Fusion of Elliptical Extended Object Estimates Parameterized with Orientation and Axes Lengths

Kolja Thormann [1,1] and Marcus Baum [2]

[1]University of Goettingen
[2]Affiliation not available

November 8, 2023

Abstract

This article considers the fusion of target estimates stemming from multiple sensors, where the spatial extent of the targets is incorporated. The target estimates are represented as ellipses parameterized with center orientation and semi-axis lengths and width. Here, the fusion faces challenges such as ambiguous parameterization and an unclear meaning of the Euclidean distance between such estimates. We introduce a novel Bayesian framework for random ellipses based on the concept of a Minimum Mean Gaussian Wasserstein (MMGW) estimator. The MMGW estimate is optimal with respect to the Gaussian Wasserstein (GW) distance, which is a suitable distance metric for ellipses. We develop practical algorithms to approximate the MMGW estimate of the fusion result. The key idea is to approximate the GW distance with a modified version of the Square Root (SR) distance. By this means, optimal estimation and fusion can be performed based on the square root of the elliptic shape matrices. We analyze different implementations using, e.g., Monte Carlo methods, and evaluate them in simulated scenarios. An extensive comparison with state-of-the-art methods highlights the benefits of estimators tailored to the Gaussian Wasserstein distances.

Fusion of Elliptical Extended Object Estimates Parameterized with Orientation and Axes Lengths

Kolja Thormann and Marcus Baum

***Abstract*—This article considers the tracking of elliptical extended targets parameterized with center, orientation, and semi-axes. The focus of this article is the fusion of extended target estimates, e.g., from multiple sensors, by handling the challenges introduced due to ambiguities in this parameterization and the unclear meaning of the Mean Square Error (MSE). For this purpose, we introduce a novel Bayesian framework for elliptic extent estimation and fusion based on two new concepts: (1) a probability density function for ellipses called Random Ellipse Density (RED) that incorporates the ambiguities that come with the ellipse parameterization, and (2) the Minimum Mean Gaussian Wasserstein (MMGW) estimate which is optimal with respect to the Gaussian Wasserstein (GW) distance – a suitable distance metric on ellipses. We develop practical algorithms for ellipse fusion and approximating the MMGW estimate. Different implementations, e.g., based on Monte Carlo simulation, are introduced and compared to state-of-the-art methods, highlighting the benefits of estimators tailored to the Gaussian Wasserstein distance.**

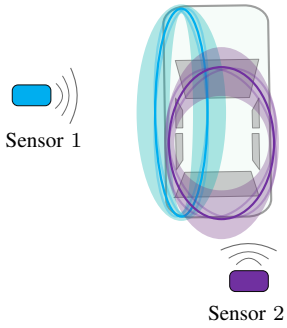


Fig. 1: A vehicle as ground truth along with two measurements from sensors 1 and 2 in light blue and purple respectively. Both measurements are ellipses and posses different uncertainties in the semi-axes and orientation represented by their more transparent versions.

I. Introduction

In many modern tracking applications the resolution of the involved sensors is high enough to resolve the spatial extent of the targets. For this reason, Extended Object Tracking (EOT) methods that estimate both the shape and kinematic parameters of a target are becoming increasingly important [1], [2]. Most EOT methods have been developed for sensors that resolve a varying number of noisy Cartesian detections from the target, e.g., based on a spatial distribution model [3]. The extent can be modeled by basic shapes like rectangles [4], [5] or ellipses [2], [6]–[9] or more detailed ones, either as a combination of multiple random matrices [10] or as a Random Hypersurface Model (RHM). The latter describes star-convex shapes and was modeled by, e.g., Fourier coefficients [11], Gaussian processes [12]–[14], or splines [15], [16].

This work focuses on (multiple) sensors (or sources) that directly produce width, length, and orientation estimates of elliptical targets, either because they generate that kind of measurements or because the fusion is conducted in a distributed fashion, with locally generated estimates which are then send to a central fusion unit [17] or shared among neighboring nodes of the sensor network [18]. The objective is to fuse the extended target densities, i.e., we consider object level fusion, and to gain an appropriate estimate from the fusion result. Thus, it is essential that the sensors not just provide their estimates, but their uncertainties as well [19].

Kolja Thormann and Marcus Baum are with the Institute of Computer Science, University of Goettingen, Germany, {kolja.thormann, marcus.baum}@cs.uni-goettingen.de

Elliptic shapes are widely-used to approximate the target extent [2], [6]–[9]. An advantage is not only the simple model, but also the usage in high noise scenarios in which the actual shape is hard to determine, as can be the case for automotive radar. A typical application scenario is tracking of traffic participants, e.g., cars or pedestrians, using multiple sensors, e.g., camera and radar, all providing ellipse estimates at each time step. The estimates have different uncertainties depending on the sensor's qualities and position relative to the target (see Figure 1). Under these conditions, there might be different uncertainties for the semi-axes or an increased uncertainty about the target's orientation due to, e.g., high noise or maneuvers like driving around a corner.

A. Contributions

The main contribution of this article is a novel Bayesian framework for estimation and (shape) fusion with elliptic extended targets parameterized by center, width, length, and orientation. The contributions include

- the promotion of the Gaussian-Wasserstein (GW) distance [20] as a risk function on elliptic shapes, defining a Minimum Mean Gaussian Wasserstein (MMGW) estimator,
- the introduction of a suitable probability density function on the explicit extent parameters of ellipses, a Random Ellipse Density (RED),
- the derivation of a MMGW estimator approximation and fusion method by replacing the GW distance via an extension of the Square Root (SR) distance [21], the Extended Square Root (ESR) distance,

- the development of practical implementations of the MMGW and RED concepts, and
- a comparison of the MMGW estimator and the fusion via REDs with state-of-the-art concepts, demonstrating improved performance in high noise scenarios.

This article is based on our previous conference publication [22]. We further develop our previous results by

- introducing the concept of RED, which allows for a meaningful Bayesian fusion of uncertain ellipses, along with an efficient mixture reduction as an improved fusion method to the one presented in the previous work,
- providing more insight on the effectiveness of approximating the GW distance with the ESR distance, and
- presenting a new and much more elaborate experimental evaluation of the fusion methods using moving targets.

B. Related Work

As elliptic targets are often represented by random matrices [6], [23], fusion methods for this representation can be found in literature as well. These include a combination of two random matrix estimates utilizing their respective Poisson rates [24], with a focus on combination of targets, also considering that their size might increase. In [18], they argue that as Random Matrices are represented by Inverse-Wishart distributions, the weighted average of the shape matrices (and of the degree of freedom) respectively minimizes the Kullback-Leibler divergence to the densities. Another approach to use measured point clouds from different sensors directly, including a method applying particle filter, can be found in [25]. The latter draws its particles from an importance distribution in the space of the explicit parameters with the orientation, length, and width of the shape matrix as mean. It then uses the random matrix likelihood to weight the particles with the point clouds generated by the sensors. There is an extension of this method for asynchronous sensors in [26], creating local estimates as particle densities, approximating them as Gaussian mixtures, and fusing via geometric mean densities. In both works, the mean of the particle density is calculated as the weighted average of the shape matrices. Another recent work [27] presents a distributed fusion method based on a variational Bayesian approach. They model the state also as a random matrix, define latent variables as noise-free measurements, and distribute their statistics across the sensors to approximate the state globally.

For rectangles, [28] provide a method which focuses on associating and fusing rectangular shapes by using the covariances of their corners. In [29], a method is presented to fuse rectangular estimates modeled by center, orientation, length, and width in Kalman fashion, but considering one of the corners as a reference point. An approach to fuse only segments, represented by points, lines, or L-shapes, is introduced in [30].

For arbitrary shapes there exists a framework by [19] to combine star-convex forms represented by Gaussian processes, again using measurements directly. There is also an extension in [31] which fuses star-convex shapes by determining a new center and then a new radial function based on the input radial functions relative to the new center. In [32], estimated Random Finite Sets (RFS) from multiple sensors are combined using the Kullback-Leibler Divergence between them.

To the best of our knowledge, there exist no fusion methods for ellipses explicitly parameterized by orientation and semi-axes in literature which go beyond applying a Kalman filter on the state densities. Additionally, we know of no other approach in tracking literature to explicitly change the risk function in estimating an extended target based on its posterior density. This idea is inspired by the Minimum Mean Optimal Subpattern Assignment (MMOSPA) estimators [33], [34], which estimate multiple point target densities [35] by minimizing the Optimal Subpattern Assignment (OSPA) distance [36]. MMOSPA estimation [34] (and this work) is also related to the concept of a Wasserstein Barycenter [37]–[40].

C. Structure

The remainder of the paper is organized as follows. The problem this paper deals with is described in Section II. Then, a novel fusion method for elliptical extended targets based on REDs is introduced in Section III, along with a more insightful recapitulation of the MMGW estimator. This is followed by approximations and implementations of fusion methods based on the newly derived RED and MMGW estimator in Section IV. Next, Section V provides an evaluation of the estimators and the fusion methods and then the results are discussed in Section VI. This article is concluded in Section VII.

II. Problem formulation

This work considers tracking of elliptic targets given ellipses either directly as measurements or as estimates provided by a sensor to the fusion center (for simplicity, we will refer only to measurements from now on). To clarify the setting, we first describe the state we want to estimate in Section II-A and the measurements we want to fuse the state with in Section II-B. In the following, lower case letters denote scalars, lower case bold letter denote vectors, and upper case bold letters denote matrices.

A. State

We model the spatial extent by a center $\mathbf{m} = \begin{bmatrix} m_1 & m_2 \end{bmatrix}^{\mathrm{T}}$, an orientation α, and semi-axis length l and width w (see also Figure 2a). This representation explicitly allows for capturing uncertainties for the different dimensions, providing important information which can be utilized for the fusion. Its usefulness has been demonstrated in, e.g., [7].

We also add a kinematic state $\mathbf{r}$ (depending on the choice of the kinematic model, $\mathbf{r}$ can also be a scalar), resulting in a state vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{m}_{\mathbf{x}}^{\mathrm{T}} & \alpha_{\mathbf{x}} & l_{\mathbf{x}} & w_{\mathbf{x}} & \mathbf{r}_{\mathbf{x}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} . \tag{1}$$

We assume the state to be a Gaussian distributed probability density with $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{C})$.

B. Measurement model

This work deals with "shape level" fusion, i.e., the measurements consist of the same parameters as the state (1). In other words, we assume that we have some estimate $\hat{\mathbf{x}}$ with covariance $\mathbf{C}$ and get an elliptic extended target measurement $\hat{\mathbf{x}}_2$ stemming from a sensor with noise $\mathbf{C}_2$ (see Figure 1). They can be related to each other with the measurement equation

$$\mathbf{x}_2 = \mathbf{H}\mathbf{x} + w_2 \ , \tag{2}$$

with $w_2 \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_2)$. We assume the entire shape is detected. If all of $\mathbf{r}$ is detected as well, we have $\mathbf{H} = \mathbf{I}_d$ with $\mathbf{I}_d$ being the d-dimensional identity matrix and d the dimension of $\mathbf{x}$. If only part or none of $\mathbf{r}$ is detected, the corresponding rows of $\mathbf{H}$ are cut. Essentially, the challenge is to effectively combine elliptic shape measurements with the current estimate, resulting in a new estimate along with a covariance. We assume w_2 to be independent of $\mathbf{x}$, ignoring cross-correlations for now.

The importance of tracking the axis and orientation separately becomes quite clear here, as one measurement from behind the target might give a precise estimate of the width, but has almost no information about the length, so that parameter should be weighted much less when fusing with the estimate. An alternative example would be scenarios with highly noisy measurements across the entire target surface from, e.g., radar. Due to the radial noise, the uncertainty orthogonal to the measurement direction can be higher than in measurement direction.

III. Bayesian Fusion and Estimation

Consider the prior ellipse estimate $\mathbf{x}$ with mean $\hat{\mathbf{x}}$ and covariance $\mathbf{C}$ and a measurement of the ellipse $\hat{\mathbf{x}}_2$ with sensor noise $\mathbf{C}_2$. The naive approach to combine them would be a linear fusion according to the Kalman filter. It finds the Minimum Mean Square Error (MMSE) estimate using the squared Euclidean distance as risk function according to

$$\hat{\mathbf{z}} = \underset{\mathbf{z}}{\operatorname{argmin}} \int ||\mathbf{z} - \mathbf{x}||_2^2 \cdot p(\mathbf{x}|\hat{\mathbf{x}}_2) \mathrm{d}\mathbf{x} = \mathrm{E}[\mathbf{x}|\hat{\mathbf{x}}_2] \ . \tag{3}$$

The fusion is conducted as

$$p(\mathbf{x}|\hat{\mathbf{x}}_2) \sim p(\hat{\mathbf{x}}_2|\mathbf{x}) \cdot p(\mathbf{x}) \ , \tag{4}$$

with prior $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{C})$ and Gaussian likelihood $p(\hat{\mathbf{x}}_2|\mathbf{x}) = \mathcal{N}(\hat{\mathbf{x}}_2; \mathbf{x}, \mathbf{C}_2)$.

To demonstrate the issues by means of an example, we consider the case of equal covariances. The estimate is then gained by averaging the two means

$$\hat{\mathbf{z}} = \frac{1}{2}(\hat{\mathbf{x}} + \hat{\mathbf{x}}_2) \ . \tag{5}$$

However, as Figure 2 shows, this approach can produce counter-intuitive results due to ambiguities in the parameterization. The two ellipses in Figures 2a and 2b are geometrically the same, just with an orientation shift of $\frac{\pi}{2}$ and the semi-axes switched, resulting in a fusion in Figure 2c which averages the intuitively wrong axes ($l_{\mathbf{x}}$ with l_2 and $w_{\mathbf{x}}$ with w_2 in Figures 2a and 2b).

The reason for this counter-intuitive behavior lies in Equation (3), namely from

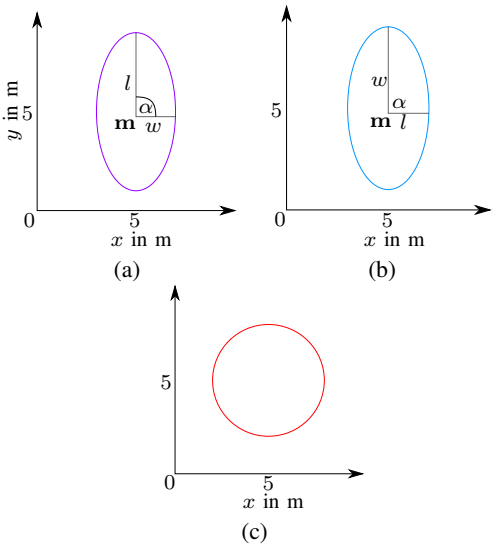


Fig. 2: Ellipses with parameters $\mathbf{m}_{\mathbf{x}} = [5 \quad 5]^{\mathrm{T}}$, $\alpha_{\mathbf{x}} = \frac{\pi}{2}$, $l_{\mathbf{x}} = 4$, and $w_{\mathbf{x}} = 2$ in Figure 2a and $\mathbf{m}_2 = [5 \quad 5]^{\mathrm{T}}$, $\alpha_2 = 0$, $l_2 = 2$, and $w_2 = 4$ in Figure 2b and their RMSE estimate using Euclidean distance in Figure 2c.

(1) the fusion $p(\mathbf{x}|\hat{\mathbf{x}}_2)$ by means of Equation (4) and
(2) the Euclidean distance $||\mathbf{z} - \mathbf{x}||_2^2$ as risk function.

Both (1) and (2) deal with real vectors and do not account for the elliptic shapes. For this reason, in this work, we propose

(1) a novel probability density for elliptic shapes in Section III-A, which allows for a sound Bayesian ellipse fusion and
(2) to use the GW distance to define an MMGW estimator to obtain intuitive point estimates in Section III-B.

As the focus is on the shape state, we will omit the kinematic part $\mathbf{r}$ in this section for readability.

A. Random Ellipse Densities

To define a density for ellipses, we need to deal with the ambiguities and the constraint that semi-axes need to be positive. For the latter, we create a truncated normal distribution by setting a lower bound of 0 for l and w,

$$p_t(\mathbf{x}) = \begin{cases} 0 & \text{if } l < 0 \vee w < 0, \\ c \cdot p(\mathbf{x}) & \text{else,} \end{cases} \tag{6}$$

with normalizing constant c. Next, to avoid the issue visualized in Figure 2, we adapt the concept of wrapped distributions [41]. We define a transformation of $\mathbf{x}$ to represent all equivalent ellipses by

$$K_k(\mathbf{x}) = \begin{bmatrix} \mathbf{m}_{\mathbf{x}}^{\mathrm{T}} & \alpha_{\mathbf{x},k} & v_k(l_{\mathbf{x}}, w_{\mathbf{x}}) & v_{k+1}(l_{\mathbf{x}}, w_{\mathbf{x}}) \end{bmatrix}^{\mathrm{T}} \ , \tag{7}$$

with

$$v_k(l, w) = \begin{cases} l & \text{if } k \text{ is even,} \\ w & \text{if } k \text{ is odd,} \end{cases} \tag{8}$$

$$\alpha_{\mathbf{x},k} = \alpha_{\mathbf{x}} + k \cdot \frac{\pi}{2} \ , \tag{9}$$

and $k \in \mathbb{Z}$. It is apparent that the orientation can be restricted between 0 and $\frac{\pi}{2}$ if the semi-axes are switched for each shift. We then define the wrapped distribution

$$\tilde{p}(\mathbf{x}) = \begin{cases} \sum_{k=-\infty}^{\infty} p_t(K_k(\mathbf{x})) & \text{if } 0 \leq \alpha_{\mathbf{x}} < \frac{\pi}{2}, \\ 0 & \text{else.} \end{cases} \tag{10}$$

We call this density a Random Ellipse Density (RED). It handles the ambiguities by representing only unique ellipses. The only exception is the case of equal semi-axes, as there is no unique representation of a circle due to the angle being chosen arbitrarily. From a mathematical point of view, a circle is a zero-probability event and does not need special consideration. However, circular point masses might be of interest, because from a practical point of view, ellipses which are close to a circle should be seen as close to each other even if their angular difference is high.

It is important to note that the Euclidean mean of this multi-modal density has no meaning and would result in a similar problem as depicted in Figure 2. To get an intuitive point estimate from the RED, the MMGW estimator described in the next section can be applied.

B. MMGW Estimator

As the Euclidean distance does not handle ambiguities in the ellipse parameterization, we suggest to use a more suitable metric. In [20], potential distance metrics on ellipses are evaluated, including Intersection-over-Union [4], the GW distance [42], Kullback-Leibler Divergence, and the Hausdorff distance, e.g., [43]. They conclude that the GW distance is the most suitable measure, as it provides a single, intuitive scalar value and can be solved in closed form. Therefore, we decided to use it as a quality measure.

The MMGW estimator changes the squared Euclidean distance on the explicit parameters in (3) with a true distance on ellipses. In the following, we will provide a more detailed explanation of this estimator and its approximation via the SR distance [21].

The GW distance [42] is defined as

$$\begin{aligned} &\mathrm{GW}(\mathbf{m_z}, \mathbf{Z}, \mathbf{m_x}, \mathbf{X}) \\ &\quad = ||\mathbf{m_z} - \mathbf{m_x}||_2^2 + \mathrm{Tr}[\mathbf{Z} + \mathbf{X} - 2(\mathbf{Z}^{\frac{1}{2}}\mathbf{X}\mathbf{Z}^{\frac{1}{2}})^{\frac{1}{2}}] \ , \end{aligned} \tag{11}$$

with shape matrix

$$\mathbf{Z} = \mathbf{R}_{\alpha_\mathbf{z}} \cdot \begin{bmatrix} l_\mathbf{z}^2 & 0 \\ 0 & w_\mathbf{z}^2 \end{bmatrix} \cdot \mathbf{R}_{\alpha_\mathbf{z}}^{\mathrm{T}} \ , \tag{12}$$

$$\mathbf{R}_\alpha = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \ , \tag{13}$$

and $\mathbf{X}$ analogous. Replacing the Euclidean distance results in the new mean estimate

$$\hat{\mathbf{z}} = \underset{\mathbf{z}}{\operatorname{argmin}} \int \mathrm{GW}(\mathbf{m_z}, \mathbf{Z}, \mathbf{m_x}, \mathbf{X}) \cdot \tilde{p}(\mathbf{x}) \mathrm{d}\mathbf{x} \ . \tag{14}$$

This gives us a MMGW estimator. Note that this would also be valid on the original density $p(\mathbf{x})$. The question is how to calculate the estimate. Already calculating the Wasserstein Barycenter from samples requires iterative optimization [44]. To obtain a closed-form solution, we utilize the Square Root (SR) distance [21]

$$\mathrm{SR}(\mathbf{Z}, \mathbf{X}) = ||\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}^{\frac{1}{2}}||_{Frobenius}^2 \ , \tag{15}$$

and extend it by including the center of the ellipse, creating the Extended Square Root (ESR) distance

$$\begin{aligned} &\mathrm{ESR}(\mathbf{m_z}, \mathbf{Z}, \mathbf{m_x}, \mathbf{X}) \\ &\quad = ||\mathbf{m_z} - \mathbf{m_x}||_2^2 + \mathrm{SR}(\mathbf{Z}, \mathbf{X}) \ . \end{aligned} \tag{16}$$

We then approximate the GW distance via the ESR distance

$$\begin{aligned} &\mathrm{GW}(\mathbf{m_z}, \mathbf{Z}, \mathbf{m_x}, \mathbf{X}) \\ &\quad \approx ||\mathbf{m_z} - \mathbf{m_x}||_2^2 + \mathrm{Tr}[(\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}^{\frac{1}{2}})(\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}^{\frac{1}{2}})] \\ &\quad = ||\mathbf{m_z} - \mathbf{m_x}||_2^2 + ||\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}^{\frac{1}{2}}||_{Frobenius}^2 \\ &\quad = \mathrm{ESR}(\mathbf{m_z}, \mathbf{Z}, \mathbf{m_x}, \mathbf{X}) \ , \end{aligned} \tag{17}$$

creating an approximated MMGW estimator. We justify the approximation in Appendix A and show that it is exact if the shape matrices would commute. The advantage of this approximation is that the MMGW estimate can be determined via averaging. This means if we define the transformation

$$T(\mathbf{x}) = \begin{bmatrix} \mathbf{m}_\mathbf{x}^{\mathrm{T}} & s_\mathbf{x}^{(11)} & s_\mathbf{x}^{(12)} & s_\mathbf{x}^{(22)} \end{bmatrix}^{\mathrm{T}} \ , \tag{18}$$

with $s^{(nm)}$ being cells of the symmetric square root matrix

$$\mathbf{R}_{\alpha_\mathbf{x}} \cdot \begin{bmatrix} l_\mathbf{x} & 0 \\ 0 & w_\mathbf{x} \end{bmatrix} \cdot \mathbf{R}_{\alpha_\mathbf{x}}^{\mathrm{T}} = \begin{bmatrix} s_\mathbf{x}^{(11)} & s_\mathbf{x}^{(12)} \\ s_\mathbf{x}^{(21)} & s_\mathbf{x}^{(22)} \end{bmatrix} = \mathbf{X}^{\frac{1}{2}} \ , \tag{19}$$

we get

$$\begin{aligned} \hat{\mathbf{z}} &= \underset{\mathbf{z}}{\operatorname{argmin}} \int \mathrm{GW}(\mathbf{m_z}, \mathbf{Z}, \mathbf{m_x}, \mathbf{X}) \cdot \tilde{p}(\mathbf{x}) \mathrm{d}\mathbf{x} \\ &\approx \underset{\mathbf{z}}{\operatorname{argmin}} \int ||T(\mathbf{z}) - T(\mathbf{x})||_2^2 \cdot \tilde{p}(\mathbf{x}) \mathrm{d}\mathbf{x} \\ &= T^{-1}(\mathrm{E}[T(\mathbf{x})]) \ , \end{aligned} \tag{20}$$

using the law of the unconscious statistician. To also include kinematic parts, the ESR distance in (16) and by that extension the transformation in (18) need to be extended by $\mathbf{r}$ analogous to $\mathbf{m}$.

IV. Implementations

With the result of the previous section, we end up with a density on ellipses in explicit parameter space and a transformation of the density to calculate the estimate with respect to a distance measure on ellipses. In this section, we present two different fusion approaches based on these concepts. First, we provide an approximation to fuse in ellipse parameter space in Section IV-A using the RED. Second, we briefly recap the fusion method used in our previous work in Section IV-B.

A. Ellipse Parameter Space

For the fusion in ellipse parameter space, each component of the prior RED needs to be multiplied with each component of the likelihood RED

$$\begin{aligned}\tilde{p}(\mathbf{x}|\hat{\mathbf{x}}_2) &\sim \tilde{p}(\hat{\mathbf{x}}_2|\mathbf{x}) \cdot \tilde{p}(\mathbf{x}) \\ &= \sum_{k=-\infty}^{\infty} p_t(K_k(\hat{\mathbf{x}}_2)|\mathbf{x}) \cdot \sum_{j=-\infty}^{\infty} p_t(K_j(\mathbf{x})) \\ &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} p_t(K_k(\hat{\mathbf{x}}_2)|\mathbf{x}) \cdot p_t(K_j(\mathbf{x})) \ , \end{aligned} \tag{21}$$

with the orientations $\alpha_{\mathbf{x}}$ and $\alpha_{\hat{\mathbf{x}}_2}$ restricted as in (10). The sums can be simplified using the 2π periodicity of the orientation to reduce the number of components to the 4 most likely ones. So, we can write (21) as

$$\tilde{p}(\mathbf{x}|\hat{\mathbf{x}}_2) \approx c_1 \cdot \sum_{j=0}^{3} \sum_{k=0}^{3} p_t(K_k(\hat{\mathbf{x}}_2)|\mathbf{x}) p_t(K_j(\mathbf{x})) \ , \tag{22}$$

with normalizing constant c_1. We further approximate the components as Gaussians, assuming the probability mass of l and w below 0 to be minor. Thus, we get the prior and likelihood as Gaussians as described in Section III

$$p_t(K_k(\hat{\mathbf{x}}_2)|\mathbf{x}) \approx p(K_k(\hat{\mathbf{x}}_2)|\mathbf{x}) \ , \tag{23}$$

$$p_t(K_j(\mathbf{x})) \approx p(K_j(\mathbf{x})) \ . \tag{24}$$

We end up with 16 components, each individual update conducted in Kalman fashion. With consecutive update steps, the number of components grows, so we apply mixture reduction. The MMGW estimate is then approximated by sampling particles from the RED, transforming them as in (18), and then averaging them. Pseudo-code can be found in Algorithm 1. As we fuse using REDs and determine the mean via MMGW estimation, this method is called Random Ellipse Density Minimum Mean Gaussian Wasserstein (RED-MMGW). A simple approximation is to fuse only the four most likely representations with the measurement and then keep the one with the highest weight. As the weight is calculated via the weighted distance between the ellipses' means, we called it the Random Ellipse Density Minimum Weighted Distance Parameterization (RED-MWDP).

B. Transformed Space

In our previous work [22], we utilized the transformations properties

$$T(K_0(\mathbf{x})) = T(K_k(\mathbf{x})) \quad \forall k \in \mathbb{Z} \ , \tag{25}$$

$$T(\hat{\mathbf{x}}) = T(\hat{\mathbf{x}}_2) \quad \text{if } l_1 = w_1 = l_2 = w_2 \wedge \mathbf{m}_1 = \mathbf{m}_2 \ , \tag{26}$$

to transform the density before fusion (as in transformed space, there are no ambiguities). To deal with the non-linearity of the transformation, we approximate the transformed density as a Gaussian with mean $\hat{\mathbf{y}}$ and covariance $\mathbf{D}$ and use a Kalman filter for the fusion. The approximation is conducted by drawing m particles

$$\mathbf{p}^{(j)} \sim \mathcal{N}(\hat{\mathbf{x}}, \mathbf{C}) \quad j \in \{1, ..., m\} \ . \tag{27}$$

input : multimodal input RED with means $\hat{\mathbf{x}}$, covariances $\mathbf{C}$, and weights $\mathbf{w}$, measurement mean $\hat{\mathbf{x}}_2$ and covariance $\mathbf{C}_2$
output: updated RED means $\hat{\mathbf{x}}^+$, covariances $\mathbf{C}^+$, and weights $\mathbf{w}^+$

for $k \in \{0, 1, 2, 3\}$ **do**
  $\hat{\mathbf{x}}_2.k \leftarrow$ `K(`$\hat{\mathbf{x}}_2$, k`)`
  $\mathbf{C}_2.k \leftarrow$ `K(`$\mathbf{C}_2$, k`)`
end
for $n \in$ `[0,len(`$\mathbf{w}$`)]` **do**
  for $k \in \{0, 1, 2, 3\}$ **do**
    $\hat{\mathbf{x}}^+.(4n+k)$, $\mathbf{C}^+.(4n+k)$, $\mathbf{w}^+.(4n+k)$ $\leftarrow$`kalman(`$\hat{\mathbf{x}}.n$, $\mathbf{C}.n$, $\hat{\mathbf{x}}_2.k$, $\mathbf{C}_2.k$`)`
    $\mathbf{w}^+.(4n+k) \leftarrow \mathbf{w}^+.(4n+k) \cdot \mathbf{w}.n$
  end
end
$\hat{\mathbf{x}}^+$, $\mathbf{C}^+$, $\mathbf{w}^+ \leftarrow$`reduce_mixture(`$\hat{\mathbf{x}}^+$, $\mathbf{C}^+$, $\mathbf{w}^+$`)`
return $\hat{\mathbf{x}}^+$, $\mathbf{C}^+$, $\mathbf{w}^+$

Algorithm 1: The RED fusion algorithm relies on the following functions. K() switches according to (7) with the last input as k when given state means and switches covariances accordingly as well, len() provides the length of the input vector, kalman() is the Kalman filter correction providing updated mean, covariance, and likelihood, and reduce_mixture() is a mixture reduction function also providing normalized weights.

input : Gaussian approximated transformed state mean $\hat{\mathbf{y}}$ and covariance $\mathbf{D}$, ellipse parameter measurement mean $\hat{\mathbf{x}}_2$ and covariance $\mathbf{C}_2$, and the number of particles m
output: updated mean $\hat{\mathbf{y}}^+$ and covariance $\mathbf{D}^+$

for $j \in [0, m]$ **do**
  $p.j \leftarrow$ `mvn(`$\hat{\mathbf{x}}_2$, $\mathbf{C}_2$`)`
end
$\hat{\mathbf{y}}_2 \leftarrow$`mean(T(`$p$`))`
$\mathbf{D}_2 \leftarrow$`mean(outer(T(`p`)`$-\hat{\mathbf{y}}_2$, `T(`p`)`$-\hat{\mathbf{y}}_2$`))`
$\hat{\mathbf{y}}^+$, $\mathbf{D}^+ \leftarrow$`kalman(`$\hat{\mathbf{y}}$, $\mathbf{D}$, $\hat{\mathbf{y}}_2$, $\mathbf{D}_2$`)`
return $\hat{\mathbf{y}}^+$, $\mathbf{D}^+$

Algorithm 2: The MC-MMGW algorithm relies on the following functions. mvn() provides a sample from a multivariate normal distribution, T() describes the transformation from (18), mean() provides a mean, outer() describes the outer product, and kalman() is a regular Kalman update.

The particles will be in state space (1). We then transform each individual particle and approximate the transformed particle density as a Gaussian distribution

$$\hat{\mathbf{y}} \approx \frac{1}{m} \sum_{j=1}^{m} T(\mathbf{p}^{(j)}) \ , \tag{28}$$

$$\mathbf{D} \approx \frac{1}{m} \sum_{j=1}^{m} (T(\mathbf{p}^{(j)}) - \hat{\mathbf{y}})(T(\mathbf{p}^{(j)}) - \hat{\mathbf{y}})^{\mathrm{T}} \ . \tag{29}$$

The estimate $\hat{\mathbf{x}}_2$ with noise $\mathbf{C}_2$ is transformed analogously. The fusion is then conducted based on the Kalman filter formulas. This method using Monte Carlo approximation of the density is called Monte Carlo Minimum Mean Gaussian Wasserstein (MC-MMGW) with pseudo-code provided in Algorithm 2. In the next section, we show that the Gaussian approximation is not enough to deal with the non-linearities in certain scenarios and that the fusion via REDs improves the results.

	Euclidean params	Euclidean shape	ESR	GW
Test 1	1.5888	1.6043	1.5834	1.5817
Test 2	4.2849	3.7328	3.6477	3.6458
Test 3	4.8260	3.7834	3.6917	3.6916

TABLE I: GW error of MMSE estimate using Euclidean distance in parameter space, Euclidean distance in shape matrix space, the approximated MMGW estimate using ESR distance, and the MMGW estimate using GW distance. The first test uses low, the second medium, and the third high orientation noise.

V. Experiments

To support our findings, we provide experiments for a comparison of the MMSE estimate using Euclidean distance, the MMGW estimate including its approximation using ESR distance, and a simple averaging of the shape matrices in Section V-A. Additionally, we evaluate the RED-MMGW and the MC-MMGW in a simulation with a moving ellipse in Section V-B, comparing them to state-of-the-art. The source code for the experiments is publicly available[1].

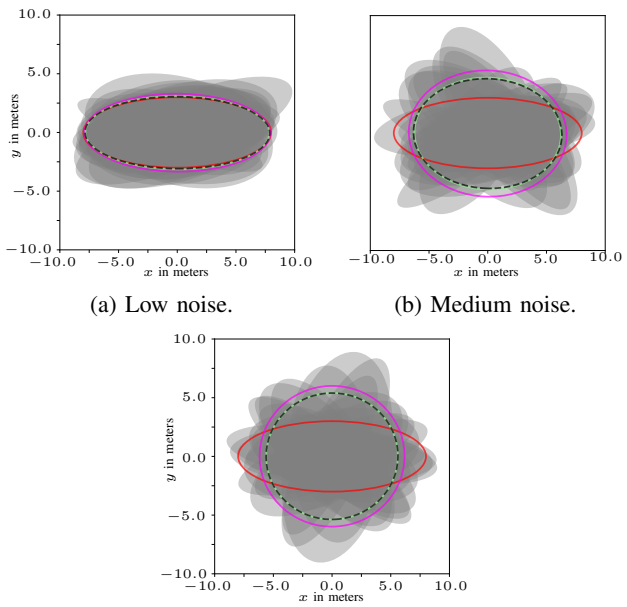
A. Ellipse estimation

In this section, we compare different approaches for determining point estimates of a Gaussian distributed posterior density on ellipses. The first naive approach is the mean as it minimizes the mean Euclidean error on the original ellipse parameter space. Our proposed method, the approximated MMGW estimate from Section III-B, minimizes the ESR distance. In essence, it approximates the state density via particles, transforms them into square root matrix space via (18), and averages. To highlight the difference of just averaging the shape matrices instead of their square roots, which might seem more intuitive given a particle density of shape matrices, we include this approach as well under the name Euclidean distance in shape space, in contrast to Euclidean distance in (ellipse) parameter space. Finally, we provide the MMGW estimate as the optimal solution, minimizing the GW distance, which we proposed as a more appropriate metric on ellipses than the Euclidean distance. For the calculation, we utilized the optimization described by [44] with initially equal weights and the mean from the ESR distance as initial guess.

We conducted three experiments. For the first, we took an ellipse estimate with mean $\hat{\mathbf{x}} = \begin{bmatrix}0 & 0 & 0 & 8 & 3\end{bmatrix}^{\mathrm{T}}$ and covariance $\mathbf{C} = diag(\begin{bmatrix}0.5 & 0.5 & 0.01\pi & 0.5 & 0.5\end{bmatrix})$, for the second one, we modified the orientation noise to 0.2π, and for the third to 0.5π, because the differences are most prominent with higher orientation uncertainty. The state consists of 2D center, orientation, length, and width. As the estimator focuses on the shape, the kinematics are omitted for these tests (see also Section III). In all cases, we drew 1000 particles to approximate the density for the estimators and to approximate the mean GW error of each estimate. The results can be found in Table I and Figure 3.

We find that the MMGW estimator approximated using ESR distance is quite accurate with respect to the actual MMGW estimate obtained via optimization. This coincides with the findings from Appendix A. Another important aspect is the relation to the Euclidean distance. For low orientation noise, the estimates are similar, but for high noise, the GW based estimators tend to a more circular form, which looks more intuitive as a high orientation noise means the true ellipse could be near 90 degree to the Euclidean estimate, making a circular estimate more reasonable. In other words, the Euclidean estimate we get as the mean of our Gaussian distributed state might seem intuitive at first, but the higher the noise, the further away it is from minimizing the GW distance.

Another interesting comparison is the Euclidean mean in shape matrix space, calculated as the mean of the particles' shape matrices. While this seems to be an intuitive estimate regarding a particle density of shape matrices, it can be seen that it is slightly larger than the MMGW estimate and even worse than the Euclidean mean in low orientation noise. This phenomenon is further elaborated in Section VI-B.



(a) Low noise. (b) Medium noise.

(c) High noise.

Fig. 3: MMSE estimates based on Euclidean distance in parameter space (red), Euclidean distance in shape space (magenta), ESR distance (green), and GW distance (dotted black) for different orientation noises. 20 sample particles to highlight the orientation uncertainty are drawn in grey.

B. Shape Fusion

For the evaluation of the fusion algorithms, the methods described in Section IV are compared with each other and with the state-of-the-art. This includes the regular fusion with Euclidean distances and random matrix based fusion approaches. Regarding the latter, a direct comparison is difficult however, as we assume sensors provide estimates parameterized as (1), not as random matrices and not as measurement points, i.e., we assume the sensor to be a black box offering us mean and covariance of an ellipse as discussed in Section II-B. To still offer a baseline, we transform the measurement ellipse into a

[1] https://github.com/Fusion-Goettingen/

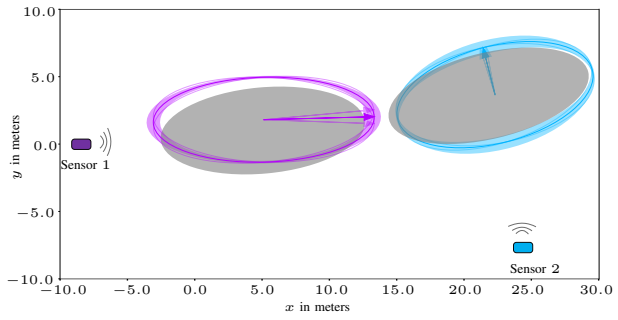


Fig. 4: Exemplary input of two consecutive time steps with ground truth in gray and imaginary sensors 1 in magenta and 2 in cyan. The measurements' orientations are represented by arrows and the uncertainties are visualized as pale ellipses. Note how the semi-axes possess different uncertainties.

shape matrix and handle all of them as Random Matrices with the same degree. We then utilize the method described by [18], giving all measurements the same weight. Each measurement random matrix gets the degree 6 and to consider a forget model, we utilize the version of the predict function from [45]. We call this method simply "shape mean". More details on the relation to random matrix based distributed fusion approaches can be found in Section VI-B.

For the MMGW-MC, we used 1000 particles. For the RED-MMGW, we apply mixture reduction, discarding components with low weights, merging those close to each other, and pruning unlikely components to ensure the number of components is below a threshold. On the result, we apply the MMGW estimator (using 1000 particles to approximate the transformed density). We also include the RED-MWDP described in Section IV-A and utilize the MMGW estimator on the posterior of the standard Kalman filter, entitled Regular-MMGW. To make the scenario more realistic, we consider moving ellipses, utilizing a Nearly Constant Velocity (NCV) model, i.e., the kinematic part is a Cartesian velocity $\mathbf{r} = \dot{\mathbf{m}}$.

The experiments provide the methods with a prior from which the ground truth is sampled in each run. We simulate sensors with different uncertainties in the semi-axis length and width, i.e., the sensor has a higher uncertainty in the direction it is facing. In addition, the sensors use different orientations for the ellipse to include the problem of ambiguous parameterization (see Figure 4 for an exemplary input). Before taking a measurement, the simulated sensor switches the parameterization by a random integer k according to (7). The sensors provide ellipse estimates (drawn from the ground truth) and their covariances. We conduct three experiments with *low*, *medium*, and *high* orientation noise. For each experiment, we conducted 1000 Monte Carlo runs and plotted the convergence of the GW error over 20 time steps, i.e., 20 measurements. Each time step is 1 second. To also provide a shape change, the ellipse will simply turn along with the velocity. The axes of the ground truth stay constant between time steps.

The prior mean $\hat{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 8 & 3 & 10 & 0 \end{bmatrix}^{\mathrm{T}}$ is used and to model that no prior knowledge is available for the orientation, the prior covariance matrix $\mathbf{C} = diag(\begin{bmatrix} 0.5 & 0.5 & 0.5\pi & 0.5 & 0.5 & 10.0 & 10.0 \end{bmatrix})$ is chosen with a relatively high orientation uncertainty. After the ground truth is drawn from the prior and after each prediction step, the orientation and velocity vector are aligned with each other. We conduct three tests with three sensor noise settings. One is sensor *low* with covariance $\mathbf{C}_2 = diag(\begin{bmatrix} 0.5 & 0.5 & 0.01\pi & 0.5 & 0.1 & 0.05 & 0.05 \end{bmatrix})$ (similar to Figure 3a). Note that before drawing the estimates from the ground truth, the orientation and semi-axes are shifted according to (7) as described above, resulting in the noise on the semi-axes being actually different between the sensors (see Figure 4). For sensor *medium* the variance of the orientation in the covariance is set to 0.2π (similar to Figure 3b) and for *high*, it is 0.5π (similar to Figure 3c). As the shape mean does not consider the measurement covariance, we modified the forget parameter τ in its prediction step, providing 0.2 in *low* noise, 5.0 in *medium* noise, and 10.0 in *high* noise, which produced the best results in our tests.

The results can be found in Figure 5. Note that the GW distance is used as an error function.

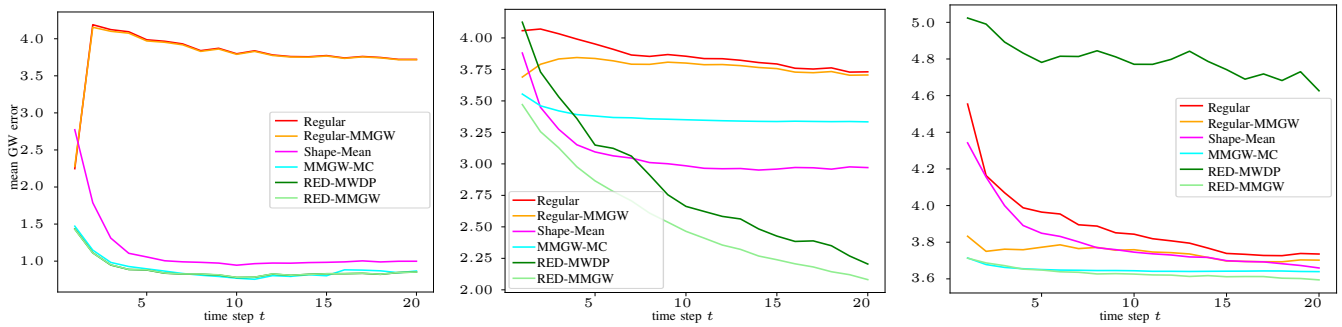
The results of test 1 in Figure 5a show the problems of the regular fusion with the ambiguous parameterization clearly. The shape mean converges, but it does not consider the different uncertainties in the semi-axes, so it is overall worse than the following approaches. The RED-MWDP, dealing with the ambiguity, demonstrates that minimization via Euclidean distance also provides good results in the case of low orientation noise, similar to the MMGW based approaches. This conforms with the findings of Figure 3a of Section V-A.

The results of test 2 in Figure 5b demonstrate the advantages of the RED fusion, dealing with the ambiguities and incorporating information about different axes uncertainties. Especially in this case, it can also be seen how MC-MMGW is outperformed by the REDs, as the approximation of the transformed density as a Gaussian fails under the increased noise. It becomes also more clear how the MMGW estimation outperforms the classical estimate.

Finally, the results of test 3 in Figure 5c demonstrate the issue of classical estimation and by extend, the issue of RED-MWDP. It keeps the shape, but due to the high noise, the orientation can be quite off, resulting in a high error. The MMGW based methods however estimate a circular target to deal with the low amount of orientation information. Two other things are of note here. First, as a clarification to avoid confusion regarding the regular fusion, we highlight that it produces more round estimates due to the axis switch (see also Figure 2). As a round target is a good estimate in this scenario, the result is good by coincidence. The same holds for the MMGW-MC, as the approximation also leads to a round estimate, explaining the improvements compared to test 2.

Given the Kalman filter-based linear fusion and the shape mean as the state-of-the-art in fusion of elliptical estimates parameterized by orientation and semi-axes, we demonstrated that our method offers improvements of up to 40% with respect to the GW distance.

Due to page constraints, we excluded tests with no ambiguous parameterizations. In these cases, where semi-axes and orientation is always correctly assigned and thus can be seen



(a) Test 1 (low α noise). (b) Test 2 (medium α noise). (c) Test 3 (high α noise).

Fig. 5: Tests comparing different ellipse fusion approaches under different sensor noise.

as simple Gaussian distributed random variables, the Kalman filter is of course the best choice. However, using the MMGW estimator on the Kalman filter posterior still provides a better estimate with respect to the GW distance.

VI. Discussion

In this section, we provide further insights on different behavior and properties of the discussed methods, comparing the MMGW estimation with the minimization of the Euclidean distance on ellipse parameters in Section VI-A and with the minimization of the Euclidean distance in shape matrix space in Section VI-B.

A. GW vs Euclidean distance on ellipse parameters

In this article, we demonstrated in simulations that in scenarios with high orientation noise, the RED-MWDP provides worse results with respect to the GW distance compared to the RED-MMGW (and the regular Kalman filter provides worse results than the Regular-MMGW Kalman filter). We further elaborate this difference by providing the MMSE estimates using Euclidean and ESR distance in low and high orientation noise scenarios. In the case of an uncertain orientation, the methods based on Euclidean distance would keep the semi-axes and average the orientation. With low information on the orientation, this can provide poor results, especially with only few measurements. The ESR based methods systematically deal with this issue by providing circular shaped estimates. Depending on the scenario, this may be more desirable. Such a situation could happen, e.g., if the semi-axis uncertainty was low (due to long tracking or prior knowledge) and the target would turn under high measurement noise. As the target's dimensions are certain, the uncertainty would be reflected in the orientation. This brings us to an important parallel between the regular estimation and the MMGW based approaches, which occurs when the shape matrices commute. It can be shown that in these cases, the ESR distance would boil down to the Euclidean distance between state vectors containing only center and the overlapping semi-axes

$$\begin{aligned}&\mathrm{ESR}(\mathbf{m}_\mathbf{z}, \mathbf{Z}, \mathbf{m}_\mathbf{x}, \mathbf{X})\\ &= || \begin{bmatrix}(\mathbf{m}_\mathbf{z} - \mathbf{m}_\mathbf{x})^\mathrm{T} & v_k(l_\mathbf{z}, w_\mathbf{z}) - l_\mathbf{x} & v_{k+1}(l_\mathbf{z}, w_\mathbf{z}) - w_\mathbf{x}\end{bmatrix} ||_2^2\\ &\forall \alpha_\mathbf{z}, k \text{ with } (\alpha_\mathbf{z} = \alpha_\mathbf{x} + k\frac{\pi}{2}) \text{ and } k \in \mathbb{Z}\ ,\end{aligned} \tag{30}$$

with $v_k()$ from (8). Now, if the ellipse orientation has a high certainty, so the corresponding value in the covariance tends to 0, most particles would nearly commute. As a result, the MMGW estimate would be almost the same as the regular estimate. Additionally, as the approximation of the GW distance via ESR distance would be exact in this case, both estimates would also be optimal in respect to the GW distance. This is further demonstrated by the results of Section V.

In this context, we also highlight that the standard Kalman filter would provide better results than the RED if the axes association was known, i.e., if there were no ambiguities. This could happen if the orientation and velocity vector were linked. However, as this assumption might not hold true, e.g., the ellipse is used to model a group target or the sensors decouple shape and kinematics or targets are stationary, the RED is still a useful tool to avoid counter-intuitive fusion results due to wrongly assigned semi-axes.

B. GW vs Euclidean distance on shape matrix space

The particles from an empirical density in matrix square root space can simply be averaged to compute the approximated MMGW estimate. Note that [25] calculate the weighted sum of their particle density using the shape matrices, not their square roots, finding the estimate which minimizes the Euclidean distance between the matrices instead. However, based on our experience, using the shape matrices' square roots provides better results with respect to the Gaussian Wasserstein distance and the ellipse parameters. This can be demonstrated with a simple example regarding commute matrices. As shown in (30), averaging the square roots in this special case boils down to averaging the lengths and widths. However, averaging the shape matrices would mean averaging the squared lengths and widths. This provides ellipses which are too large, which can be seen in Figure 3. For example, if two equally likely estimates with lengths 2 and 10 are drawn, averaging the square root matrices results in a length estimate $\frac{2+10}{2} = 6$ and averaging the shape matrices produces $\sqrt{\frac{4+100}{2}} \approx 7$.

We also want to discuss the difference to random matrix based fusion approaches. In general, as these approaches employ an Inverse-Wishart distribution to represent an ellipse, their applicability on ellipses parameterized with orientation and semi-axes is limited.

The method from [24] is made for merging of shapes, also increasing their size if both shapes are further apart which is a different type of application. The weighted average to minimize the Kullback-Leibler divergence [18] does not use the measurement covariances, i.e., it is not suitable if they differ between measurements. The method from [25] utilizes the measurements directly, which we do not consider is this article. [26] provide a method sharing Gaussian mixture approximations of the density transformed into the same parameterization we use between sensors. However, the Gaussian mixtures are generated from sampling from Random matrices based on an importance distribution and weighting them with their likelihoods, while our approach assumes elliptical estimates and their covariances are given. These estimates cannot be directly translated to those Gaussian mixtures. Hence, a direct comparison with our work is not feasible. The same holds for the latent variable approximations in [27].

VII. Conclusion and Future Work

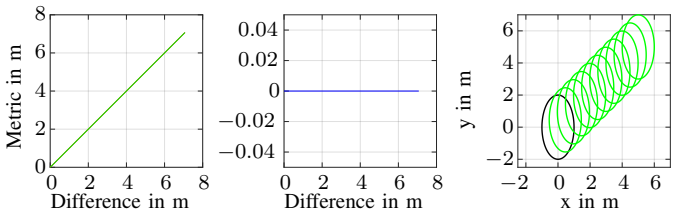
In this work, we proposed a novel Bayesian framework for estimation and fusion of ellipse densities represented by center, orientation, and semi-axis length and width. The main concepts are the identification of a density on ellipses, the RED, and the utilization of the GW distance to create a MMGW estimator. Approximating the GW distance with the ESR distance, we derived an estimator which can be calculated via averaging of particles. In simulations, we demonstrated the robustness of this approximation and compared our methods to state-of-the-art algorithms in respect to estimation and fusion. We highlighted the advantages of our approach, those being the inclusion of the ellipse parameter state's covariance, the dealing with ambiguous parameterizations of ellipses, and more intuitive estimates in scenarios with high orientation noise. In summary, the RED provides a suitable way of avoiding the ambiguous parameterization of ellipses during fusion, if the parameterization of the ellipse is unclear. The MMGW estimator provides an intuitive estimate based on the posterior density.

For future work, we intend to further improve the mixture reduction of RED based fusion. Regarding the Monte Carlo-based approaches for fusion in transformed space, we seek to better preserve the transformed density by means of direct multiplication of particle densities [46]. As for other shapes, we want to investigate appropriate metrics for rectangles and test if the MMGW and RED concepts also work for them due to the equal parameterization. Regarding more complex contours, like star-convex shapes, we wish to determine whether we can apply the same principle of finding an appropriate distance measure and density representation here as well.

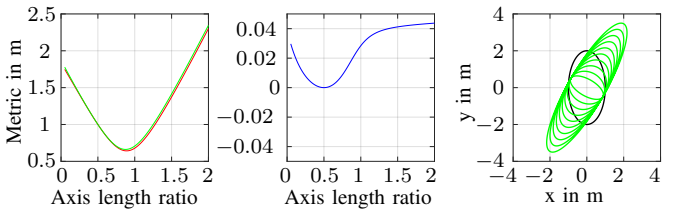
Appendix

A. Comparison of GW and ESR distance

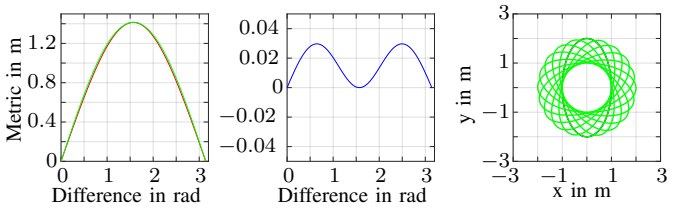
For the comparison, we utilized the experiments from [20] with the GW and the ESR distances only (see Figure 6). For the first experiment, the ellipses' orientations are the same. With the approximation of the GW distance by the ESR distance being exact in this case, the two metrics behave the same (see Figure 6a). For the second experiment, the second ellipse is tilted slightly. A difference can thus be seen here, getting larger with the ellipse's length increasing beyond the ground truth's length, but except for that shift, they still behave similar (see Figure 6b). Finally, for the rotation experiment, the general behavior is also the same, with the difference between the metrics being greatest when the angle offset between the ellipses is $\frac{\pi}{4}$ shifted by a multiple of $\frac{\pi}{2}$ (see Figure 6c), so when the angle is in the middle of two orientations which would make the estimate and the ground truth commute.



(a) Same ellipse moving away.



(b) Slightly tilted and smaller ellipse growing in length.



(c) Same ellipse rotating.

Fig. 6: Comparison of GW and ESR distance with metric output on the left (red: GW, green: ESR), metric difference in the middle, and ground truth with every tenth iteration of the estimate on the right (black: ground truth, green: estimates).

References

[1] L. Mihaylova, A. Carmi, F. Septier, A. Gning, S. Pang, and S. Godsill, "Overview of Bayesian Sequential Monte Carlo Methods for Group and Extended Object Tracking," *Digital Signal Processing*, vol. 25, pp. 1–16, Feb. 2014.

[2] K. Granström, M. Baum, and S. Reuter, "Extended Object Tracking: Introduction, Overview and Applications," *ISIF Journal of Advances in Information Fusion*, vol. 12, no. 2, Dec. 2017.

[3] K. Gilholm and D. Salmond, "Spatial Distribution Model for Tracking Extended Objects," *IEE Proceedings on Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 364–371, Oct. 2005.

[4] K. Granström, C. Lundquist, and U. Orguner, "Tracking Rectangular and Elliptical Extended Targets Using Laser Measurements," in *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, Chicago, Illinois, USA, Jul. 2011.

[5] K. Schueler, T. Weiherer, E. Bouzouraa, and U. Hofmann, "360 Degree Multi Sensor Fusion for Static and Dynamic Obstacles," in *IEEE Intelligent Vehicles Symposium*, 2012, pp. 692–697.

[6] M. Feldmann, D. Fränken, 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.

[7] S. Yang and M. Baum, "Tracking the Orientation and Axes Lengths of an Elliptical Extended Object," *IEEE Transactions on Signal Processing*, 2019.

[8] P. Brosseit, B. Duraisamy, and J. Dickmann, "The Volcanormal Density for Radar-Based Extended Target Tracking," in *IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Oct 2017, pp. 1–6.

[9] J. Lan and X. R. Li, "Tracking of Extended Object or Target Group Using Random Matrix: New Model and Approach," *IEEE Transactions on Aerospace and Electrical Systems*, vol. 52, no. 6, pp. 2973–2988, Dec. 2016.

[10] ——, "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.

[11] M. Baum and U. D. Hanebeck, "Extended Object Tracking with Random Hypersurface Models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, pp. 149–159, Jan. 2014.

[12] N. Wahlström and E. Özkan, "Extended Target Tracking Using Gaussian Processes," *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4165–4178, 2015.

[13] T. Hirscher, A. Scheel, S. Reuter, and K. Dietmayer, "Multiple Extended Object Tracking Using Gaussian Processes," in *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, July 2016, pp. 868–875.

[14] X. Tang, M. Li, R. Tharmarasa, and T. Kirubarajan, "Seamless Tracking of Apparent Point and Extended Targets Using Gaussian Process PMHT," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4825–4838, 2019.

[15] H. Kaulbersch, J. Honer, and M. Baum, "A Cartesian B-Spline Vehicle Model for Extended Object Tracking," in *21st International Conference on Information Fusion (FUSION 2018)*, July 2018.

[16] B. Naujoks, P. Burger, and H.-J. Wuensche, "Fast 3D Extended Target Tracking using NURBS Surfaces," *arXiv preprint arXiv:1909.00767*, 2019.

[17] H. Durrant-Whyte and T. C. Henderson, "Multi Sensor Data Fusion," *Springer handbook of robotics*, pp. 585–610, 2008.

[18] W. Li, Y. Jia, D. Meng, and J. Du, "Distributed Tracking of Extended Targets Using Random Matrices," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 3044–3049.

[19] M. Michaelis, P. Berthold, D. Meissner, and H.-J. Wuensche, "Heterogeneous Multi-Sensor Fusion for Extended Objects in Automotive Scenarios Using Gaussian Processes and a GMPHD-Filter," in *IEEE Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, 2017, pp. 1–6.

[20] S. Yang, M. Baum, and K. Granström, "Metrics for Performance Evaluation of Elliptic Extended Object Tracking Methods," in *Proceedings of the 2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2016)*, Baden-Baden, Germany, Sept. 2016.

[21] I. L. Dryden, A. Koloydenko, and D. Zhou, "Non-Euclidean Statistics for Covariance Matrices, with Applications to Diffusion Tensor Imaging," *The Annals of Applied Statistics*, vol. 3, no. 3, pp. 1102–1123, 2009.

[22] K. Thormann and M. Baum, "Optimal Fusion of Elliptic Extended Target Estimates Based on the Wasserstein Distance," in *22nd International Conference on Information Fusion (FUSION 2019)*, Ottawa, Canada, July 2019. [Online]. Available: https://arxiv.org/abs/1904.00708

[23] 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, July 2008.

[24] K. Granström and U. Orguner, "On Spawning and Combination of Extended/Group Targets Modeled With Random Matrices," *IEEE Transactions on Signal Processing*, vol. 61, no. 3, pp. 678–692, 2013.

[25] G. Vivone, K. Granström, P. Braca, and P. Willett, "Multiple Sensor Measurement Updates for the Extended Target Tracking Random Matrix Model," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 5, pp. 2544–2558, 2017.

[26] J. Liu and G. Guo, "Distributed Asynchronous Extended Target Tracking Using Random Matrix," *IEEE Sensors Journal*, vol. 20, no. 2, pp. 947–956, 2019.

[27] J. Hua and C. Li, "Distributed Variational Bayesian Algorithms for Extended Object Tracking," *arXiv preprint arXiv:1903.00182*, 2019.

[28] S. Nilsson and A. Klekamp, "Object Level Fusion of Extended Dynamic Objects," in *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, Sept 2016, pp. 251–258.

[29] M. Herrmann, J. Müller, J. Strohbeck, and M. Buchholz, "Environment Modeling Based on Generic Infrastructure Sensor Interfaces Using a Centralized Labeled-Multi-Bernoulli Filter," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 2414–2420.

[30] B. Duraisamy, M. Gabb, A. V. Nair, T. Schwarz, and T. Yuan, "Track Level Fusion of Extended Objects from Heterogeneous Sensors," in *IEEE Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, 2016, pp. 876–885.

[31] M. Michaelis, P. Berthold, T. Luettel, D. Meissner, and H.-J. Wuensche, "A Merging Strategy for Gaussian Process Extended Target Estimates in Multi-Sensor Applications," in *IEEE Intelligent Vehicles Symposium (IV)*, 2019, pp. 1803–1808.

[32] M. Fröhle, K. Granström, and H. Wymeersch, "Decentralized Poisson Multi-Bernoulli Filtering for Extended Target Tracking," *arXiv preprint arXiv:1901.04518*, 2019.

[33] M. Guerriero, L. Svensson, D. Svensson, and P. Willett, "Shooting Two Birds with Two Bullets: How to Find Minimum Mean OSPA Estimates," *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*, 2010.

[34] M. Baum, P. Willett, and U. D. Hanebeck, "On Wasserstein Barycenters and MMOSPA Estimation," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1511–1515, Oct. 2015.

[35] B.-N. Vo, M. Mallick, Y. Bar-Shalom, S. Coraluppi, R. Osborne, R. Mahler, B.-T. Vo, and J. G. Webster, *Multitarget Tracking*. John Wiley & Sons, Inc., 2015.

[36] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A Consistent Metric for Performance Evaluation of Multi-Object Filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447 –3457, Aug. 2008.

[37] M. Agueh and G. Carlier, "Barycenters in the Wasserstein Space," *SIAM Journal on Mathematical Analysis*, vol. 43, no. 2, pp. 904–924, 2011.

[38] J. Rabin, G. Peyre, J. Delon, and M. Bernot, "Wasserstein Barycenter and its Application to Texture Mixing," in *Scale Space and Variational Methods in Computer Vision*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, vol. 6667, pp. 435–446.

[39] G. Carlier, A. Oberman, and E. Oudet, "Numerical Methods for Matching for Teams and Wasserstein Barycenters," Tech. Rep., May 2014.

[40] M. Cuturi and A. Doucet, "Fast Computation of Wasserstein Barycenters," in *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, Oct. 2014.

[41] K. V. Mardia and P. E. Jupp, *Directional Statistics*. John Wiley & Sons, 2009, vol. 494.

[42] 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.

[43] R. Veltkamp and M. Hagedoorn, "Shape Similarity Measures, Properties and Constructions," in *Advances in Visual Information Systems*, ser. Lecture Notes in Computer Science, R. Laurini, Ed. Springer Berlin / Heidelberg, 2000, vol. 1929, pp. 133–153.

[44] G. Puccetti, L. Rüschendorf, and S. Vanduffel, "On the Computation of Wasserstein Barycenters," *Available at SSRN 3276147*, 2018.

[45] K. Granström, M. Fatemi, and L. Svensson, "Gamma Gaussian inverse-Wishart Poisson multi-Bernoulli Filter for Extended Target Tracking," in *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, Heidelberg, Germany, July 2016.

[46] U. D. Hanebeck, "Bayesian Fusion of Empirical Distributions Based on Local Density Reconstruction," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2015, pp. 277–282.