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

Kolja Thormann and Marcus Baum

***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.**

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 elliptic targets. The objective is to fuse the extended target estimates in order to gain an improved estimate, i.e., we consider object level fusion. Note that the estimate's uncertainty, which is usually provided by the sensor, is essential in this case [17].

Elliptic shapes are widely-used to approximate the target extent. 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.

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

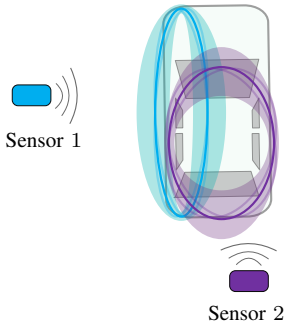


Fig. 1: Estimates of a vehicle from two sensors 1 and 2 as a light blue and a purple ellipse. Both estimates posses different uncertainties in the semi-axes and orientation represented by the more transparent versions of the estimates.

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. Thus, the challenges this article aims to tackle lie in the parameterization of the ellipses and the best way to combine estimates from multiple sensors to get an improved estimate of the actual target.

A. Contributions

The main contribution of this article is a novel systematic Bayesian approach to object-level fusion of extended target estimates containing width, length, and orientation information. It includes

- the definition of a suitable probability density function on the explicit extent parameters of ellipses (consisting of center, width, length, and orientation),
- the promotion of the Gaussian-Wasserstein (GW) distance as a risk function on elliptic shapes, defining a Minimum Mean Gaussian Wasserstein (MMGW) estimator,
- the derivation of an Approximated Minimum Mean Gaussian Wasserstein (AMMGW) estimator and fusion methods by approximating the GW distance via an extension of the Square Root (SR) distance [18], the Extended Square Root (ESR) distance,
- the development of practical implementations of the AMMGW concept, and

- a comparison of the AMMGW estimator and state-of-the-art concepts, demonstrating improved performance in high noise scenarios, and a discussion of problems and advantages.

This article is based on our previous conference publication [19]. We further develop our previous results by introducing the concept of Random Ellipse Density (RED), highlighting the effectiveness of approximating the GW distance with the ESR distance, introducing an improved fusion method using a particle filter, and providing a new and much more elaborate experimental evaluation.

B. Related Work

As elliptic targets are often represented by random matrices [6], [20], 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 [21] and an approach to use the measured point clouds from different sensors directly, including a method applying particle filter [22]. 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 [23], 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.

For rectangles, [24] provide a method which focuses on associating and fusing rectangular shapes by using the covariances of their corners. An approach to fuse only segments, represented by points, lines, or L-shapes, is introduced in [25].

For arbitrary shapes there exists a framework by [17] to combine star-convex forms represented by Gaussian processes, again using measurements directly. There is also an extension in [26] 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 [27], estimated Random Finite Sets (RFS) from multiple sensors are combined using the Kullback-Leibler Divergence between them.

This work is also inspired by the Minimum Mean Optimal Subpattern Assignment (MMOSPA) estimators [28], [29], which track multiple targets [30] by minimizing the Optimal Subpattern Assignment (OSPA) distance [31]. MMOSPA estimation [29] (and this work) is also related to the concept of a Wasserstein Barycenter [32]–[35].

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 Bayesian estimator is introduced in Section III, followed by approximations and implementations of fusion methods based on the newly derived estimator in Section IV. Next, Section V provides an evaluation of the estimators and then the results are discussed in Section VI. This article is concluded in Section VII.

II. Problem formulation

This work considers the fusion of elliptic extended target estimates stemming from multiple sensors (see Figure 1). The extent estimates are represented 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.

Assume there are multiple ellipse estimates with densities $p(\mathbf{x}_i)$ and $\mathbf{x}_i = \begin{bmatrix} \mathbf{m}_{\mathbf{x}_i}^{\mathrm{T}} & \alpha_{\mathbf{x}_i} & l_{\mathbf{x}_i} & w_{\mathbf{x}_i} \end{bmatrix}^{\mathrm{T}}$, $i \in \mathbb{N}$, the challenge is to find an appropriate way to combine them and determine a suitable point estimate.

We assume the densities to be Gaussian distributions with $\mathbf{x}_i \sim \mathcal{N}(\hat{\mathbf{x}}_i, \mathbf{C}_i)$ and that the sensors provide us with means $\hat{\mathbf{x}}_i$ and covariance matrices $\mathbf{C}_i$. We also consider the sensor measurements to be independent, ignoring cross-correlations.

III. Bayesian Estimation with Random Ellipses

Consider a prior ellipse estimate $\mathbf{x}_1$ with mean $\hat{\mathbf{x}}_1$ and covariance $\mathbf{C}_1$ 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 Euclidean distance as error

$$\hat{\mathbf{z}} = \underset{\mathbf{z}}{\operatorname{argmin}} \int ||\mathbf{z} - \mathbf{x}_1||_2^2 \cdot p(\mathbf{x}_1|\hat{\mathbf{x}}_2) \mathrm{d}\mathbf{x}_1 = \mathrm{E}[\mathbf{x}_1|\hat{\mathbf{x}}_2] \ . \quad (1)$$

The fusion is conducted as

$$p(\mathbf{x}_1|\hat{\mathbf{x}}_2) \sim p(\hat{\mathbf{x}}_2|\mathbf{x}_1) \cdot p(\mathbf{x}_1) \ , \quad (2)$$

with prior $p(\mathbf{x}_1) = \mathcal{N}(\mathbf{x}_1; \hat{\mathbf{x}}_1, \mathbf{C}_1)$ and Gaussian likelihood $p(\hat{\mathbf{x}}_2|\mathbf{x}_1) = \mathcal{N}(\hat{\mathbf{x}}_2; \mathbf{x}_1, \mathbf{C}_2)$. In the case of equal covariances, the estimate is gained by averaging the two means

$$\hat{\mathbf{z}} = \frac{1}{2}(\hat{\mathbf{x}}_1 + \hat{\mathbf{x}}_2) \ . \quad (3)$$

However, as the example in Figure 2 shows, this approach can produce counter-intuitive results. The two ellipses in Figures 2a and 2b are the same, just with an orientation shift of $\frac{\pi}{2}$ and the semi-axes switched, resulting in a fusion in Figure 2c which looks different. The reason is that the intuitively wrong axes are averaged (l_1 with l_2 and w_1 with w_2 in Figures 2a and 2b).

To solve this issue, we first introduce a new density type on ellipses in Section III-A and then provide the GW distance as a risk function for more intuitive results in Section III-B.

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,

$$p_t(\mathbf{x}) = \begin{cases} 0 & \text{if } l < 0 \vee w < 0, \\ c \cdot p(\mathbf{x}) & \text{else,} \end{cases} \quad (4)$$

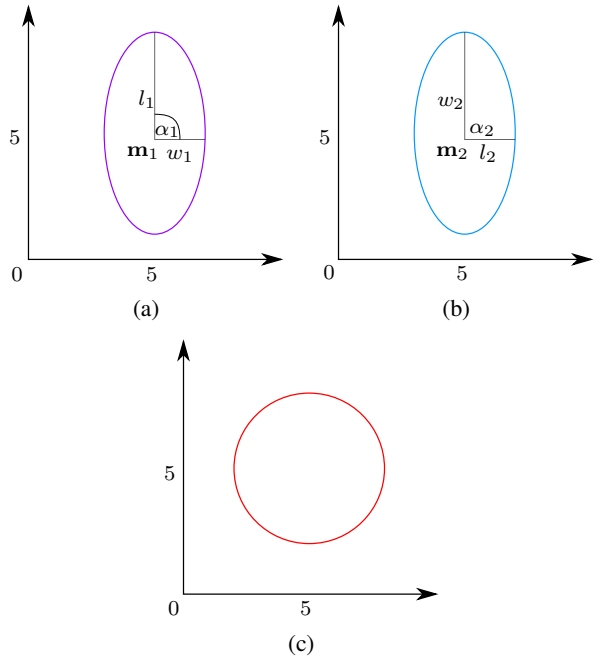


Fig. 2: Ellipses with parameters $\mathbf{m}_1 = [5 \quad 5]^{\mathrm{T}}$, $\alpha_1 = \frac{\pi}{2}$, $l_1 = 4$, and $w_1 = 2$ in Figure 2a and $\mathbf{m}_2 = [5 \quad 5]^{\mathrm{T}}$, $\alpha_2 = 0$, $l_2 = 2$, and $w_2 = 4$ in Figure 2b and their RMSE estimate using Euclidean distance in Figure 2c.

with normalizing constant c. Next, to avoid the issue visualized in Figure 2, we apply the concept of wrapped distributions. Defining a representation of $\mathbf{x}$ for all equal ellipses

$$\mathbf{x}^{(k)} = \begin{bmatrix} \mathbf{m}_{\mathbf{x}}^{\mathrm{T}} & \alpha_{\mathbf{x}}^{(k)} & v^{(k)}(l_{\mathbf{x}}, w_{\mathbf{x}}) & v^{(k+1)}(l_{\mathbf{x}}, w_{\mathbf{x}}) \end{bmatrix}^{\mathrm{T}} \ , \quad (5)$$

with

$$v^{(k)}(l, w) = \begin{cases} l & \text{if } k \text{ is even,} \\ w & \text{if } k \text{ is odd,} \end{cases} \quad (6)$$

$$\alpha^{(k)} = \alpha + k \cdot \frac{\pi}{2} \ , \quad (7)$$

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

$$\tilde{p}(\mathbf{x}) = \begin{cases} \sum_{k=-\infty}^{\infty} p_t(\mathbf{x}^{(k)}) & \text{if } 0 \leq \alpha_{\mathbf{x}} < \frac{\pi}{2}, \\ 0 & \text{else.} \end{cases} \quad (8)$$

We call this density a Random Ellipse Density (RED). However, the Euclidean mean of this multi-modal density has no meaning and would result in a similar problem as depicted in Figure 2. Furthermore, there is no unique representation of a circle as the angle is 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.

B. MMGW Estimator

To provide a more reasonable mean, we propose to change the squared Euclidean distance on the explicit parameters in (1) with a true distance on ellipses.

In [36], potential distance metrics on ellipses are evaluated, including Intersection-over-Union [4], the GW distance [37], Kullback-Leibler Divergence, and the Hausdorff distance, e.g., [38]. 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. For this reason, we employ the GW distance in this work, which is defined as

$$\begin{aligned} &GW(\mathbf{m}_{\mathbf{z}}, \mathbf{Z}, \mathbf{m}_{\mathbf{x}_1}, \mathbf{X}_1) \\ &\quad = ||\mathbf{m}_{\mathbf{z}} - \mathbf{m}_{\mathbf{x}_1}||_2^2 + \mathrm{Tr}[\mathbf{Z} + \mathbf{X}_1 - 2(\mathbf{Z}^{\frac{1}{2}}\mathbf{X}_1\mathbf{Z}^{\frac{1}{2}})^{\frac{1}{2}}] \ , \end{aligned} \quad (9)$$

with shape matrix

$$\mathbf{Z} = \mathbf{R}_{\alpha_{\mathbf{z}}} \cdot \begin{bmatrix} l_{\mathbf{z}}^2 & 0 \\ 0 & w_{\mathbf{z}}^2 \end{bmatrix} \cdot \mathbf{R}_{\alpha_{\mathbf{z}}}^{\mathrm{T}} \ , \quad (10)$$

$$\mathbf{R}_{\alpha} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \ , \quad (11)$$

and $\mathbf{X}_1$ analogous. Replacing the Euclidean distance results in the new mean estimate

$$\hat{\mathbf{z}} = \underset{\mathbf{z}}{\mathrm{argmin}} \int \mathrm{GW}(\mathbf{m}_{\mathbf{z}}, \mathbf{Z}, \mathbf{m}_{\mathbf{x}_1}, \mathbf{X}_1) \cdot \tilde{p}(\mathbf{x}_1) \mathrm{d}\mathbf{x}_1 \ . \quad (12)$$

This gives us a MMGW estimator. The question is how to solve the estimator. Even calculating the Wasserstein Barycenter from samples requires iterative optimization [39]. To obtain a closed-form solution, we utilize the SR distance [18]

$$SR(\mathbf{Z}, \mathbf{X}_1) = ||\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}_1^{\frac{1}{2}}||_{Frobenius}^2 \ , \quad (13)$$

and extend it by including the center of the ellipse, creating the ESR distance

$$\begin{aligned} &ESR(\mathbf{m}_{\mathbf{z}}, \mathbf{Z}, \mathbf{m}_{\mathbf{x}_1}, \mathbf{X}_1) \\ &\quad = ||\mathbf{m}_{\mathbf{z}} - \mathbf{m}_{\mathbf{x}_1}||_2^2 + SR(\mathbf{Z}, \mathbf{X}_1) \ . \end{aligned} \quad (14)$$

We then approximate the GW distance via the ESR distance

$$\begin{aligned} &GW(\mathbf{m}_{\mathbf{z}}, \mathbf{Z}, \mathbf{m}_{\mathbf{x}_1}, \mathbf{X}_1) \\ &\quad \approx ||\mathbf{m}_{\mathbf{z}} - \mathbf{m}_{\mathbf{x}_1}||_2^2 + \mathrm{Tr}[(\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}_1^{\frac{1}{2}})(\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}_1^{\frac{1}{2}})] \\ &\quad = ||\mathbf{m}_{\mathbf{z}} - \mathbf{m}_{\mathbf{x}_1}||_2^2 + ||\mathbf{Z}^{\frac{1}{2}} - \mathbf{X}_1^{\frac{1}{2}}||_{Frobenius}^2 \\ &\quad = ESR(\mathbf{m}_{\mathbf{z}}, \mathbf{Z}, \mathbf{m}_{\mathbf{x}_1}, \mathbf{X}_1) \ , \end{aligned} \quad (15)$$

creating the AMMGW 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 AMMGW estimate can be determined via averaging. This means if we define the transformation

$$T(\mathbf{x}) = \begin{bmatrix} \mathbf{m}_{\mathbf{x}}^{\mathrm{T}} & s_{\mathbf{x}}^{(11)} & s_{\mathbf{x}}^{(12)} & s_{\mathbf{x}}^{(22)} \end{bmatrix}^{\mathrm{T}} \ , \quad (16)$$

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

$$\mathbf{R}_{\alpha_{\mathbf{x}}} \cdot \begin{bmatrix} l_{\mathbf{x}} & 0 \\ 0 & w_{\mathbf{x}} \end{bmatrix} \cdot \mathbf{R}_{\alpha_{\mathbf{x}}}^{\mathrm{T}} = \begin{bmatrix} s_{\mathbf{x}}^{(11)} & s_{\mathbf{x}}^{(12)} \\ s_{\mathbf{x}}^{(21)} & s_{\mathbf{x}}^{(22)} \end{bmatrix} = \mathbf{X}^{\frac{1}{2}} \ , \quad (17)$$

we get

$$\begin{aligned}\hat{\mathbf{z}} &= \underset{\mathbf{z}}{\operatorname{argmin}} \int \mathrm{GW}(\mathbf{m}_\mathbf{z}, \mathbf{Z}, \mathbf{m}_{\mathbf{x}_1}, \mathbf{X}_1) \cdot \tilde{p}(\mathbf{x}_1)\mathrm{d}\mathbf{x}_1 \\ &\approx \underset{\mathbf{z}}{\operatorname{argmin}} \int ||T(\mathbf{z}) - T(\mathbf{x}_1)||_2^2 \cdot \tilde{p}(\mathbf{x}_1)\mathrm{d}\mathbf{x}_1 \\ &= T^{-1}(\mathrm{E}[T(\mathbf{x}_1)]) \ , \end{aligned} \tag{18}$$

using the law of the unconscious statistician.

IV. Bayesian Fusion with Random Ellipses

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 mean with respect to a distance measure for ellipses. To fuse estimates using these concepts, we provide an approximation to fuse in explicit parameter space in Section IV-A, while Section IV-B presents a method to apply the transformation before the fusion.

A. Explicit Parameter Space

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

$$\begin{aligned}\tilde{p}(\mathbf{x}_1|\hat{\mathbf{x}}_2) &\sim \tilde{p}(\hat{\mathbf{x}}_2|\mathbf{x}_1) \cdot \tilde{p}(\mathbf{x}_1) \\ &= \sum_{k=-\infty}^{\infty} p_t(\hat{\mathbf{x}}_2^{(k)}|\mathbf{x}_1) \cdot \sum_{j=-\infty}^{\infty} p_t(\mathbf{x}_1^{(j)}) \\ &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} p_t(\hat{\mathbf{x}}_2^{(k)}|\mathbf{x}_1) p_t(\mathbf{x}_1^{(j)}) \ , \end{aligned} \tag{19}$$

with the orientations $\alpha_{\mathbf{x}_1}$ and $\alpha_{\hat{\mathbf{x}}_2}$ restricted as in (8). The sums can be simplified using the 2π periodicity of the orientation to reduce the number of components to the 4 most likely ones. Thus, we can write (19) as

$$\tilde{p}(\mathbf{x}_1|\hat{\mathbf{x}}_2) \approx c_1 \cdot \sum_{j=0}^{3} \sum_{k=0}^{3} p_t(\hat{\mathbf{x}}_2^{(k)}|\mathbf{x}_1) p_t(\mathbf{x}_1^{(j)}) \ , \tag{20}$$

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

$$p_t(\hat{\mathbf{x}}_2^{(k)}|\mathbf{x}_1) \approx p(\hat{\mathbf{x}}_2^{(k)}|\mathbf{x}_1), \ p_t(\mathbf{x}_1^{(j)}) \approx p(\mathbf{x}_1^{(j)}) \ , \tag{21}$$

with rows and columns for l and w of $\mathbf{C}_1^{(j)}$ and $\mathbf{C}_2^{(k)}$ switched accordingly regarding j and k. Thus, we end up with 16 components. We then reduce the prior to one representation, so $j = 0$, and after the fusion, we only keep the component with the highest likelihood, corresponding to a k_{opt}, ending up with 1 component. Minimizing the negative log-likelihood, we get

$$\begin{aligned}k_{opt} = \underset{k}{\operatorname{argmin}} \frac{1}{2} \cdot (&\boldsymbol{\nu}_k^\mathrm{T} \mathbf{S}_k^{-1} \boldsymbol{\nu}_k + \log(\det(\mathbf{S}_k)) \\ &+ 5\log(2\pi)) \ , \end{aligned} \tag{22}$$

with

input : explicit state mean $\hat{\mathbf{x}}$ and covariance $\mathbf{C}$, measurement mean $\hat{\mathbf{x}}_i$ and covariance $\mathbf{C}_i$
output: updated mean $\hat{\mathbf{x}}^+$ and covariance $\mathbf{C}^+$

```
ν ← x̂_i − x̂
S ← C + C_i
d_min ← mvn_pdf(ν, 0, S)
k_opt ← 0
for k ∈ {1, 2, 3} do
    x̂_i^(k) ← switch_params(x̂_i, k)
    C_i^(k) ← switch_params(C_i, k)
    ν ← x̂_i^(k) − x̂
    S ← C + C_i^(k)
    d_cur ← mvn_pdf(ν, 0, S)
    if d_cur ≤ d_min then
        d_min ← d_cur
        k_opt ← k
    end
end
x̂^+, C^+ ←kalman_k(x̂, C, x̂_i, C_i, k_opt)
return x̂^+, C^+
```

Algorithm 1: The MWDP algorithm relies on the following functions. mvn_pdf() is the likelihood for a multivariate normal distribution, switch_params() switches according to (5) with the last input as k when given state means and switches covariances accordingly as well, and kalman_k() is the Kalman filter fusion using the measurement switched according to (5) with the last input as k.

$$\boldsymbol{\nu}_k = \hat{\mathbf{x}}_2^{(k)} - \hat{\mathbf{x}}_1, \ \mathbf{S}_k = \mathbf{C}_1 + \mathbf{C}_2^{(k)} \ , \tag{23}$$

and $k \in \{0, 1, 2, 3\}$. We then get the posterior

$$\tilde{p}(\mathbf{x}_1|\hat{\mathbf{x}}_2) \approx c_2 \cdot p(\hat{\mathbf{x}}_2^{(k_{opt})}|\mathbf{x}_1) p(\mathbf{x}_1) \ , \tag{24}$$

with normalizing constant c_2. As this method essentially uses the likelihood to find the parameterization which minimizes the Euclidean distance between prior and measurement, their dimensions weighted with the covariances, we call it the Minimum Weighted Distance Parameterization (MWDP). Pseudocode can be found in Algorithm 1.

B. Transformed Space

The RED concept from the previous section provides a suitable representation of random ellipses, but there are practical problems. First, in respect to the GW distance, a circle is a singularity, so for shapes close to a circle, the orientation should have only minor influence on the distance. Second, there is the problem of $\tilde{p}(\mathbf{x})$ containing an infinite number of components, with the method in Section IV-A only approximating the fusion by discarding the unlikely ones. To deal with this issue, we utilize the transformation's properties

$$T(\mathbf{x}^{(0)}) = T(\mathbf{x}^{(k)}) \quad \forall k \in \mathbb{Z} \ , \tag{25}$$

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

and transform before the fusion. The transformed density $\overline{p}(\mathbf{y})$ with $T(\mathbf{x}) = \mathbf{y}$ has one component and no ambiguities, thus

$$\begin{aligned}\hat{\mathbf{z}} &\approx \underset{\mathbf{z}}{\operatorname{argmin}} \int ||T(\mathbf{z}) - \mathbf{y}_1||_2^2 \cdot \overline{p}(\mathbf{y}_1|\hat{\mathbf{y}}_2)\mathrm{d}\mathbf{y}_1 \\ &= T^{-1}(\mathrm{E}[\mathbf{y}_1|\hat{\mathbf{y}}_2]) \ . \end{aligned} \tag{27}$$

This can be justified as follows. Given the definition of $\tilde{p}(\mathbf{x})$, the transformation $T(\mathbf{x})$ is bijective, except for $\mathbf{x}$ describing circles (which is a zero-probability event). Thus, we can write the transformed density as

$$\begin{aligned}\overline{p}(\mathbf{y}) &= \int \delta(\mathbf{y} - T(\mathbf{x}))\tilde{p}(\mathbf{x})\mathrm{d}\mathbf{x} \\ &= |\det(T_{\mathbf{y}}^{-1})|\tilde{p}(\mathbf{x}_{\mathbf{y}}) \text{ if } s_{\mathbf{x}_{\mathbf{y}}}^{(11)} \neq s_{\mathbf{x}_{\mathbf{y}}}^{(22)} \wedge s_{\mathbf{x}_{\mathbf{y}}}^{(12)} \neq 0 \ , \end{aligned} \tag{28}$$

with $\mathbf{x}_{\mathbf{y}} = T^{-1}(\mathbf{y})$ having $0 \leq \alpha_{\mathbf{x}_{\mathbf{y}}} < \frac{\pi}{2}$ and $\det(T_{\mathbf{y}}^{-1})$ as the determinant of the Jacobian of the transformation's inverse at $\mathbf{y}$. Given a prior $\tilde{p}(\mathbf{x}_1)$ and a measurement with likelihood $\tilde{p}(\hat{\mathbf{x}}_2|\mathbf{x}_1)$, we get

$$\overline{p}(\mathbf{y}_1|\hat{\mathbf{y}}_2) = |\det(T_{\mathbf{y}_1}^{-1})|\tilde{p}(\mathbf{x}_1|\hat{\mathbf{y}}_2) \tag{29}$$

$$= |\det(T_{\mathbf{y}_1}^{-1})|\frac{\overline{p}(\hat{\mathbf{y}}_2|\mathbf{x}_1)\tilde{p}(\mathbf{x}_1)}{\overline{p}(\hat{\mathbf{y}}_2)} \tag{30}$$

$$= |\det(T_{\mathbf{y}_1}^{-1})|\frac{|\det(T_{\hat{\mathbf{y}}_2}^{-1})|\tilde{p}(\hat{\mathbf{x}}_2|\mathbf{x}_1)\tilde{p}(\mathbf{x}_1)}{|\det(T_{\hat{\mathbf{y}}_2}^{-1})|\tilde{p}(\hat{\mathbf{x}}_2)} \tag{31}$$

$$= |\det(T_{\mathbf{y}_1}^{-1})|\tilde{p}(\mathbf{x}_1|\hat{\mathbf{x}}_2) \ . \tag{32}$$

This means that except for the special case of a circle, it is the same if the transformation is performed before (29) or after (32) fusion. To fuse in transformed space, we need to deal with the nonlinear transformation $T(\cdot)$. In the following, we present two approaches, approximating the density from sampled particles as a Gaussian in Section IV-B1 and using a particle filter with a likelihood based on $\tilde{p}(\mathbf{x})$ in Section IV-B2.

1) Kalman filter via Gaussian approximation: The goal of this method is to approximate the transformed density of estimate i as a Gaussian with mean $\hat{\mathbf{y}}_i$ and covariance $\mathbf{D}_i$. For each estimate $\mathbf{x}_i$, we draw m particles

$$\mathbf{p}_i^{(j)} \sim \mathcal{N}(\hat{\mathbf{x}}_i, \mathbf{C}_i) \quad j \in \{1, ..., m\} \ . \tag{33}$$

We then transform each individual particle and approximate the transformed particle density as a Gaussian distribution

$$\hat{\mathbf{y}}_i \approx \frac{1}{m}\sum_{j=1}^{m} T(\mathbf{p}_i^{(j)}) \ , \tag{34}$$

$$\mathbf{D}_i \approx \frac{1}{m}\sum_{j=1}^{m} (T(\mathbf{p}_i^{(j)}) - \hat{\mathbf{y}}_i)(T(\mathbf{p}_i^{(j)}) - \hat{\mathbf{y}}_i)^{\mathrm{T}} \ . \tag{35}$$

The fusion is then conducted based on the Kalman filter formulas. This method using Monte Carlo approximation of the density is called MMGW-MC with pseudo-code provided in Algorithm 2.

2) Particle filter: Further improvements at the cost of higher computational power, can be achieved with a particle filter implementation (without a Gaussian assumption). The particles are drawn from the prior and then transformed, while the measurements can remain in the explicit parameter space. For the likelihood calculation, the particles are transformed back. Due to the aforementioned ambiguity in ellipse parameterization, we utilize the wrapped distribution from (8). To make the calculation feasible, we use the 2π periodicity of the orientation to reduce the components to the 4 most likely ones similar to the MWDP, meaning the Gaussian likelihood is calculated for $\hat{\mathbf{x}}_2^{(0)}$ and $\mathbf{x}_1^{(j)}$ with $j \in \{0, 1, 2, 3\}$ and the sum is used to weight the particle

$$\tilde{p}(\hat{\mathbf{x}}_2|\mathbf{y}_1) \approx \sum_{j=0}^{3} p(\hat{\mathbf{x}}_2|\mathbf{x}_1^{(j)}) \ , \tag{36}$$

with $T(\mathbf{x}_1^{(j)}) = \mathbf{y}_1$ and $0 \leq \alpha_{\mathbf{x}_1^{(0)}} < \frac{\pi}{2}$. Note that this definition would still be problematic for circles as this is the only case in which the restriction of the orientation would not make $T(\cdot)$ bijective. Furthermore, as the target is stationary, particle degeneration can pose a problem in high noise scenarios. To avoid this issue, we dismissed resampling and increased the number of particles. This method is named MMGW-PF with pseudo-code provided in Algorithm 3.

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

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

Algorithm 2: The MMGW-MC algorithm relies on the following functions. mvn() provides a sample from a multivariate normal distribution, T() describes the transformation from (16), mean() provides a mean, outer() describes the outer product, and kalman() is a regular Kalman update.

input : transformed particle density $\mathbf{p}$ with weights $\mathbf{w}$ and measurement mean $\hat{\mathbf{x}}_i$ and covariance $\mathbf{C}_i$
output: updated particle weights $\mathbf{w}^+$

for $i \in [0, \mathbf{p}$`.length()`$]$ **do**
  $\mathbf{p}_{inv} \leftarrow$`T_inv(`$\mathbf{p}.i$`)`
  $w_{tmp} \leftarrow 0$
  for $j \in \{0, 1, 2, 3\}$ **do**
    $\mathbf{p}_{inv}^{(j)} \leftarrow$`switch_params(`$\mathbf{p}_{inv}$, j`)`
    $w_{tmp} \leftarrow w_{tmp} +$`mvn_pdf(`$\mathbf{p}_{inv}^{(j)}$, $\hat{\mathbf{x}}_i$, $\mathbf{C}_i$`)`
  end
  $\mathbf{w}^+.i \leftarrow w_{tmp}$
end
`normalize(`$\mathbf{w}^+$`)`
return $\mathbf{w}^+$

Algorithm 3: The MMGW-PF algorithm relies on the following functions. length() gives the length of an array, T_inv() is the inverse of (16), returning the parameterization with $0 \leq \alpha < 0.5\pi$, switch_params() switches according to (5) with the last input as k when given state means and switches covariances accordingly as well, mvn_pdf() is the likelihood for a multivariate normal distribution, and normalize() normalizes the input.

	Euclidean	ESR	GW
Test 1	1.8299	1.8249	1.8237
Test 2	4.9062	3.8092	3.8091

TABLE I: GW error of MMSE estimate using Euclidean distance and MMGW estimate using ESR and GW distance. The first test uses low, the second high orientation noise.

V. Experiments

To support our findings, we provide experiments for a comparison of the MMSE estimate using Euclidean distance and the MMGW estimate as well as its approximation using ESR distance in Section V-A and an evaluation of the fusion methods in Section V-B. The source code for the experiments can be found online[1].

A. MMGW estimate

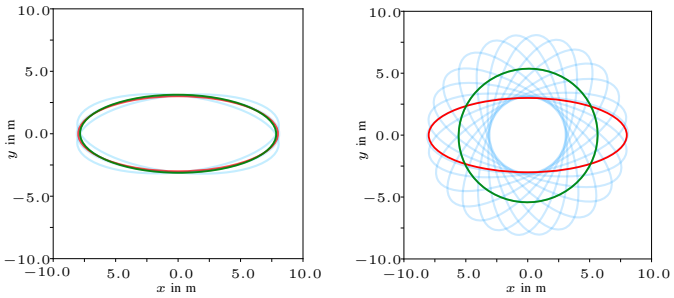
For assessing the quality of the MMGW estimate, we conducted two experiments. For the first, we took an ellipse estimate with mean $\hat{\mathbf{x}} = \begin{bmatrix}0 & 0 & 0 & 8 & 3\end{bmatrix}^{\mathrm{T}}$ and covariance $\mathbf{C} = diag(\begin{bmatrix}0.5 & 0.5 & 0.01\pi & 0.5 & 0.1\end{bmatrix})$, for the second one, we modified the orientation noise to 0.5π. In both cases, we drew 1000 particles to determine the MMGW estimate. The given mean was obviously the estimate using Euclidean distance. For the ESR distance, we transformed each particle using (16) and then calculated their mean. For the GW distance, we utilized the optimization described by [39] with initially equal weights and the mean from the ESR distance as initial guess. We then used the particles to calculate the estimates' quality with respect to the GW distance. The results can be found in Table I.

We find that the ESR and GW optimization do not differ significantly in error, with the GW estimate being slightly better. This coincides with the findings from Appendix A. Another important aspect is the relation to the Euclidean distance, which can be observed in Figure 3. 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.

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. To highlight the difference to this fusion method outside of the problem of ambiguous parameterization, the MWDP method (see Section IV-A) is utilized as well. In addition, as random matrices are a common way to represent ellipses, we added the method described by [21] for comparison, because it takes ellipse estimates and not point clouds as input (here called RM Mean; we treat each ellipse equally by setting the Poisson rate of each one to 1). For the MMGW-MC, we used 1000 particles. As resampling is not used in MMGW-PF, we increased the number of particles to 100000. Increasing the number of particles for the MMGW-MC did not provide significant improvements.

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



(a) Low orientation noise. (b) High orientation noise.

Fig. 3: MMSE estimate based on Euclidean distance (red) and ESR distance (green) for low orientation noise in Figure 3a and high orientation noise in Figure 3b. Sample particles to highlight the orientation uncertainty are drawn in cyan.

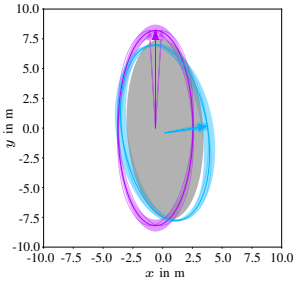


Fig. 4: Exemplary input with ground truth in gray, sensor 1 in magenta, and sensor 2 in cyan. The ellipses' respective orientation is represented by an arrow and the uncertainties are visualized as pale ellipses. Note how the measurements possess different uncertainties for the semi-axis.

The experiments provide the methods with a prior from which the ground truth is sampled in each run. We simulate two sensors with different uncertainties in the semi-axis length and width, e.g., one sensor tracking from the back, the other from the side. In addition, the second sensor uses a different orientation for the ellipse to include the problem of ambiguous parameterization (see Figure 4 for an exemplary input). The sensors provide ellipse estimates (drawn from the ground truth) and their covariances. The fusion methods are taking these estimates alternating between the sensors. We conduct four experiments, two with an elongated prior and two with a round prior and in both cases, there is one with low and one with high orientation noise. For each experiment, we conducted 100 Monte Carlo runs and plotted the convergence of the GW error over 20 measurements (10 alternating from the sensors).

In summary, the prior always has $\mathbf{m} = \begin{bmatrix}0 & 0\end{bmatrix}^{\mathrm{T}}$ and $\alpha = 0$ and for the semi-axes, there are two settings, prior *long* with $l = 8$ and $w = 3$ and prior *round* with $l = w = 5$, both with covariance $\mathbf{C} = diag(\begin{bmatrix}0.5 & 0.5 & 0.5\pi & 0.5 & 0.5\end{bmatrix})$. The sensor noise has two settings. One is sensor *low* with covariance $\mathbf{C}_1 = diag(\begin{bmatrix}0.5 & 0.5 & 0.01\pi & 0.5 & 0.1\end{bmatrix})$ and

	Prior *long*	Prior *round*
Sensor *low*	Test 1	Test 3
Sensor *high*	Test 2	Test 4

TABLE II: Experiment settings.

$\mathbf{C}_2$ the same. Note that before drawing the estimates of sensor 2 from the ground truth, the orientation is shifted by $\frac{\pi}{2}$ and l and w are switched to simulate the problems with ambiguous parameterization, resulting in the noise on the semi-axes being actually different between the sensors (see Figure 4). For sensor *high* the variance of the orientation in both covariances is set to 0.5π.

The experiments' settings can be found in Table II and the results in Figure 5. Note that as the goal is to minimize the GW distance, the Root Mean Square Error (RMSE) is calculated with the GW distance as error.

The results of test 1 in Figure 5a show the problems of the regular fusion with the ambiguous parameterization clearly. The RM mean converges, but it does not consider the different uncertainties in the semi-axes, so it is overall worse than the following approaches. The 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 AMMGW based approaches. This conforms with the findings of Figure 3a of Section V-A.

The results of test 2 in Figure 5b demonstrate the problems of the Euclidean distance in high orientation noise scenarios. The MWDP keeps the shape, but due to the high noise, the orientation can be quite off, resulting in a high error. The AMMGW 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. Second, it can be seen in the comparison between the MMGW-MC and the better MMGW-PF that the Gaussian approximation does not capture the transformed density precise enough.

For the case of round target priors in test 3, the insufficient Gaussian approximation can be seen more clearly in the results in Figure 5c, with the MMGW-MC being worse than the MWDP. The MMGW-PF, however, demonstrates that by keeping the transformed density intact, the AMMGW estimator provides the best results in respect to the GW distance. The results of test 4 in Figure 5d show the same problem of MWDP and advantage of MMGW-PF as test 2.

Given the RM Mean as the state-of-the-art in ellipse fusion, we demonstrated that our method offers improvements of up to 50%, especially with varying noise on the semi-axes.

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 explicit parameters in Section VI-A and with the minimization of the Euclidean distance between the shape matrices in Section VI-B.

A. GW vs Euclidean distance on explicit parameters

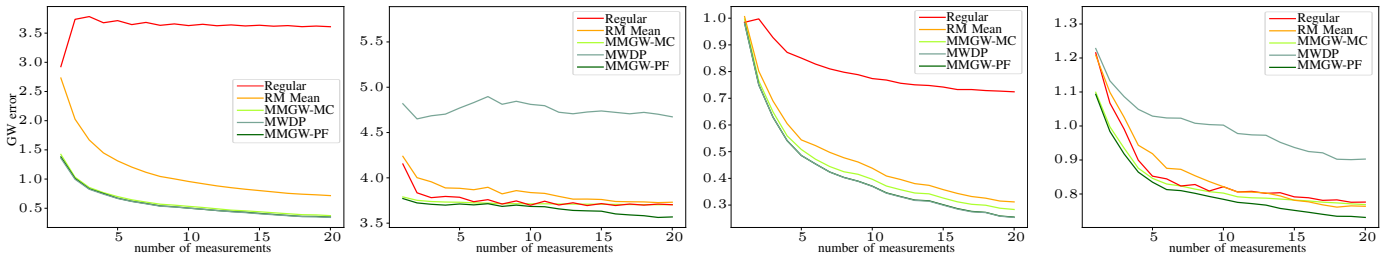
In this article, we demonstrated in simulations that in scenarios with elongated ellipses and high orientation noise, the MWDP provides worse results with respect to the GW distance compared to the AMMGW based methods. We further elaborate this difference by providing the MMSE estimates using Euclidean 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. However, should the angle noise be low and the estimates have similar orientations, the MWDP provides good estimates in respect to the GW distance. This brings us to an important parallel between the MWDP and the AMMGW 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 semi-axis length and width

$$\begin{aligned}&\mathrm{ESR}(\mathbf{m}_\mathbf{z}, \mathbf{Z}, \mathbf{m}_{\mathbf{x}_1}, \mathbf{X}_1)\\ &\quad = || \begin{bmatrix}(\mathbf{m}_\mathbf{z} - \mathbf{m}_{\mathbf{x}_1})^\mathrm{T} & l_\mathbf{z} - l_{\mathbf{x}_1} & w_\mathbf{z} - w_{\mathbf{x}_1}\end{bmatrix}^\mathrm{T} ||_2^2\\ &\quad \forall \alpha_\mathbf{z} \text{ with } (\alpha_\mathbf{z} \bmod \frac{\pi}{2}) = (\alpha_{\mathbf{x}_1} \bmod \frac{\pi}{2}) \ . \end{aligned} \tag{37}$$

Now, if the ellipse orientation has a high certainty, so the corresponding value in the covariance tends to 0, MWDP would choose the k_{opt} for which the orientation difference is 0. In this case, the overlapping semi-axes would be compared, resulting in the same fusion estimate as the AMMGW estimator. As the shape matrices would also commute in this case, the MWDP and also the AMMGW estimation would provide an exact estimate with respect to the GW distance as well. This is further demonstrated by the results of Section V.

B. GW vs Euclidean distance on shape matrix

The particles from an empirical density in ESR space can simply be averaged to compute the AMMGW estimate. Note that [22] 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 explicit shape parameters. This can be demonstrated with a simple example regarding commute matrices. As shown in (37), 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. For



(a) Test 1 (long prior and low orientation noise).

(b) Test 2 (long prior and high orientation noise).

(c) Test 3 (round prior and low orientation noise).

(d) Test 4 (round prior and high orientation noise).

Fig. 5: Tests comparing different ellipse fusion approaches. Refer to Table II for details on the test setups. In Figure 5a and Figure 5c, MWDP and MMGW-PF lie over each other.

example, if the length is Gaussian and two estimates, 2 and 10, are drawn, averaging the square root matrices results in a matrix with length $\frac{2+10}{2} = 6$ and averaging the shape matrices produces a matrix with length $\sqrt{\frac{4+100}{2}} \approx 7$.

VII. Conclusion and Future Work

In this work, we proposed a novel method for the fusion of ellipses 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 a distance measure on ellipse, replacing the Euclidean distance in a MMSE estimator with the GW distance to create a MMGW estimator. Approximating the GW distance with the ESR distance, we derived an AMMGW estimator. In simulations, we demonstrated the robustness of this approximation and compared our methods to state-of-the-art algorithms. We highlighted the advantages of our approach, those being the inclusion of the explicit parameter state's covariance, the dealing with ambiguous parameterizations of ellipses, and more intuitive estimates in scenarios with high angular noise. In summary, the MWDP provides a suitable way of avoiding the ambiguous parameterization of ellipses, but only in scenarios with low orientation noise and (nearly) commute matrices is it a good approximation of the MMGW estimator. Otherwise, the MMGW-PF would be more appropriate.

For future work, we intend to further improve the MMGW-PF by finding an appropriate kernel noise to tackle particle degeneration and increase its efficiency. To also increase accuracy, we seek to find better ways to preserve the transformed density by means of direct multiplication of particle densities [40]. Another topic of interest are alternative fusion methods for the mixture densities in explicit parameter space, like geometric mean densities [23], and Gaussian mixture reductions to preserve more information than the MWDP. As for other shapes, we want to find an appropriate metric for rectangles and test if the MMGW concept also works for them as the parameterization is the same as for ellipses. 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 here as well.

Appendix

A. Comparison of GW and ESR distance

For the comparison, we utilized the experiments from [36] 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 the same (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 exactly between two orientations which would make the estimate and the ground truth commute.

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.

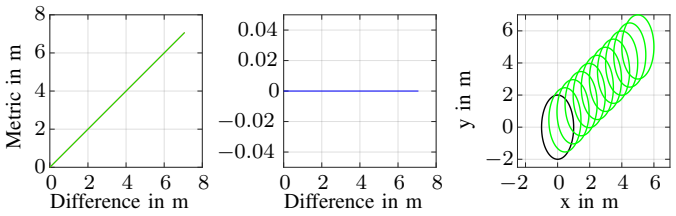
[5] K. Schueler, T. Weiherer, E. Bouzouraa, and U. Hofmann, "360 Degree Multi Sensor Fusion for Static and Dynamic Obstacles," in *IEEE Intelligent Vehicles Symposium*, 2012, pp. 692–697.

[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.

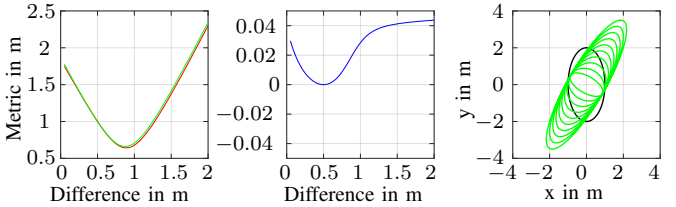
[7] S. Yang and M. Baum, "Tracking the Orientation and Axes Lengths of an Elliptical Extended Object," *IEEE Transactions on Signal Processing*, 2019.

[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.

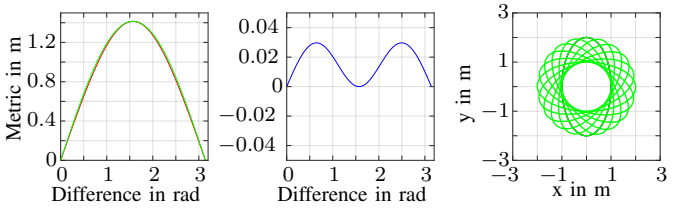
[9] J. Lan and X. R. Li, "Tracking of Extended Object or Target Group Using Random Matrix: New Model and Approach," *IEEE Transactions on Aerospace and Electrical Systems*, vol. 52, no. 6, pp. 2973–2988, Dec. 2016.



(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 (GW in red and ESR in green), metric difference in the middle, and ground truth with every tenth iteration of the estimate on the right (ground truth in black, estimates in green).

[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] 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.

[18] 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.

[19] 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

[20] 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.

[21] 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.

[22] 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.

[23] J. Liu and G. Guo, "A Random Matrix Approach for Extended Target Tracking Using Distributed Measurements," *IEEE Sensors Journal*, 2019.

[24] 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.

[25] 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.

[26] 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.

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

[28] 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.

[29] 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.

[30] 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.

[31] 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.

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

[33] 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.

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

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

[36] 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.

[37] 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.

[38] 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.

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

[40] 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.