

Multi-Ellipsoidal Extended Target Tracking with Variational Bayes Inference

BARKIN TUNCER ¹, Emre Özkan ², and Umut Orguner ²

¹middle east technical university

²Affiliation not available

October 30, 2023

Abstract

In this work, we propose a novel extended target tracking algorithm, which is capable of representing a target or a group of targets with multiple ellipses. Each ellipse is modeled by an unknown symmetric positive-definite random matrix. The proposed model requires solving two challenging problems. First, the data association problem between the measurements and the sub-objects. Second, the inference problem that involves non-conjugate priors and likelihoods which needs to be solved within the recursive filtering framework. We utilize the variational Bayes inference method to solve the association problem and to approximate the intractable true posterior. The performance of the proposed solution is demonstrated in simulations and real-data experiments. The results show that our method outperforms the state-of-the-art methods in accuracy with lower computational complexity.

Multi-Ellipsoidal Extended Target Tracking with Variational Bayes Inference

Barkın Tuncer, *Student Member, IEEE*, Umut Orguner, *Member, IEEE*, Emre Özkan, *Member, IEEE*,

Abstract—In this work, we propose a novel extended target tracking algorithm, which is capable of representing a target or a group of targets with multiple ellipses. Each ellipse is modeled by an unknown symmetric positive-definite random matrix. The proposed model requires solving two challenging problems. First, the data association problem between the measurements and the sub-objects. Second, the inference problem that involves non-conjugate priors and likelihoods which needs to be solved within the recursive filtering framework. We utilize the variational Bayes inference method to solve the association problem and to approximate the intractable true posterior. The performance of the proposed solution is demonstrated in simulations and real-data experiments. The results show that our method outperforms the state-of-the-art methods in accuracy with lower computational complexity.

Index Terms—Extended target tracking, random matrix, variational Bayes.

I. INTRODUCTION

Recent advances in autonomous vehicles, robotics, and intelligent systems entail the need to not only estimating the position of an object but also recognizing its shape. This requirement is usually fulfilled with short-range sensor systems, where it is possible to collect multiple instantaneous measurements from a single target. In contrast to traditional point target tracking methods, one can extract more information from the measurements, such as the shape, size, or orientation of the target. The special algorithms, which are capable of estimating these unknowns together with the target's kinematic state, are called Extended Target/Object Tracking (ETT/EOT) algorithms.

A primitive approach to ETT is to represent the target's extent as a simple shape and estimate the relevant parameters. These simple shapes can be a line [1], a circle [2] or a rectangle [3]. More complex shapes can be defined using *random hyper-surface models* which assume that the measurements are generated from an unknown random surface [4]–[6]. More recently, Gaussian Process (GP) based models [7], [8] were proposed to estimate the extent of targets with unknown shapes. GP based ETT algorithms define the target's contour as an unknown radial function with a GP prior. An advantage of these models is that the estimated contours are descriptive, i.e., the contour representations can be utilized in further purposes such as classification of targets [9], [10].

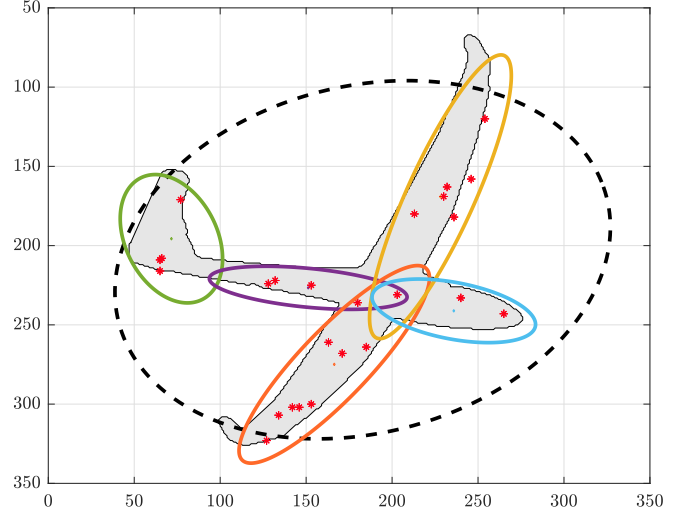


Fig. 1: An extended target representation with multiple ellipses (solid lines). Classical RM models represent the same target extent as a single ellipse (dashed black line). The measurements are shown with red stars.

One of the most common approaches in the ETT literature is the random matrix model (RM) [6], [11]–[16], which was pioneered by Koch [11]. In RM based methods, the extent is approximated by an ellipse which is represented by a symmetric positive definite (SPD) unknown matrix. The inference in [11] neglects the measurement noise covariance matrix in order to meet the conjugacy requirement. This approach was later improved in [12], where the measurement noise covariance is incorporated into the updates. More recent studies consider the orientation angle of the target together with the RM model [6], [14]–[16].

A single ellipsoidal representation of the target can provide only lumped information about the shape of the target, which may result in over-simplified representations. To remedy this problem, more recent works on random matrices focus on multi-ellipsoidal (ME) models [17]–[20], where the target extent is represented with more than one ellipse. Each ellipse is called a sub-object, and as the number of sub-objects increases, a finer representation of the target's extent can be obtained. A representative example of using multiple ellipses for target extent is given in Figure 1.

ME models in the literature use different inference methods for estimating a mixture of Gaussian inverse Wishart (GIW) distributions which represent the kinematic state together with the extent. In [17], a particle filter is used for inference which becomes computationally heavy when the number of particles

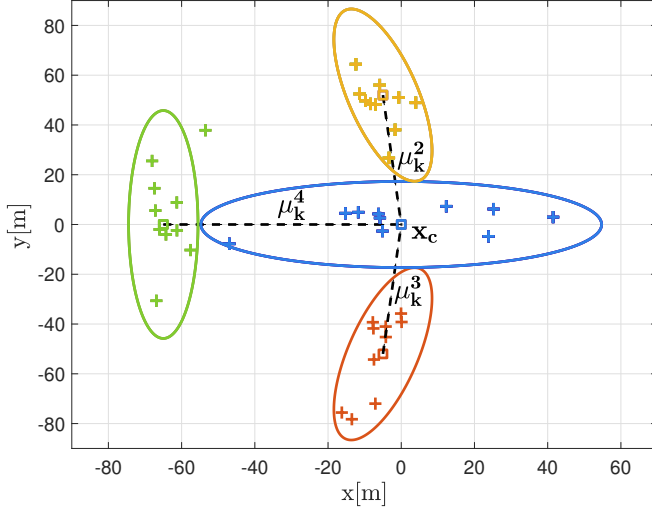


Fig. 2: Illustration of the extent estimation for $L = 4$ where $+$ sign denotes the measurements.

is increased. The ME model proposed in [20] does not contain a unified kinematic model for different sub-objects. In [19], the kinematics of the sub-objects are unified however, the method requires computationally heavy partitioning algorithms to associate the measurements with the sub-objects. Furthermore, mixture reduction algorithms are required to manage the number of components resulting from different association hypotheses. In [18], an approach that can handle varying numbers of sub-objects in time is proposed. In this work, we present a ME-ETT approach utilizing variational Bayes inference for solving the measurement association problem and obtaining an approximate distribution for the intractable posterior. Our approach does not require any clustering [20], partitioning [19], mixture reduction [19] and merging methods [20]. The resulting algorithm has low computational complexity, and outperforms the state-of-the-art methods in accuracy.

The rest of the paper is organized as follows. First, the problem formulation is given in Section II. Then, the variational inference for the measurement update is explained in Section III. Section IV presents the time update. In Section V the simulation results will be shown and discussed. Finally, we will conclude the article in Section VI.

II. PROBLEM FORMULATION

We consider an extended target composed of $L \geq 1$ sub-objects. At any time k , the extent state of the ℓ th sub-object is represented by the SPD matrix $\mathbf{X}_k^\ell \in \mathbb{R}^{n_y \times n_y}$, $\ell = 1, \dots, L$. We define the kinematic state of the extended object as

$$\mathbf{x}_k \triangleq [(\mathbf{x}_k^c)^T \quad \tilde{\mathbf{x}}_k^T \quad (\boldsymbol{\mu}_k^2)^T \quad \dots \quad (\boldsymbol{\mu}_k^L)^T]^T \quad (1)$$

where $\mathbf{x}_k^c \in \mathbb{R}^{n_y}$ denotes a reference point on the target and $\{\boldsymbol{\mu}_k^\ell \in \mathbb{R}^{n_y}\}_{\ell=1}^L$ denote the displacement vectors of the sub-object centers from the reference point \mathbf{x}_k^c [17]. We assume that $\boldsymbol{\mu}_k^1 = \mathbf{0}$, i.e., the reference point of the extended target is at the center of the first sub-object without loss of generality. The vector $\tilde{\mathbf{x}}_k$ contains all non-positional kinematic

TABLE I: NOTATIONS

- Set of real matrices of size $m \times n$ is represented with $\mathbb{R}^{m \times n}$.
- $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents the multivariate Gaussian distribution with mean vector $\boldsymbol{\mu} \in \mathbb{R}^{n_x}$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{S}_{++}^{n_x}$.
- $\mathcal{IW}(\mathbf{X}; v, \mathbf{V})$ represents the inverse Wishart distribution over the real-valued positive definite matrix \mathbf{X} with degrees of freedom v , and positive definite scale matrix \mathbf{V}

$$\mathcal{IW}(\mathbf{X}; v, \mathbf{V}) \triangleq \frac{\text{etr}(-\frac{1}{2}\mathbf{X}^{-1}\mathbf{V})}{|\mathbf{X}|^{v/2}},$$

where $\text{etr}(\cdot) \triangleq e^{\text{tr}(\cdot)}$.

- For the number of measurements $M_k \in \mathbb{Z}^+$, $\mathbf{y}_k^{1:M_k}$ represents the measurement set $\{\mathbf{y}_k^1, \dots, \mathbf{y}_k^{M_k}\}$ at time k .
- \mathbf{r}_k represents the vector $[r_k^1, \dots, r_k^a]^T$ with size $a \in \mathbb{Z}^+$.
- $|\mathbf{A}|$ denotes the determinant of the matrix \mathbf{A} .
- c_{ϕ} is a generic constant that denotes the constant terms with respect to variable ϕ in an equation.
- $\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}$ where a_{ii} is the i th diagonal element of $\mathbf{A} \in \mathbb{R}^{n \times n}$.
- The iterate numbers are shown with parenthesized superscripts, e.g., $q^{(i)}(\cdot)$.
- The expectation with respect to a specific variable is shown with a subscript in the expectation sign, e.g., $\mathbb{E}_{\mathbf{z}}$ will show an expectation with respect $\mathbf{z}_k^{1:M_k}$. When it is clear from the context which variable we are taking the expectation with respect to, we will show the expectation with an overline, e.g., $\overline{(\mathbf{X}_k)^{-1}}$ will denote $\mathbb{E}_{\mathbf{X}}[(\mathbf{X}_k)^{-1}]$.
- When it is necessary to take expectations with respect to all random variables except for one of them, we will use a backslash “\” in the subscript of the expectation operator, e.g., $\mathbb{E}_{\setminus \mathbf{z}}$ will denote expectation with respect to all random variables except $\mathbf{z}_k^{1:M_k}$.
- In the derivations, the joint distribution $p(\mathbf{y}_k^{1:M_k}, \mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k}, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_{k-1})$ is abbreviated as $p(\cdot | \mathcal{Y}_{k-1})$ for the sake of brevity.
- The quadratic forms $\mathbf{x}^T \mathbf{A} \mathbf{x}$ and outer products $\mathbf{x} \mathbf{x}^T$ are written as $\mathbf{x}^T \mathbf{A}(\cdot)$ and $\mathbf{x}(\cdot)^T$, respectively to avoid unnecessarily duplicating long expressions.

information, such as velocity and acceleration. An illustration of this kinematic state vector for $L = 4$ is shown in Figure 2. The kinematic state vector $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is assumed to evolve with the following linear Gaussian dynamics

$$\mathbf{x}_k = \mathbf{F} \mathbf{x}_{k-1} + \mathbf{w}_k, \quad (2)$$

where $\mathbf{F} \in \mathbb{R}^{n_x \times n_x}$ denotes the state transition matrix and $\mathbf{w}_k \in \mathbb{R}^{n_x}$ denotes zero-mean white Gaussian process noise vector with covariance matrix \mathbf{Q} , i.e., $\mathbf{w}_k \sim \mathcal{N}(\mathbf{w}_k; \mathbf{0}, \mathbf{Q})$.

Suppose at time k , a set of M_k target-originated measurements are captured by the sensor which is denoted as $\mathbf{y}_k^{1:M_k} \triangleq \{\mathbf{y}_k^j\}_{j=1}^{M_k}$. The measurements are assumed to be conditionally i.i.d. and distributed according to the following Gaussian mixture.

$$p(\mathbf{y}_k^j | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \sum_{\ell=1}^L \pi_k^\ell \mathcal{N}(\mathbf{y}_k^j; \mathbf{H}_\ell \mathbf{x}_k, \mathbf{S} \mathbf{X}_k^\ell + \mathbf{R}) \quad (3)$$

for $j = 1, \dots, M_k$, where

- $\pi_k^{1:L} \triangleq \{\pi_k^\ell\}_{\ell=1}^L$ is the set of time-varying, unknown and random mixture probabilities π_k^ℓ , $\ell = 1, \dots, L$, satisfying $\pi_k^\ell \geq 0$ and $\sum_{\ell=1}^L \pi_k^\ell = 1$.
- $\mathbf{H}_\ell \in \mathbb{R}^{n_y \times n_x}$, $\ell = 1, \dots, L$ are measurement matrices defined such that $\mathbf{H}_\ell \mathbf{x}_k \triangleq \mathbf{x}_k^c + \boldsymbol{\mu}_k^\ell$. Note that \mathbf{H}_1 should be defined such that $\mathbf{H}_1 \mathbf{x}_k \triangleq \mathbf{x}_k^c$ since $\boldsymbol{\mu}_k^1 = \mathbf{0}$.

- $\mathbf{X}_k^{1:L} \triangleq \{\mathbf{X}_k^\ell\}_{\ell=1}^L$ is the set of sub-object extent matrices \mathbf{X}_k^ℓ , $\ell = 1, \dots, L$.
- $\mathbf{R} \in \mathbb{R}^{n_y \times n_y}$ is the SPM measurement noise covariance matrix.
- $s \in \mathbb{R}^+$ is a known positive scaling constant.
- Other necessary notations can be found in Table I.

The likelihood of the set of measurements $\mathbf{y}_k^{1:M_k}$ is then given as

$$p(\mathbf{y}_k^{1:M_k} | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \prod_{j=1}^{M_k} p(\mathbf{y}_k^j | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) \quad (4a)$$

$$= \prod_{j=1}^{M_k} \sum_{\ell=1}^L \pi_k^\ell \mathcal{N}(\mathbf{y}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s \mathbf{X}_k^\ell + \mathbf{R}). \quad (4b)$$

The aim is to find the posterior distribution $p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_k)$ recursively where $\mathcal{Y}_k \triangleq \{\mathbf{y}_t^{1:M_k}\}_{t=0}^k$ denotes the cumulative set of measurements obtained from the sensor up to and including time k . At each time step k , the posterior distribution $p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_k)$ is assumed to be in the form given below.

$$\begin{aligned} p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_k) &= \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k}, \mathbf{P}_{k|k}) \\ &\times \left(\prod_{\ell=1}^L \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k}^\ell, \mathbf{V}_{k|k}^\ell) \right) \\ &\times \mathcal{D}(\pi_k^{1:L}; \alpha_{k|k}^{1:L}), \end{aligned} \quad (5)$$

where $\mathbf{m}_{k|k}$, $\mathbf{P}_{k|k}$ are the mean and covariance of the kinematic state vector \mathbf{x}_k , respectively. The notations $\mathcal{IW}(\mathbf{X}; v, \mathbf{V})$ and $\mathcal{D}(\pi^{1:L}; \alpha^{1:L})$ denote the inverse Wishart and Dirichlet distributions, respectively

$$\mathcal{IW}(\mathbf{X}; v, \mathbf{V}) \propto \frac{\text{etr}(-\frac{1}{2} \mathbf{X}^{-1} \mathbf{V})}{|\mathbf{X}|^{v/2}}, \quad (6a)$$

$$\mathcal{D}(\pi^{1:L}; \alpha^{1:L}) \propto \prod_{\ell=1}^L (\pi^\ell)^{\alpha^\ell - 1}. \quad (6b)$$

The recursive calculation of the posterior $p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_k)$ will be carried out in two steps, namely, the measurement update and time update, which will be investigated separately in the following sections.

III. MEASUREMENT UPDATE

Suppose now that the predicted distribution $p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_{k-1})$ has the form in (5) with the following parameters

$$\begin{aligned} p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_{k-1}) &= \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ &\times \left(\prod_{\ell=1}^L \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) \right) \\ &\times \mathcal{D}(\pi_k^{1:L}; \alpha_{k|k-1}^{1:L}). \end{aligned} \quad (7)$$

In the measurement update, we would like to update the predicted distribution in (7) with the likelihood in (4b). Unfortunately, such an update is not analytically tractable, and furthermore, it would not result in a posterior in the form (5). In order to ensure analytical tractability and to preserve the

form of the posterior, first, we are going to define some latent variables and then resort to variational approximation.

We first define the latent variables \mathbf{z}_k^j , $j = 1, \dots, M_k$ which represent the noise-free measurements satisfying

$$p(\mathbf{z}_k^j | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \sum_{\ell=1}^L \pi_k^\ell \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s \mathbf{X}_k^\ell), \quad (8a)$$

$$p(\mathbf{y}_k^j | \mathbf{z}_k^j) = \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}). \quad (8b)$$

Note that the conditional joint distribution for $\mathbf{y}_k^j, \mathbf{z}_k^j$ is given as

$$\begin{aligned} p(\mathbf{y}_k^j, \mathbf{z}_k^j | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) &= \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \\ &\times \sum_{\ell=1}^L \pi_k^\ell \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s \mathbf{X}_k^\ell). \end{aligned} \quad (9)$$

We also define the association/responsibility vector \mathbf{r}_k^j for each measurement \mathbf{y}_k^j (or \mathbf{z}_k^j), which is a binary vector of size L defined as

$$\mathbf{r}_k^j \triangleq [r_k^{j,1} \ r_k^{j,2} \ \dots \ r_k^{j,L}]^T \quad (10)$$

where the elements $r_k^{j,\ell} \in \{0, 1\}$, $\ell = 1, \dots, L$, which are called responsibilities in the literature [21, Ch. 10]. These binary variables satisfy

$$\sum_{\ell=1}^L r_k^{j,\ell} = 1. \quad (11)$$

Note that, with these properties, the elements $r_k^{j,\ell}$ are all equal to zero except for one of them which is unity. The index ℓ^* for which r_k^{j,ℓ^*} is equal to unity, i.e., $r_k^{j,\ell^*} = 1$, is the index of the sub-object which the measurement \mathbf{y}_k^j (or \mathbf{z}_k^j) is associated to. The distribution of $r_k^{j,\ell}$ is defined as follows.

$$P\{r_k^{j,\ell} = 1 | \pi_k^{1:L}\} \triangleq \pi_k^{\ell} \quad (12)$$

for $\ell = 1, \dots, L$. The expression (12) can be written as

$$P\{\mathbf{r}_k^j | \pi_k^{1:L}\} \triangleq \prod_{\ell=1}^L (\pi_k^\ell)^{r_k^{j,\ell}} \quad (13)$$

for $j = 1, \dots, M_k$. Note that given the association vector \mathbf{r}_k^j , the noisy and noiseless measurements \mathbf{y}_k^j and \mathbf{z}_k^j are distributed as

$$p(\mathbf{y}_k^j | \mathbf{r}_k^j, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \mathcal{N}(\mathbf{y}_k^j; \mathbf{H}_{\ell^*} \mathbf{x}_k, s \mathbf{X}_k^{\ell^*} + \mathbf{R}), \quad (14a)$$

$$p(\mathbf{z}_k^j | \mathbf{r}_k^j, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_{\ell^*} \mathbf{x}_k, s \mathbf{X}_k^{\ell^*}), \quad (14b)$$

for $j = 1, \dots, M_k$ where ℓ^* is the index for which the element r_k^{j,ℓ^*} is equal to unity, i.e., $r_k^{j,\ell^*} = 1$. The expressions (14) can conveniently be written as

$$p(\mathbf{y}_k^j | \mathbf{r}_k^j, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \prod_{\ell=1}^L \mathcal{N}^{r_k^{j,\ell}}(\mathbf{y}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s \mathbf{X}_k^\ell + \mathbf{R}), \quad (15a)$$

$$p(\mathbf{z}_k^j | \mathbf{r}_k^j, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \prod_{\ell=1}^L \mathcal{N}^{r_k^{j,\ell}}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s \mathbf{X}_k^\ell), \quad (15b)$$

for $j = 1, \dots, M_k$. Using these expressions, we can write the conditional joint distribution for $\mathbf{y}_k^j, \mathbf{z}_k^j, \mathbf{r}_k^j$ as

$$p(\mathbf{y}_k^j, \mathbf{z}_k^j, \mathbf{r}_k^j | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = p(\mathbf{y}_k^j | \mathbf{z}_k^j) \times p(\mathbf{z}_k^j | \mathbf{r}_k^j, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) \times P\{\mathbf{r}_k^j | \pi_k^{1:L}\}, \quad (16a)$$

$$= \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \prod_{\ell=1}^L \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \prod_{\ell=1}^L (\pi_k^\ell)^{r_k^{j,\ell}}, \quad (16b)$$

$$= \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \prod_{\ell=1}^L (\pi_k^\ell)^{r_k^{j,\ell}} \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell). \quad (16c)$$

The overall conditional joint distribution for $\mathbf{y}_k^{1:M_k}, \mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k}$ can then be written as

$$p(\mathbf{y}_k^{1:M_k}, \mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k} | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) = \prod_{j=1}^{M_k} \left(\mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \prod_{\ell=1}^L (\pi_k^\ell)^{r_k^{j,\ell}} \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \right). \quad (17)$$

Since the posterior for $\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}$ given \mathcal{Y}_k is intractable, we aim for approximating the joint posterior for $\mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k}, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}$ given \mathcal{Y}_k in the following form.

$$p(\mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k}, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_k) \approx q_{\mathbf{z}}(\mathbf{z}_k^{1:M_k}) q_{\mathbf{r}}(\mathbf{r}_k^{1:M_k}) q_{\mathbf{x}}(\mathbf{x}_k) q_{\mathbf{X}}(\mathbf{X}_k^{1:L}) q_{\pi}(\pi_k^{1:L}). \quad (18)$$

We calculate the terms of the approximation above iteratively using variational approximation with the following true joint density.

$$\begin{aligned} & p(\mathbf{y}_k^{1:M_k}, \mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k}, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_{k-1}) \\ &= p(\mathbf{y}_k^{1:M_k}, \mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k} | \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L}) \\ & \quad \times p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_{k-1}), \\ &= \prod_{j=1}^{M_k} \left(\mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \prod_{\ell=1}^L (\pi_k^\ell)^{r_k^{j,\ell}} \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \right) \\ & \quad \times \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ & \quad \times \left(\prod_{\ell=1}^L \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) \right) \\ & \quad \times \mathcal{D}(\pi_k^{1:L}; \{\alpha_{k|k-1}^\ell\}_{\ell=1}^L). \end{aligned} \quad (19)$$

The logarithm of the joint density above is given as

$$\begin{aligned} & \log p(\mathbf{y}_k^{1:M_k}, \mathbf{z}_k^{1:M_k}, \mathbf{r}_k^{1:M_k}, \mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_{k-1}) \\ &= \sum_{j=1}^{M_k} \left(\log \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \right. \\ & \quad + \sum_{\ell=1}^L r_k^{j,\ell} \left(\log \pi_k^\ell + \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \right) \\ & \quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ & \quad + \sum_{\ell=1}^L \log \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) \\ & \quad \left. + \log \mathcal{D}(\pi_k^{1:L}; \{\alpha_{k|k-1}^\ell\}_{\ell=1}^L) \right). \end{aligned} \quad (20)$$

With this log-distribution, the factors $q_{\mathbf{z}}(\cdot)$, $q_{\mathbf{r}}(\cdot)$, $q_{\mathbf{x}}(\cdot)$, $q_{\mathbf{X}}(\cdot)$, and $q_{\pi}(\cdot)$ can be calculated using variational Bayes approach [21, Ch. 10] as

$$q_{\mathbf{z}}(\mathbf{z}_k^{1:M_k}) = \prod_{j=1}^{M_k} \mathcal{N}(\mathbf{z}_k^j; \hat{\mathbf{z}}_{k|k}^j, \mathbf{P}_{k|k}^{z,j}), \quad (21a)$$

$$q_{\mathbf{r}}(\mathbf{r}_k^{1:M_k}) = \prod_{j=1}^{M_k} \prod_{\ell=1}^L (\gamma_{k|k}^{j,\ell})^{r_k^{j,\ell}}, \quad (21b)$$

$$q_{\mathbf{x}}(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k}, \mathbf{P}_{k|k}), \quad (21c)$$

$$q_{\mathbf{X}}(\mathbf{X}_k^{1:L}) = \prod_{\ell=1}^L \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k}^\ell, \mathbf{V}_{k|k}^\ell), \quad (21d)$$

$$q_{\pi}(\pi_k^{1:L}) = \mathcal{D}(\pi_k^{1:L}; \alpha_{k|k}^{1:L}), \quad (21e)$$

where the parameters $\hat{\mathbf{z}}_{k|k}^j, \mathbf{P}_{k|k}^{z,j}, \gamma_{k|k}^{j,\ell}, \mathbf{m}_{k|k}, \mathbf{P}_{k|k}, v_{k|k}^\ell, \mathbf{V}_{k|k}^\ell, \alpha_{k|k}^\ell$, $j = 1, \dots, M_k$, $\ell = 1, \dots, L$, are found iteratively. Once the updated distributions in (21) are obtained we can approximate the updated posterior $p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_k)$ as

$$p(\mathbf{x}_k, \mathbf{X}_k^{1:L}, \pi_k^{1:L} | \mathcal{Y}_k) \approx q_{\mathbf{x}}(\mathbf{x}_k) q_{\mathbf{X}}(\mathbf{X}_k^{1:L}) q_{\pi}(\pi_k^{1:L}), \quad (22)$$

which is in the form (5) as required. The derivations for the distributions in (21) and for the iterations of their parameters are given in the following subsections. A summary of the resulting iterative update procedure can be found in Algorithm 1.

A. Calculation of $q_{\mathbf{z}}^{(i+1)}(\cdot)$

The density $q_{\mathbf{z}}^{(i+1)}(\cdot)$ is given as [21, Ch. 10]

$$\log q_{\mathbf{z}}^{(i+1)}(\mathbf{z}_k^{1:M_k}) = \mathbb{E}_{\mathbf{z}} [\log p(\cdot | \mathcal{Y}_{k-1})] + c_{\mathbf{z}} \quad (23)$$

where $c_{\mathbf{z}}$ denotes any constant term(s) with respect to the variables $\mathbf{z}_k^{1:M_k}$. The joint log-distribution in the expectation above is given as

$$\begin{aligned} & \log p(\cdot | \mathcal{Y}_{k-1}) \\ &= \sum_{j=1}^{M_k} \left(\log \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \right. \\ & \quad \left. + \sum_{\ell=1}^L r_k^{j,\ell} \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \right) + c_{\mathbf{z}}, \end{aligned} \quad (24a)$$

$$\begin{aligned} &= \sum_{j=1}^{M_k} \left(\log \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \right. \\ & \quad \left. - \frac{1}{2} \sum_{\ell=1}^L r_k^{j,\ell} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k)^\top (s\mathbf{X}_k^\ell)^{-1} (\cdot) \right) + c_{\mathbf{z}}. \end{aligned} \quad (24b)$$

Taking the expectation of both sides, we get

$$\begin{aligned} & \mathbb{E}_{\mathbf{z}} [\log p(\cdot | \mathcal{Y}_{k-1})] \\ &= \sum_{j=1}^{M_k} \left(\log \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) \right. \\ & \quad \left. - \frac{1}{2} \sum_{\ell=1}^L \overline{r_k^{j,\ell}} (\mathbf{z}_k^j - \mathbf{H}_\ell \overline{\mathbf{x}_k})^\top \overline{(s\mathbf{X}_k^\ell)^{-1}} (\cdot) \right) + c_{\mathbf{z}}, \end{aligned} \quad (25a)$$

$$= \sum_{j=1}^{M_k} \left(\log \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) - \frac{1}{2} \sum_{\ell=1}^L (\mathbf{z}_k^j - \mathbf{H}_\ell \bar{\mathbf{x}}_k)^T \overline{r_k^{j,\ell} (s\mathbf{X}_k^\ell)^{-1}} (\cdot) \right) + c_{\setminus \mathbf{z}}, \quad (25b)$$

$$= \sum_{j=1}^{M_k} \left(\log \mathcal{N}(\mathbf{y}_k^j; \mathbf{z}_k^j, \mathbf{R}) + \sum_{\ell=1}^L \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \bar{\mathbf{x}}_k, \left(\overline{r_k^{j,\ell} (s\mathbf{X}_k^\ell)^{-1}} \right)^{-1}) \right) + c_{\setminus \mathbf{z}}, \quad (25c)$$

$$= \sum_{j=1}^{M_k} \log \mathcal{N}(\mathbf{z}_k^j; \hat{\mathbf{z}}_{k|k}^{j,(i+1)}, \mathbf{P}_{k|k}^{z,j,(i+1)}) + c_{\setminus \mathbf{z}}, \quad (25d)$$

where

$$\hat{\mathbf{z}}_{k|k}^{j,(i+1)} \triangleq \mathbf{P}_{k|k}^{z,j,(i+1)} \left(\mathbf{R}^{-1} \mathbf{y}_k^j + \sum_{\ell=1}^L \overline{r_k^{j,\ell} (s\mathbf{X}_k^\ell)^{-1}} \mathbf{H}_\ell \bar{\mathbf{x}}_k \right), \quad (26a)$$

$$\mathbf{P}_{k|k}^{z,j,(i+1)} \triangleq \left(\mathbf{R}^{-1} + \sum_{\ell=1}^L \overline{r_k^{j,\ell} (s\mathbf{X}_k^\ell)^{-1}} \right)^{-1}, \quad (26b)$$

for $j = 1, \dots, M_k$. Hence we have

$$q_{\mathbf{z}}^{(i+1)}(\mathbf{z}_k^{1:M_k}) = \prod_{j=1}^{M_k} \mathcal{N}(\mathbf{z}_k^j; \hat{\mathbf{z}}_{k|k}^{j,(i+1)}, \mathbf{P}_{k|k}^{z,j,(i+1)}). \quad (27)$$

B. Calculation of $q_{\mathbf{r}}^{(i+1)}(\cdot)$

The density $q_{\mathbf{r}}^{(i+1)}(\cdot)$ is given as

$$\log q_{\mathbf{r}}^{(i+1)}(\mathbf{r}_k^{1:M_k}) = \mathbb{E}_{\setminus \mathbf{r}} [\log p(\cdot | \mathcal{Y}_{k-1})] + c_{\setminus \mathbf{r}} \quad (28)$$

where $c_{\setminus \mathbf{r}}$ denotes any constant term(s) with respect to the variables $\mathbf{r}_k^{1:M_k}$. The joint log-distribution in the expectation above is given as

$$\begin{aligned} & \log p(\cdot | \mathcal{Y}_{k-1}) \\ &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \left(\log \pi_k^\ell + \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \right) + c_{\setminus \mathbf{r}}, \end{aligned} \quad (29a)$$

$$\begin{aligned} &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \left(\log \pi_k^\ell - \frac{1}{2} \log |2\pi s\mathbf{X}_k^\ell| \right. \\ &\quad \left. - \frac{1}{2} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k)^T (s\mathbf{X}_k^\ell)^{-1} (\cdot) \right) + c_{\setminus \mathbf{r}}, \end{aligned} \quad (29b)$$

$$\begin{aligned} &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \left(\log \pi_k^\ell - \frac{1}{2} \log |2\pi s\mathbf{X}_k^\ell| \right. \\ &\quad \left. - \frac{1}{2} \text{tr} \left((s\mathbf{X}_k^\ell)^{-1} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \right) + c_{\setminus \mathbf{r}}. \end{aligned} \quad (29c)$$

Taking the expectation of both sides, we get

$$\begin{aligned} & \mathbb{E}_{\setminus \mathbf{r}} [\log p(\cdot | \mathcal{Y}_{k-1})] \\ &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \left(\overline{\log \pi_k^\ell} - \frac{1}{2} \overline{\log |2\pi s\mathbf{X}_k^\ell|} \right. \end{aligned}$$

$$\left. - \frac{1}{2} \text{tr} \left((s\mathbf{X}_k^\ell)^{-1} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \right) + c_{\setminus \mathbf{r}}, \quad (30a)$$

$$\begin{aligned} &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \log \exp \left(\overline{\log \pi_k^\ell} - \frac{1}{2} \overline{\log |2\pi s\mathbf{X}_k^\ell|} \right. \\ &\quad \left. - \frac{1}{2} \text{tr} \left((s\mathbf{X}_k^\ell)^{-1} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \right) + c_{\setminus \mathbf{r}}, \end{aligned} \quad (30b)$$

$$\begin{aligned} &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L \log \left(\left(\exp \left(\overline{\log \pi_k^\ell} - \frac{1}{2} \overline{\log |2\pi s\mathbf{X}_k^\ell|} \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{1}{2} \text{tr} \left((s\mathbf{X}_k^\ell)^{-1} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \right) \right)^{r_k^{j,\ell}} \right) + c_{\setminus \mathbf{r}}, \end{aligned}$$

$$= \sum_{j=1}^{M_k} \sum_{\ell=1}^L \log \left((\gamma_{k|k}^{j,\ell,(i+1)})^{r_k^{j,\ell}} \right) + c_{\setminus \mathbf{r}}, \quad (30c)$$

where

$$\gamma_{k|k}^{j,\ell,(i+1)} \triangleq \frac{\tilde{\gamma}_{k|k}^{j,\ell,(i+1)}}{\sum_{\ell=1}^L \tilde{\gamma}_{k|k}^{j,\ell,(i+1)}}, \quad (31a)$$

$$\begin{aligned} \tilde{\gamma}_{k|k}^{j,\ell,(i+1)} &\triangleq \exp \left(\overline{\log \pi_k^\ell} - \frac{1}{2} \overline{\log |s\mathbf{X}_k^\ell|} \right. \\ &\quad \left. - \frac{1}{2} \text{tr} \left((s\mathbf{X}_k^\ell)^{-1} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \right), \end{aligned} \quad (31b)$$

for $j = 1, \dots, M_k$, $\ell = 1, \dots, L$. Hence, we have

$$q_{\mathbf{r}}^{(i+1)}(\mathbf{r}_k^{1:M_k}) = \prod_{j=1}^{M_k} \prod_{\ell=1}^L (\gamma_{k|k}^{j,\ell,(i+1)})^{r_k^{j,\ell}}. \quad (32)$$

Note that the expression above corresponds to the following probabilistic characterization.

$$P_{q_{\mathbf{r}}^{(i+1)}} \{r_k^{j,\ell} = 1\} = \gamma_{k|k}^{j,\ell,(i+1)} \quad (33)$$

for $j = 1, \dots, M_k$, $\ell = 1, \dots, L$.

C. Calculation of $q_{\mathbf{x}}^{(i+1)}(\cdot)$

The density $q_{\mathbf{x}}^{(i+1)}(\cdot)$ is given as

$$\log q_{\mathbf{x}}^{(i+1)}(\mathbf{x}_k) = \mathbb{E}_{\setminus \mathbf{x}} [\log p(\cdot | \mathcal{Y}_{k-1})] + c_{\setminus \mathbf{x}} \quad (34)$$

where $c_{\setminus \mathbf{x}}$ denotes any constant terms with respect to the variable \mathbf{x}_k . The joint log-distribution in the expectation above is given as

$$\begin{aligned} & \log p(\cdot | \mathcal{Y}_{k-1}) \\ &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \\ &\quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (35a)$$

$$\begin{aligned} &= \sum_{\ell=1}^L \sum_{j=1}^{M_k} r_k^{j,\ell} \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \\ &\quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (35b)$$

$$\begin{aligned} &= -\frac{1}{2} \sum_{\ell=1}^L \sum_{j=1}^{M_k} r_k^{j,\ell} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k)^T (s\mathbf{X}_k^\ell)^{-1} (\cdot) \\ &\quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (35c)$$

$$= -\frac{1}{2} \sum_{\ell=1}^L \text{tr} \left((s\mathbf{X}_k^\ell)^{-1} \sum_{j=1}^{M_k} r_k^{j,\ell} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}. \quad (35d)$$

Taking the expectation of both sides, we get

$$\begin{aligned} & \mathbb{E}_{\setminus \mathbf{x}} [\log p(\cdot | \mathcal{Y}_{k-1})] \\ &= -\frac{1}{2} \sum_{\ell=1}^L \text{tr} \left(\overline{(s\mathbf{X}_k^\ell)^{-1}} \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} (\overline{\mathbf{z}_k^j} - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \\ & \quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (36a)$$

$$\begin{aligned} &= -\frac{1}{2} \sum_{\ell=1}^L \text{tr} \left(\overline{(s\mathbf{X}_k^\ell)^{-1}} \left(\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} \right) \right. \\ & \quad \times \left. \frac{\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} (\overline{\mathbf{z}_k^j} - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T}{\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}}} \right) \\ & \quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (36b)$$

$$\begin{aligned} &= -\frac{1}{2} \sum_{\ell=1}^L \text{tr} \left(\overline{(s\mathbf{X}_k^\ell)^{-1}} \left(\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} \right) \right. \\ & \quad \times \left. \left(\frac{\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} \overline{\mathbf{z}_k^j}}{\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}}} - \mathbf{H}_\ell \mathbf{x}_k \right) (\cdot)^T \right) \\ & \quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (36c)$$

$$\begin{aligned} &= -\frac{1}{2} \sum_{\ell=1}^L \text{tr} \left(\overline{(s\mathbf{X}_k^\ell)^{-1}} \left(\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} \right) (\mathbf{u}_k^\ell - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \\ & \quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \\ &= \sum_{\ell=1}^L \log \mathcal{N}(\mathbf{u}_k^\ell; \mathbf{H}_\ell \mathbf{x}_k, (\mathbf{\Lambda}_k^\ell)^{-1}) \\ & \quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (36d)$$

$$\begin{aligned} &= \log \mathcal{N}(\mathbf{u}_K^{1:L}; \mathbf{H}_{1:L} \mathbf{x}_k, \mathbf{\Lambda}_{1:L}^{-1}) \\ & \quad + \log \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (36e)$$

where

$$\mathbf{u}_k^\ell \triangleq \frac{\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} \overline{\mathbf{z}_k^j}}{\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}}}, \quad (37a)$$

$$\mathbf{\Lambda}_k^\ell \triangleq \overline{(s\mathbf{X}_k^\ell)^{-1}} \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}}, \quad (37b)$$

$$\mathbf{u}_k^{1:L} \triangleq \begin{bmatrix} (\mathbf{u}_k^1)^T & (\mathbf{u}_k^2)^T & \cdots & (\mathbf{u}_k^L)^T \end{bmatrix}^T, \quad (37c)$$

$$\mathbf{H}_{1:L} \triangleq \begin{bmatrix} \mathbf{H}_1^T & \mathbf{H}_2^T & \cdots & \mathbf{H}_L^T \end{bmatrix}^T, \quad (37d)$$

$$\mathbf{\Lambda}_{1:L} \triangleq \text{blkdiag}(\mathbf{\Lambda}_k^1, \mathbf{\Lambda}_k^2, \dots, \mathbf{\Lambda}_k^L). \quad (37e)$$

for $\ell = 1, \dots, L$. Hence we have

$$q_{\mathbf{x}}^{(i+1)}(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k}^{(i+1)}, \mathbf{P}_{k|k}^{(i+1)}), \quad (38)$$

where

$$\mathbf{m}_{k|k}^{(i+1)} = \mathbf{P}_{k|k}^{(i+1)} (\mathbf{P}_{k|k-1}^{-1} \mathbf{m}_{k|k-1} + \mathbf{H}_{1:L}^T \mathbf{\Lambda}_{1:L} \mathbf{u}_k^{1:L}), \quad (39a)$$

$$\mathbf{P}_{k|k}^{(i+1)} = (\mathbf{P}_{k|k-1}^{-1} + \mathbf{H}_{1:L}^T \mathbf{\Lambda}_{1:L} \mathbf{H}_{1:L})^{-1}. \quad (39b)$$

D. Calculation of $q_{\mathbf{x}}^{(i+1)}(\cdot)$

The density $q_{\mathbf{x}}^{(i+1)}(\cdot)$ is given as

$$\log q_{\mathbf{x}}^{(i+1)}(\mathbf{X}_k^{1:L}) = \mathbb{E}_{\setminus \mathbf{x}} [\log p(\cdot | \mathcal{Y}_{k-1})] + c_{\setminus \mathbf{x}} \quad (40)$$

where $c_{\setminus \mathbf{x}}$ denotes any constant terms with respect to the variables $\mathbf{X}_k^{1:L}$. The joint log-distribution in the expectation above is given as

$$\begin{aligned} & \log p(\cdot | \mathcal{Y}_{k-1}) \\ &= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \\ & \quad + \sum_{\ell=1}^L \log \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (41a)$$

$$\begin{aligned} &= \sum_{\ell=1}^L \left(\sum_{j=1}^{M_k} r_k^{j,\ell} \log \mathcal{N}(\mathbf{z}_k^j; \mathbf{H}_\ell \mathbf{x}_k, s\mathbf{X}_k^\ell) \right. \\ & \quad \left. + \log \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) \right) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (41b)$$

$$\begin{aligned} &= \sum_{\ell=1}^L \left(-\frac{1}{2} \left(\sum_{j=1}^{M_k} r_k^{j,\ell} \right) \log |\mathbf{X}_k^\ell| \right. \\ & \quad \left. - \frac{1}{2} \text{tr} \left((\mathbf{X}_k^\ell)^{-1} \frac{1}{s} \sum_{j=1}^{M_k} r_k^{j,\ell} (\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \right. \\ & \quad \left. + \log \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) \right) + c_{\setminus \mathbf{x}}. \end{aligned} \quad (41c)$$

Taking the expectation of both sides, we get

$$\begin{aligned} & \mathbb{E}_{\setminus \mathbf{x}} [\log p(\cdot | \mathcal{Y}_{k-1})] \\ &= \sum_{\ell=1}^L \left(-\frac{1}{2} \left(\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} \right) \log |\mathbf{X}_k^\ell| \right. \\ & \quad \left. - \frac{1}{2} \text{tr} \left((\mathbf{X}_k^\ell)^{-1} \frac{1}{s} \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} (\overline{\mathbf{z}_k^j} - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \right. \\ & \quad \left. + \log \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) \right) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (42a)$$

$$\begin{aligned} &= \sum_{\ell=1}^L \log \mathcal{IW} \left(\mathbf{X}_k^\ell; \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}}, \frac{1}{s} \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} (\overline{\mathbf{z}_k^j} - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T \right) \\ & \quad + \log \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell) + c_{\setminus \mathbf{x}}, \\ &= \sum_{\ell=1}^L \log \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k}^{\ell,(i+1)}, \mathbf{V}_{k|k}^{\ell,(i+1)}) + c_{\setminus \mathbf{x}}, \end{aligned} \quad (42b)$$

where

$$v_{k|k}^{\ell,(i+1)} \triangleq v_{k|k-1}^\ell + \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}}, \quad (43a)$$

$$\mathbf{V}_{k|k}^{\ell,(i+1)} \triangleq \mathbf{V}_{k|k-1}^\ell + \frac{1}{s} \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} (\overline{\mathbf{z}_k^j} - \mathbf{H}_\ell \mathbf{x}_k) (\cdot)^T, \quad (43b)$$

for $\ell = 1, \dots, L$. Hence, we have

$$q_X^{(i+1)}(\mathbf{X}_k^{1:L}) = \prod_{\ell=1}^L \mathcal{IW}(\mathbf{X}_k^\ell; v_{k|k}^{\ell,(i+1)}, \mathbf{V}_{k|k}^{\ell,(i+1)}). \quad (44)$$

E. Calculation of $q_\pi^{(i+1)}(\cdot)$

The density $q_\pi^{(i+1)}(\cdot)$ is given as

$$\log q_\pi^{(i+1)}(\pi_k^{1:L}) = \mathbb{E}_{\setminus \pi} [\log p(\cdot | \mathbf{Y}_{k-1})] + c_{\setminus \pi} \quad (45)$$

where $c_{\setminus \pi}$ denotes any constant terms with respect to the variables $\pi_k^{1:L}$. The joint log-distribution in the expectation above is given as

$$\begin{aligned} \log p(\cdot | \mathbf{Y}_{k-1}) &= \left(\sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \log \pi_k^\ell \right) \\ &+ \log \mathcal{D}(\pi_k^{1:L}; \{\alpha_{k|k-1}^\ell\}_{\ell=1}^L) + c_{\setminus \pi}, \end{aligned} \quad (46a)$$

$$= \sum_{j=1}^{M_k} \sum_{\ell=1}^L r_k^{j,\ell} \log \pi_k^\ell + \sum_{\ell=1}^L (\alpha_{k|k-1}^\ell - 1) \log \pi_k^\ell + c_{\setminus \pi}, \quad (46b)$$

$$= \sum_{\ell=1}^L \left(\sum_{j=1}^{M_k} r_k^{j,\ell} + \alpha_{k|k-1}^\ell - 1 \right) \log \pi_k^\ell + c_{\setminus \pi}. \quad (46c)$$

Taking the expectation of both sides, we get

$$\begin{aligned} \mathbb{E}_{\setminus \pi} [\log p(\cdot | \mathbf{Y}_{k-1})] &= \sum_{\ell=1}^L \left(\sum_{j=1}^{M_k} \overline{r_k^{j,\ell}} + \alpha_{k|k-1}^\ell - 1 \right) \log \pi_k^\ell + c_{\setminus \pi}. \end{aligned} \quad (47)$$

Hence we have

$$q_\pi^{(i+1)}(\pi_k^{1:L}) = \mathcal{D}(\pi_k^{1:L}; \{\alpha_{k|k}^{\ell,(i+1)}\}_{\ell=1}^L) \quad (48)$$

where

$$\alpha_{k|k}^{\ell,(i+1)} \triangleq \alpha_{k|k-1}^\ell + \sum_{j=1}^{M_k} \overline{r_k^{j,\ell}}, \quad (49)$$

for $\ell = 1, \dots, L$.

F. Expected Values Required for Iterations

The expected values required in the previous subsections can be calculated as follows.

$$\begin{aligned} \overline{(s\mathbf{X}_k^\ell)^{-1}} &\triangleq \mathbb{E}_{\mathbf{X}}[(s\mathbf{X}_k^\ell)^{-1}] \\ &= (v_{k|k}^{\ell,(i)} - n_y - 1)(s\mathbf{V}_{k|k}^{\ell,(i)})^{-1}, \end{aligned} \quad (50a)$$

$$\overline{\mathbf{x}_k} \triangleq \mathbb{E}_{\mathbf{X}}[\mathbf{x}_k] = \mathbf{m}_{k|k}^{(i)}, \quad (50b)$$

$$\overline{r_k^{j,\ell}} \triangleq \mathbb{E}_{\mathbf{r}}[r_k^{j,\ell}] = P_{q_r^{(i)}}\{r_k^{j,\ell} = 1\} = \gamma_{k|k}^{j,\ell,(i)}, \quad (50c)$$

$$\begin{aligned} \overline{\log \pi_k^\ell} &\triangleq E_\pi[\log \pi_k^\ell] \\ &= \psi(\alpha_{k|k}^{\ell,(i)}) - \psi\left(\sum_{\ell=1}^L \alpha_{k|k}^{\ell,(i)}\right), \end{aligned} \quad (50d)$$

$$\overline{\log |\mathbf{X}_k^\ell|} \triangleq \mathbb{E}_{\mathbf{X}}[\log |\mathbf{X}_k^\ell|]$$

$$\begin{aligned} &= \log |\mathbf{V}_{k|k}^{\ell,(i)}| - n_y \log 2 \\ &\quad - \sum_{d=1}^{n_y} \psi\left(\frac{v_{k|k}^{\ell,(i)} - n_y - d}{2}\right), \end{aligned} \quad (50e)$$

$$\overline{\mathbf{z}_k^j} \triangleq \mathbb{E}_{\mathbf{z}}[\mathbf{z}_k^j] = \mathbf{z}_{k|k}^{j,(i)}, \quad (50f)$$

$$\begin{aligned} \overline{(\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k)(\cdot)^T} &\triangleq \mathbb{E}_{\mathbf{z}, \mathbf{x}}[(\mathbf{z}_k^j - \mathbf{H}_\ell \mathbf{x}_k)(\cdot)^T] \\ &= (\mathbf{z}_{k|k}^{j,(i)} - \mathbf{H}_\ell \mathbf{m}_{k|k}^{(i)})(\cdot)^T + \mathbf{P}_{k|k}^{z,j,(i)} \\ &\quad + \mathbf{H}_\ell \mathbf{P}_{k|k}^{(i)} \mathbf{H}_\ell^T, \end{aligned} \quad (50g)$$

for $j = 1, \dots, M_k$, $\ell = 1, \dots, L$ where $\psi(\cdot)$ denotes the digamma function (i.e., the logarithmic derivative of the Gamma function or polygamma function of order zero.).

Algorithm 1 Measurement Update for ME-ETT Model

1: **Given:** $\mathbf{m}_{k|k-1}$, $\mathbf{P}_{k|k-1}$, $\{v_{k|k-1}^\ell, \mathbf{V}_{k|k-1}^\ell, \alpha_{k|k-1}^\ell\}_{\ell=1}^L$ and $\mathbf{y}_k^{1:M_k}$, calculate $\mathbf{m}_{k|k}$, $\mathbf{P}_{k|k}$, $\{v_{k|k}^\ell, \mathbf{V}_{k|k}^\ell, \alpha_{k|k}^\ell\}_{\ell=1}^L$ as follows.

2: **Initialization at time $k = 0$:**

3: $\mathbf{m}_{k|k}^{(0)} \leftarrow \mathbf{m}_{k|k-1}$

4: $\mathbf{P}_{k|k}^{(0)} \leftarrow \mathbf{P}_{k|k-1}$

5: **for** $\ell = 1, \dots, L$ **do**

6: $v_{k|k}^{\ell,(0)} \leftarrow v_{k|k-1}^\ell$

7: $\mathbf{V}_{k|k}^{\ell,(0)} \leftarrow \mathbf{V}_{k|k-1}^\ell$

8: $\alpha_{k|k}^{\ell,(0)} \leftarrow \alpha_{k|k-1}^\ell$

9: **for** $j = 1, \dots, M_k$ **do**

10: $\gamma_{k|k}^{j,\ell,(0)} \leftarrow \frac{\alpha_{k|k-1}^\ell}{\sum_{\ell'=1}^L \alpha_{k|k-1}^{\ell'}}$

11: **end for**

12: **end for**

13: **for** $j = 1, \dots, M_k$ **do**

14: $\mathbf{z}_{k|k}^{j,(0)} \leftarrow \mathbf{y}_k^j$

15: $\mathbf{P}_{k|k}^{z,j,(0)} \leftarrow \mathbf{R}$

16: **end for**

17: **Iterations:**

18: $\mathbf{H}_{1:L} = [\mathbf{H}_1^T \ \mathbf{H}_2^T \ \dots \ \mathbf{H}_L^T]^T$

19: **for** $i = 0, \dots, i_{max} - 1$ **do**

20: **for** $j = 1, \dots, M_k$ **do**

21: **for** $\ell = 1, \dots, L$ **do**

22: $\mathbf{W}_{j\ell}^{(i)} \triangleq (\mathbf{z}_{k|k}^{j,(i)} - \mathbf{H}_\ell \mathbf{m}_{k|k}^{(i)})(\mathbf{z}_{k|k}^{j,(i)} - \mathbf{H}_\ell \mathbf{m}_{k|k}^{(i)})^T + \mathbf{P}_{k|k}^{z,j,(i)} + \mathbf{H}_\ell \mathbf{P}_{k|k}^{(i)} \mathbf{H}_\ell^T$

23: $\gamma_{k|k}^{j,\ell,(i+1)} = \exp\left(\psi(\alpha_{k|k}^{\ell,(i)}) - \frac{1}{2} \log |\mathbf{V}_{k|k}^{\ell,(i)}| + \frac{1}{2} \sum_{d=1}^{n_y} \psi\left(\frac{v_{k|k}^{\ell,(i)} - n_y - d}{2}\right) - \frac{v_{k|k}^{\ell,(i)} - n_y - 1}{2} \text{tr}((s\mathbf{V}_{k|k}^{\ell,(i)})^{-1} \mathbf{W}_{j\ell}^{(i)})\right)$

24: **end for**

25: **for** $\ell = 1, \dots, L$ **do**

26: $\gamma_{k|k}^{j,\ell,(i+1)} = \frac{\gamma_{k|k}^{j,\ell,(i+1)}}{\sum_{\ell'=1}^L \gamma_{k|k}^{j,\ell',(i+1)}}$

27: **end for**

28: **end for**

29: **for** $j = 1, \dots, M_k$ **do**

30: $\mathbf{P}_{k|k}^{z,j,(i+1)} = \left(\mathbf{R}^{-1} + \sum_{\ell=1}^L \left[\gamma_{k|k}^{j,\ell,(i)} (v_{k|k}^{\ell,(i)} - n_y - 1)(s\mathbf{V}_{k|k}^{\ell,(i)})^{-1}\right]\right)^{-1}$

31: $\mathbf{z}_{k|k}^{j,(i+1)} = \mathbf{P}_{k|k}^{z,j,(i+1)} \left(\mathbf{R}^{-1} \mathbf{y}_k^j + \sum_{\ell=1}^L \left[\gamma_{k|k}^{j,\ell,(i)} (v_{k|k}^{\ell,(i)} - n_y - 1)(s\mathbf{V}_{k|k}^{\ell,(i)})^{-1} \mathbf{H}_\ell \mathbf{m}_{k|k}^{(i)}\right]\right)$

32: **end for**

33: **for** $\ell = 1, \dots, L$ **do**

34: $v_{k|k}^{\ell,(i+1)} = v_{k|k-1}^\ell + \sum_{j=1}^{M_k} \gamma_{k|k}^{j,\ell,(i)}$

```

35:  $\mathbf{V}_{k|k}^{\ell,(i+1)} = \mathbf{V}_{k|k-1}^{\ell} + \frac{1}{s} \sum_{j=1}^{M_k} [\gamma_{k|k}^{j,\ell,(i)} \mathbf{W}_{j\ell}^{(i)}]$ 
36:  $\alpha_{k|k}^{\ell,(i+1)} = \alpha_{k|k-1}^{\ell} + \sum_{j=1}^{M_k} \gamma_{k|k}^{j,\ell,(i)}$ 
37:  $\mathbf{u}_k^{\ell} = \frac{\sum_{j=1}^{M_k} [\gamma_{k|k}^{j,\ell,(i)} \mathbf{z}_{k|k}^{j,\ell,(i)}]}{\sum_{j=1}^{M_k} \gamma_{k|k}^{j,\ell,(i)}}$ 
38:  $\Lambda_k^{\ell} = (v_{k|k}^{\ell,(i)} - n_y - 1) (s \mathbf{V}_{k|k}^{\ell,(i)})^{-1} \sum_{j=1}^{M_k} \gamma_{k|k}^{j,\ell,(i)}$ 
39: end for
40:  $\mathbf{u}_k^{1:L} = [(\mathbf{u}_k^1)^T \ (\mathbf{u}_k^2)^T \ \dots \ (\mathbf{u}_k^L)^T]^T$ 
41:  $\Lambda_{1:L} = \text{blkdiag}(\Lambda_k^1, \Lambda_k^2, \dots, \Lambda_k^L)$ 
42:  $\mathbf{P}_{k|k}^{(i+1)} = \left( \mathbf{P}_{k|k-1}^{-1} + \mathbf{H}_{1:L}^T \Lambda_{1:L} \mathbf{H}_{1:L} \right)^{-1}$ 
43:  $\mathbf{m}_{k|k}^{(i+1)} = \mathbf{P}_{k|k}^{(i+1)} \left( \mathbf{P}_{k|k-1}^{-1} \mathbf{m}_{k|k-1} + \mathbf{H}_{1:L}^T \Lambda_{1:L} \mathbf{u}_k^{1:L} \right)$ 
44: end for
45: Set final estimates:
46:  $\mathbf{m}_{k|k} = \mathbf{m}_{k|k}^{(i_{\max})}$ 
47:  $\mathbf{P}_{k|k} = \mathbf{P}_{k|k}^{(i_{\max})}$ 
48: for  $\ell = 1, \dots, L$  do
49:  $v_{k|k}^{\ell} = v_{k|k}^{\ell,(i_{\max})}$ 
50:  $\mathbf{V}_{k|k}^{\ell} = \mathbf{V}_{k|k}^{\ell,(i_{\max})}$ 
51:  $\alpha_{k|k}^{\ell} = \alpha_{k|k}^{\ell,(i_{\max})}$ 
52: end for

```

IV. TIME UPDATE

With the target dynamics given in (2), the time update of the kinematic state is performed following the regular Kalman filter time update equations

$$\mathbf{m}_{k|k-1} = \mathbf{F} \mathbf{m}_{k-1|k-1}, \quad (51a)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F} \mathbf{P}_{k-1|k-1} \mathbf{F}^T + \mathbf{Q}. \quad (51b)$$

The parameters of the inverse Wishart distribution of each sub-object are updated with a forgetting factor λ_{IW} ,

$$v_{k|k-1}^{\ell} = \lambda_{IW} v_{k-1|k-1}^{\ell}, \quad (52a)$$

$$\mathbf{V}_{k|k-1}^{\ell} = \lambda_{IW} \mathbf{V}_{k-1|k-1}^{\ell}, \quad \text{for } \ell = 1, \dots, L. \quad (52b)$$

The forgetting factor is used to obtain the maximum entropy prediction density of the extent states when the dynamics of the extent state is slowly varying and unknown [22, Theorem 1]. Similarly, the sufficient statistics of the Dirichlet distributed mixture weights $\pi_k^{1:L}$ are propagated with a forgetting factor λ_D ,

$$\alpha_{k|k-1}^{\ell} = \lambda_D \alpha_{k-1|k-1}^{\ell} \quad \text{for } \ell = 1, \dots, L. \quad (53)$$

If the true parameter evolution model is slowly varying, the time update equations (52) and (53) will not underestimate the uncertainty by maximizing the entropy. Note that, for the special case of stationary parameters forgetting factors are set to 1, i.e., $\lambda_{IW} = \lambda_D = 1$.

The proposed algorithm is also versatile to perform under alternative random matrix time update schemes.

V. EXPERIMENTAL RESULTS

In this section, we will demonstrate the performance of the proposed algorithm in various experiments and compare its performance with the algorithms presented in [6], [12], [17], [20]. In the sequel, we will denote the proposed algorithm as VB and the methods in [17] and [20] as VPF and JL, respectively.

To assess the performance of the algorithms, we consider the Intersection-Over-Union (IOU) similarity measure [3], [7], [23] together with Gaussian Wasserstein (GW) distance [24], [25] for the extent estimates and the root-mean-square error (RMSE) for the kinematic state estimates. The RMSE between the true and the estimated states is defined as

$$\text{RMSE}(\mathbf{x}^{\text{true}}, \mathbf{x}) = \sqrt{\frac{1}{N} \sum_{k=1}^N \|\mathbf{x}_k^{\text{true}} - \mathbf{x}_k\|_2^2}, \quad (54)$$

where N is the number of time steps in a single run.

The IOU measure between the estimated extent and the true extent is calculated as

$$\text{IOU}(\mathbf{X}_k^{\text{true}}, \mathbf{X}_{k|k}) = \frac{\text{area}(\mathbf{X}_k^{\text{true}} \cap \mathbf{X}_{k|k})}{\text{area}(\mathbf{X}_k^{\text{true}} \cup \mathbf{X}_{k|k})} \in [0, 1], \quad (55)$$

$$\text{IOU}(\mathbf{X}^{\text{true}}, \mathbf{X}) = \frac{1}{N} \sum_{k=1}^N \text{IOU}(\mathbf{X}_k^{\text{true}}, \mathbf{X}_{k|k}), \quad (56)$$

where $\mathbf{X}_k^{\text{true}}$ and $\mathbf{X}_{k|k}$ denotes the true and estimated extent matrices at time k , respectively. Note that IOU takes values between 0 and 1, where 1 corresponds to the perfect match, while 0 indicates no intersection between the true and the estimated extents. The GW distance [25] can be expressed as

$$\text{GW}(\mathbf{c}_k^{\text{true}}, \mathbf{X}_k^{\text{true}}, \mathbf{c}_k, \mathbf{X}_{k|k})^2 \triangleq \left\| \mathbf{c}_k^{\text{true}} - \mathbf{c}_{k|k} \right\|_2^2 + \text{tr} \left[\mathbf{X}_k^{\text{true}} + \mathbf{X}_{k|k} - 2(\mathbf{X}_k^{\text{true}})^{\frac{1}{2}} \mathbf{X}_{k|k} (\mathbf{X}_k^{\text{true}})^{\frac{1}{2}} \right]^{\frac{1}{2}}. \quad (57)$$

$$\begin{aligned} &\text{GW}(\mathbf{c}^{\text{true}}, \mathbf{X}^{\text{true}}, \mathbf{c}, \mathbf{X}) \\ &= \frac{1}{N} \sum_{k=1}^N \text{GW}(\mathbf{c}_k^{\text{true}}, \mathbf{X}_k^{\text{true}}, \mathbf{c}_{k|k}, \mathbf{X}_{k|k}), \end{aligned} \quad (58)$$

where $\mathbf{c}_k^{\text{true}}$, $\mathbf{c}_{k|k}$ and $\mathbf{X}_k^{\text{true}}$, $\mathbf{X}_{k|k}$ stand for true and estimated center locations and true and estimated extent matrices at time k , respectively.

The comparison metrics above are formulated for a single simulation run. In the sequel, all simulation experiments are performed 100 times with different realizations of the measurement noise at each run. The presented numbers in the simulations are the averages of the 100 Monte Carlo (MC) runs.

A. Simulations

The first simulation experiment consists of two co-centered elliptical objects moving according to the near constant velocity model with parameters given in Table II. Throughout the simulation, 7 measurements are generated from each object per time step. We compare the performances of VB, VPF (with $N=100$ particles), and JL algorithms over 100 MC runs. An example MC run is illustrated in Figure 3. The comparison metrics and the computation time of the algorithms are given in Table III. VB algorithm estimates both the kinematic and extent states better than other approaches in terms of RMSE, IOU, and GW distance.

Note that the number of association events in the JL algorithm grows exponentially as the number of the sub-objects

TABLE II: The simulation parameters for the first scenario.

\mathbf{X}_{true}	$\begin{bmatrix} 800 & 700 \\ 700 & 800 \end{bmatrix}$	$\begin{bmatrix} 800 & -700 \\ -700 & 800 \end{bmatrix}$
\mathbf{m}_0	$[0, 0, 200, 0, 0, 0]^T$	
\mathbf{P}_0	$\text{blkdiag}(50\mathbf{I}_2, 10\mathbf{I}_4)$	
\mathbf{V}_0	$1500\mathbf{I}_2$	
v_0	10	
\mathbf{F}_c	$\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \otimes \mathbf{I}_2$	
$\mathbf{F}_{\mu,\ell}$	\mathbf{I}_2	
\mathbf{F}	$\text{blkdiag}(\mathbf{F}_c, \mathbf{F}_{\mu,2})$	
\mathbf{Q}_c	$\sigma^2 \begin{bmatrix} T^3 & T^2 \\ 3T^2 & 2T \end{bmatrix} \otimes \mathbf{I}_2$	
$\mathbf{Q}_{\mu,\ell}$	\mathbf{I}_2	
\mathbf{Q}	$\text{blkdiag}(\mathbf{Q}_c, \mathbf{Q}_{\mu,2})$	
\mathbf{R}	$10\mathbf{I}_2$	
σ^2	0.1	
T	0.1	

TABLE III: The RMSE, IOU, GW values and the computation times for the first scenario.

	Sub-obj	IOU	RMSE	GW	Comp. Time Per Time Step
JL	1 st	0.70	3.89	9.68	13.96 s
	2 nd	0.70	4.26	9.69	
VPF	1 st	0.70	1.95	8.64	298.56 ms
	2 nd	0.70	2.12	8.57	
VB	1 st	0.77	1.64	5.06	25.91 ms
	2 nd	0.77	1.92	5.08	

and measurements increases [20]. Hence the JL algorithm's computation time is significantly higher than the other methods as shown in Table III. Being a sequential Monte Carlo method, VPF algorithm is also computationally costly. On the other hand, the computation time of VB is linear in both number of measurements M_k and number of sub-objects L . The computational complexity of the proposed solution is $\mathcal{O}(LM_k)$ at time k . Consequently, the computation time of the VB algorithm is significantly lower than the alternatives as shown in Table III.

In the second simulation scenario, we have a target with a shape similar to an airplane (see: Figure 4.). The object performs a constant velocity motion on a straight line.

The relevant parameters of the scenario are given in Table IV. Throughout the simulation, the number of measurements is set to 2 for each sub-object, yielding 8 measurements in total per time step. The performance evaluation metrics are given in Table V. The VB algorithm outperforms JL and VPF algorithms in terms of accuracy in the extent estimates and provides results which exhibit smaller variation over time.

B. Occlusion Scenario

In this section, we illustrate the capabilities of the algorithm in the presence of an occlusion problem which is frequently encountered in various target tracking applications. For instance, aerial objects might be partly or fully occluded by thick clouds while tracking with a day camera. Many practical systems resort to multiple complementary sensors, such as a

TABLE IV: The simulation parameters for the second scenario.

\mathbf{X}_{true}	Body:	$\begin{bmatrix} 6000 & 0 \\ 0 & 500 \end{bmatrix}$	Bottom Wing:	$\begin{bmatrix} 300 & 500 \\ 500 & 2000 \end{bmatrix}$
	Top Wing:	$\begin{bmatrix} 300 & -500 \\ -500 & 2000 \end{bmatrix}$	Tail:	$\begin{bmatrix} 150 & 0 \\ 0 & 3000 \end{bmatrix}$
\mathbf{m}_0	$[-10, 0, 400, 0, 0, -120, 0, 120, -200, 0]^T$			
\mathbf{P}_0	$\text{blkdiag}(10\mathbf{I}_4, 100\mathbf{I}_6)$			
\mathbf{V}_0	$3000\mathbf{I}_2$			
v_0	30			
\mathbf{F}_c	$\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \otimes \mathbf{I}_2$			
$\mathbf{F}_{\mu,\ell}$	\mathbf{I}_2			
\mathbf{F}	$\text{blkdiag}(\mathbf{F}_c, \mathbf{F}_{\mu,2}, \dots, \mathbf{F}_{\mu,5})$			
\mathbf{Q}_c	$\sigma^2 \begin{bmatrix} T^3 & T^2 \\ 3T^2 & 2T \end{bmatrix} \otimes \mathbf{I}_2$			
$\mathbf{Q}_{\mu,\ell}$	\mathbf{I}_2			
\mathbf{Q}	$\text{blkdiag}(\mathbf{Q}_c, \mathbf{Q}_{\mu,2}, \dots, \mathbf{Q}_{\mu,5})$			
\mathbf{R}	$10\mathbf{I}_2$			
σ^2	0.1			
T	0.1			

TABLE V: The RMSE, IOU, GW values and the computation times for the second scenario.

	Sub-obj:	1 st	2 nd	3 rd	4 th	Comp. Time Per Time Step[s]
JL	RMSE	6.98	12.46	11.54	14.24	15.97 s
	IOU	0.65	0.65	0.65	0.65	
	GW	26.98	16.15	16.13	18.04	
VPF	RMSE	5.35	11.46	9.86	6.44	351.44 ms
	IOU	0.65	0.70	0.71	0.76	
	GW	32.27	16.76	15.75	11.48	
VB	RMSE	2.58	8.61	7.71	6.73	27.87 ms
	IOU	0.71	0.75	0.75	0.75	
	GW	18.60	11.73	11.02	12.37	

thermal and a day camera, to prevent track loss during such occlusions.

In this simulation scenario, we simulated a group of co-ordinated targets, which contains 5 individuals moving on a straight path in a V-shape formation¹. During a certain part of the simulated trajectory, the line of sight of the sensor is partly blocked. It is assumed that no measurements can be obtained from the top two targets of the formation at the corresponding time instances. During the occlusion period, the group of targets gradually slows down. Then, they increase their velocity incrementally back to their regular pace. A visualization which describes the corresponding scenario is depicted in Figure 5.

We compare the performance of the proposed approach, VB, against the algorithm in [12], referred to as FFK in the sequel. The VB algorithm incorporates a unified kinematic model, i.e., the sub-objects depend on a common kinematic state as described in Section II. However, FFK treats each sub-object as a different target and tracks them individually without considering the interactions between them. During the occlusion period, FFK relies only on time update equations to estimate the kinematic and extent state of the occluded targets.

¹Note that, the data generated in this scenario can also be interpreted to originate from a single solid object whose extent is composed of five ellipses.

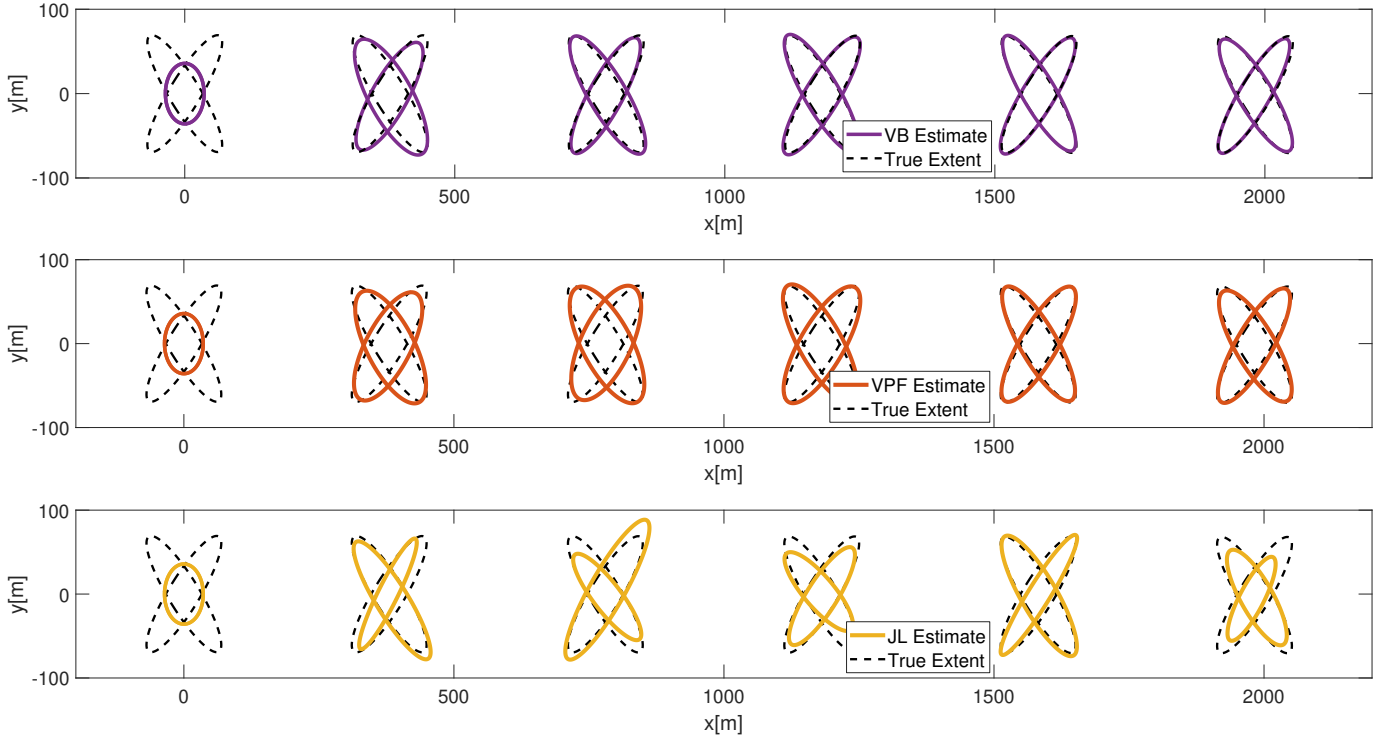


Fig. 3: The estimation results for the first scenario where two co-centered objects move along a straight line. The extent estimates of VB, VPF, and JL are shown in purple, orange, and yellow lines, respectively. The ground truth is depicted by black dashed lines.

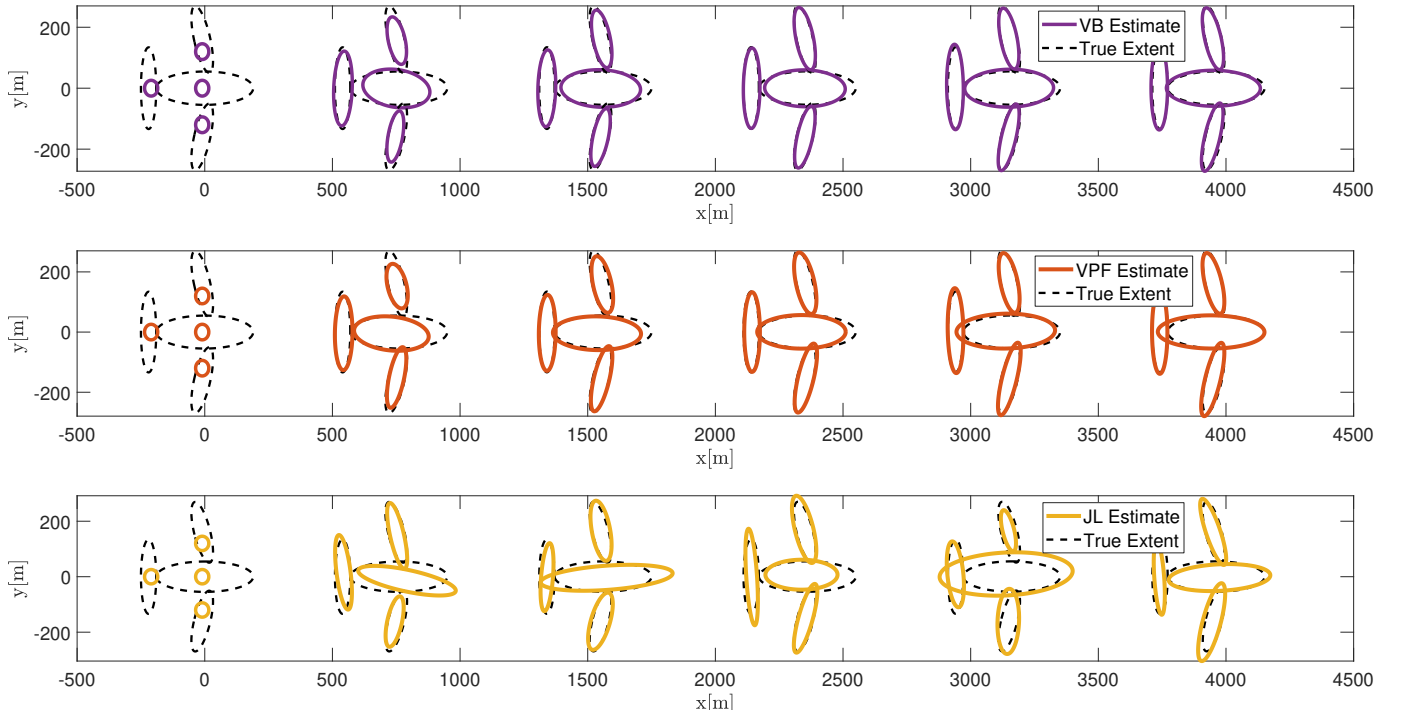


Fig. 4: The posterior means of the estimation results for the second scenario where four sub-objects with distinct extents move along a straight line. The extent estimates of VB, VPF, and JL are shown in purple, orange, and yellow lines, respectively. The ground truth is depicted by black dashed lines.

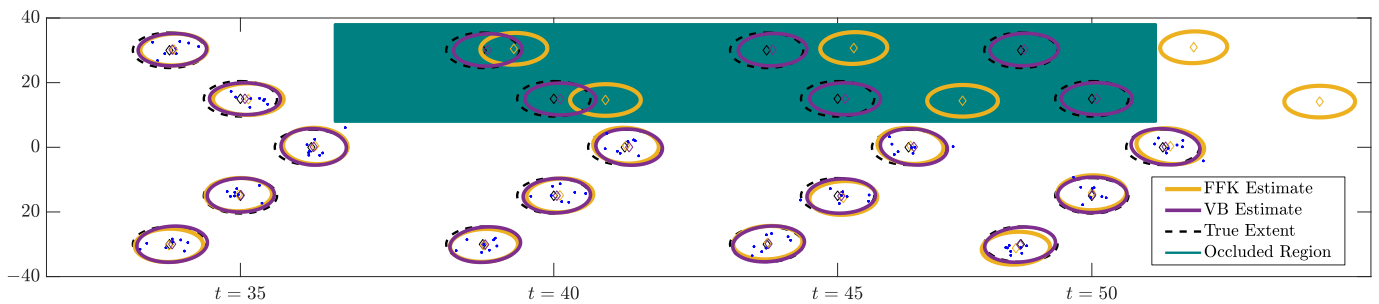


Fig. 5: The extent estimates of VB (purple solid line) and FFK (yellow solid line) together with the ground truth (black dashed line). The top two targets are occluded between $t = 36$ and $t = 51$. During this interval, no measurements are acquired from these targets. The occlusion region is represented by the green area.

On the contrary, VB can utilize the measurements that are collected from the visible targets to extract information about occluded targets' state. Consequently, it can provide much better performance as shown in Figure 5.

C. Experiments with Real Data

In this subsection, the performance of the proposed algorithm is demonstrated with real data. The data are extracted from aerial footages of a delta-wing aircraft and a glider. Throughout the experiments, each frame of the videos is processed to generate point measurements. In order to distinguish the airplanes from the background, a thresholding is performed in the HSV color space and a median filter is used to reduce the clutters and to eliminate distant measurements. After this pre-processing step, the pixel coordinates belonging to the airplanes are uniformly sampled to obtain the measurements.

The first real data experiment consists of a delta wing airplane [26]. A representative example of VB algorithm's performance is illustrated in Figure 6a. Note that, the camera zoom is not constant throughout the scenario, which introduces additional challenges for the algorithm.

The second scenario is an air-footage of a glider in motion [27]. The performance of VB algorithm is illustrated in Figure 6b.

In Figure 6a and 6b, we also depict the extent estimates of two different single ellipsoidal target tracking algorithms [6], [12]. For the sake of clarity, only the estimates in the final frames of the video sequences are plotted. Both results demonstrate that modeling the extent with multiple ellipses provides a more accurate representation of the target extent.

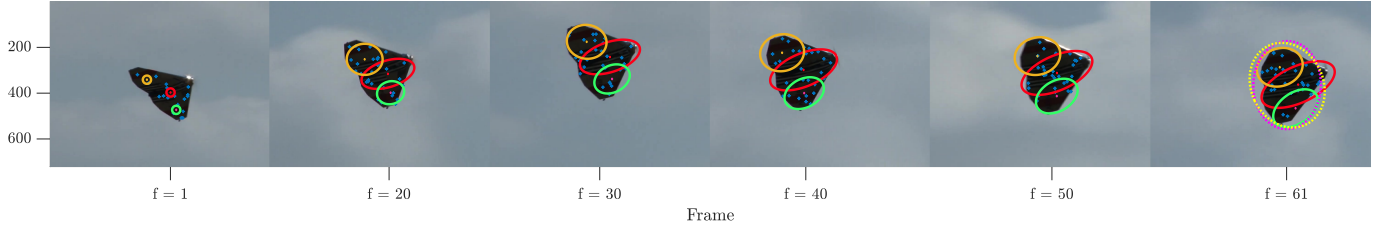
VI. CONCLUSION AND FUTURE WORK

In this study, a novel extended target tracking algorithm that is capable of representing a target or a group of targets with multiple ellipses while simultaneously estimating the kinematics is presented. The proposed solution involves approximating the intractable posterior distribution with the variational Bayes method to estimate kinematic and extent states. We demonstrated the performance of the approach in simulations and real data experiments. The results of the experiments show that the proposed method significantly improves the computation time while achieving better kinematic and

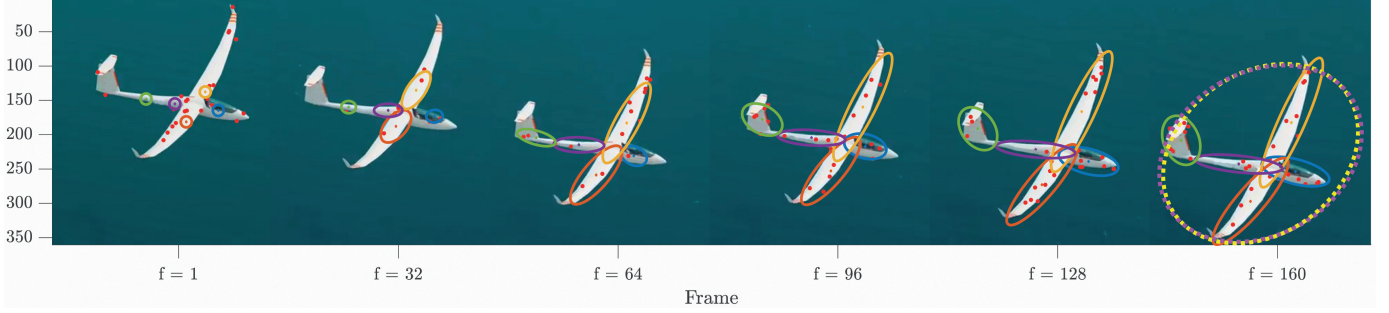
extent state estimation performance compared to the existing approaches in the literature.

REFERENCES

- [1] K. Granström and C. Lundquist, "On the use of multiple measurement models for extended target tracking," in *16th International Conference on Information Fusion (FUSION)*. IEEE, 2013, pp. 1534–1541.
- [2] M. Baum, V. Klumpp, and U. D. Hanebeck, "A novel Bayesian method for fitting a circle to noisy points," in *13th International Conference on Information Fusion (FUSION)*. IEEE, 2010, pp. 1–6.
- [3] K. Granström, C. Lundquist, and U. Orguner, "Tracking rectangular and elliptical extended targets using laser measurements," in *14th International Conference on Information Fusion (FUSION)*. IEEE, 2011, pp. 1–8.
- [4] M. Baum and U. D. Hanebeck, "Random hypersurface models for extended object tracking," in *International Symposium on Signal Processing and Information Technology*. IEEE, 2009, pp. 178–183.
- [5] —, "Shape tracking of extended objects and group targets with star-convex RHMs," in *14th International Conference on Information Fusion (FUSION)*. IEEE, 2011, pp. 1–8.
- [6] S. Yang and M. Baum, "Tracking the orientation and axes lengths of an elliptical extended object," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4720–4729, Jul 2019.
- [7] N. Wahlström and E. Özkan, "Extended target tracking using Gaussian processes," *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4165–4178, 2015.
- [8] M. Kumru and E. Özkan, "3D extended object tracking using recursive Gaussian processes," in *2018 21st International Conference on Information Fusion (FUSION)*. IEEE, 2018, pp. 1–8.
- [9] B. Tuncer, M. Kumru, E. Özkan, and A. A. Alatan, "Extended object tracking and shape classification," in *21st International Conference on Information Fusion (FUSION)*. IEEE, 2018, pp. 1–5.
- [10] B. Tuncer, M. Kumru, and E. Özkan, "Extended target tracking and classification using neural networks," in *2019 22th International Conference on Information Fusion (FUSION)*. IEEE, 2019, pp. 1–7.
- [11] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 1042–1059, 2008.
- [12] M. Feldmann, D. Franken, and W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1409–1420, 2011.
- [13] U. Orguner, "A variational measurement update for extended target tracking with random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3827–3834, 2012.
- [14] S. Yang and M. Baum, "Second-order extended Kalman filter for extended object and group tracking," in *2016 19th International Conference on Information Fusion (FUSION)*. IEEE, 2016, pp. 1178–1184.
- [15] —, "Extended Kalman filter for extended object tracking," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 4386–4390.
- [16] B. Tuncer and E. Özkan, "Random matrix based extended target tracking with orientation: A new model and inference," *arXiv preprint arXiv:2010.08820*, 2020.
- [17] S. F. Kara and E. Özkan, "Multi-ellipsoidal extended target tracking using sequential Monte Carlo," in *2018 21st International Conference on Information Fusion (FUSION)*. IEEE, 2018, pp. 1–8.



(a) Delta wing airplane tracking results.



(b) Glider tracking results.

Fig. 6: The results of experiments with real data: (a) Delta wing airplane (b) Glider. In the last frame of each sequence, the result of the single ellipsoidal target tracking algorithms [6] (magenta dashed line) and [12] (yellow dashed line) are also presented.

- [18] Q. Hu, H. Ji, and Y. Zhang, "Tracking of maneuvering non-ellipsoidal extended target with varying number of sub-objects," *Mechanical Systems and Signal Processing*, vol. 99, pp. 262–284, 2018.
- [19] K. Granström, P. Willett, and Y. Bar-Shalom, "An extended target tracking model with multiple random matrices and unified kinematics," in *2015 18th International Conference on Information Fusion (Fusion)*. IEEE, 2015, pp. 1007–1014.
- [20] J. Lan and X. R. Li, "Tracking of maneuvering non-ellipsoidal extended object or target group using random matrix," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2450–2463, 2014.
- [21] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [22] E. Özkan, V. Šmídl, S. Saha, C. Lundquist, and F. Gustafsson, "Marginalized adaptive particle filtering for nonlinear models with unknown time-varying noise parameters," *Automatica*, vol. 49, no. 6, pp. 1566 – 1575, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0005109813001350>
- [23] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.
- [24] C. R. Givens and R. M. Shortt, "A class of Wasserstein metrics for probability distributions," *The Michigan Mathematical Journal*, vol. 31, no. 2, pp. 231–240, 1984.
- [25] S. Yang, M. Baum, and K. Granström, "Metrics for performance evaluation of elliptic extended object tracking methods," in *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2016, pp. 523–528.
- [26] Marc Talloen. "Verhees Delta Flight and LANDING of that little Plane Sanicole 2012 part II." Dec. 30, 2018. [YouTube video, used with permission]. [Online]. Available: <https://youtu.be/rEwIv2jX08c?t=19>
- [27] Joyplanes RC. "How can can gliders fly without propulsion | The most complete explanation" Aug. 23, 2019. [YouTube video, used with permission]. [Online]. Available: <https://youtu.be/b4YrpmhgNGs?t=14>