

Let the weighted particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_{kl}^{(p,j)})\}_{j=1}^{J}$ and the scalar $\underline{\tilde{\alpha}}_{kl}^{(p)}$ be a particle-based representation of $\underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, r_k)$. For $p = 1$, we set $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_{kl}^{(1,j)})\}_{j=1}^{J} \triangleq \{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_k^{(1,j)})\}_{j=1}^{J}$ and $\underline{\tilde{\alpha}}_{kl}^{(p)} \triangleq \underline{\tilde{\alpha}}_k$; for $p > 1$ this representation is calculated as discussed in Section V-C. An approximation $\underline{\tilde{\beta}}_{kl}^{(p)}$ of the message value $\underline{\beta}_{kl}^{(p)} \triangleq \underline{\beta}_{kl}^{(p)}(b_l = k)$ in (19), can now be obtained as

$$\underline{\tilde{\beta}}_{kl}^{(p)} = \frac{1}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_l) \, \underline{\tilde{\alpha}}_{kl}^{(p)}} \sum_{j=1}^{J} \underline{w}_{kl}^{(p,j)} \mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}) f(\boldsymbol{z}_l | \underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}).$$

Here, $\sum_{j=1}^{J} \underline{w}_{kl}^{(p,j)} \mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}) f(\boldsymbol{z}_l | \underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)})$ is the Monte Carlo integration [60] of $\iint \mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k) f(\boldsymbol{z}_l | \underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k) \underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1) \mathrm{d}\underline{\boldsymbol{x}}_k \mathrm{d}\underline{\boldsymbol{e}}_k$ in (17). This Monte Carlo integration is based on the proposal distribution $\underline{\alpha}_{kl}^{(p)}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k = 1)$. Similarly, an approximation $\overline{\tilde{\beta}}_{kl}^{(p)}$ of the message values $\overline{\beta}_{kl}^{(p)}$ related to new POs can be obtained. Here, for Monte Carlo integration and further particle-based processing, it was found useful to use a proposal distribution $f_{\mathrm{p}}(\overline{\boldsymbol{x}}_k, \overline{\boldsymbol{e}}_k; \boldsymbol{z}_k)$ that is calculated from the measurement $\boldsymbol{z}_k$ and

its uncertainty characterization, e.g., by means of the unscented transformation [61].

Note that calculation of these messages relies on the likelihood function $f(\boldsymbol{z}_l|\boldsymbol{x}_k, \boldsymbol{e}_k)$ introduced in (6), which involves the integration $\int d\boldsymbol{v}_{k,n}^{(l)}$. For general nonlinear and non-Gaussian measurement models, evaluation of the likelihood function $f(\boldsymbol{z}_l|\boldsymbol{x}_k, \boldsymbol{e}_k)$ can potentially also be performed by means of Monte Carlo integration [60]. Alternatively, if the measurement model $\boldsymbol{d}(\cdot)$ is invertible in the sense that we can reformulate (5) as

$$\boldsymbol{d}^{-1}\left(\boldsymbol{z}_l + \mathbf{u}_l\right) = \mathbf{p}_k + \mathbf{v}_k^{(l)} \tag{36}$$

then an approximated linear-Gaussian measurement model (7) can be obtained. In particular, the PDF of $\boldsymbol{d}^{-1}(\boldsymbol{z}_l + \mathbf{u}_l)$ (for observed $\boldsymbol{z}_l$) is approximated by a Gaussian with mean $\tilde{z}_l = \boldsymbol{d}^{-1}(\boldsymbol{z}_l + \boldsymbol{\mu}_{\mathbf{u}_l})$ and covariance matrix $\boldsymbol{\Sigma}_{\tilde{z}_l}$. This approximation can be obtained, e.g., by linearizing $\boldsymbol{d}^{-1}(\cdot)$ or by applying the unscented transformation [61]. As a result, closed-form expressions for $f(\boldsymbol{z}_l|\boldsymbol{x}_k, \boldsymbol{e}_k)$ discussed in Section II-B and [56, Section. 2] can be used.

C. Measurement Update, Belief Calculation, and Extrinsic Informations

The approximate measurement evaluation messages discussed in Section V-B are used for the subsequent approximation of $\underline{\xi}_{kl}^{(p)}$ and $\overline{\xi}_{kl}^{(p)}$ (cf. (26)) required in the measurement update step (cf. (25) and (27)). The calculation of the weighted particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_k^{(j)})\}_{j=1}^J$ that represent the legacy PO belief in (30) is discussed next. Weighted particles representing new PO beliefs (32) and extrinsic information messages in (28) and (29) can be obtained by performing similar steps.

The measurement update step (25) and the belief calculation step (30) are implemented by means of importance sampling [59], [60]. To that end, we first rewrite the belief $\tilde{f}(\underline{\boldsymbol{y}}_k) = \tilde{f}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)$ in (30) by using (15), (16), and (25), i.e.,

$$\underline{\tilde{f}}_k \propto \underline{\alpha}_k^{\mathrm{n}} \prod_{l=1}^M \tilde{\underline{\xi}}_{kl}^{(P)}$$

$$\tilde{f}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, 1) \propto \underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, 1)$$
$$\times \prod_{l=1}^M \left(\frac{\mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k) f(\boldsymbol{z}_l|\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k)}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_l)} + \tilde{\underline{\xi}}_{kl}^{(P)} \right). \tag{37}$$

Here, we also replaced $\xi_{kl}^{(P)}$ introduced in (26) by its particle-based approximation $\tilde{\xi}_{kl}^{(P)}$ (cf. Section V-B), even though we do not indicate this additional approximation in our notation $\tilde{f}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, \underline{r}_k)$.

We now calculate nonnormalized weights corresponding to (37) for each particle $j \in \{1, \ldots, J\}$ as

$$\underline{w}_k^{\mathrm{A}(j)} = \underline{w}_k^{(1,j)} \prod_{l=1}^M \left(\frac{\mu_{\mathrm{m}}(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}) f(\boldsymbol{z}_l|\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)})}{\mu_{\mathrm{fa}} f_{\mathrm{fa}}(\boldsymbol{z}_l)} + \tilde{\underline{\xi}}_{kl}^{(P)} \right). \tag{38}$$

Note that here we perform importance sampling with proposal density $\underline{\alpha}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, 1)$. This proposal density is represented by the weighted particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_k^{(1,j)})\}_{j=1}^J$. Similarly, we calculate a single nonnormalized weight corresponding to (37) as

$$\underline{w}_k^{\mathrm{B}} = \tilde{\underline{\alpha}}_k^{\mathrm{n}} \prod_{l=1}^M \tilde{\underline{\xi}}_{kl}^{(P)} \tag{39}$$

in which $\tilde{\underline{\alpha}}_k^{\mathrm{n}}$ has been calculated in (35).

Next, weighted particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_k^{(j)})\}_{j=1}^J$ representing the belief $\tilde{f}(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{e}}_k, r_k)$ are obtained by reusing the particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)})\}_{j=1}^J$ and calculating the corresponding new weights as

$$\underline{w}_k^{(j)} = \frac{\underline{w}_k^{\mathrm{A}(j)}}{\underline{w}_k^{\mathrm{B}} + \sum_{j'=1}^J \underline{w}_k^{\mathrm{A}(j')}}. \tag{40}$$

Here, $\underline{w}_k^{\mathrm{B}} + \sum_{j=1}^J \underline{w}_k^{\mathrm{A}(j)}$ is a particle-based approximation of the normalization constant $\underline{C}_k$ in (31). We recall that $p_k^{\mathrm{e}} = \sum_{j=1}^J \underline{w}_k^{(j)}$.

Next, we discuss the calculation of the particle representations of the extrinsic information messages in (28) and (29). For example, weighted particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)}, \underline{w}_{kl'}^{(p+1,j)})\}_{j=1}^J$ representing $\underline{\alpha}_{kl'}^{(p+1)}(\underline{\boldsymbol{y}}_k)$ can be obtained as follows. The particle locations and extents $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)})\}_{j=1}^J$ are again reused and the corresponding new weights $\underline{w}_{kl'}^{(p+1,j)}$ are obtained by replacing in (38) $\prod_{l=1}^M$ with $\prod_{\substack{l=1 \\ l \neq l'}}^M$ and $\tilde{\underline{\xi}}_{kl}^{(P)}$ with $\tilde{\underline{\xi}}_{kl}^{(p)}$. This is equal to plugging (25) into (28) and evaluating (28) at the particles $\{(\underline{\boldsymbol{x}}_k^{(j)}, \underline{\boldsymbol{e}}_k^{(j)})\}_{j=1}^J$. Here, normalization of the weights in (40) can be avoided and the constant $\tilde{\underline{\alpha}}_{kl'}^{(p+1)}$ is obtained as $\tilde{\underline{\alpha}}_{kl'}^{(p+1)} = \tilde{\underline{\alpha}}_k \prod_{\substack{l=1 \\ l \neq l'}}^M \tilde{\underline{\xi}}_{kl}^{(p)}$ (cf. (39)).

D. Detection, Estimation, Pruning, and Resampling

The weighted particles $\{(\boldsymbol{x}_k^{(j)}, \boldsymbol{e}_k^{(j)}, w_k^{(j)})\}_{j=1}^J$ can now be used for object detection and estimation. First, for each (legacy or new) PO k, an approximation p_k^{e} of the existence probability $p(r_k = 1|\boldsymbol{z})$ is calculated from the particle weights $\{w_k^{(j)}\}_{j=1}^J$ as in (33). PO k is then detected (i.e., considered to exist) if p_k^{e} is above a threshold P_{th} (cf. Section III-A). For the detected objects k, an approximation of the MMSE estimate $\hat{\boldsymbol{x}}_k^{\mathrm{MMSE}}$ of the kinematic state in (8) is calculated according to

$$\hat{\boldsymbol{x}}_k = \frac{1}{p_k^{\mathrm{e}}} \sum_{j=1}^J w_k^{(j)} \boldsymbol{x}_k^{(j)}. \tag{41}$$

Similarly, an MMSE estimate $\hat{\boldsymbol{e}}_k$ of the extent state can be obtained by replacing $\boldsymbol{x}_k^{(j)}$ in (41) by $\boldsymbol{e}_k^{(j)}$.

As discussed in Section II-C, the number of POs would grow with time k. Therefore, legacy and new POs whose approximate existence probabilities p_k^{e} are below a threshold P_{pr} are pruned [4], [15] from the state space. In addition, a resampling step may be performed to avoid particle degeneracy [59], [60].
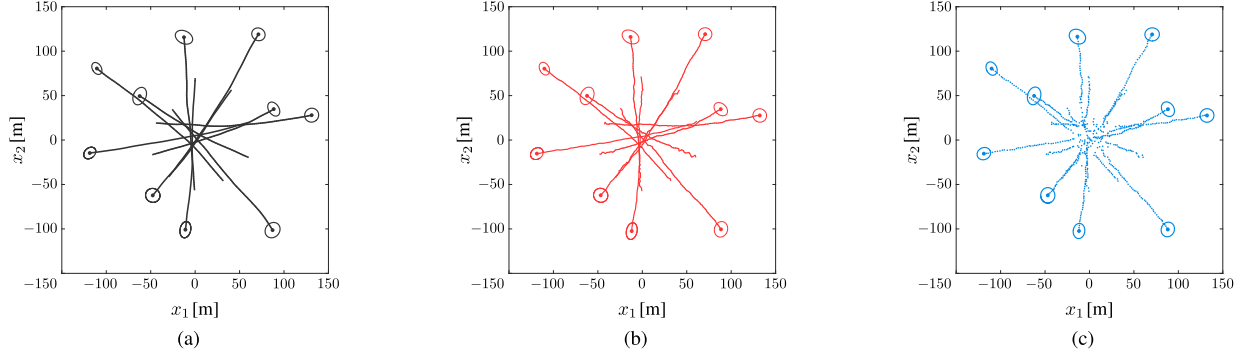
Fig. 3. Example realization of true object tracks are depicted in (a). Estimation results of the proposed SPA and the PMBM-FR are depicted in (b) and (c), respectively. True and estimated object shapes at the last time step of an object's existence are also shown. In (a), track estimates provided by the proposed SPA method are shown as red line. In (b), state estimates provided by PMBM-FR are shown as blue dots.

VI. NUMERICAL RESULTS

Next, we report simulation results evaluating the performance of our method and comparing it with that of the PMBM filter implementation in [1]. Note that a performance comparison with other data association algorithms based on the SPA has been presented in [55].

A. Simulation Scenario

We simulated ten extended objects whose states consist of two-dimensional (2-D) position and velocity, i.e., $\mathbf{x}_{k,n} = [\mathsf{p}_{k,n}^{(1)} \ \mathsf{p}_{k,n}^{(2)} \ \dot{\mathsf{p}}_{k,n}^{(1)} \ \dot{\mathsf{p}}_{n,k}^{(2)}]^{\mathsf{T}}$. The objects move in a region of interest (ROI) defined as $[-150\,\mathrm{m},\,150\,\mathrm{m}] \times [-150\,\mathrm{m},\,150\,\mathrm{m}]$ and according to the nearly constant-velocity motion model, i.e., $\mathbf{x}_{k,n} = \mathbf{A}\mathbf{x}_{k,n-1} + \mathbf{W}\mathbf{c}_{k,n}$, where $\mathbf{A} \in \mathbb{R}^{4 \times 4}$ and $\mathbf{W} \in \mathbb{R}^{4 \times 2}$ are chosen as in [62, Section 6.3.2] with $T = 0.2\,\mathrm{s}$, and $\mathbf{c}_{k,n} \sim \mathcal{N}(\mathbf{c}_{k,n}; \mathbf{0}, \sigma_{\mathsf{c}}^2 \mathbf{I}_2)$ with $\sigma_{\mathsf{c}} = 1\,\mathrm{m/s^2}$ is an independent and identically distributed (iid) sequence of 2-D Gaussian random vectors.

We considered a challenging scenario where the ten object tracks intersect at the ROI center. The object tracks were generated by first assuming that the ten objects start moving toward the ROI center from positions uniformly placed on a circle of radius 75m about the ROI center, with an initial speed of 10m/s, and then letting the objects start to exist (i.e., generate measurements) in pairs at times $n \in \{3, 6, 9, 12, 15\}$. Object tracks intersect at the ROI center at around time 40 and disappear in pairs at times $n \in \{83, 86, 89, 92, 95\}$. The extent of each object is obtained by drawing a sample from the inverse Wishart distribution with mean matrix $3\mathbf{I}_2$ and 100 degrees of freedom. The extent state of objects does not evolve in time, i.e., it remains unchanged for all time steps. The survival probability is $p_{\mathsf{s}} = 0.99$. Example realizations of object tracks and extents are shown in Fig. 3(a).

Since the PMBM filter is based on the Gaussian inverse Wishart model [5], we consider elliptical object shapes and a linear-Gaussian measurement model. In particular, a measurement l that was originated by object k, is modeled as

$$\mathbf{z}_{l,n} = \mathbf{p}_{k,n} + \mathbf{v}_{k,n}^{(l)} + \mathbf{u}_{l,n} \qquad (42)$$

where $\mathbf{u}_{l,n} \sim \mathcal{N}(\boldsymbol{u}_{l,n}; \mathbf{0}, \sigma_{\mathsf{u}}^2 \mathbf{I}_2)$ is the measurement noise with $\sigma_{\mathsf{u}} = 1\mathrm{m}$. In addition, $\mathbf{v}_{k,n}^{(l)} \sim \mathcal{N}(\boldsymbol{v}_{k,n}^{(l)}; \mathbf{0}, \boldsymbol{\Sigma}_{\mathsf{v}})$ is the random relative position of the reflection point with $\boldsymbol{\Sigma}_{\mathsf{v}}$ determined by the extent state. The mean of the number of measurements $\mathsf{L}_{k,n}$ is $\mu_{\mathsf{m}} = 8$ for all objects and the mean number of false alarm measurements is $\mu_{\mathsf{fa}} = 10$. The false alarm PDF $f_{\mathsf{fa}}(\boldsymbol{z}_{l,n})$ is uniform on the ROI.

The performance of all simulated methods is measured by the optimal subpattern assignment (OSPA) [63] and the generalized OSPA (GOSPA) [64] metrics. Both metrics are based on the Gaussian-Wasserstein distance with parameters $p = 1$ and $c = 20$. The threshold for object declaration is $P_{\mathsf{th}} = 0.5$ and the threshold for pruning POs is $P_{\mathsf{th}} = 10^{-3}$. These parameters are used for all simulated methods.

B. Performance Comparison With the PMBM Filter

For the proposed SPA-based method we use the following settings. Newly detected objects are modeled as $f_{\mathsf{n}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}) = f_{\mathsf{n}}(\overline{\boldsymbol{p}}_{k,n}) f_{\mathsf{n}}(\overline{\boldsymbol{m}}_{k,n}) f_{\mathsf{n}}(\overline{\boldsymbol{e}}_{k,n})$, with $f_{\mathsf{n}}(\overline{\boldsymbol{p}}_{k,n})$ uniformly distributed on the ROI, $f_{\mathsf{n}}(\overline{\boldsymbol{m}}_{k,n})$ following a zero-mean Gaussian PDF with covariance matrix $10^2\,\mathrm{m^2/s^2}\,\mathbf{I}_2$, and $f_{\mathsf{n}}(\overline{\boldsymbol{e}}_{k,n})$ distributed according to an inverse Wishart distribution with mean matrix $3\mathbf{I}_2$ and 100 degrees of freedom. The proposed method uses the state transition model in (4) with $\mathbf{V}(\boldsymbol{m}_{k,n-1})$ given by $\mathbf{I}_2$ and $q_{k,n} = 20000$. The mean number of newly detected objects is set to $\mu_{\mathsf{n}} = 10^{-2}$. To facilitate track initialization, we perform message censoring and ordering of measurements as discussed in [55]. The number of SPA iterations was set to $P \in \{2, 3\}$ and the number of particles was set to $J = \{1000, 10000\}$. The resulting parameter combinations are denoted as "SPA-2-1000," "SPA-3-1000," "SPA-2-10000," and "SPA-3-10000," respectively.

For the PMBM filter, we set the Poisson point process that represents object birth consistent with the newly detected object representation introduced above. Furthermore, the probability of detection is set to 1. The Gamma distribution has an a priori mean of $\mu_{\mathsf{m}} = 8$ and a variance of 10^{-4}. Its parameters remain unchanged at all time steps. The transformation matrix and maneuvering correlation constant (see [5, Table III]) used for extent prediction are set to $\mathbf{I}_2$ and 10^5, respectively. The PMBM
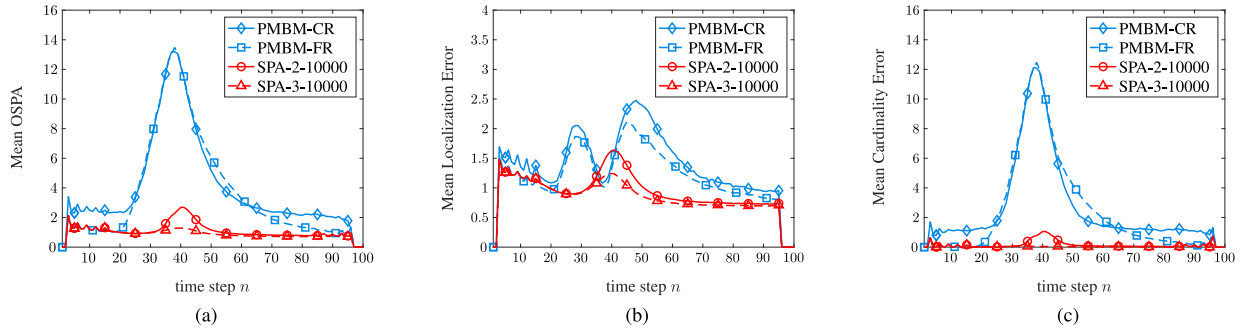
Fig. 4. Mean total OSPA (a) mean state error, (b) and mean cardinality error, (c) versus time of the proposed SPA method compared to the PMBM filter in a scenario with 10 closely-spaced extended objects.

implementation described in [1] relies on measurement gating and clustering as well as pruning of global association events [5]. The gate threshold is chosen such that the probability that an object-oriented measurement is in the gate is 0.999.

Clusters of measurements and likely association events are obtained by using the density-based spatial clustering (DB-SCAN) and Murthy's algorithm, respectively. We simulated two different settings for measurement clustering and event pruning. Coarse clustering "PMBM-C" calculates measurement partitions by using the 50 different distance values equally spaced between 0.1 and 5 as well as a maximum number of 20 assignments for each partition of measurements. Fine clustering "PMBM-F" clusters with 2000 different distance values equally spaced between 0.01 and 20 as well as uses a maximum number of 200 assignments for each partition of measurements. Clustering is performed for each distance value individually. All resulting clusters are then combined into one joint set of clusters. In this way, a diverse set of overlapping clusters is obtained. We also simulated variants of the PMBM that perform recycling of pruned Bernoulli components [65]. These variants are denoted as "PMBM-CR" and "PMBM-FR".

Fig. 3(b) and (c) shows estimation results of SPA-10000 and PMBM-FR. By comparing Fig. 3(c) with Fig. 3(a), it can be seen that PMBM-FR is unable to accurately estimate the state of objects that are in close proximity. Fig. 4 shows the mean OSPA error and its state and cardinality error contributions—averaged over 1200 simulation runs—of four simulated methods versus time. It can be seen that the proposed SPA-based methods outperform the PMBM at those time steps where objects are in close proximity. This can be explained by the fact that, due to its excellent scalability, the proposed SPA-based method can avoid clustering as performed by the PMBM filter implementations. In Fig. 4(c), it is shown that the main reason for the increased OSPA error of PMBM is an increased cardinality error. This is because large clusters that consist of measurements generated by multiple objects are associated to a single object. Thus, to certain other objects, no measurements are assigned, their probability of existence is reduced, and they are declared to be non-existent. The reduced state error of the PMBM methods compared to the proposed SPA method around time step 40 in Fig. 4(b) can be explained as follows. Since in this scenario PMBM methods tend to underestimate the number of objects, the optimum assignment step performed for OSPA calculation tends to find a solution with

TABLE I
MEAN GOSPA AND RUNTIME PER TIME STEP IN SECONDS FOR THE CONSIDERED SIMULATION. THE TOTAL GOSPA ERROR AS WELL AS INDIVIDUAL ERROR CONTRIBUTIONS ARE SHOWN

Method	Total	State	Missed	False	Runtime
SPA-2-1000	13.3	11.4	1.9	0.05	1.05
SPA-2-10000	8.7	8.0	0.7	0.04	8.39
SPA-3-1000	11.4	10.3	1.1	0.07	1.36
SPA-3-10000	7.5	7.3	0.1	0.03	11.67
PMBM-FR	22.5	10.3	11.8	0.41	50.59
PMBM-F	22.7	10.1	12.3	0.37	52.71
PMBM-CR	25.3	11.1	10.2	4.01	3.75
PMBM-C	25.2	10.8	10.8	3.55	4.29

a lower state error. Table I shows the mean GOSPA error and corresponding individual error contributions as well as runtimes per time step for MATLAB implementations on a single core of an Intel Xeon Gold 5222 CPU. Notably, despite not using gating, measurement clustering, and pruning of association events, the proposed SPA method has a runtime that is comparable with the PMBM filter.

C. Size, Orientation, and Scalability

To investigate the capability of the methods to estimate size and orientation, we considered a scenario similar to that discussed in Section VI-A, changing the prior distribution of object extent to an inverse Wishart distribution with mean $\text{diag}(3\text{m}, 1.5\text{m})$. Estimated object sizes are calculated from estimated extent states $\hat{\mathbf{E}}_{k,n}$ as the area of the represented ellipses, i.e., as $\pi \hat{\lambda}_{1,\hat{\mathbf{E}}_{k,n}} \hat{\lambda}_{2,\hat{\mathbf{E}}_{k,n}}$. Similarly, orientation is obtained by restricting the larger of the two eigenvectors $\hat{\boldsymbol{\lambda}}_{1,\hat{\mathbf{E}}_{k,n}}$ of $\hat{\mathbf{E}}_{k,n}$ to the upper half plane and calculating its angle. To compute mean size errors and mean orientation errors, we use the optimum assignments of the OSPA [63] metric calculated as discussed in Section VI-A for each time steps and each of the 120 simulation runs. Then, we use these optimum assignments to calculate mean size errors and mean orientation errors. Fig. 5 shows the resulting mean errors of the four simulated methods versus time. It can be seen that the proposed particle-based SPA method can estimate
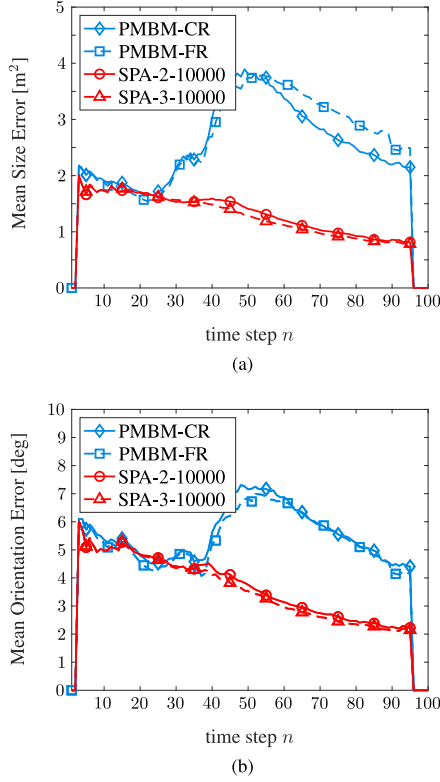
Fig. 5. Mean size error (a) and mean orientation error (b) of the proposed SPA method compared to the PMBM filter in a scenario with 10 closely-spaced extended objects.
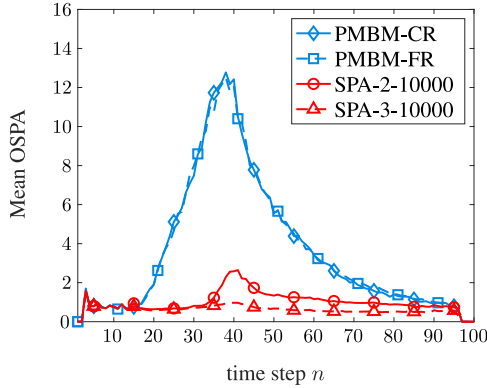

Fig. 6. Mean total OSPA of the proposed SPA method compared to the PMBM filter in a scenario with 20 closely-spaced extended objects.

size and orientation more reliably than the PMBM filter when objects are in close proximity.

To demonstrate scalability, we furthermore simulated the same scenario discussed in Section VI-A but with the number of objects increased to 20. Fig. 6 shows the mean OSPA error—averaged over 120 simulation runs—of four simulated methods versus time. Also in this larger scenario the performance advantages of the SPA methods over the PMBM methods are comparable to the smaller scenario with 10 objects. Plots for individual error contributions are similar to the ones for the scenario with 10 objects and are thus omitted. The average runtimes per time step for MATLAB implementations on a single core

of an Intel Xeon Gold 5222 CPU were measured as 48.1s for SPA-2-10000, 76.0s for SPA-3-10000, 6.7s for PMBM-CR, and 171.4s for PMBM-FR. We emphasize that this simulation is an extreme case, where all 20 objects come into close proximity at the mid-point in time. As previously discussed, additional, well-known techniques can be utilized to limit the growth in complexity for well-spaced objects to linear by decoupling the joint EOT problem into smaller separate ones. The unique benefit of the proposed method is its ability to scale in problems where object proximity precludes trivial decoupling.

VII. REAL-DATA PROCESSING

To validate our method, we present results in an urban autonomous driving scenario. The observations are part of the *nuScenes* datasets [66] and were collected by a lidar sensor mounted on the roof of an autonomous vehicle. The dataset consists of 850 scenes. Every scene has a duration of 20s and consists of approximately 400 sets of lidar observations (or "point clouds") and corresponding ground truth annotations for vehicles and pedestrians. A single lidar measurement is given by the 3-D Cartesian coordinates of a reflecting obstacle in the environment. Annotations are 3-D cubes defined by a position, dimensions, and heading.

We focused on tracking of vehicles in the environment. To extract measurement points related to reflections on vehicles, we used the supervised learning method presented in [67]. In particular, we employed the point clouds and annotations of 700 scenes for the training of the deep neural network (see [67] for details). After training, measurement extraction provides cubes that indicate vehicle locations and corresponding confidence scores on (0, 1]. We only used lidar observations inside vehicle-related detected cubes that have a confidence score larger than 0.25 as measurements for the EOT methods. Measurements and ground truth annotations are converted to a global reference frame. The ROI covers an area of 62,500m^2. The four scenes S1–S4[8] considered for performance evaluation, have not been used for the training.

We represent the extent state of vehicles by using the 2-D version of the cubical model introduced in Section II-B and [56, Section 2]. The kinematic states of vehicles is modeled by their 2-D position and velocity, as well as their turn rate $t_{k,n}$ and their mean number of generated measurements $s_{k,n}$, i.e., $\mathbf{x}_{k,n} = [\mathsf{p}_{k,n}^{(1)} \ \mathsf{p}_{k,n}^{(2)} \ \dot{\mathsf{p}}_{k,n}^{(1)} \ \dot{\mathsf{p}}_{n,k}^{(2)} \ t_{k,n} \ s_{k,n}]^{\mathrm{T}}$. Since the number of measurements that a vehicle generates strongly depends on its distance to the lidar sensor, the mean number of measurements is part of the random vehicle state, i.e., $\mu_{\mathrm{m}}(\mathbf{x}_{k,n}, \mathbf{e}_{k,n}) = s_{k,n}$. The mean number of measurements $s_{k,n}$ follows the state transition model in [5, Table III] with parameter $\eta = 10$. The kinematic vehicle state and the extent state follow the state transition model in (4) with $f(\cdot)$ and $\mathbf{V}(\cdot)$ as given in [12]. The parameters of the state transition model are $q = 2000$, $\sigma_{\mathbf{c}} = 10\,\mathrm{m/s}^2$, and $\sigma_{\mathrm{t}} = 0.003\,\mathrm{rad/s}^2$. The survival probability is $p_{\mathrm{s}} = 0.999$.

---

[8]These four scenes can be identified by their tokens `8b31aa5cac3846a6` `a197c507523926f9`, `9e79971671284ff482c63a1f658e703d`, `ee 4979d9ef9e417f9337d09d62072692`, and `06b1e705b90641ce83 500b71c0dbf037`, respectively (see [66] for details).
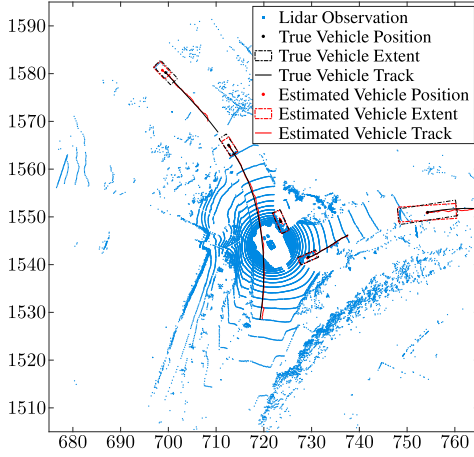
Fig. 7. Time step $n = 105$ of scene S4 in the *nuScenes* urban autonomous driving dataset. Lidar observations, ground truth, and estimation results are shown.

TABLE II
MEAN GOSPA AND RUNTIME PER TIME STEP IN SECONDS FOR THE CONSIDERED VEHICLE TRACKING SCENARIO. THE TOTAL GOSPA ERROR AS WELL AS INDIVIDUAL ERROR CONTRIBUTIONS ARE SHOWN

Method	Total	State	Missed	False	Runtime
SPA-1000	7.63	4.33	1.40	1.89	0.07
SPA-10000	7.21	4.13	1.37	1.71	0.43
PMBM-FR	8.60	4.99	1.82	1.80	2.93
PMBM-F	8.60	4.98	1.82	1.81	2.34
PMBM-CR	8.57	5.03	1.76	1.78	0.16
PMBM-C	8.61	5.07	1.76	1.77	0.14

In addition, the linear-Gaussian measurement model in (42) with $\sigma_{\mathrm{u}} = 5 \cdot 10^{-3}$ m was used. Extracted lidar observations are voxelized with a resolution of 0.5m and their z-coordinate is ignored. We model newly detected vehicles as $f_{\mathrm{n}}(\overline{\boldsymbol{x}}_{k,n}, \overline{\boldsymbol{e}}_{k,n}) = f_{\mathrm{n}}(\overline{\boldsymbol{p}}_{k,n}) f_{\mathrm{n}}(\overline{\boldsymbol{e}}_{k,n}) f_{\mathrm{n}}(\overline{\boldsymbol{m}}_{k,n})$, with $f_{\mathrm{n}}(\overline{\boldsymbol{p}}_{k,n})$ uniformly distributed on the ROI and $f_{\mathrm{n}}(\overline{\boldsymbol{e}}_{k,n})$ following an inverse Wishart distribution with mean matrix $3\mathbf{I}_2$ and 30 degrees of freedom. In addition, we set $f_{\mathrm{n}}(\overline{\boldsymbol{m}}_{k,n}) = f_{\mathrm{n}}(\dot{\overline{\boldsymbol{p}}}_{k,n}, \overline{t}_{k,n}) f_{\mathrm{n}}(\overline{s}_{k,n})$ where $f_{\mathrm{n}}(\dot{\overline{\boldsymbol{p}}}_{k,n}, \overline{t}_{k,n})$ is zero-mean Gaussian distributed with covariance matrix $\mathrm{diag}(5^2\mathrm{m}^2/\mathrm{s}^2, 5^2\mathrm{m}^2/\mathrm{s}^2, 0.01^2\,\mathrm{rad}^2/\mathrm{s}^2)$ and $f_{\mathrm{n}}(\overline{s}_{k,n})$ follows a Gamma distribution with mean 30 and variance 5. The number of SPA iterations is $P = 3$ and the number of particles was again either set to $J = 1000$ or to $J = 10000$ (denoted as SPA-1000 and SPA-10000, respectively). All the other parameters are set as described in Section VI-B.

Fig. 7 shows the lidar observations, true vehicle positions, true vehicle extents, true vehicle tracks, estimated vehicle positions, estimated vehicle extents, and estimated vehicle tracks for time step $n = 105$ of scene S4. It can seen that all five actual vehicles are detected and tracked reliably. Videos of estimation results are available on `fmeyer.ucsd.edu`.

We use the PMBM filter based on the same system model described in Section VI-A as a reference method. The driving noise standard deviation of the PMBM was set to $\sigma_{\mathbf{c}} = 30\,\mathrm{m/s}^2$. Furthermore, the transformation matrix and maneuvering correlation constant (see [5, Table III]) used for extent prediction were set to $\mathbf{I}_2$ and 20 respectively. All other model parameters were chosen as for the proposed method. The parameters for measurement clustering, event pruning, and recycling were set as discussed in Section VI-B (denoted as PMBM-C, PMBM-F, PMBM-CR, and PMBM-FR).

As performance metric we use the generalized OSPA (GOSPA) [64] based on the 2-norm with parameters $p = 1$ and $c = 5$. Only the kinematic state was used for the evaluation of the GOSPA. The mean GOSPA for the individual methods—averaged over the four considered scenes and all time steps—as well as runtimes per time step on a single core of an Intel Xeon Gold 5222 CPU are summarized in Table II. It can again

be seen that the proposed SPA-based method outperforms the PMBM in terms of mean GOSPA and mean individual error contributions. These performance advantages of the proposed SPA-based method are related to its more accurate system model as well as its particle-based implementation. Interestingly, all PMBM variants perform very similarly.

VIII. CONCLUSION

Detection, localization, and tracking of multiple objects is a key task in a variety of applications including autonomous navigation and applied ocean sciences. This paper introduced a scalable method for EOT. The proposed method is based on a factor graph formulation and the SPA. It dynamically introduces states of newly detected objects, and efficiently performs probabilistic data association, allowing for multiple measurements per object. Scalable inference of object tracks and their geometric shape is enabled by modeling association uncertainty by measurement-oriented association variables and newly detected objects by a Poisson birth process. The fully particle-based approach can represent the extent of objects by different geometric shapes. In addition, it yields a computational complexity that only scales quadratically in the number of objects and the number of measurements. This excellent scalability translates to an improved EOT performance compared to existing methods because it makes it possible to (i) generate and maintain a very large number of POs and (ii) avoid clustering of measurements, allowing problems with closely-spaced extended objects to be addressed.

Simulation results in a challenging scenario with intersecting object tracks showed that the proposed method can outperform the recently introduced Poisson multi-Bernoulli (PMBM) filter despite yielding a reduced computational complexity. The application of the proposed method to real measurement data captured by a lidar sensor in an urban autonomous driving scenario also demonstrated performance advantages compared to the PMBM filter. For the extraction of vehicle-related lidar measurements, supervised learning based on a deep neural network was used. This motivates future research on multiobject tracking methods that closely integrate neural networks and iterative message passing. Other promising directions for future research are the development of highly parallelized variants of the proposed

method that exploit particle flow [68] and are suitable for real-time implementations on graphical processing units (GPUs).

ACKNOWLEDGMENT

The authors would like to thank Dr. E. Leitinger and W. Zhang for carefully reading the manuscript.

REFERENCES

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

[2] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT: Y. Bar-Shalom, 2011.

[3] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.

[4] F. Meyer et al., "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.

[5] K. Granström, M. Fatemi, and L. Svensson, "Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 208–225, Feb. 2020.

[6] M. Schuster, J. Reuter, and G. Wanielik, "Probabilistic data association for tracking extended targets under clutter using random matrices," in *Proc. Int. Conf. Inf. Fusion*, Washington DC, 2015, pp. 961–968.

[7] K. Granström, C. Lundquist, and U. Orguner, "Extended target tracking using a Gaussian-mixture PHD filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3268–3285, Oct. 2012.

[8] M. Beard, S. Reuter, K. Granström, B. Vo, B. Vo, and A. Scheel, "Multiple extended target tracking with labeled random finite sets," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1638–1653, Apr. 2016.

[9] Y. Xia, K. Granström, L. Svensson, F. Ángel García-Fernández, and J. L. Williams, "Extended target Poisson multi-Bernoulli mixture trackers based on sets of trajectories," in *Proc. Int. Conf. Inf. Fusion*, Ottawa, Canada, 2019, pp. 1-8.

[10] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 1042–1059, Jul. 2008.

[11] C.-Y. Feldmann, D. Fränken, and J. W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1409–1420, Apr. 2011.

[12] K. Granström and U. Orguner, "New prediction for extended targets with random matrices," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 1577–1589, 2014.

[13] S. Yang and M. Baum, "Tracking the orientation and axes lengths of an elliptical extended object," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4720–4729, Sep. 2019.

[14] G. Vivone, K. Granström, P. Braca, and P. Willett, "Multiple sensor measurement updates for the extended target tracking random matrix model," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 5, pp. 2544–2558, 2017.

[15] J. L. Williams, "Marginal multi-Bernoulli filters: RFS derivation of MHT, JIPDA and association-based MeMBer," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 1664–1687, Jul. 2015.

[16] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Autom. Control*, vol. 24, no. 6, pp. 843–854, Dec. 1979.

[17] T. Kurien, "Issues in the design of practical multitarget tracking algorithms," in *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed. Norwood, MA, USA: Artech-House, 1990, pp. 43–83.

[18] S. Coraluppi and C. Carthel, "Distributed tracking in multistatic sonar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 3, pp. 1138–1147, Jul. 2005.

[19] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1224–1245, Oct. 2005.

[20] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3553–3567, Jul. 2007.

[21] B.-N. Vo, B.-T. Vo, and D. Phung, "Labeled random finite sets and the Bayes multi-target tracking filter," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6554–6567, Dec. 2014.

[22] B. K. Habtemariam, R. Tharmarasa, T. Kirubarajan, D. Grimmett, and C. Wakayama, "Multiple detection probabilistic data association filter for multistatic target tracking," in *Proc. Int. Conf. Inf. Fusion*, 2011, pp. 1-6.

[23] 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.

[24] B. Habtemariam, R. Tharmarasa, T. Thayaparan, M. Mallick, and T. Kirubarajan, "A multiple-detection joint probabilistic data association filter," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 461–471, Jun. 2013.

[25] F. Meyer and M. Z. Win, "Scalable data association for extended object tracking," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 6, pp. 491–507, May 2020.

[26] S. P. Coraluppi and C. A. Carthel, "Multiple-hypothesis tracking for targets producing multiple measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 3, pp. 1485–1498, Jun. 2018.

[27] G. Gennarelli, G. Vivone, P. Braca, F. Soldovieri, and M. G. Amin, "Multiple extended target tracking for through-wall radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6482–6494, Dec. 2015.

[28] G. Vivone and P. Braca, "Joint probabilistic data association tracker for extended target tracking applied to X-band marine radar data," *IEEE J. Ocean. Eng.*, vol. 41, no. 4, pp. 1007–1019, Oct. 2016.

[29] K. Granström, L. Svensson, S. Reuter, Y. Xia, and M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods," *IEEE Trans. Intell. Veh.*, vol. 3, no. 1, pp. 30–45, 2018.

[30] J. Böhler, T. Baur, S. Wirtensohn, and J. Reuter, "Stochastic partitioning for extended object probability hypothesis density filters," in *Proc. SDF-19*, Bonn, Germany, 2019, pp. 1-6.

[31] T. Kropfreiter, F. Meyer, and F. Hlawatsch, "A fast labeled multi-Bernoulli filter using belief propagation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 3, pp. 2478–2488, Jun. 2020.

[32] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA, USA: MIT Press, 2009.

[33] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[34] J. Yedidia, W. Freemand, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.

[35] M. P. C. Fossorier, M. Mihaljevic, and H. Imai, "Reduced complexity iterative decoding of low-density parity check codes based on belief propagation," *IEEE Trans. Commun.*, vol. 47, no. 5, pp. 673–680, May 1999.

[36] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[37] F. Meyer, B. Etzlinger, Z. Liu, F. Hlawatsch, and M. Z. Win, "A scalable algorithm for network localization and synchronization," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4714–4727, Dec. 2018.

[38] J. L. Williams and R. Lau, "Approximate evaluation of marginal association probabilities with belief propagation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 2942–2959, Oct. 2014.

[39] T. Kropfreiter, F. Meyer, and F. Hlawatsch, "Sequential Monte Carlo implementation of the track-oriented marginal multi-Bernoulli/Poisson filter," in *Proc. Int. Conf. Inf. Fusion*, 2016, pp. 972–979.

[40] F. Meyer, P. Braca, P. Willett, and F. Hlawatsch, "A scalable algorithm for tracking an unknown number of targets using multiple sensors," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3478–3493, Jul. 2017.

[41] G. Soldi, F. Meyer, P. Braca, and F. Hlawatsch, "Self-tuning algorithms for multisensor-multitarget tracking using belief propagation," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 3922–3937, Aug. 2019.

[42] R. A. Lau and J. L. Williams, "A structured mean field approach for existence-based multiple target tracking," in *Proc. Int. Conf. Inf. Fusion*, Heidelberg, Germany, 2016, pp. 1111–1118.

[43] W. Zhang and F. Meyer, "Graph-based multiobject tracking with embedded particle flow," in *Proc. IEEE RadarConf-21*, Atlanta, GA, USA, 2021.

[44] E. Leitinger, F. Meyer, P. Meissner, K. Witrisal, and F. Hlawatsch, "Belief propagation based joint probabilistic data association for multipath-assisted indoor navigation and tracking," in *Proc. ICL-GNSS-16*, Barcelona, Spain, 2016. pp. 1-6.

[45] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, "A belief propagation algorithm for multipath-based SLAM," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5613–5629, Nov. 2019.

[46] R. Mendrzik, F. Meyer, G. Bauch, and M. Z. Win, "Enabling situational awareness in millimeter wave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1196–1211, Sep. 2019.

[47] X. Li, E. Leitinger, A. Venus, and F. Tufvesson, "Sequential detection and estimation of multipath channel parameters using belief propagation," 2021, *arXiv:2109.05623*

[48] F. Meyer and M. Z. Win, "Joint navigation and multitarget tracking in networks," in *Proc. IEEE Int. Conf. Commun.*, Kansas City, MO, USA, 2018. pp. 1-6.

[49] F. Meyer, O. Hlinka, H. Wymeersch, E. Riegler, and F. Hlawatsch, "Distributed localization and tracking of mobile networks including noncooperative objects," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 2, no. 1, pp. 57–71, Mar. 2016.

[50] F. Meyer, Z. Liu, and M. Z. Win, "Network localization and navigation using measurements with uncertain origin," in *Proc. Int. Conf. Inf. Fusion-18*, Cambridge, U.K., 2018, pp. 2237–2243.

[51] F. Meyer, Z. Liu, and M. Z. Win, "Scalable probabilistic data association with extended objects," in *Proc. IEEE Int. Conf. Commun.*, pp. 1-6, Shanghai, China, pp. 1–6, 2019.

[52] Y. Li, P. Wei, Y. Chen, Y. Wei, and H. Zhang, "Message passing based extended objects tracking with measurement rate and extension estimation," in *Proc. IEEE Radar Conf-21*, Atlanta, GA, USA, 2021, pp. 1–6.

[53] S. Yang, K. Thormann, and M. Baum, "Linear-time joint probabilistic data association for multiple extended object tracking," in *Proc. IEEE SAM-18*, Sheffield, U.K., 2018, pp. 6–10.

[54] S. Yang, L. M. Wolf, and M. Baum, "Marginal association probabilities for multiple extended objects without enumeration of measurement partitions," in *Proc. Int. Conf. Inf. Fusion-20*, Pretoria, South Africa, 2020, pp. 1-8.

[55] F. Meyer and J. L. Williams, "Scalable detection and tracking of extended objects," in *Proc. IEEE ICASSP-20*, Barcelona, Spain, 2020, pp. 8916–8920.

[56] F. Meyer and J. L. Williams, "Scalable detection and tracking of geometric extended objects: Supplementary material," 2021, *arXiv:2103.11279.*

[57] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer, 1994.

[58] J. L. Williams and R. A. Lau, "Multiple scan data association by convex variational inference," *IEEE Trans. Signal Process.*, vol. 66, no. 8, pp. 2112–2127, Apr. 2018.

[59] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[60] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. Berlin, Germany: Springer, 2001.

[61] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[62] Y. Bar-Shalom, T. Kirubarajan, and X.-R. Li, *Estimation With Applications to Tracking and Navigation*. New York, NY, USA: Wiley, 2002.

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

[64] A. S. Rahmathullah, A. F. Garcia-Fernandez, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *Proc. Int. Conf. Inf. Fusion-17*, Xi'an, China, 2017, pp. 1-8.

[65] J. L. Williams, "Hybrid Poisson and multi-Bernoulli filters," in *Proc. Int. Conf. Inf. Fusion-12*, Singapore, 2012, pp. 1103–1110.

[66] "The nuScenes Dataset," [Online]. Available: https://www.nuscenes.org/

[67] B. Zhu, Z. Jiang, X. Zhou, Z. Li, and G. Yu, "Class-balanced grouping and sampling for point cloud 3D object detection," *arXiv:1908.09492*, 2019.

[68] Y. Li and M. Coates, "Particle filtering with invertible particle flow," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4102–4116, Aug. 2017.

Florian Meyer (Member, IEEE) received the Dipl.-Ing. (M.Sc.) and Ph.D. degrees (with highest Hons.) in electrical engineering from TU Wien, Vienna, Austria, in 2011 and 2015, respectively. He is an Assistant Professor with the University of California San Diego, La Jolla, CA, jointly between the Scripps Institution of Oceanography and the Electrical and Computer Engineering Department. From 2017 to 2019, he was a Postdoctoral Fellow and Associate with Laboratory for Information & Decision Systems, the Massachusetts Institute of Technology, Cambridge, MA, USA. From 2016 to 2017, he was a Research Scientist with NATO Centre for Maritime Research and Experimentation, La Spezia, Italy.

Dr. Meyer was a Keynote Speaker at the IEEE Aerospace Conference in 2020. He was on the technical program committees of several IEEE conferences and as the Co-Chair of IEEE Workshop on Advances in Network Localization and Navigation at IEEE International Conference on Communications in 2018, 2019, and 2020. He is an Associate Editor for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS and the *ISIF Journal of Advances in Information Fusion* as well as an Erwin Schrödinger Fellow.

Jason L. Williams (Senior Member, IEEE) received the degrees of B.E.(Electronics)/BInfTech from the Queensland University of Technology, Australia, the MSEE degree from the United States Air Force Institute of Technology, Dayton, OH, USA, and the Ph.D. degree from the Massachusetts Institute of Technology, Cambridge, MA, USA. He is currently a Senior Research Scientist in robotic perception, with the Robotics and Autonomous Systems Group of Commonwealth Scientific and Industrial Research Organisation, Brisbane, Australia. His research interests include SLAM, computer vision, multiple object tracking and motion planning. He previously worked in sensor fusion and resource management as a Senior Research Scientist with the Defence Science and Technology Group, Australia, and in electronic warfare as an Engineering Officer with the Royal Australian Air Force.